(Almost) 200 Years of News-Based Economic Sentiment*

J. H. van Binsbergen[†] S. Bryzgalova[‡] M. Mukhopadhyay[§] V. Sharma[¶]

December 2023

Abstract

Using text from 200 million pages of 13,000 US local newspapers and machine learning methods, we construct a 170-year-long measure of economic sentiment at the country and state levels, that expands existing measures in both the time series (by more than a century) and the cross-section. Our measure predicts GDP (both nationally and locally), consumption, and employment growth, even after controlling for commonly-used predictors, as well as monetary policy decisions. Our measure is distinct from the information in expert forecasts and leads its consensus value. Interestingly, news coverage has become increasingly negative across all states in the past half-century.

Keywords: Business cycle, macroeconomic news, economic sentiment, monetary policy, textual analysis, machine learning, big data, neural networks
JEL codes: G1, G4, E2.

*Any errors or omissions are the responsibility of the authors. For helpful comments, discussions, and suggestions, we thank Pedro Bordalo (discussant), Will Cong, Jingyu He (discussant), Manish Jha (discussant), Zhengyang Jiang, Asger Lunde (discussant), Alan Moreira (discussant), Maximilian Schleritzko (discussant), and Laura Veldkamp. We also thank the seminar and conference participants at London Business School, The Chinese University of Hong Kong, University of Surrey, 2023 Adam Smith Workshop, Macro Finance Society 22nd Workshop, Fifth International Workshop in Financial Econometrics, Georgia State FinTech Conference, Eastern Finance Association 2023, the CFA Society, and the Jacobs Levy Equity Management Center.

†Wharton and NBER; julesv@wharton.upenn.edu

‡London Business School; sbryzgalova@london.edu

§London Business School; mmukhopadhyay@london.edu

¶Kelley School of Business, Indiana University; vs12@iu.edu

Introduction

Although measures of economic sentiment have been widely studied,[1] their applications have been somewhat limited because of a short time series span and a lack of granularity. This paper uses machine learning methods to construct a new measure of economic sentiment from US local newspaper pages published in the past 170 years. We show that economic sentiment is important for understanding the business cycle at both local and national levels. We find that economic sentiment a) predicts economic fundamentals at the country and state level, b) leads consensus GDP forecasts, and c) exhibits large cross-sectional variation across states, with the nationwide component driving only 35% of the state variation.

To this end, we use a historical collection of 170 years of digitized newspapers, which includes the text of 200 million newspaper pages from 13,000 local newspapers. Our corpus includes approximately one billion newspaper articles, a major increment over the Wall Street Journal corpus, a much-used source of text data in economics and finance that contains about 1 million articles. In fact, our data is about 95 times larger than the total number of English-language Wikipedia entries combined. By leveraging the collection of local newspapers, we can measure sentiment at a higher degree of granularity, such as the county or state level.

To measure text-based economic sentiment, we customize the machine learning technique pioneered by Singla and Mukhopadhyay (2022). We create a fully automated topic-specific dictionary by leveraging Word2vec, a neural-network-based algorithm, that allows us to capture the meaning of words and phrases from the context in which they are used (Mikolov, Sutskever, Chen, Corrado, and Dean (2013) and Mikolov, Chen, Corrado, and Dean (2013)). Instead of using a binary positive/negative connotation, we produce a continuous measure of sentiment for each word and phrase in the dictionary, following Hamilton, Clark, Leskovec, and Jurafsky (2016). As a result, our method automatically overcomes many common challenges faced by simple word-count techniques, such as detecting negation and measuring word/phrase intensity.

Our measure exhibits large variation in both time series and cross-section. Furthermore, our approach detects a notable downward trend in the overall news sentiment since the 1970s. We show that the coverage of economic news does not drive this trend. Instead, it relates to the general coverage of both economic and non-economic news. Separating this overall sentiment from the economic news allows us to construct a measure of economic sentiment that is not affected by the changes in how newspapers have changed their overall tone.[2]

[1]See Lopez-Salido, Stein, and Zakrajek (2017), Bordalo, Gennaioli, and Shleifer (2018), and Bordalo, Gennaioli, Ma, and Shleifer (2020).

[2]Following the coverage of the U.S. elections of 1992, many previous studies focused on the apparent negativity bias in media, see Damstra and Boukes (2021), Fogarty (2005), Soroka (2006), and Soroka (2014)).

Before proceeding, one important qualification is in order. Consistent with recent NLP advances, we measure sentiment as the (broadly defined) positivity/negativity of economic news coverage in the mass media. As a result, our measure could reflect (a) pure factual coverage of economic news, (b) irrational exuberance or pessimism, or (c) some combination of (a) and (b). Putting this in the language of recent asset pricing papers, our measure can reflect both variations in cash flow expectations and expected returns. In that sense, the word "sentiment" in our paper relates more to the general meaning of the English word rather than the more specific interpretation it has recently been receiving in behavioral economics and finance.

We embed our sentiment measure in predictive regressions alongside various information sets to see whether they span it. Our tests do not establish causality. Instead, they provide diagnostics on the informational content of the news-based economic sentiment and its empirical relationship with the integral aspects of the economy: Business cycle fluctuations in fundamentals and agents expectations.

Our measure predicts future GDP per capita growth even after controlling for current GDP growth: a one standard deviation increase in sentiment corresponds to 2% additional GDP growth over the next year in the sample period from 1850 to 2017. Furthermore, over the sample period 1947Q1 - 2019Q4, when we also observe the slope of the yield curve (a standard recession predictor), we show that a one standard deviation increase in sentiment leads to 0.29% of additional GDP growth one quarter ahead (corresponding to 1.1% annualized growth). In addition, this predictability remains even after controlling for the consensus forecast in the Survey of Professional Forecasters, implying that our measure captures important information not spanned by such consensus forecasts. Notably, as we show, our measure actually leads the survey in that it predicts the next quarter's consensus value. Overall, our results indicate the importance of sentiment in understanding the business cycle.

The flexibility of the method allows us further to explore the sources of predictability of our measure. Newspaper articles cover current events but also opinions and views about the future. To differentiate between these two sources of news, we create topic-specific dictionaries of words related to the present and future and measure the intensity of this language on each newspaper page to classify them as related to either current or future events. We use this classification to create current and future economic sentiment. We find that all our predictability results are driven by the forward-looking, future component of the sentiment, while the measure created from current events has little to no predictive power for macroeconomic variables. This suggests that our measure reflects forward-looking news about economic fundamentals.

Given that our measure predicts GDP growth, one could wonder which inputs to GDP are predicted by our economic sentiment measure. We show that our measure operates mainly through the labor channel rather than the capital one. It predicts employment, consumption, and services but neither investment nor industrial production. Similarly, we show that our measure of economic sentiment relates to the real economy and does not predict inflation.

Next, we evaluate the extent to which news-based economic sentiment is reflected in monetary policy decisions. To this end, we quantify the importance of sentiment in explaining the changes in the federal funds rate relative to what the forward-looking Taylor rule specification proposed by Romer and Romer (2004) would imply. We find that news-based economic sentiment has an important connection to the key policy rate: A one-standard-deviation decrease in sentiment over the past two quarters corresponds to a 25 basis point (5 basis point) decrease in the policy rate during recessions (expansions). Furthermore, we find that sentiment has significant predictive power for the federal funds rate during recessions even after controlling for its predictive power on the ex-ante (measured by Tealbook projections) and ex-post (realized) GDP growth.

Our methodology also allows us to measure sentiment at a more granular level, which we implement at the state level, revealing significant cross-sectional heterogeneity across states. The common component across states drives only approximately 35% of the variation in state-level sentiment. State-level sentiment predicts state GDP growth even after controlling for national sentiment and national and state GDP growth. Furthermore, using the dispersion in sentiment across states as a measure of heterogeneity, we find that higher dispersion predicts lower national future GDP growth after controlling for national sentiment and current national GDP growth.

Our paper relates to Shapiro, Sudhof, and Wilson (2022), who study a national time series of economic sentiment derived from economic and financial newspaper articles between 1980 and 2015. This sample period broadly coincides with the availability of the Michigan Survey of Consumer Sentiment. Using a methodology different from ours, the authors find that over this period, their sentiment measure has predictive power for consumer sentiment. These sentiment shocks produce positive impulse responses for consumption, output, interest rates, and inflation. We find little predictive power of our sentiment measure for industrial production and inflation but do find significant predictive power for employment, consumption, and services.

Another recent related paper in this area is Macaulay and Song (2022), which studies the relationship between media portrayals of inflation and consumer sentiment over the sample period 2014-2021. These authors also use natural language processing to contrast inflation news related to financial variables and inflation related to the pricing of real (rather than financial) variables. Linking inflation news to social network data from Twitter, they find that exposure to articles emphasizing the connection between inflation and the real economy significantly reduces sentiment, particularly in periods of high inflation.

Our paper also relates to the literature that studies expectational distortions of fundamentals and/or financials and evaluates their importance for real outcomes. For example, Maxted (2022) incorporates diagnostic expectations into a general equilibrium macroeconomic model with a financial intermediary sector. In his model, the interaction of sentiment with financial frictions generates a short-run amplification effect followed by a long-run reversal effect, which he refers to as the feedback from behavioral frictions to financial frictions.

A related contribution in this area is Krishnamurthy and Li (2021), who develop a model of financial crises with both a financial amplification mechanism (via frictional intermediation) and a role for sentiment through time-varying beliefs about an illiquidity state. They argue that the sentiment channel is particularly important for the *pre*-crisis behavior of asset values.

Finally, a third contribution in this space is that of Hirshleifer, Li, and Yu (2015) who introduce extrapolative bias into a standard production-based model with recursive preferences and reconcile salient stylized facts about business cycles (e.g., low consumption volatility and high investment volatility relative to output) and financial markets (high equity premium, volatile stock returns, and a low and smooth risk-free rate) with plausible levels of risk aversion and intertemporal elasticity of substitution. Furthermore, their model captures return predictability based on dividend yield, Q, and investment. The key driver of their model is that the extrapolative bias increases the variation in the wealth-consumption ratio, which is heavily priced under recursive preferences and thus helps explain the financial market effects.

These contributions make up the important context related to the findings in our paper. There are two channels through which sentiment could manifest: An expected returns channel and/or an (expected) cash-flow channel. The former operates through financial markets in which higher (lower) firm valuations are accompanied by lower (higher) expected returns, and thus lower (higher) returns going forward. This in turn affects firms' hiring and investment decisions today, caused by persistent changes in the expected returns, not only changes induced by short-term sentiment. On the contrary, the second channel works only through the cash flows and their expectations and thus does not require a change in the expected returns. Higher (lower) sentiment about future economic activity spurs (dampens) hiring and, potentially, investment in the present, which ex-post can rationalize the current improved sentiment.

In contrast to the abovementioned contributions, our economic sentiment measure does not have much predictive power for variables such as stock returns. Instead, economic sentiment about the future, as reported by newspaper articles and captured by our measure, has important predictive power for future fundamentals such as GDP and labor. As such, our paper does not provide direct evidence in favor of behavioral biases that stimulate real economic activity through inflated asset valuations (lower expected returns) and thereby contribute to boom-bust cycles. That said, our paper does not rule out the notion that sentiment can become a self-fulfilling prophecy (rational sentiment) operating through the cash flow channel: Fluctuations in expectations lead to cycles in hiring, which in turn lead to fluctuations in GDP growth. Given that we find little effect on firm investment, our findings are most consistent with short-term sentiment fluctuations that spur short-term hiring.

Related Literature

News-based measures of sentiment have been used in various applications in economics and finance, most notably its relationship with asset returns (for an excellent overview, see Baker and Wurgler (2007)). Availability of text data and methods from computational linguistics opened a new avenue of research and allowed to analyze new patterns of investor behavior and return predictability. Antweiler and Frank (2004) analyze the tone of messages on Internet boards on stock returns and volatility. Bollen, Mao, and Zeng (2011) and Agrawal, Azar, Lo, and Singh (2018) use Twitter data to predict stock market returns, while Heston and Sinha (2017) rely on a proprietary model of news sentiment developed by Thompson Reuters to establish short-term predictability of the markets. Da, Engelberg, and Gao (2014) use Internet search queries to create a measure of investor sentiment and establish its impact on return reversal and investor flows. Our data is unique in both the length of the time series and the availability of the cross-sectional coverage. We focus on the relationship of the news-based economic sentiment with the business cycle and economic fundamentals.

There is growing empirical evidence regarding the role of economic news, expectations, and sentiment over the business cycle (Greenwood and Hanson (2013), Lopez-Salido, Stein, and Zakrajek (2017), Mian, Sufi, and Verner (2017), Bordalo, Gennaioli, and Shleifer (2018), Bordalo, Gennaioli, Ma, and Shleifer (2020)). To understand the role of sentiment and the role of news coverage, it is essential to measure them first. Most of the literature uses surveys to measure economic sentiment. However, surveys are available only for short periods and lack cross-sectional variation. For example, the Michigan Survey of Consumer Sentiment, the oldest and most prominent survey, is available only from 1966 and has no state-level variation. At the same time, there is a growing strand of literature that studies the content and structure of economic news and shows that text data is a rich and powerful source of information (Gentzkow and Shapiro (2006), Gentzkow and Shapiro (2008), Gentzkow and Shapiro (2010)), Bybee, Kelly, Manela, and Xiu (2021), and Bybee, Kelly, Manela, and Xiu (2020)). Our paper combines these strands of literature to measure economic sentiment from text data and investigate its influence on business cycles.

With the advancement in natural language processing techniques, text data has been leveraged extensively in the last few years to measure various hard-to-quantify aspects of economics and finance. For example, Baker, Bloom, and Davis (2016) use newspaper articles to measure economy-wide political risk. Hassan, Hollander, Van Lent, and Tahoun (2019) further extend the literature on political risk by measuring it at a firm level using firm-specific text from conference calls. Similarly, using firms' annual reports (10-Ks), other studies have created measures that have been instrumental in better understanding various essential phenomena in economics and finance such as product differentiation, competition, and offshoring (Hoberg and Phillips (2010), Hoberg and Phillips (2016), and Hoberg and Moon (2017)).[3] We add to this growing body of literature by creating the first measure of economic sentiment over a long period and at a much higher geographic granularity.

While the higher granularity and longer historical coverage of text data allow for constructing a more useful measure, the associated increase in granularity and time span presents its own challenges. Natural language processing techniques require large corpora of text data, usually at the level at which the measure is being constructed. One contribution of our paper is to show how text data can be used to create new state, county, or city measures. Another challenge arises due to the changes in the context and meaning of the words over time. A proper historical text analysis requires a dictionary that captures the historical meaning and context of the relevant words. In this paper, we apply a methodology that enables the construction of an automated dictionary that addresses this issue. We further discuss our methodology in the context of the broader literature on text analysis in Section II.3.

The remainder of the paper is organized as follows. Section I describes the coverage and granularity of the newspaper text data we use to construct the measures of economic and non-economic sentiment. Section II describes the various components of our methodology. Section III presents the time series and cross-sectional dynamics of the national and state-level economic and non-economic sentiment measures. Section IV highlights the relationship between national and state economic sentiment, economic activity, and the business cycle. Section V explores various aspects of the interaction between economic sentiment and monetary policy. The final section concludes.

[3]Also see Tetlock (2007) and Tetlock, Saar-Tsechansky, and Macskassy (2008) for how media coverage can affect firms' stock prices.

I Data

One of the key strengths of the paper is the size and granularity of the text we use to construct the measures of economic and non-economic sentiment. We use a historical collection of 170 years of digitized newspapers, first introduced in Singla and Mukhopadhyay (2022). The data includes the full text of 200 million newspaper pages of 13,000 local newspapers, corresponding to approximately 1 billion newspaper articles.

Table A1 in the Appendix, Panel A, shows the times series of the national newspaper coverage characteristics, conditional on the pages presenting information about economic news. The median number of newspapers in our sample in a given year is 1,365, with the coverage varying over time. The 10th percentile corresponds to 339 outlets, while the 90th is 1,722. As the newspaper coverage varies over time, we adopt a sampling approach to ensure that all time periods are adequately represented in the data.[4]

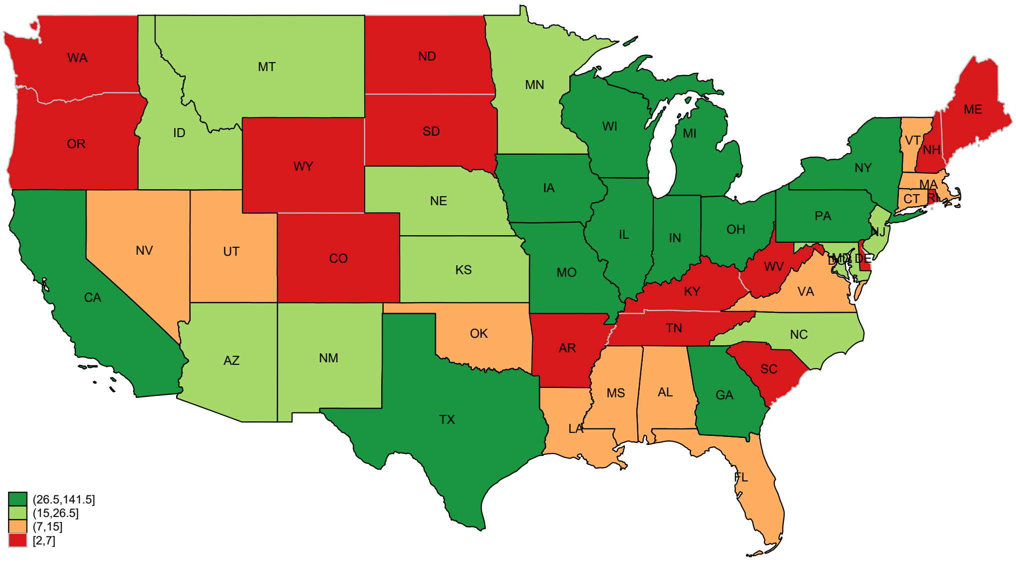
Another crucial aspect of our data is that it covers local newspapers, and, as a result, it provides an ideal setup to measure sentiment at a granular level (for example, city, county, or state). Our corpus includes local newspapers from all 50 US states. These articles were published across various cities within a given state, allowing us to measure county- or city-level sentiment. Figure 1 shows the median number of newspapers for the U.S. states and territories (averaged across time). Table A1 in the Appendix, Panel B, further provides summary statistics for the distribution of newspapers across years for all U.S states and territories. Within the entire sample of 170 years, we observe data for 47 states for at least 100 years. As expected, we have fewer newspapers from smaller states such as New Hampshire and Rhode Island. However, even for these states, we have an average of three newspapers per year, which again highlights the granularity level of this large data sample.

We rely on the University of Michigan Consumer Sentiment survey to validate our measure. We use information from the Federal Reserve Economic Data (FRED) database for additional macroeconomic data. In particular, we use real GDP per capita (A939RX0Q048SBEA), the consumer price index for all urban consumers (CPIAUCSL), total non-farm payroll employment (PAYEMS), gross private domestic investment (GPDICTPI), industrial production (INDPRO), and real personal consumption expenditures (PCECC96). To understand how sentiment influences different sub-components of consumption, we use real personal consumption expenditures per capita for non-durable goods (A796RX0Q048SBEA), durable goods (A795RX0Q048SBEA), and services (A797RX0Q048SBEA).

In addition, we leverage the Maddison Project Database for historical GDP data (at an

[4]See Section II for details.

Figure 1: Newspaper Coverage Across States



The figure shows the median number of available newspapers per year for the U.S. states and territories.

annual frequency) and the Survey of Professional Forecasters for consensus forecasts of macroeconomic variables. We use real GDP in chained dollars (SAGDP9) from the US Bureau of Economic Analysis (BEA) for state-level analysis.

To explore the interaction between economic sentiment and monetary policy, we leverage the extended Romer and Romer (2004) monetary policy shock series provided by Wieland and Yang (2020).[5]

II Methodology

Sentiment analysis in economics and finance typically follows three broad steps. First, creating a dictionary of relevant words for a given topic, that is, economy-related terms and phrases in our setup. Second, identifying words that convey a positive or negative sentiment, for example, *"growth"* versus *"recession"*. Third, using the dictionary to measure the intensity of positive and negative sentiment in a particular document, for example, an article or a newspaper page. We follow these broad steps and rely on the various machine-learning techniques at each stage.

[5]The sample size for the extended series is from Jan 1969 to Dec 2007. See: Updated Romer-Romer Monetary Policy Shocks.

II.1 Word Embeddings

To construct sentiment measures, we rely on a natural language processing technique called *word embeddings*, which uses the co-occurrence of words to produce vector representations of those terms and phrases (see Mikolov, Sutskever, Chen, Corrado, and Dean (2013), Mikolov, Chen, Corrado, and Dean (2013), and Pennington, Socher, and Manning (2014)). This procedure captures semantic information about every term or phrase based on the context in which the latter is used. In essence, the algorithm we use (Word2vec) is based on the fundamental principle of linguistics that "you shall know a word by the company it keeps" (Firth (1957)). The algorithm's objective function ensures that similar words have similar vector representations, which implies similar meanings. Note that Word2vec applies to phrases and standalone words and is known to outperform existing alternatives (Mikolov, Chen, Corrado, and Dean (2013)). Once the algorithm has produced these vector representations, we can use distances between them to measure their similarities (cosine similarity). Because the algorithm looks at the context in which a word/phrase is used, not the physical distance between words in a given sentence, it can identify two similar terms that do not appear together as long as their neighboring context words and phrases are similar enough.

Compared to word counts, the traditional approach in economics and finance that treats words as distinct objects, word embedding allows us to infer relationships between words/phrases by directly studying the context in which they are used. Unsurprisingly, although still infrequently used, word embedding techniques have recently come to the forefront of NLP-based applications in the economics and finance literature (for recent examples, see Cao, Kim, Wang, and Xiao (2020) and Ash, Chen, and Ornaghi (2021)).

II.2 Automated Dictionary and Sentiment Scores: Implementation

In the first step of our approach, we create a fully automated, topic-specific dictionary of the words/phrases related to the economy using Word2vec. To train our word vectors, we use the skip-gram implementation of Word2vec, which has been shown to be the most efficient variation of the algorithm (see Mikolov, Sutskever, Chen, Corrado, and Dean (2013) and Mikolov, Chen, Corrado, and Dean (2013)). Two key hyperparameters are crucial for training a Word2vec model: The dimension of the vector embeddings and the window length to determine the neighborhood of each word. Previous research has shown that a vector dimension size of 300 is sufficient for the efficient capture of word meanings in downstream NLP tasks (see Mikolov, Sutskever, Chen, Corrado, and Dean (2013), Mikolov, Chen, Corrado, and Dean (2013), and Pennington, Socher, and Manning (2014)). Therefore, we follow the literature and use a vector dimension of 300. In line with the model specification used in applications of Word2vec in the economics literature (Ash, Chen, and Ornaghi (2021)), we rely on the standard context window of length 10.

We select 1 billion words at random for each of the five years in our sample from 1850 to 2020 to ensure that the corpus is balanced across history and that each of the time periods has consistent representation. Using Word2vec-trained vectors, we produce a dictionary of words and phrases related to the economy. In particular, the word *"economy"* turns out to be most closely associated with the following terms and phrases: *"Economic growth"*, *"inflation*, *"economic*, *"recession"*, *"economic recovery"*, *"consumer spending"*, and others, which also serves as initial validation of our method. The entire automatic dictionary of words and phrases is quite large and diverse, with the top 1000 words and phrases displayed as the word cloud in Figure 2.

Naturally, having only a stand-alone dictionary of words related to the economy is not enough to measure economic sentiment. To reflect the latter, each word in our dictionary must be classified as positive or negative. Therefore, we rely on the dictionary generation literature using word vectors and, specifically, *Sentprop*, which produces a sentiment dictionary of positive and negative words, starting from a few initial seed words (Hamilton, Clark, Leskovec, and Jurafsky (2016)). Sentprop is a label propagation algorithm that classifies all the words in the dictionary, beginning with a few initial seeds. Another benefit of the algorithm lies in its ability to produce a continuous polarity score for each word in the corpus, which signifies its association with the positive and negative dimensions. As a result, it can reflect not only whether the word or phrase has a positive or negative connotation but also the latter's strength. This method achieves state-of-the-art performance in measuring sentiment precisely in economics and finance contexts (see Hamilton, Clark, Leskovec, and Jurafsky (2016)).

We rely on the following seed words that reflect positive sentiment: *"Expansion"*, *"boom"*, *"growth"*, *"profit"*, *"optimistic"*, *"optimism"*, *"opportunity"*, *"success"*, *"successful"*, *"profitable"*, *"prosperity"*, *"profitability*, and *"bullish"*. Similarly, for the negative words, we start with the following seeds: *"Recession"*, *"bankrupt"*, *"shrinking"*, *"unemployment"*, *"loss"*, *"bankruptcy"*, *"cutback"*, *"layoff"*, *"redundancy"*, *"pessimism"*, *"contraction"*, *"unsuccessful"*, *"failure"*, *"insolvent"*, *"insolvency"*, and *"bearish"*. Beginning with these seed words, Sentprop classifies all the words in our sentiment dictionary and assigns a continuous score to them. Figure 3 illustrates the polarity of the most positive and negative terms; see Panels (a) and (b) correspondingly. Based on their polarity score, it reflects the words and phrases that belong to the top/bottom 0.1%. In particular, the top 10 most positive words/phrases are *"successful"*, *"success"*, *"profit"*, *"opportunity"*, *"profitable"*, *"cess"*, *"proven"*, *"expansion"*, *"interview"*, and

Figure 2: Word Cloud of Words related to Economy

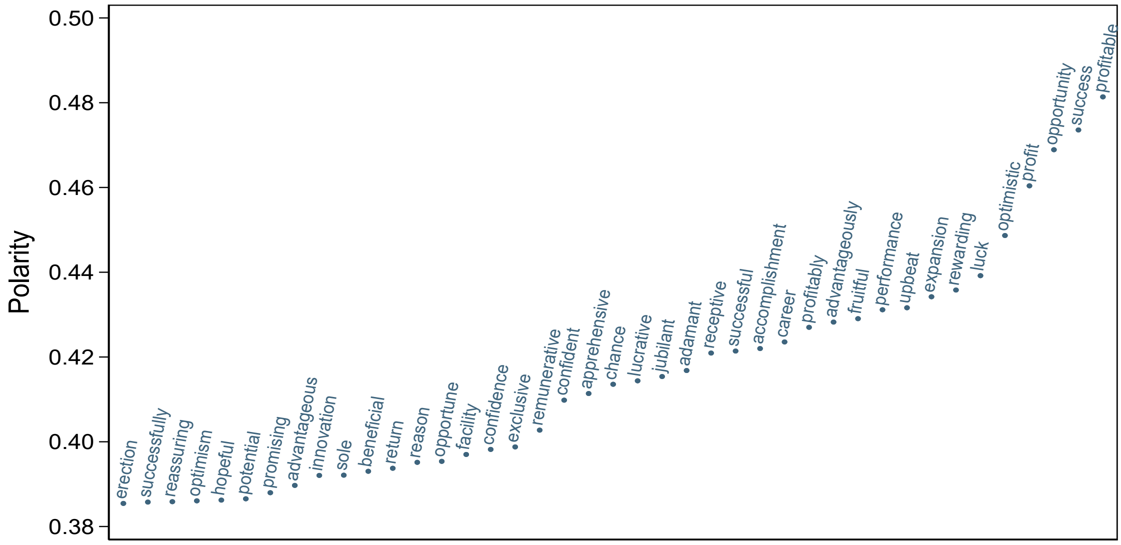
The figure shows a word cloud of the 1000 most similar words to the word economy calculated from the Word2vec trained vectors.

"rewarding". At the same time, the top 10 negative words/phrases are *"failure"*, *"unsuccessful"*, *"insolvent"*, *"earnest"*, *"redundancy'*, *"bankruptcy"*, *"bankrupt"*, *"insolvent debtor"*, and *"contraction"*. Figure A1 in the Appendix further illustrates the nature of negative and positive words in the resulting dictionary via the word cloud.

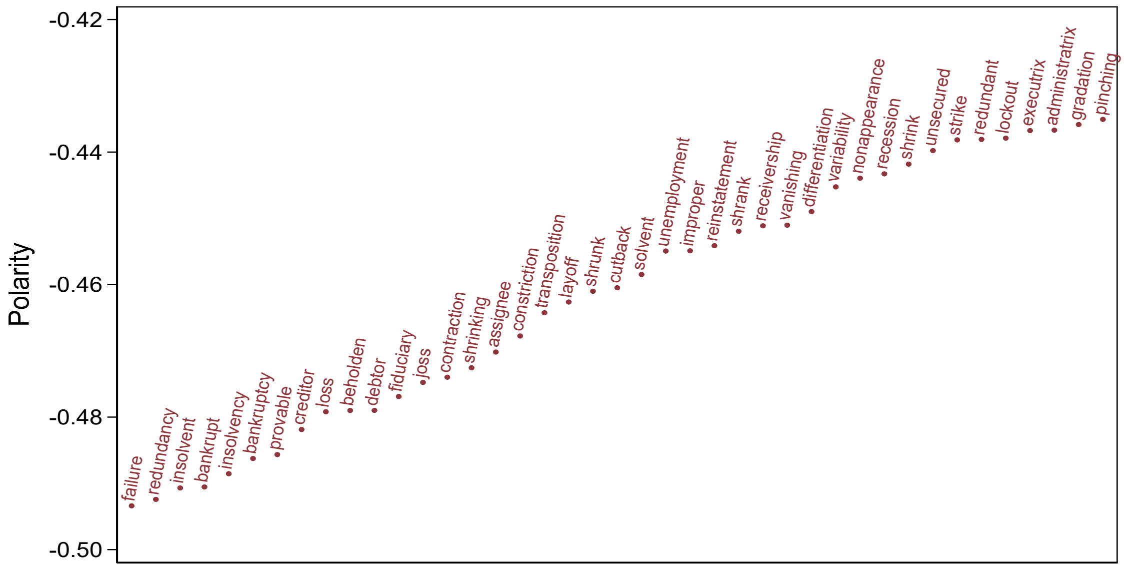
To construct the final dictionary, we retain words and phrases with a polarity of three standard deviations away from the mean. Next, we intersect them with the 1,000 most associated items for each target word and phrase and include only these in our dictionary. This step ensures that the final dictionary contains only words with high polarity related to the economy. Furthermore, we add all the positive word vectors and subtract the negative ones multiplied by their absolute polarity scores to create a vector score for the entire dictionary.

The final step of our approach is producing a vector representation of the document, following Arora, Liang, and Ma (2017). We compute the cosine similarity between our economic dictionary and the document vector to measure the sentiment from any newspaper page. To do this, we use the *Smooth Inverse-Frequency* (SIF) method, which outperforms more sophisticated supervised methods in producing document vectors (Arora, Liang, and Ma (2017)). This method removes the first Principal Component (PC) of the weighted average of the word vectors. We train an SIF model using the same randomly selected sentences as our Word2vec model. Finally, we subtract the first Principal Component calculated from the above models from the document vector (each newspaper page) to obtain the overall page embedding.

Figure 3: Sentiment Polarity



(a) 0.1% of the most positive words



(b) 0.1% of the most negative words

The figure plots the polarity of the most positive and most negative words in panel (a) and panel (b), respectively. The most positive (negative) words are defined as words in the top (bottom) 0.1 percentile in terms of their polarity score.

Once we have the entire page embedding, we measure the similarity between the vector representation of each newspaper page and our economic dictionary (weighted by the word sentiment score) and aggregate our measure to the city, county, state, and country levels, weighting the measures for each page by the number of words. Finally, we standardize and seasonally adjust the output.

The language used in newspapers to describe the economy might change over time. Thus, the seed words we provide to Word2vec may not be able to capture the economic content in newspapers across different periods correctly. To intuitively check if our model can accurately identify the words related to the economy, we plot the word cloud of three major recessions across our sample. Figure A2 in the Appendix plots the word cloud of newspaper coverage of the recession of 1866 in Panel (a), the Great Depression in Panel (b), and the Global Financial Crisis in Panel (c). We find that the news coverage uniformly captured the words identified by our methodology to describe economic sentiment during all three recessions.

II.3 Advantages of the Methodology and Relation to the Literature

Textual analysis is becoming an increasingly important tool for economics and finance in measuring the informational content of unstructured news articles, regulatory filings, central bank reports and statements, and narratives of corporate executives. Although the literature relying on textual analysis is vast (and fast-growing), this section aims to highlight a few of the most closely related contributions and methodological advantages of our approach to measuring sentiment compared to those existing in the literature.

Numerous studies examine the fraction of negative and positive sentiment words in news articles, 10-K reports, and conference calls to understand firms' disclosure choices, as well as the nature of information that investors process and the speed with which it is incorporated into asset prices (see Tetlock (2007), Tetlock, Saar-Tsechansky, and Macskassy (2008), Loughran and McDonald (2016), Hanley and Hoberg (2019) and Cohen, Malloy, and Nguyen (2020)). Other studies develop text-based measures to quantify various qualitative constructs. Gentzkow and Shapiro (2010) and Martin and Yurukoglu (2017) use natural language processing techniques to identify media slant. Baker, Bloom, and Davis (2016) create economic policy uncertainty indexes using historical news articles. Hoberg and Phillips (2010) and Hoberg and Phillips (2016) use 10K forms to identify product-market competitors based on the textual similarity between product descriptions. Hassan, Hollander, Van Lent, and Tahoun (2019) create a firm-level measure of political risk using textual analysis of quarterly earnings conference calls. Li, Mai, Shen, and Yan (2021) use Word2vec trained on earnings call transcripts to measure corporate culture. Hansen, McMahon, and Prat (2017) use text analysis of transcripts from the Federal Reserve's monetary policy committee meetings to study how transparency affects their communication. Cong, Liang, and Zhang (2019) combine word embedding and topic modeling (LDA) to generate textual factors for prediction and inference.[6] Bellstam, Bhagat, and Cookson (2021) use LDA to measure corporate innovation. Arteaga-Garavito, Croce, Farroni, and Wolfskeil (2022) rely on the topic analysis of Twitter announcements to build a high-frequency measure of the market's exposure to shocks about the global contagion risk during the time of COVID19.

The traditional approach to measuring sentiment in the economics and finance literature relies on manually created sentiment dictionaries and simple counts of these words and phrases in a text document. One of the key assumptions in this approach is that the order and context of the word usage are unimportant (for a discussion, see Loughran and McDonald (2016)). As a result, this method may omit more complex patterns of language by ignoring negation, context, sequence of words, and relationships between them (see Chen and Manning (2014) and Cao, Kim, Wang, and Xiao (2020)). To partially address this problem, some studies attempt to identify context-specific information by using bi-grams (e.g., Hassan, Hollander, Van Lent, and Tahoun (2019)) or a set of words or phrases (e.g., Baker, Bloom, and Davis (2016)). However, such rule-based approaches still do not fully account for the complex patterns observed in human natural language (see Cao, Kim, Wang, and Xiao (2020)).

Recognizing many limitations of the traditional word count approach, we build up on the recent literature that moved away from word counts and relies on word vectors instead. Ash, Chen, and Ornaghi (2021) use word vectors to measure gender bias in US judicial opinions. Similarly, Cao, Kim, Wang, and Xiao (2020) use word embeddings and a natural language parser to measure the tone of financial narratives in earnings conference calls. We contribute to this growing body of literature using the word vector approach to construct our measures. Our methodology has several advantages. First, it allows for an almost fully automated construction of topic-specific dictionaries. Because the word vectors we train are generally applicable, to do this, we only need to change the list of input words to create a new dictionary of related terms. These dictionaries can be readily constructed for other contexts and topics. As a result, it avoids the problem of subjectivity and the need for manual intervention in the creation of the dictionary. Second, our method reflects the context in which words/phrases are used and automatically addresses negation, a common challenge for word counts (that is, it automatically captures the difference between *"good"* and *"not good"*). Finally, our approach generates an automatic and continuous measure of word sentiment scores. It can, therefore, distinguish many shades of the underlying language intensifiers (*"good"*, *"very good"*, or *"excellent"*). Together, these aspects yield a new flexible and robust method for measuring the topic-specific sentiment of the textual data.

[6]For the general introduction of topic modeling in text analysis, see also Blei, Ng, and Jordan (2003).

II.4 Our methodology versus BERT and ChatGPT

Innovations in the field of natural language processing have been very rapid with the development of language models such as BERT and ChatGPT in the last few years (see Chen, Kelly, and Xiu (2022), Hiew, Huang, Mou, Li, Wu, and Xu (2022), and Lopez-Lira and Tang (2023)). However, these techniques are not yet feasible in our context due to the size of the text corpus and the focus on historical data. For example, while BERT is a more recent technology, it requires significantly more computational resources to process a text corpus as large as ours. A back-of-the-envelope estimate suggests it will take around \$2,500 to 50,000\$ to train the smallest model on Wikipedia data (see Sharir, Peleg, and Shoham (2020)). As our corpus is 95 times larger than Wikipedia, scaling up the estimate linearly would give us a range of around \$237,500 to \$4,750,000 to use the BERT on our corpus. ChatGPT is the latest and very powerful generative AI tool in text analysis. Again, it is not yet possible to use ChatGPT for a corpus as big as ours.

Furthermore, compared to BERT and ChatGPT, which are essentially black-box methodologies, our approach allows us to validate and examine the output of intermediate steps (as shown in the word cloud and word polarity figures). Thus, our methodology, which uses Word2vec and Sentprop, is well suited for a large-scale empirical application as ours, requiring custom-built word embeddings from scratch for a specific task in a parsimonious and transparent manner.

III A New Measure of Economic Sentiment

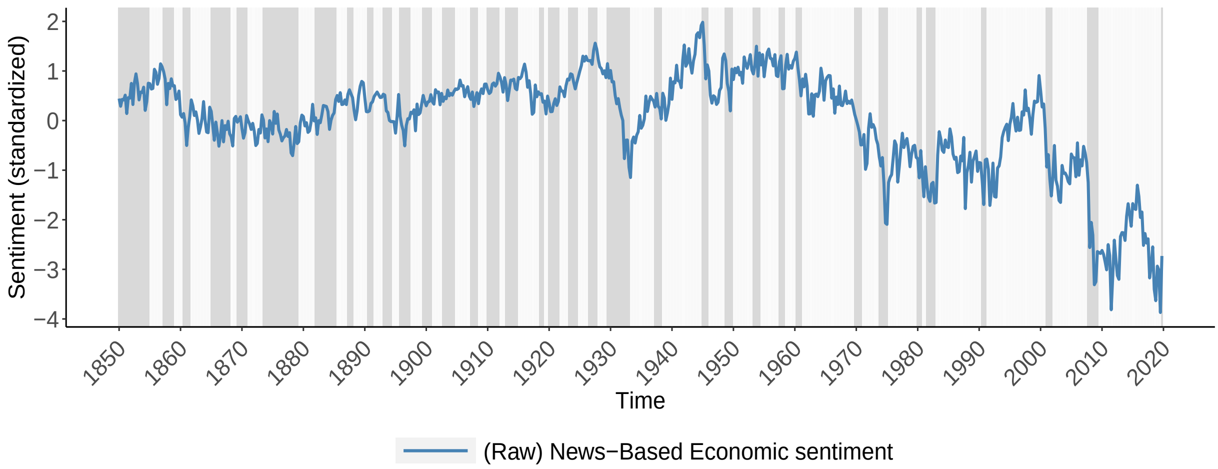
In this section we present a new text-based measure of economic sentiment. We first present a national measure in Section III and then introduce a state-level measure in Section III.2.

III.1 National economic sentiment

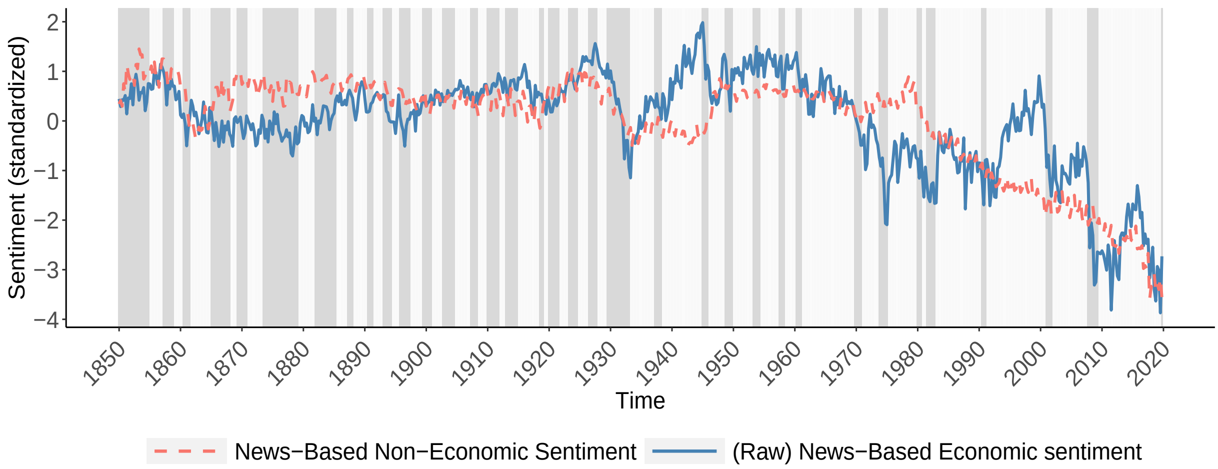
Having laid out the methodology in the previous section, we now present our main economic sentiment measures. First, Figure 4a displays the dynamics of the national news-based economic sentiment over the 170 years under analysis. Panel (a) reports the gross level of sentiment, measured by the content of local newspapers. Economic sentiment exhibits substantial variation over the business cycle, consistent with many prominent economic crises throughout U.S. history, such as the Great Depression, the Great Recession, and the recession of 1973–1975.

One of the important features of the text-based sentiment plotted in Figure 4a is the downward trend that begins in the 1970s, roughly coinciding with the secular stagnation period studied in the macroeconomic literature. Although there is extensive variation around this

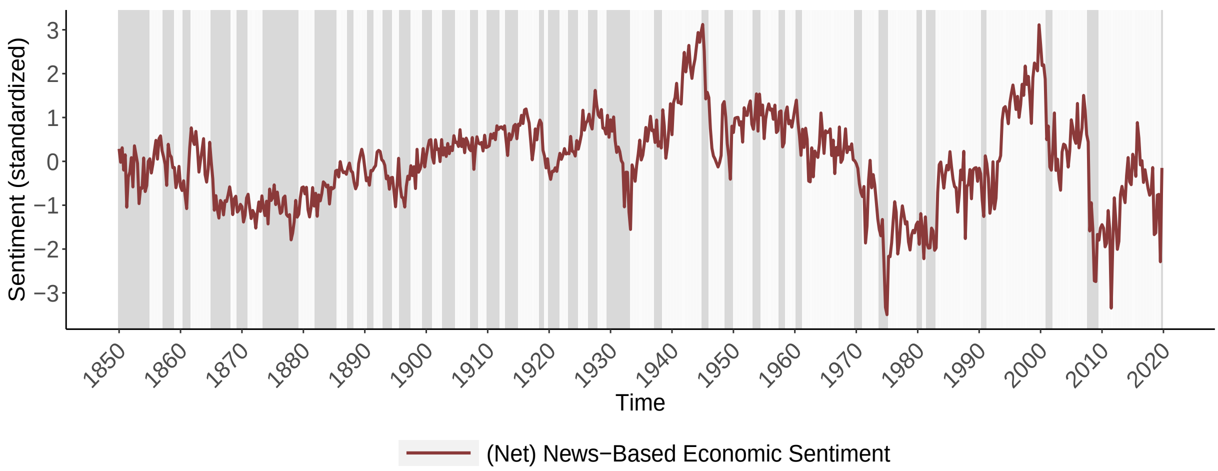
Figure 4: National Economic and Non-economic Sentiment



(a) Economic Sentiment



(b) Economic and Non-economic Sentiment



(c) Net Economic Sentiment

The figure shows the levels of national economic and non-economic sentiment and the net economic sentiment at a quarterly frequency. Shaded areas indicate NBER recessions. Sample size is 1850Q1–2019Q4.

trend, it still explains a sizeable fraction of the variation in sentiment in the past half-century. Goidel and Langley (1995) provided the first systematic study on the relationship between news coverage and actual economic conditions and famously concluded that the media disproportionately focuses on negative economic news.[7] Our results are particularly consistent with the previous studies that documented a general negativity bias in reporting macroeconomic news (Damstra and Boukes (2021), Fogarty (2005), Soroka (2006), and Soroka (2014)), and, particularly, about unemployment rates (Soroka (2012)).

Because the public needs to be aware of important risks in society, a certain level of general negative bias in news coverage is expected, particularly if one adheres to the view that traditional media fulfills a watchdog/surveillance function. Although this argument could explain the average level of negativity in news reporting, it does not address its increasingly downward trend. What factors could speak to it? The world of news, especially that of printed newspapers, has become increasingly competitive over the years. Therefore, It is natural to expect that to attract a larger audience, many outlets have been increasingly focusing on negative news. It is well-known that people are more responsive to negative information, see, e.g., Holbrook, Krosnick, Visser, Gardner, and Cacioppo (2001), Soroka (2006), and Tversky and Kahneman (1974)). Fighting the media negativity bias, the local Russian newspaper City Reporter decided in 2014 to report only positive news for a day and lost two-thirds of its readers.[8]

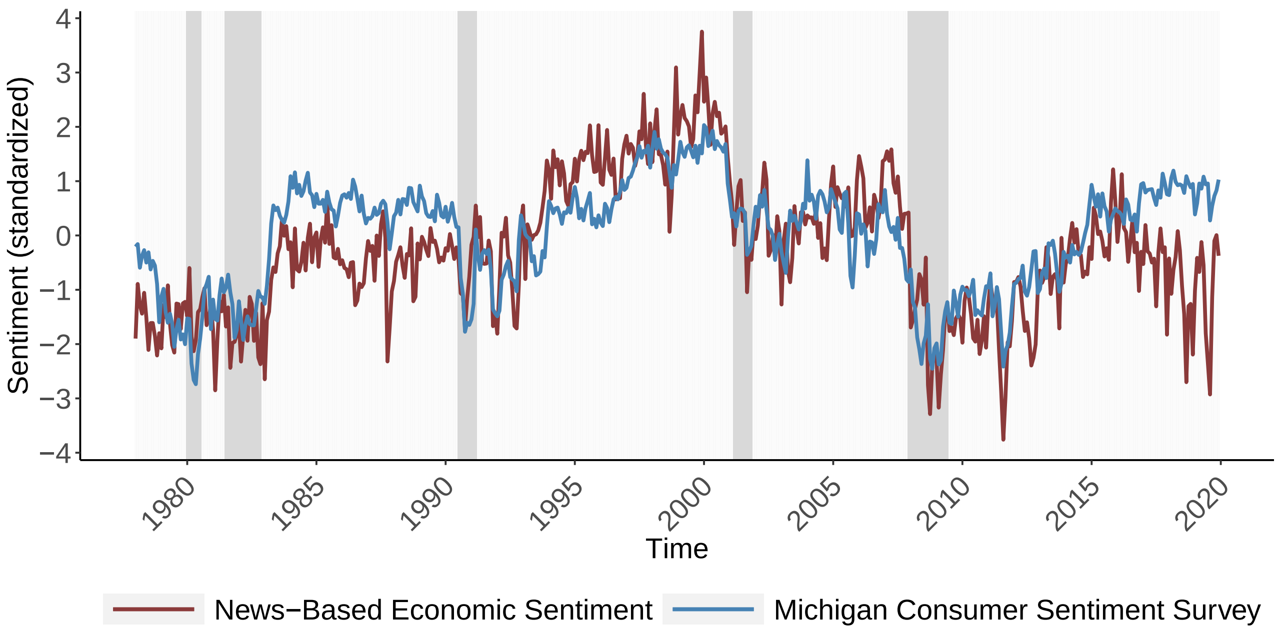
To directly test the hypothesis of an overall increasing negativity bias – not solely in the coverage of *economic* news – we measure the sentiment of news unrelated to the economy. To do this, we isolate pages that do not contain any words or phrases related to the economy (according to the output of our *Word2vec* approach) and build a similar measure of their general, non-economic sentiment. If an increasing negativity bias reflects the general trend in news coverage and/or language (likely caused by increasing competition among media outlets), we expect this non-economic measure to capture it.

Figure 4b presents our findings. Although the level of general sentiment also exhibits some variation (e.g., it is lower during the period of the Great Depression), overall, it does not reflect the same business-cycle-related variation as the level of economic sentiment. Nevertheless, it reveals the same negative trend that began in the 70s and affected the overall level of national sentiment going forward. This implies that most text-based sentiment and/or news coverage measures that rely on a fairly long time series of data may be biased. This, however, implies that we can use the non-economic sentiment to debias our original economic sentiment measure.

[7]For earlier evidence on the negativity of media bias, see Harrington (1989), Stein (1975), and Wattenberg (1984)). See also "Is the economy suffering from media malady?", Washington Post, 1990.

[8]See https://www.independent.co.uk/news/world/europe/website-reports-only-good-news-for-a-day-loses-two-thirds-of-its-readers-9905916.html.

Figure 5: National Economic Sentiment and Michigan Consumer Survey



The figure plots the levels of (net) news-based economic sentiment and the Michigan Consumer Sentiment Survey at a monthly frequency. Shaded areas indicate NBER recessions. Sample size is 1978M1–2020M1.

Figure 4c shows the dynamics of net economic sentiment obtained by orthogonalizing the original measure with respect to the non-economic sentiment.Although it retains most of the business cycle variation from our original sentiment measure, it no longer exhibits a significant downward trend over the last several decades. Throughout most of the empirical analysis, we will use this *net* economic sentiment measure (hereinafter, "news-based economic sentiment"). Furthermore, we show that all the empirical results on predictability originate from the original measure of economic sentiment and are unrelated to the non-economic sentiment.

Table A2 in the Appendix presents the summary statistics for all three measures of news-based sentiment: The net economic sentiment, its original, unresidualized version, and the non-economic sentiment measure. It also reports summary statistics of the changes in each of these variables.

To validate our economic sentiment measure, we focus on the period with the availability of its valid counterpart, the Michigan Consumer Sentiment Survey (see Figure 5). There are important similarities between the two measures, and they have a relatively high correlation of 0.71: Our news-based sentiment measure picks up similar variation as the Michigan survey. This makes our measure an important candidate variable to use when constructing a longer, backward-extended time series of the Michigan survey through projection methods.[9]

[9]As argued previously, in addition to extending the time series, our sentiment measure has the benefit of

In the overlapping sample, our measure deviates from the Michigan survey in three important episodes. Most importantly, our measure suggests substantially more negativity during and since the COVID-19 pandemic. Second, our measure displays more optimism during the dot-com bubble than the Michigan survey. Finally, our measure was more negative during the mid- and late 80s, when the savings and loans crisis unfolded.

In addition, we also check how our news-based economic sentiment correlates with the Organisation for Economic Co-operation and Development (OECD) measures of Consumer Confidence and the Business Confidence. Figure A4 in the Appendix plots the national economic sentiment and the OECD consumer confidence index in panel (a) and national economic sentiment, the OECD business confidence index, and the OECD composite lending index in panel (b). Our measure correlates with the OECD Consumer Confidence Index but not the OECD Business Confidence Index or the OECD Composite Lending Index. This provides further evidence that our measure is more related to the sentiment of households than businesses.

To further provide the context of our measure in light of important historical events, we plot the net economic sentiment and key historical events at an annual frequency in Figure 6. As expected, we find that the economic sentiment declines after key events that trigger economic and financial crises, such as the American Civil War, the Panic of 1893, World War I, the Great Depression, the Oil Crisis, and the Global Financial Crisis.

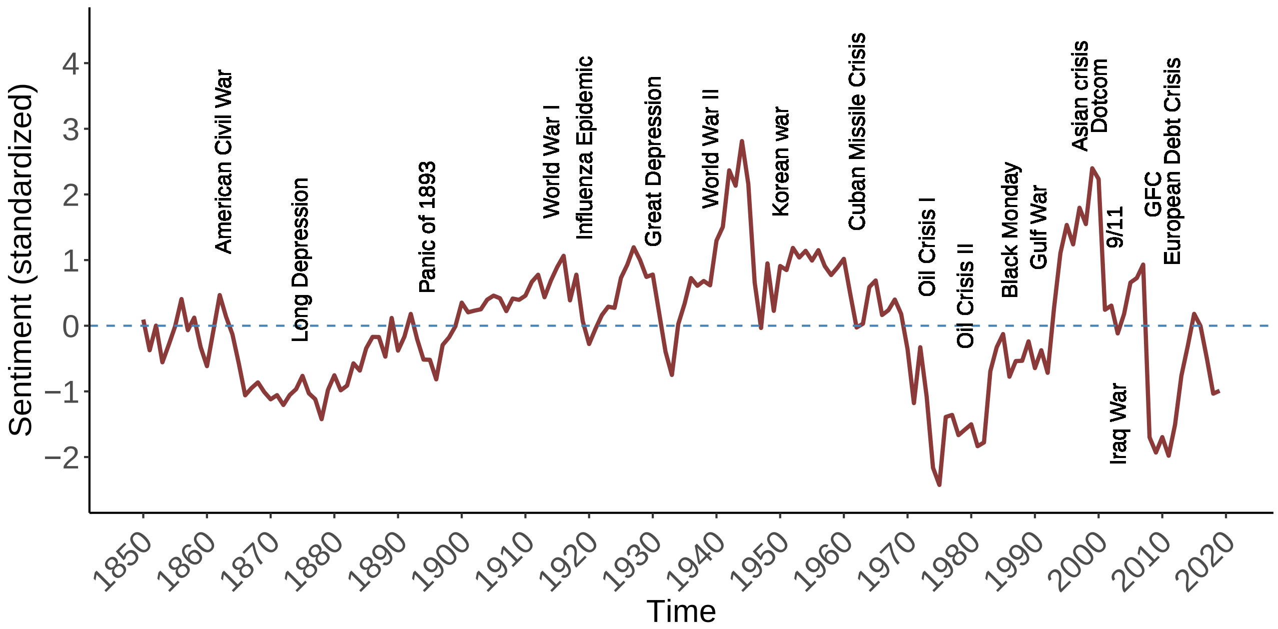
III.2 Local Economic Sentiment

Local newspaper coverage is one crucial aspect of our data. As a result, it provides us with an ideal setup to measure sentiment at a granular level (e.g., city, county, or state). Our sample includes local newspapers from all 50 US states. As these local newspapers are published across various cities within the states and cover local and national news, we can measure sentiment at the state level and test the effect of local sentiment on local economic variables. By applying our methodology to newspapers that are published in a given U.S. state, we construct the measure of state-level economic sentiment. Similar to our previous analysis, we first construct the state-level measures of raw economic and non-economic sentiment and standardize both to have a mean of zero and a standard deviation of one. We find the same negative bias in state-level non-economic sentiment across almost all the U.S. states, as illustrated by Figure 7.

Following our approach of removing the negative bias in news coverage from national economic sentiment, we orthogonalize state-level economic sentiment with the non-economic ones and take the outcomes as a net measure of economic sentiment. There is a large variation

providing granularity across states.

Figure 6: National Economic Sentiment and Key Historical Events



The figure plots the levels of news-based economic sentiment and key historical events at an annual frequency. The sample size is 1850–2019.

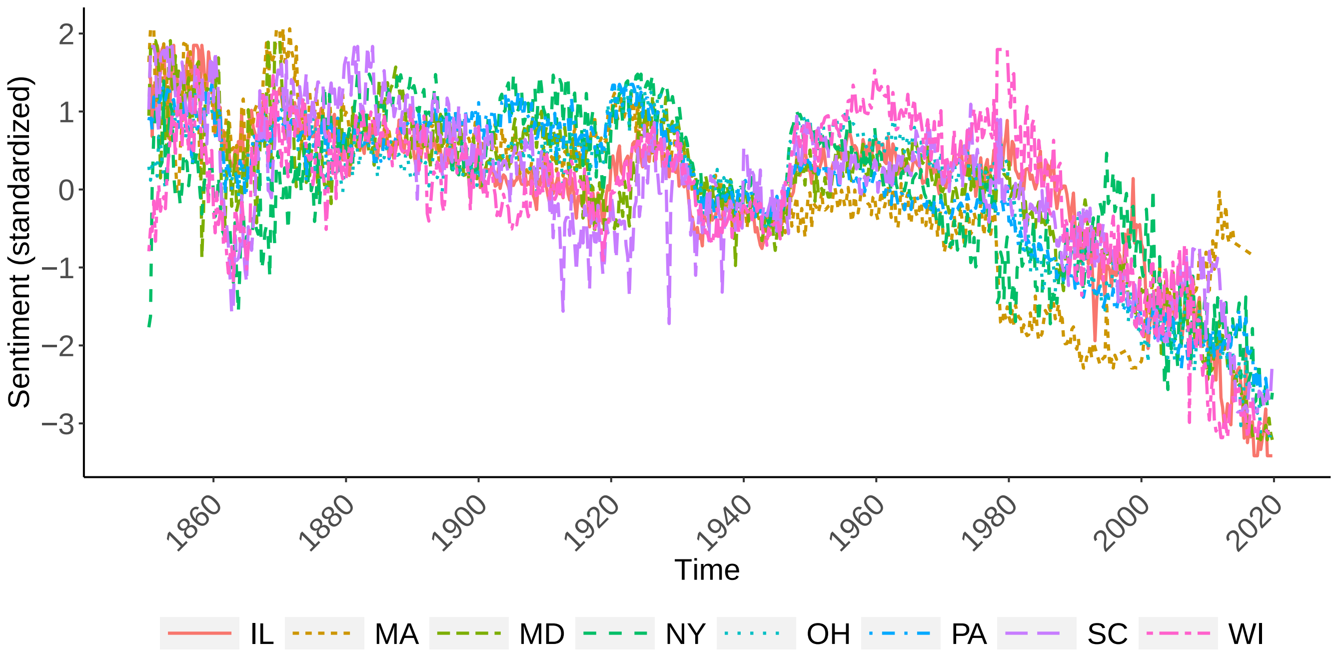
in economic sentiment across states, with the common component driving only approximately 35% of the state-level sentiment.

To intuitively check the dynamics of local sentiment, we plot the sentiment across U.S. states during important economic downturns with specific geographic epicenters. Given California's large concentration of technology-focused firms, we expect to see a significant decline in its state-level economic sentiment during the dot-com crash. Figure 8a plots the changes in the quarterly economic sentiment during the dot-com crash (2002 Q4). As expected, the figure shows that the economic sentiment declined more in California during the dot-com crash than in other U.S. states.

Next, we investigate how the economic sentiment evolved during the Global Financial Crisis (GFC). Because the Global Financial Crisis originated in the banking and insurance sector, we expect the economic sentiment to decline more in New York, which has a high concentration of financial services firms. Figure 8b plots the quarterly state economic sentiment changes during the Global Financial Crisis (2008 Q4). The figure shows that, indeed, during the GFC, the economic sentiment declined more in New York than in other states.

Thus far, we have seen a lot of heterogeneity in the evolution of local economic sentiment across regions during two major recessions with a specific geographic epicenter. To further investigate the heterogeneity in sentiment across states over the entire sample of 170 years,

Figure 7: State-level News-Based Non-Economic Sentiment

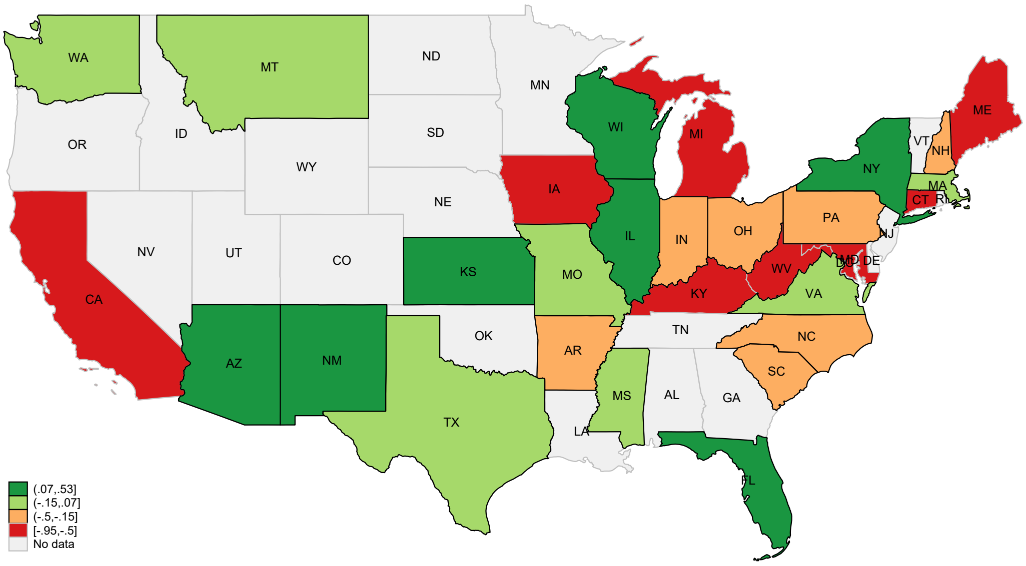


The figure shows state-level news-based non-economic sentiment for Illiois (IL), Massachusetts (MA), Maryland (MD), New York (NY), Ohio (OH), Pennsylvania (PA), South Carolina (SC), and Wisconsin (WI).

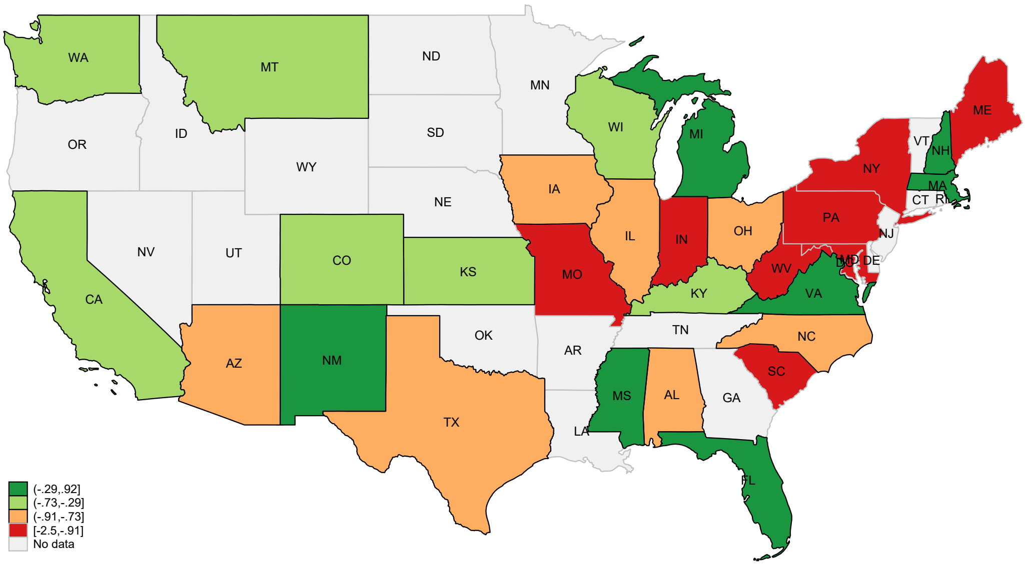
we measure its time-varying dispersion by computing the cross-sectional standard deviation of state-wide measures in each quarter. Figure 9 plots the cross-sectional dispersion in economic sentiment across states and key historical events. It reveals substantial average and, crucially, time-varying heterogeneity in economic sentiment across states.

Our default approach constructed the measure of national economic sentiment by pooling all newspapers across states. Thus, it is possible that we may have overweighted the contribution of certain states because they have a larger number of available newspapers. Because there is no available historical data on newspaper circulation for all the outlets in our sample, we construct alternative proxies of the importance of statewide news coverage. We construct alternative aggregate measures of national economic sentiment by taking the average mean of state-level economic sentiment measures, weighted using state GDP (whenever available) and population. Figure A3 in the Appendix plots the national economic sentiment and state-aggregated population-weighted national economic sentiment. The figure reveals that both measures have very similar dynamics. This alleviates the concern that our main measure may be driven by specific states and not reflect the overall country-wide dynamics.

Figure 8: Changes in Local Economic Sentiment



(a) Dot-com crash



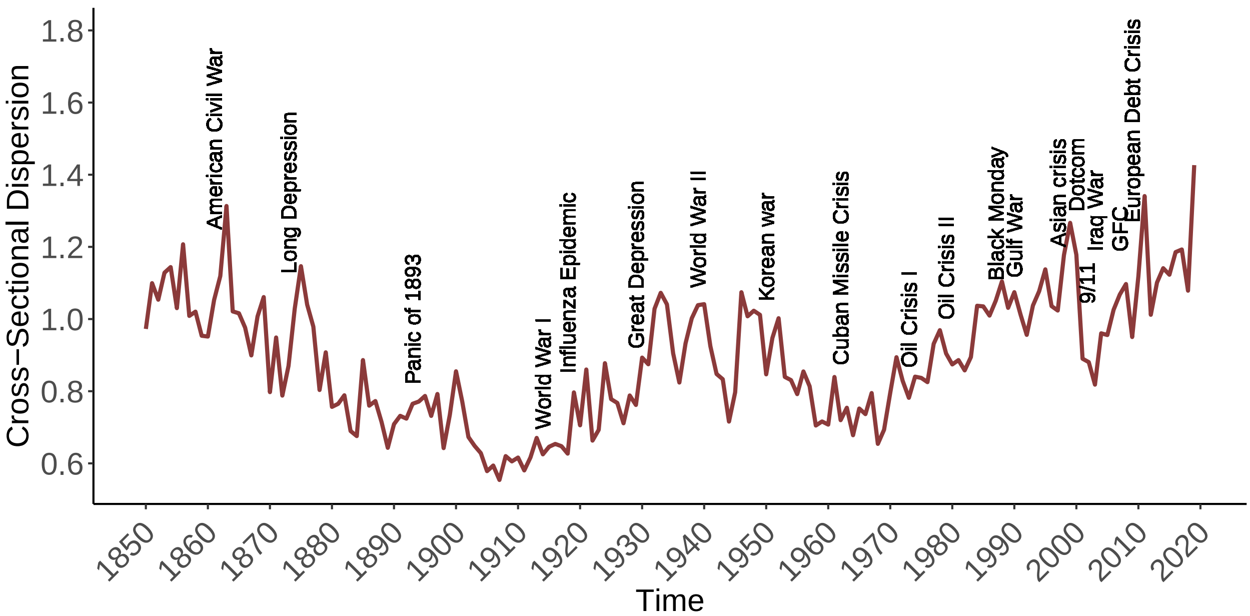
(b) Global Financial Crisis

The figure plots the changes in the quarterly state economic sentiment during the dot-com crash (2002Q4, panel (a)) and the Global Financial Crisis (200 Q4, panel (b)).

IV Economic sentiment and the economy

After establishing the time series and cross-sectional properties of the national and local sentiment measures, we now analyze their informational content. In the following subsections, we augment traditional predictive models for economic growth (and other specifications) with economic sentiment and check whether existing macroeconomic indicators span the latter's

Figure 9: Cross-Sectional Dispersion in Economic Sentiment



The figure shows annual cross-sectional dispersion in state-level (net) news-based economic sentiment.

dynamics and potential predictive impact.

IV.1 National Economic Sentiment and Aggregate Business Cycles

We begin our analysis by exploring the predictive ability of economic sentiment for national GDP growth rates. In particular, we run a predictive regression of real GDP growth per capita on changes in sentiment, as follows:

$$\Delta ln(GDP)_t = \sum_{n=1,2,\ldots,6} \theta_n \Delta Sent_{t-n} + \gamma \Delta ln(GDP)_{t-1} + \beta Term\ Spread_{t-1} + \epsilon_t,$$

where $\Delta ln(GDP)_t$ is the real growth rate of US GDP per capita and $\Delta Sent_{t-n}$, $n = 1, \ldots, 6$ are the previous quarter changes in the level of economic sentiment (up to six lags). In addition, we control for the previous quarter's GDP growth and the yield curve slope, one of the standard leading business cycle indicators (Harvey (1989)).

Table 1 presents our findings. Columns (I) and (II) establish the baseline results for the predictability of economic growth with fundamentals. Because GDP growth has some degree of persistence, its previous quarter performance carries over to the current quarter, leading to an overall R-squared of approximately 12%.

Economic sentiment is included as a regressor in columns (III) – (IX) of Table 1. These

Table 1: News-Based Economic Sentiment and GDP Growth, 1947Q1–2019Q4

	$\Delta ln(GDP_t)$								
	I	II	III	IV	V	VI	VII	VIII	IX
$\Delta ln(GDP_{t-1})$	0.355***	0.356***		0.333***	0.335***			0.327***	0.330***
	(5.91)	(5.79)		(5.57)	(5.48)			(5.17)	(5.15)
$Term\ Spread_{t-1}$		0.072*			0.063*			0.060	0.060
		(1.97)			(1.69)			(1.59)	(1.58)
$\Delta Sent_{t-1}$			0.333***	0.255***	0.242***	0.421***	0.390***	0.288***	0.250**
			(3.88)	(2.78)	(2.60)	(4.05)	(3.53)	(2.73)	(2.23)
$\Delta Sent_{t-2}$						0.167*	0.176*	-0.015	-0.006
						(1.67)	(1.73)	(-0.16)	(-0.06)
$\Delta Sent_{t-3}$						0.317***	0.316**	0.225**	0.223**
						(2.60)	(2.58)	(2.04)	(2.01)
$\Delta Sent_{t-4}$						0.267**	0.267**	0.152	0.151
						(2.23)	(2.20)	(1.37)	(1.35)
$\Delta Sent_{t-5}$						0.074	0.055	0.007	-0.017
						(0.74)	(0.54)	(0.08)	(-0.18)
Q4 Dummy							-0.107		-0.131
							(-0.90)		(-1.05)
$\Delta Sent_{t-1} \times Q4Dummy$							0.051		0.057
							(0.25)		(0.30)
F statistic ($\Delta Sent$ (t-2 to t-n))						2.235	2.242	2.499	2.435
P-value (all $\Delta Sent$ (t-2 to t-n))						0.065	0.065	0.043	0.048
Observations	290	290	290	290	290	290	290	290	290
Adjusted R-squared	0.12	0.13	0.04	0.14	0.15	0.06	0.06	0.16	0.16

The table reports a linear predictive model for the GDP growth per capita. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

specifications show that economic sentiment has significant predictive power for GDP statistically and economically. On average, adding economic sentiment leads to an improvement in R-squared of up to 4% relative to a benchmark model based solely on economic fundamentals. A one standard deviation increase in the previous quarter's sentiment growth corresponds to a 0.25%–0.42% increase in GDP growth in the following quarter. Furthermore, while most of the effect is coming from the sentiment change of this previous quarter, some of the additional lags in sentiment growth have an impact as well (see columns (VI)–(IX)). The bottom section of Table 1 also reports the joint significance test for all six lags of economic sentiment used in specifications (VI)–(IX), and finds them jointly significant (although the bulk of this effect comes from the first quarter).

Because our measure of (net) economic sentiment has been orthogonalized relative to the non-economic sentiment (see the discussion in Section III), one may wonder which of its two components is driving the overall results. Table A3 in the Appendix demonstrates how we decompose the GDP growth predictability to separately study the predictability driven by gross economic and non-economic sentiments. We find that all predictability comes from the former rather than the latter. Though the general, non-economic sentiment is important for identifying the negativity bias in the level of the sentiment, it does not reflect much information regarding future economic growth in the upcoming quarters.

To further test the ability of our news-based economic sentiment measure to predict GDP, we add more controls to our baseline estimation (Table 1). Table A5 in the Appendix controls for twelve lags of GDP Growth and Table A6 controls for the lagged dividend-to-price ratio. We find that these additional controls do not affect the GDP predictability of our news-based economic sentiment measure. Furthermore, Table A4 in the Appendix reports nonlinear estimation of key specifications of Table 1, conducted via Random Forest and associated premutation-based p-values and variable importance (see Figure A5 in the Appendix). Despite having a relatively short sample and low signal-to-noise ratio, data confirms results of the linear specification, with economic sentiment (particualrly the first lag) being a significant predictor of the GDP growth.

As discussed earlier, we construct the national economic sentiment measure by pooling newspapers across states. Thus, our measure may be biased towards states with more newspapers or states with greater economic news coverage. To address this concern, we take the population-weighted average of the state-level sentiment to construct the state-aggregated national economic sentiment. Table A7 in the Appendix presents the linear predictive model for GDP per capita using state-aggregated national economic sentiment, displaying the same predictive power as the national economic sentiment.

The previous analysis has established that economic sentiment can be successfully used to predict quarterly GDP growth, even after controlling for the fundamentals. However, the sample in this analysis was restricted, as the U.S. quarterly GDP growth data is available only for the past 72 years. Therefore, we repeat the analysis using annual data, using the full sample of economic sentiment (167 years). The results are reported in Table 2, which shows that economic sentiment has significant predictive power for the entire period. A one-standard-deviation shock of economic sentiment leads to approximately a 2% increase in next year's GDP growth per capita, so the impact is statistically significant and economically meaningful. Using annual frequency and a longer sample period, sentiment explains a substantial fraction of GDP growth in R-squared terms. Adding sentiment triples the R-squared value relative to a linear predictive regression model that only includes the previous year's GDP growth.

Next, we explore the predictive power of economic sentiment over various subsamples. In particular, we divide the overall data into three subsamples in columns (IV)–(VI). The first

Table 2: News-Based Economic Sentiment and GDP Growth, 1850-2017

	$\Delta \ln(GDP_t)$					
	I	II	III	IV	V	VI
$\Delta \ln(GDP_{t-1})$	0.147		0.072	-0.223*	0.203	0.148
	(1.29)		(0.65)	(-1.79)	(1.42)	(1.17)
$\Delta Sent_{t-1}$		2.129***	1.929***	2.262	2.334**	1.105**
		(3.58)	(3.73)	(1.51)	(2.01)	(2.65)
Period	Full	Full	Full	1850-1914	1915-1980	1981-2017
Observations	167	167	167	63	66	38
Adjusted R-squared	0.02	0.06	0.06	0.02	0.02	0.25

The table reports results of estimating a linear predictive model for the GDP growth per capita. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

period, 1850-1914, exhibited many economic and financial crises in the absence of a central banking system. While the coefficient on economic sentiment is insignificant over this subsample period (arguably due to a lack of power), the estimated coefficient is somewhat larger than the coefficient corresponding to the full sample. It is important to note that this was also the period with limited specialized economy-specific news coverage. As such, we expect a larger measurement error in the sentiment index, leading to an attenuation bias. Over the second subsample, which spans the period of 1914–1980, the coefficient on sentiment is both economically very large and statistically significant. This could reflect the increased quality and quantity of economic news coverage and the increasing role of a previously excluded set of economic agents in the economy and financial sector. Finally, over the last subsample, which covers the years 1981-2017, economic sentiment is a particularly strong predictor of economic growth, which includes the Great Moderation, the dot-com bubble, and the Great Recession. Over this subsample, sentiment appears to drive a quarter of the annual variation in economic growth. Potential explanations for these strong results include the increased quality of our sentiment measure and the democratization of economic activity.

Table A8 in the Appendix presents the results of GDP predictability by economic or non-economic sentiment using annual data. We find strong support for the predictive power of economic sentiment, consistent with the previous findings using the quarterly observations. Interestingly, over our full sample (167 years), the two types of sentiment seem to have opposing effects on GDP growth: Economic sentiment positively predicts GDP growth. In contrast, non-economic sentiment predicts it negatively in all three subsamples, although the latter coefficients are insignificant. Furthermore, the predictive power of non-economic sentiment has declined over the years, with an estimated coefficient in the last subsample that is particularly small. Overall, we find that the role of economic sentiment has become even more pronounced over the years, while that of the non-economic one seems to have been on a steady decline. This finding could be consistent with the time-varying importance of different sentiment scores and improved economic news coverage (and, hence, measurement).

IV.2 Local Economic Sentiment and Local Business Cycles

We now focus on the predictability of local GDP by local sentiment and the relationship between the cross-sectional dispersion in local sentiment across states and aggregate economic growth.

We start by investigating the predictability of local GDP by local sentiment. Table 3 presents the results. We find that state-level sentiment predicts state GDP growth even after controlling for national sentiment and national and state GDP growth. We also find that local economic sentiment has more persistent predictability for local GDP as it can predict local GDP growth up to six years in the future.

Thus far, our findings show that local and national sentiment changes predict GDP growth at both local and national levels. State-level economic activity and trade across states influence the overall national GDP. Thus, the dispersion in state-level sentiment may also predict national GDP growth. Our state-level sentiment allows us to measure the heterogeneity in economic sentiment across states to test this hypothesis. In Table 4, we use the cross-sectional dispersion in economic sentiment to measure heterogeneity across states. Higher dispersion predicts lower national GDP growth after controlling for national sentiment and current national GDP growth (columns I and II). We also test the effect of dispersion in economic sentiment on inflation (columns III and IV) and find no evidence of any predictability.

Having established the key link between national economic sentiment and economic growth, we now turn to another vital aspect of predictability: its relationship to the GDP expectations and forecasts.

IV.3 Economic Sentiment and GDP Forecasts

How does economic sentiment relate to professional GDP forecasts? To answer this question, we consider the data from the Survey of Professional Forecasters, which has been available since 1968 and is the oldest quarterly survey of macroeconomic forecasts in the United States. We estimate the following equation:

$$\Delta ln(GDPForecast)_t = \sum_{n=1,2..6} \theta_n \Delta Sentiment_{t-n} + \gamma \Delta ln(GDPForecast)_{t-1} + \epsilon_t,$$

Table 5 demonstrates that economic sentiment leads the survey of professional forecasters. In every specification we estimate, an increase in economic sentiment by one standard deviation leads to an increase in forecasted GDP growth by 0.08%–0.20%. Economic sentiment contributes 5%–11% of the total time series variation (R-squared) in professional GDP growth forecasts.

Even though text data could reflect the drivers of the economic growth expectations, it does not necessarily imply that sentiment on its own could have predictive power beyond what is already formulated by professional forecasters. We test this hypothesis in Table 6, which focuses

Table 3: State-Level Economic Sentiment and Local GDP

	$\Delta ln(GSP_{s,t})$					
	I	II	III	IV	V	VI
$\Delta Sentiment_{s,t-1}$	0.263	0.648**	0.723**	-0.153	0.182	0.247
	(0.268)	(0.273)	(0.287)	(0.320)	(0.342)	(0.357)
$\Delta Sentiment_{s,t-2}$		1.244***	1.362***		0.777***	0.901***
		(0.246)	(0.273)		(0.274)	(0.302)
$\Delta Sentiment_{s,t-3}$		1.156***	1.321***		1.017***	1.173***
		(0.289)	(0.329)		(0.289)	(0.330)
$\Delta Sentiment_{s,t-4}$		1.301***	1.455***		1.281***	1.424***
		(0.283)	(0.288)		(0.289)	(0.293)
$\Delta Sentiment_{s,t-5}$			0.334			0.353
			(0.257)			(0.266)
$\Delta Sentiment_{s,t-6}$			0.740**			0.611*
			(0.319)			(0.321)
$\Delta ln(GSP_{s,t-1})$				0.028	0.013	0.010
				(0.022)	(0.024)	(0.025)
$\Delta ln(GDP_{t-1})$				0.488***	0.421***	0.404***
				(0.055)	(0.071)	(0.075)
$\Delta NationalSentiment_{t-1}$				0.229	0.276	0.305
				(0.186)	(0.192)	(0.193)
Controls	No	No	No	Yes	Yes	Yes
Observations	1,172	1,172	1,172	1,172	1,172	1,172
R-squared	0.00	0.03	0.03	0.03	0.04	0.05

Table reports a linear predictive model for state-level GDP growth per capita. State-level clustered standard errors are reported in parentheses. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

Table 4: Statewide Dispersion in National Economic Sentiment and GDP

	$\Delta ln(GDP_t)$		$\Delta ln(CPI_t)$	
	I	II	III	IV
$\Delta Sentiment_{t-1}$	0.339***	0.248***	0.021	0.067
	(3.95)	(2.69)	(0.32)	(1.57)
$Dispersion_{t-1}$	-0.551*	-0.497*	-0.473	-0.083
	(-1.68)	(-1.90)	(-1.51)	(-0.57)
$\Delta ln(GDP_{t-1})$		0.334***		
		(5.40)		
Term Spread_{t-1}		0.069*		-0.074***
		(1.82)		(-3.06)
$\Delta ln(CPI_{t-1})$				0.695***
				(10.24)
Controls	No	Yes	No	Yes
Observations	290	290	290	290
Adjusted R-squared	0.05	0.16	0.00	0.55

The table reports a linear predictive model for national GDP and CPI. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

on the extended specification of the GDP forecast equation considered earlier. In addition to the economic fundamentals and sentiment, we now include the Survey of Professional Forecasters data to test whether these forecasts span all the relevant information in our economic sentiment measure.

Table 6 reveals that economic sentiment is not driven out as a predictor by including the GDP forecasts as a regressor. Interestingly, past economic growth and the impact of the term spread seem to lose their predictive relevance once we control for professional GDP forecasts. Instead, the effect of economic sentiment stays robust, with a one standard deviation change in the latter leading to 0.2%–0.24% change in economic growth. Although there is some evidence that sentiment over the past several quarters seems to impact GDP, most of the effect comes from the last quarter. Overall, the incremental power of economic sentiment in predicting GDP is about 2%–4%, which is still economically sizeable and consistent with our previous findings.

IV.4 Sentiment over Current and Future Events

While newspapers cover current economic events, they also opine about the impact of recent events on future economic activity. For example, during the Global Financial Crisis, newspapers

Table 5: News-Based Economic Sentiment and the Professional Forecast Predictability

	$GDP\ Forecast_{t,t+1}$						
	I	II	III	IV	V	VI	VII
$GDP\ Forecast_{t-1}$	0.786***	0.727***		0.770***	0.716***		0.726***
	(13.15)	(12.52)		(11.90)	(11.82)		(12.43)
$Term\ Spread_{t-1}$		0.057***			0.054***		0.054***
		(3.85)			(3.68)		(3.73)
$\Delta Sent_{t-1}$			0.154***	0.094***	0.084***	0.205***	0.082***
			(3.35)	(3.10)	(3.05)	(3.98)	(2.89)
$\Delta Sent_{t-2}$						0.151***	-0.038
						(3.39)	(-1.06)
$\Delta Sent_{t-3}$						0.163***	0.020
						(3.30)	(0.47)
$\Delta Sent_{t-4}$						0.118**	0.004
						(2.31)	(0.12)
$\Delta Sent_{t-5}$						0.058	-0.001
						(1.28)	(-0.04)
F statistic ($\Delta Sent$ (t-2 to t-n))						3.708	0.971
P-value ($\Delta Sent$ (t-2 to t-n))						0.006	0.425
Observations	204	204	204	204	204	204	204
Adjusted R-squared	0.62	0.65	0.05	0.63	0.67	0.11	0.67

The table shows the predictive regression of the GDP forecast by its previous quarter values, term spread, and changes in economic sentiment. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

covered both current events, such as the failure of Lehman Brothers, and the impact of these events on future GDP growth, as well as the possible monetary policy response by central banks. To better understand the source of predictability of our news-based economic sentiment measure, we decompose it into a current and a future component.

To do this, we first create a topic-specific dictionary of words related to the present and the future and measure the intensity of this language on each newspaper page to classify them as related to current or future events. Subsequently, we use this classification to create two separate economic sentiment measures: current and future. Figure 10 plots these two measures. We find that current economic sentiment does not seem to have much variation at a business cycle frequency and exhibits a fairly persistent behavior. On the other hand, the measure of economic sentiment related to the future varies a lot over the business cycle and, therefore, is a likely driver of our predictability findings.

To formally test this, we re-estimate the specification in Table 1, replacing the single original

Table 6: News-Based Economic Sentiment and GDP Forecast

	$\Delta ln(GDP_t)$			
	I	II	III	IV
$\Delta ln(GDP_{t-1})$	0.298***	0.054	0.025	-0.002
	(3.14)	(0.59)	(0.28)	(-0.02)
$Term\ Spread_{t-1}$	0.083**	0.024	0.015	0.018
	(2.37)	(0.75)	(0.49)	(0.58)
$GDPForecast_{t-1}$		0.870***	0.875***	0.817***
		(5.02)	(5.19)	(5.07)
$\Delta Sent_{t-1}$			0.201***	0.244***
			(2.93)	(3.08)
$\Delta Sent_{t-2}$				0.050
				(0.57)
$\Delta Sent_{t-3}$				0.180*
				(1.71)
$\Delta Sent_{t-4}$				0.234**
				(2.35)
$\Delta Sent_{t-5}$				0.072
				(0.83)
F statistic ($\Delta Sent$ (t-2 to t-n))				1.941
P-value ($\Delta Sent$ (t-2 to t-n))				0.105
Observations	204	204	204	204
Adjusted R-squared	0.11	0.23	0.25	0.27

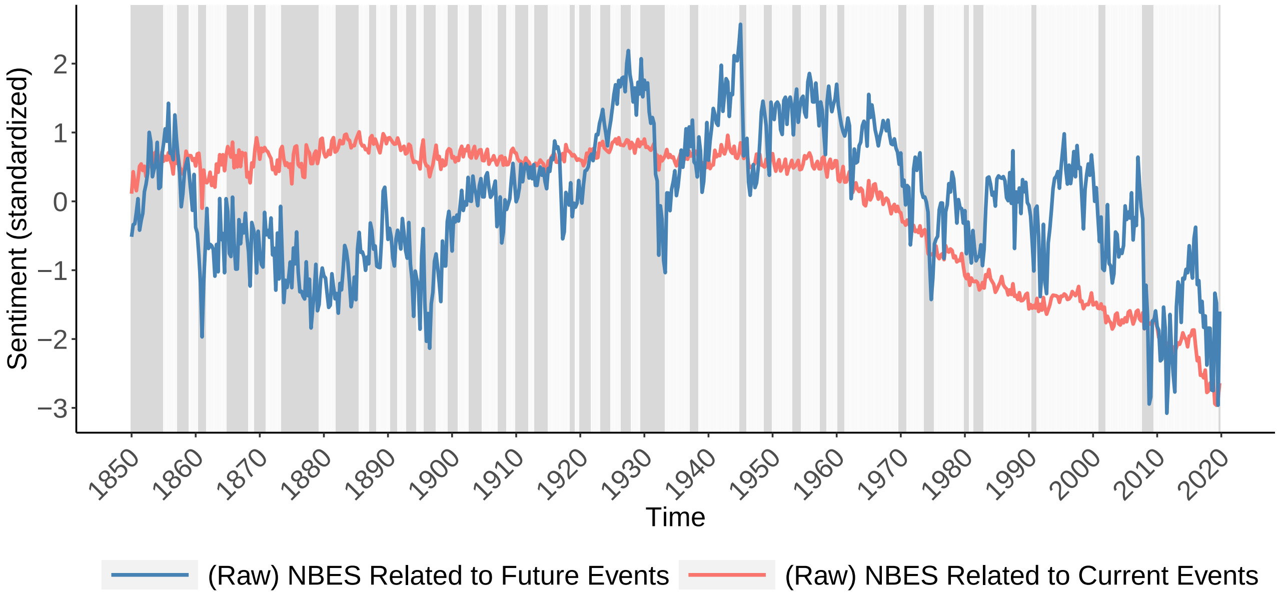
The table shows the predictive regression of GDP growth on its previous quarter value, term spread, professional GDP forecast, and economic sentiment. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

measure of economic sentiment with two regressors: Current and future economic sentiment. In particular, to understand the source of predictability, we estimate the following equation:

$$\Delta ln(GDP)_t = \theta \Delta FutureNewsSent_{t-1} + \omega \Delta CurrentNewsSent_{t-1} +$$
$$\phi \Delta Non - EconSent_{t-1} + \gamma \Delta ln(GDP)_{t-1} + \beta Term\ Spread_{t-1} + \epsilon_t.$$

Table 7 presents the results. Our predictability findings are driven by the measure's forward-looking, future component. The current events coverage has little to no predictive power for macroeconomic variables. This suggests that our measure can capture forward-looking news about economic fundamentals.

Figure 10: National Economic Sentiment about Current and Future



The figure plots the levels of the (raw) news-based economic sentiment computed using the newspaper text that covers current and future events. Shaded areas indicate NBER recessions. The sample size is 1850Q1–2019Q4.

Table 7: News-Based Economic Sentiment and GDP Growth: Current and Future Sentiments

	$\Delta ln(GDP_t)$				
	I	II	III	IV	V
$\Delta ln(GDP_{t-1})$	0.355***	0.356***		0.335***	0.338***
	(5.91)	(5.79)		(5.51)	(5.43)
Term Spread$_{t-1}$		0.072*			0.065*
		(1.97)			(1.73)
$\Delta Current\ News\ Sent\ (gross)_{t-1}$			-0.232	-0.091	-0.111
			(-0.45)	(-0.20)	(-0.25)
$\Delta Future\ News\ Sent\ (gross)_{t-1}$			0.416***	0.295**	0.278**
			(3.41)	(2.45)	(2.28)
$\Delta Non-Econ\ Sent_{t-1}$			0.031	-0.080	-0.104
			(0.12)	(-0.31)	(-0.40)
Observations	290	290	290	290	290
Adjusted R-squared	0.12	0.13	0.02	0.13	0.14

Table reports a linear predictive model for national GDP growth and economic sentiment measured from text about current and future events. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

Table 8: National Economic Sentiment, Investment, and Industrial Production

	$\Delta ln(Investment_t)$					$\Delta ln(IndustrialProduction_t)$				
	I	II	III	IV	V	VI	VII	VIII	IX	X
$\Delta ln(Investment_{t-1})$	0.738***			0.738***	0.743***					
	(13.70)			(13.79)	(13.83)					
$Term\ Spread_{t-1}$	-0.046**			-0.044**	-0.045**	0.225*			0.226*	0.215*
	(-2.08)			(-1.99)	(-2.08)	(1.88)			(1.89)	(1.78)
$\Delta Sent_{t-1}$		-0.073	-0.073	-0.038	-0.019		0.472	0.826*	-0.039	0.133
		(-0.90)	(-0.66)	(-0.72)	(-0.34)		(0.99)	(1.74)	(-0.09)	(0.33)
$\Delta Sent_{t-2}$			0.016		0.108			0.739*		0.213
			(0.12)		(1.54)			(1.97)		(0.58)
$\Delta Sent_{t-3}$			-0.043		-0.024			1.152***		0.702*
			(-0.32)		(-0.38)			(3.27)		(1.81)
$\Delta Sent_{t-4}$			-0.018		0.032			0.450		-0.060
			(-0.14)		(0.44)			(1.03)		(-0.15)
$\Delta Sent_{t-5}$			-0.030		-0.006			0.092		0.061
			(-0.24)		(-0.10)			(0.25)		(0.20)
$\Delta ln(IndustrialProduction_{t-1})$						0.409***			0.410***	0.402***
						(3.92)			(4.13)	(4.04)
F statistic ($\Delta Sent$ (t-2 to t-n))			0.175		1.062			3.144		0.873
P-value ($\Delta Sent$ (t-2 to t-n))			0.951		0.376			0.015		0.480
Observations	290	290	290	290	290	399	399	399	399	399
Adjusted R-squared	0.61	-0.00	-0.01	0.61	0.61	0.17	0.00	0.01	0.17	0.17

The table shows the linear regression of investment and industrial production growth on the term spread and economic sentiment. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. Significance: * 10%; ** 5%; *** 1%.

IV.5 Economic Sentiment and GDP Components

We now explore the sources of economic predictability by the sentiment through the lens of the inputs to GDP provided by the agents in the economy: Capital and labor. To distinguish their separate impacts on consumer and firm decisions, we focus on the individual components of GDP to see what economic sentiment affects.

First, we evaluate whether sentiment affects firms' investment decisions (the capital input) using the following predictive regressions for investment and industrial production growth:

$$\Delta ln(Inv/Prod)_t = \sum_{n=1,2..6} \theta_n \Delta Sent_{t-n} + \gamma \Delta ln(Inv/Prod)_{t-1} + \beta\ Term\ Spread_{t-1} + \epsilon_t.$$

Table 8 presents our results. Panel A shows that controlling for the previous level of investment and the term spread, there is no evidence that investment growth responds to changes in economic sentiment (or lags thereof). Panel B reports similar results for the impact of economic sentiment on industrial production growth: We do not find any significant impact of the former, neither statistically nor economically. In other words, we find no evidence that sentiment

Table 9: News-Based Economic Sentiment, Employment, and Consumption

	$\Delta ln(Employment_t)$					$\Delta ln(Consumption_t)$				
	I	II	III	IV	V	VI	VII	VIII	IX	X
$\Delta ln(Employment_{t-1})$	0.757***			0.748***	0.748***					
	(14.73)			(14.65)	(16.61)					
$Term\ Spread_{t-1}$	0.042**			0.036*	0.036*	0.068			0.064	0.059
	(2.29)			(1.94)	(1.85)	(1.50)			(1.40)	(1.27)
$\Delta Sent_{t-1}$		0.249***	0.339***	0.154***	0.159***		0.135**	0.202**	0.106	0.173*
		(3.11)	(3.67)	(3.06)	(2.92)		(2.04)	(2.55)	(1.34)	(1.74)
$\Delta Sent_{t-2}$			0.315***		0.005			0.185**		0.164*
			(2.71)		(0.06)			(2.13)		(1.76)
$\Delta Sent_{t-3}$			0.258*		-0.014			0.234**		0.216*
			(1.79)		(-0.16)			(2.21)		(1.75)
$\Delta Sent_{t-4}$			0.241*		0.049			0.126		0.115
			(1.76)		(0.70)			(1.11)		(1.18)
$\Delta Sent_{t-5}$			0.107		-0.029			0.097		0.100
			(0.97)		(-0.52)			(1.19)		(1.14)
$\Delta ln(Consumption_{t-1})$						0.083			0.071	0.039
						(0.52)			(0.44)	(0.25)
F statistic ($\Delta Sent$ (t-2 to t-n))			2.234		0.736			1.581		1.075
P-value ($\Delta Sent$ (t-2 to t-n))			0.065		0.568			0.179		0.369
Observations	322	322	322	322	322	290	290	290	290	290
Adjusted R-squared	0.56	0.02	0.06	0.57	0.57	0.01	0.01	0.02	0.02	0.02

The table shows the linear predictive model for employment and total consumption growth on term spread and economic sentiment. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

affects economic growth through the investment (capital) channel.

If economic sentiment does not affect investment or industrial production, how does it predict economic growth? Table 9 investigates whether the labor channel drives the results. Panel A demonstrates that national economic sentiment significantly predicts employment growth. A one-standard-deviation shock to sentiment leads to a 0.15%-0.16% increase in employment growth rate over the next quarter. This effect remains even after controlling for past employment growth and term spread. Panel (b) reports similar findings for the total consumption growth, which we also find to be significantly predicted by the changes in economic sentiment. While the overall degree of predictability for these time series is not as large as for overall GDP, their relative contribution are clear.

Which of the consumption components are reacting to the changes in economic sentiment? Table 10 provides an answer to this question. Panel (a) reports estimation results of a linear predictive model for nondurable consumption growth, where we see a relatively small effect of economic sentiment on nondurable consumption. At the same time, as reported in Panel (b), services seem to have a significantly more robust response to changes in the economic senti-

Table 10: News-Based Economic Sentiment, Nondurable Consumption, and Services

	$\Delta ln(Cons^{nondur}_{t-1})$					$\Delta ln(Cons^{serv}_{t-1})$				
	I	II	III	IV	V	VI	VII	VIII	IX	X
$\Delta ln(Cons^{nondur}_{t-1})$	0.096			0.086	0.057					
	(1.06)			(0.95)	(0.67)					
$Term\ Spread_{t-1}$	0.049			0.047	0.043	0.026			0.025	0.020
	(1.41)			(1.35)	(1.25)	(0.88)			(0.85)	(0.62)
$\Delta Sent_{t-1}$		0.097	0.153*	0.061	0.118		0.037	0.104**	0.008	0.061
		(1.40)	(1.94)	(0.91)	(1.52)		(0.87)	(2.13)	(0.17)	(1.20)
$\Delta Sent_{t-2}$			0.183*		0.161			0.201***		0.155***
			(1.81)		(1.54)			(3.26)		(2.64)
$\Delta Sent_{t-3}$			0.198**		0.177*			0.169***		0.094
			(2.08)		(1.84)			(2.71)		(1.60)
$\Delta Sent_{t-4}$			0.103		0.090			0.153**		0.099*
			(1.04)		(0.95)			(2.54)		(1.75)
$\Delta Sent_{t-5}$			0.114		0.116			0.046		0.008
			(1.32)		(1.32)			(0.70)		(0.13)
$\Delta ln(Cons^{serv}_{t-1})$						0.360***			0.359***	0.323***
						(5.11)			(5.16)	(4.56)
F statistic ($\Delta Sent$ (t-2 to t-n))			1.383		1.101			3.460		2.269
P-value ($\Delta Sent$ (t-2 to t-n))			0.240		0.356			0.009		0.062
Observations	290	290	290	290	290	290	290	290	290	290
Adjusted R-squared	0.01	0.00	0.01	0.01	0.02	0.12	-0.00	0.05	0.12	0.14

The table shows the linear predictive model for nondurable consumption and services per capita on term spread and economic sentiment. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

ment, with several lags over the previous quarters being jointly significant. Both nondurable consumption and services appear to react to the economic sentiment with a lag, consistent with the models of a slow-moving component in consumption growth and adjustment costs.

Table A9 in the Appendix presents the results of a similar model for durable consumption growth, where we find no evidence of its response to the changes in economic sentiment.

To summarize, we find employment to have the most potent response to changes in the news-based economic sentiment. At the same time, it also has a slow-moving impact on nondurable consumption growth and services. We find no evidence that economic sentiment affects the economy through an investment channel. Instead, all the empirical evidence points to the role of labor. These findings are essential for developing macroeconomic models with sentiment, indicating the role of labor, not capital/investment, in the overall propagation mechanism.

Table 11: News-Based Economic Sentiment and Taylor Rule

	ΔFFR_t											
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
$\Delta Sent_{t-1,t-7}$		0.077***	0.050**		0.072***	0.044**		0.123***	0.065***		0.070***	0.049***
		(3.07)	(2.57)		(2.86)	(2.19)		(3.38)	(3.54)		(3.05)	(2.65)
$\Delta Sent_{t-1,t-7} \times Res_{t-1}$			0.197**			0.205**			0.337**			0.145
			(2.04)			(2.14)			(2.38)			(1.42)
Specification	R&R	R&R	R&R	-Unemp	-Unemp	-Unemp	-GDP	-GDP	-GDP	-Inf	-Inf	-Inf
Observations	346	346	346	346	346	346	346	346	346	346	346	346
Adjusted R-squared	0.23	0.25	0.27	0.22	0.24	0.26	0.04	0.09	0.15	0.21	0.22	0.23

The table shows the linear predictive model for changes in the federal funds rate (FFR) on economic sentiment. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. The sample size is from January 1969 to December 2007. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

V Economic Sentiment and Monetary Policy

This section explores three aspects of the interaction between economic sentiment and monetary policy. First, we check if central banks account for economic sentiment in the interest-rate-setting process. Second, we quantify the extent to which economic sentiment influences the decision-making process compared to other macroeconomic variables. Finally, we explore the channel through which economic sentiment influences monetary policy.

We quantify the importance of sentiment in explaining the changes in the federal funds rate relative to what the forward-looking Taylor rule estimation proposed by Romer and Romer (2004) would imply. Table 11 presents the results. News-based economic sentiment has a large influence on the key policy rate, particularly during recessions: a one-standard-deviation decrease in sentiment over the past two quarters leads to a 25 basis point decrease in the policy rate (see column III of Table 11).

Furthermore, these effects are economically large and are on par with the importance of macroeconomic fundamentals, the forecasts of GDP, inflation, and unemployment, which are the central factors affecting policy rates. Adding the lagged change in economic sentiment over the past two quarters increases the adjusted R-squared by two to four percentage points (see column I vs columns II and III in Table 11). In comparison, adding unemployment increases the adjusted R-squared by one percentage point (column I vs column V), and adding inflation increases the adjusted R-squared by two percentage points (column I vs column X). This further highlights that economic sentiment has a considerable influence on monetary policy.

What is the source of our findings? On the one hand, sentiment could affect interest rates

Table 12: Incremental Effect of News-Based Economic Sentiment on Taylor Rule

	ΔFFR_t			
	I	II	III	IV
$\Delta Sent_{t-1,t-7}$		0.063**	0.037*	0.032
		(2.13)	(1.67)	(1.41)
$\Delta Sent_{t-1,t-7} \times Recession_{t-1}$			0.195*	0.198**
			(1.95)	(1.98)
Predicted GDP Growth	0.197***	0.119	0.113	0.072
	(2.68)	(1.34)	(1.29)	(0.86)
Predicted GDP Growth (Two Quarters Ahead)				0.073
				(1.10)
Specification	R&R	R&R	R&R	R&R
Observations	346	346	346	346
Adjusted R-squared	0.24	0.25	0.27	0.27

The table shows the linear predictive model for changes in the Fed Funds Rate (FFR) on economic sentiment after controlling for sentiment-predicted GDP growth. The sentiment-predicted GDP growth is estimated using the specification of column V of Table 1. The specifications include all Romer and Romer (2004) controls. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. The sample size is from January 1969 to December 2007. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

due to its incremental predictive power for GDP growth. On the other hand, sentiment could induce the Fed to deviate from the Taylor rule. To test these two channels, we re-estimate the standard Taylor rule specification, including changes in economic sentiment (i.e., column II and III of Table 11) after controlling for a sentiment-augmented GDP forecast, computed following specification of column V of Table 1.

Table 12 presents the results. Economic sentiment is reflected in monetary policy even after controlling for its incremental predictive power for GDP. Note that since the sentiment-augmented economic growth prediction relies on the actual ex-post realization of GDP, it contains a significant look-ahead bias and likely overestimates the potential predictive ability of sentiment on the latter. Nevertheless, even after controlling for the forecast, news-based economic sentiment significantly impacts the FFR during economic recessions.

VI Conclusion

Using 193 million pages of US local newspaper articles and state-of-the-art machine learning methods, we construct state- and national-level economic sentiment measures for almost two centuries. Our measure is highly correlated with existing survey-based economic sentiment measures in the sample period where both measures exist (1978 – 2017). Our measure significantly extends the time series availability and granularity at which we can measure economic sentiment. Sentiment exhibits significant time series variation and declines considerably during major US recessions in the 19th, 20th, and 21st centuries. We show that our measure predicts economic fundamentals such as GDP, consumption, and employment growth, even after controlling for the macroeconomic information available at the time of the forecast. It is distinct from the information in expert forecasts and also predicts their consensus value.

Sentiment also exhibits large cross-sectional variation across states, with their common component driving only 35% of the total variation. State-level sentiment predicts local GDP growth even after controlling for both national sentiment and GDP growth. Furthermore, cross-sectional dispersion in sentiment across states is associated with lower national GDP growth. In conclusion, our results indicate the importance of sentiment at both local and national levels in understanding the business cycle.

Our paper does not provide any evidence for the causal relationship between sentiment and business cycle. That said, it offers insights into the informational content of the news-based economic sentiment and its empirical relationship with the business cycle fluctuations in fundamentals and agents expectations. Our empirical findings have important implications for research in macrofinance and behavioral economics.

References

AGRAWAL, S., P. D. AZAR, A. W. LO, AND T. SINGH (2018): "Momentum, Mean-reversion, and Social Media: Evidence from Stocktwits and Twitter," *Journal of Portfolio Management*, 44, 85–95.

ALTMANN, A., L. TOLOSI, O. SANDER, AND T. LENGAUER (2010): "Permutation Importance: A Corrected Feature Importance Measure," *Bioinformatics*, 26, 1340–1347.

ANTWEILER, W., AND M. Z. FRANK (2004): "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards," *Journal of Finance*, 59(3), 1259–1294.

ARORA, S., Y. LIANG, AND T. MA (2017): "A Simple but Tough-to-Beat Baseline for Sentence Embeddings," *International Conference on Learning Representations*.

ARTEAGA-GARAVITO, M. J., M. M. CROCE, P. FARRONI, AND I. WOLFSKEIL (2022): "When the Markets Get CO.V.I.D: Contagion, Viruses, and Information Diffusion," available at SSRN: `https://ssrn.com/abstract=3560347`.

ASH, E., D. L. CHEN, AND A. ORNAGHI (2021): "Gender Attitudes in the Judiciary: Evidence from U.S. Circuit Courts," QAPEC Discussion Papers 08.

BAKER, M., AND J. WURGLER (2007): "Investor Sentiment in the Stock Market," *Journal of Economic Perspectives*, 21, 129–152.

BAKER, S. R., N. BLOOM, AND S. J. DAVIS (2016): "Measuring Economic Policy Uncertainty," *Quarterly Journal of Economics*, 131(4), 1593–1636.

BELLSTAM, G., S. BHAGAT, AND J. A. COOKSON (2021): "A Text-Based Analysis of Corporate Innovation," *Management Science*, 67(7), 4004–4031.

BLEI, D. M., A. Y. NG, AND M. I. JORDAN (2003): "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 3(Jan), 993–1022.

BOLLEN, J., H. MAO, AND X. ZENG (2011): "Twitter Mood Predicts the Stock Market," *Journal of Computational Science*, 2(1), 1–8.

BORDALO, P., N. GENNAIOLI, Y. MA, AND A. SHLEIFER (2020): "Overreaction in Macroeconomic Expectations," *American Economic Review*, 110(9), 2748–2782.

BORDALO, P., N. GENNAIOLI, AND A. SHLEIFER (2018): "Diagnostic Expectations and Credit Cycles," *Journal of Finance*, 73(1), 199–227.

BYBEE, L., B. T. KELLY, A. MANELA, AND D. XIU (2020): "Narrative Asset Pricing: Interpretable Systematic Risk Factors from News Text," available at `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3895277`.

——— (2021): "Business News and Business Cycles," available at `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3446225`.

CAO, S., Y. KIM, A. WANG, AND H. XIAO (2020): "From Words to Syntax: Identifying Context-specific Information in Textual Analysis," available at `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3568504`.

Chen, D., and C. Manning (2014): "A Fast and Accurate Dependency Parser using Neural Networks," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 740–750, Doha, Qatar. Association for Computational Linguistics.

Chen, Y., B. T. Kelly, and D. Xiu (2022): "Expected Returns and Large Language Models," available at SSRN: `https://ssrn.com/abstract=4416687`.

Cohen, L., C. Malloy, and Q. Nguyen (2020): "Lazy Prices," *Journal of Finance*, 75(3), 1371–1415.

Cong, L. W., T. Liang, and X. Zhang (2019): "Textual Factors: A Scalable, Interpretable, and Data-Driven Approach to Analyzing Unstructured Information," *Interpretable, and Data-driven Approach to Analyzing Unstructured Information (September 1, 2019)*.

Da, Z., J. Engelberg, and P. Gao (2014): "The Sum of All FEARS Investor Sentiment and Asset Prices," *Review of Financial Studies*, 28, 1–32.

Damstra, A., and M. Boukes (2021): "The Economy, the News, and the Public: A Longitudinal Study of the Impact of Economic News on Economic Evaluations and Expectations," *Communication Research*, 48(1), 26–50.

Firth, J. R. (1957): "A Synopsis of Linguistic Theory, 1930–1955," *Studies in linguistic analysis*.

Fogarty, B. J. (2005): "Determining Economic News Coverage," *International Journal of Public Opinion Research*, 17(2), 149–172.

Gentzkow, M., and J. Shapiro (2006): "Media Bias and Reputation," *Journal of Political Economy*, 114(2), 280–316.

Gentzkow, M., and J. M. Shapiro (2008): "Competition and Truth in the Market for News," *Journal of Economic Perspectives*, 22(2), 133–154.

——— (2010): "What Drives Media Slant? Evidence from U.S. Daily Newspapers," *Econometrica*, 78(1), 35–71.

Goidel, R. K., and R. E. Langley (1995): "Media Coverage of the Economy and Aggregate Economic Evaluations: Uncovering Evidence of Indirect Media Effects," *Political Research Quarterly*, 48(2), 313–328.

Greenwood, R., and S. G. Hanson (2013): "Issuer Quality and Corporate Bond Returns," *Review of Financial Studies*, 26(6), 1483–1525.

Hamilton, W. L., K. Clark, J. Leskovec, and D. Jurafsky (2016): "Inducing Domain-Specific Sentiment Lexicons from Unlabeled Corpora," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing*, vol. 2016, p. 595. NIH Public Access.

Hanley, K. W., and G. Hoberg (2019): "Dynamic Interpretation of Emerging Risks in the Financial Sector," *Review of Financial Studies*, 32(12), 4543–4603.

Hansen, S., M. McMahon, and A. Prat (2017): "Transparency and Deliberation Within the

FOMC: A Computational Linguistics Approach," *Quarterly Journal of Economics*, 133(2), 801–870.

Harrington, D. E. (1989): "Economic News on Television: The Determinants of Coverage," *Public Opinion Quarterly*, 53(1), 17–40.

Harvey, C. R. (1989): "Forecasts of Economic Growth from the Bond and Stock Markets," *Financial Analysts Journal*, 45(5), 38–45.

Hassan, T. A., S. Hollander, L. Van Lent, and A. Tahoun (2019): "Firm-level Political Risk: Measurement and Effects," *Quarterly Journal of Economics*, 134(4), 2135–2202.

Heston, S. L., and N. R. Sinha (2017): "News vs. Sentiment: Predicting Stock Returns from News Stories," *Financial Analysts Journal*, 73(3), 67–83.

Hiew, J. Z. G., X. Huang, H. Mou, D. Li, Q. Wu, and Y. Xu (2022): "BERT-based Financial Sentiment Index and LSTM-based Stock Return Predictability," available at arXiv: https://arxiv.org/abs/1906.09024.

Hirshleifer, D., J. Li, and J. Yu (2015): "Asset Pricing in Production Economies with Extrapolative Expectations," *Journal of Monetary Economics*, 76, 87–106.

Hoberg, G., and S. K. Moon (2017): "Offshore Activities and Financial vs. Operational Hedging," *Journal of Financial Economics*, 125(2), 217–244.

Hoberg, G., and G. Phillips (2010): "Product Market Synergies and Competition in Mergers and Acquisitions: A Text-Based Analysis," *Review of Financial Studies*, 23(10), 3773–3811.

——— (2016): "Text-Based Network Industries and Endogenous Product Differentiation," *Journal of Political Economy*, 124(5), 1423–1465.

Holbrook, A., J. Krosnick, P. Visser, W. Gardner, and J. Cacioppo (2001): "Attitudes toward Presidential Candidates and Political Parties: Initial Optimism, Inertial First Impressions, and a Focus on Flaws," *American Journal of Political Science*, 45(4), 930–950.

Krishnamurthy, A., and W. Li (2021): "Dissecting Mechanisms of Financial Crises: Intermediation and Sentiment," *Available at SSRN 3554788*.

Li, K., F. Mai, R. Shen, and X. Yan (2021): "Measuring Corporate Culture Using Machine Learning," *Review of Financial Studies*, 34(7), 3265–3315.

Lopez-Lira, A., and Y. Tang (2023): "Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models," available at SSRN: https://ssrn.com/abstract=4412788.

Lopez-Salido, D., J. C. Stein, and E. Zakrajek (2017): "Credit-Market Sentiment and the Business Cycle," *Quarterly Journal of Economics*, 132(3), 1373–1426.

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

Macaulay, A., and W. Song (2022): "Narrative-Driven Fluctuations in Sentiment: Evidence

Linking Traditional and Social Media," available at SSRN: https://ssrn.com/abstract=4150087.

Martin, G. J., and A. Yurukoglu (2017): "Bias in Cable News: Persuasion and Polarization," *American Economic Review*, 107(9), 2565–2599.

Maxted, P. (2022): "A Macro-Finance Model with Sentiment," *Review of Economic Studies*, forthcoming.

Mian, A., A. Sufi, and E. Verner (2017): "Household Debt and Business Cycles Worldwide," *Quarterly Journal of Economics*, 132(4), 1755–1817.

Mikolov, T., K. Chen, G. Corrado, and J. Dean (2013): "Efficient Estimation of Word Representations in Vector Space," available at https://arxiv.org/abs/1301.3781.

Mikolov, T., I. Sutskever, K. Chen, G. S. Corrado, and J. Dean (2013): "Distributed Representations of Words and Phrases and their Compositionality," in *Advances in Neural Information Processing Systems*, pp. 3111–3119.

Newey, K. K., and K. D. West (1987): "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Pennington, J., R. Socher, and C. Manning (2014): "GloVe: Global Vectors for Word Representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532–1543, Doha, Qatar. Association for Computational Linguistics.

Romer, C. D., and D. H. Romer (2004): "A New Measure of Monetary Shocks: Derivation and Implications," *American Economic Review*, 94(4), 1055–1084.

Shapiro, A. H., M. Sudhof, and D. J. Wilson (2022): "Measuring News Sentiment," *Journal of Econometrics*, 228(2), 221–243.

Sharir, O., B. Peleg, and Y. Shoham (2020): "The Cost of Training NLP Models: A Concise Overview," *arXiv preprint arXiv:2004.08900*.

Singla, S., and M. Mukhopadhyay (2022): "Gender Norms Do Not Persist But Converge Across Time," available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4183488.

Soroka, S. N. (2006): "Good News and Bad News: Asymmetric Responses to Economic Information," *Journal of Politics*, 68(2), 372–385.

——— (2012): "The Gatekeeping Function: Distributions of Information in Media and the Real World," *Journal of Politics*, 74(2), 514–528.

——— (2014): *Negativity in Democratic Politics: Causes and Consequences*. Cambridge University Press.

Stein, H. (1975): "Media Distortions: A Former Official's View," *Columbia Journalism Review*, 13(6), 37.

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

Tetlock, P. C., M. Saar-Tsechansky, and S. Macskassy (2008): "More than Words: Quantifying Language to Measure Firms' Fundamentals," *Journal of Finance*, 63(3), 1437–1467.

TVERSKY, A., AND D. KAHNEMAN (1974): "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185(4157), 1124–1131.

WATTENBERG, M. P. (1984): "The Decline of American Political Parties, 1952–1980," *American Political Science Review*, 49, 871–873.

WIELAND, J. F., AND M.-J. YANG (2020): "Financial Dampening," *Journal of Money, Credit and Banking*, 52(1), 79–113.

A Appendix

Figure A1: Sentiment Word Cloud

(a) Positive Sentiment

(b) Negative Sentiment

The figure shows a word cloud of the most positive and most negative words in panel (a) and panel (b), respectively.

Figure A2: Recession Word Cloud

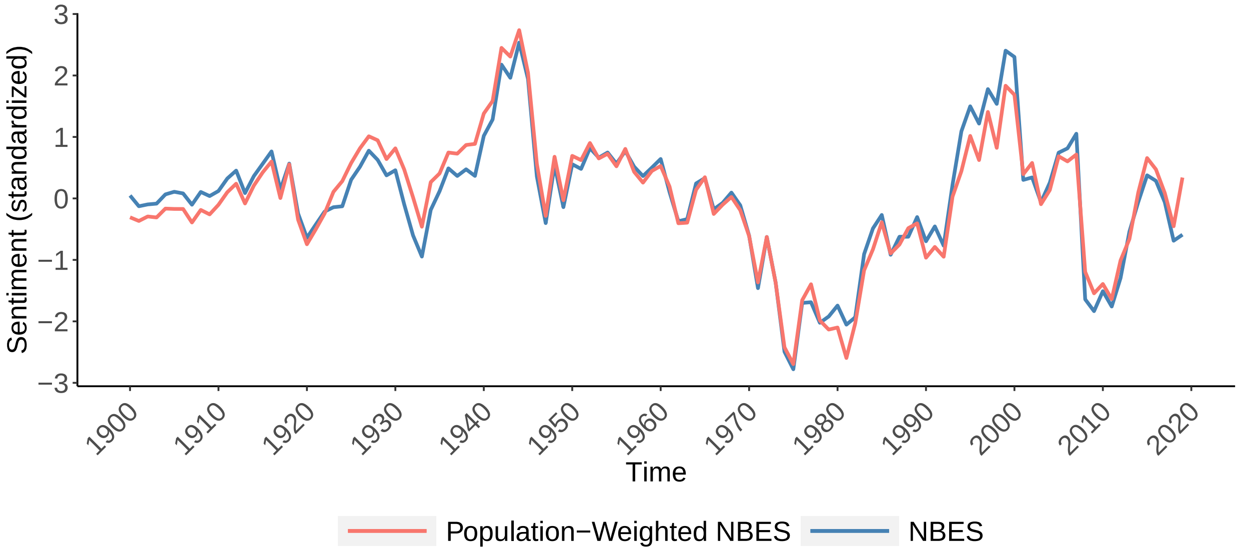
(a) Recession of 1866

(b) Great Depression

(c) Global Financial Crisis

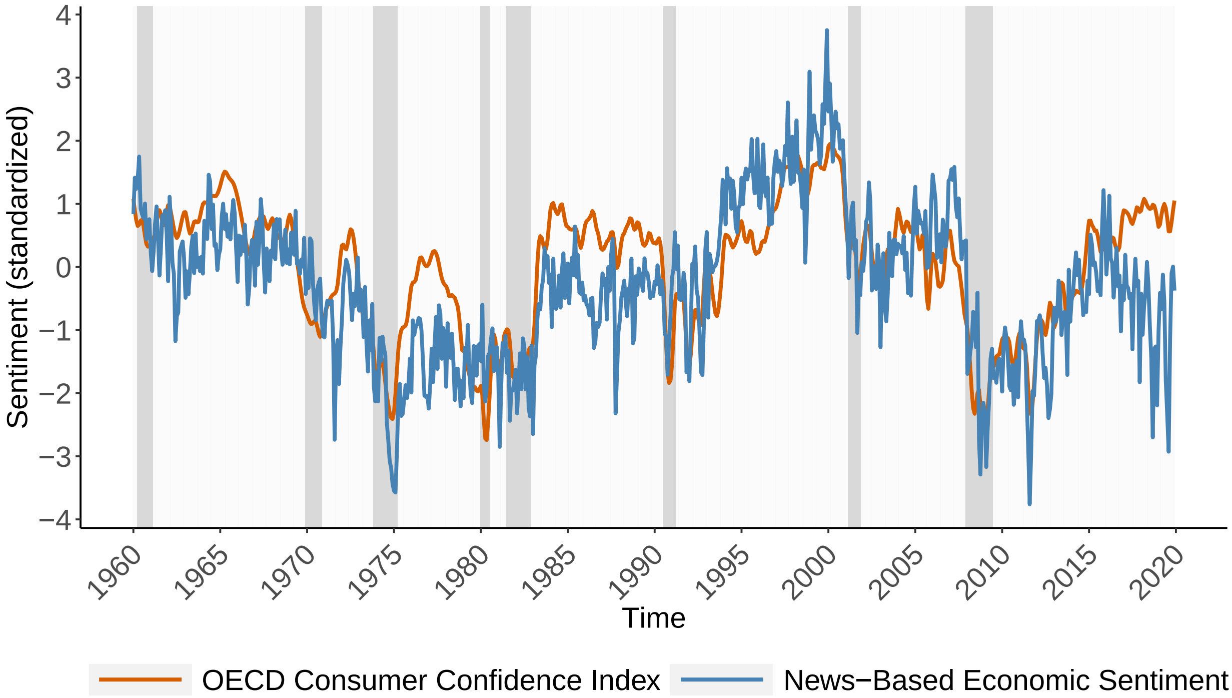
The figure shows the word cloud of three major recessions in our sample. The recession of 1866 in panel (a), the Great Depression in panel (b), and the Global Financial Crisis in panel (c).

Figure A3: National versus State-Aggregated National Economic Sentiment

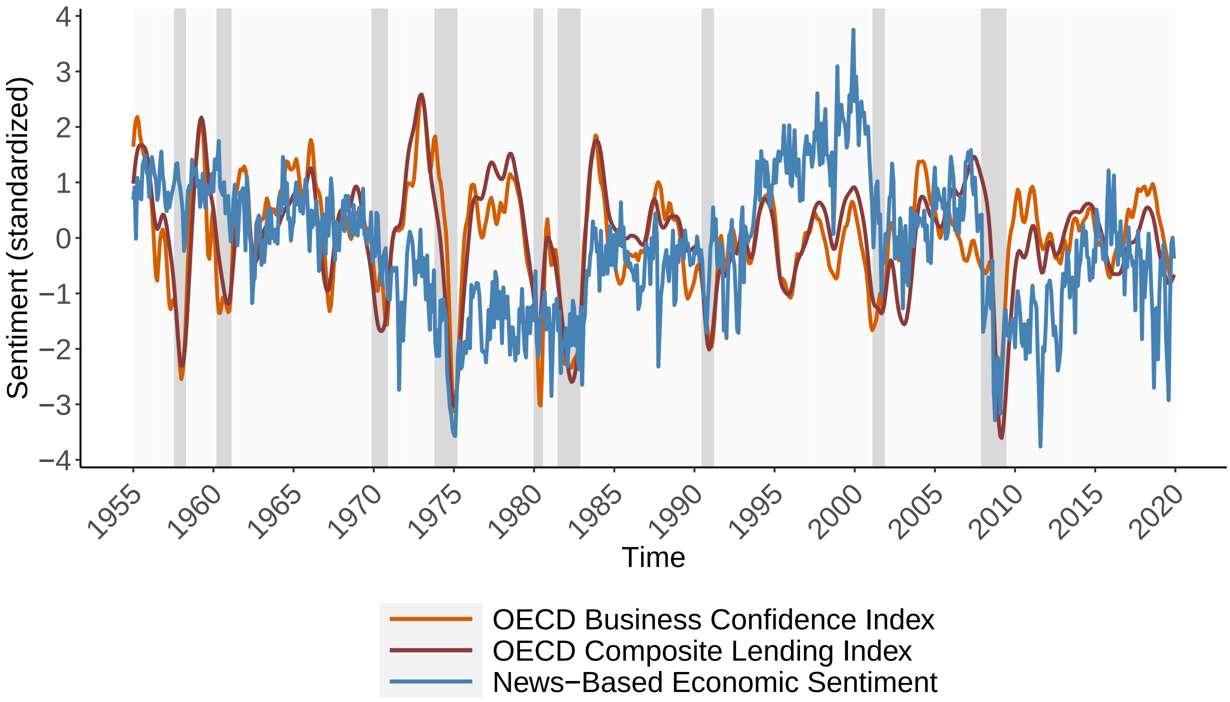


The figure shows the national (net) news-based economic sentiment (NBER) and population-weighted news-based national economic sentiment. The latter is calculated by aggregating economic sentiment across states, weighted by their population.

Figure A4: National Economic Sentiment and OECD Confidence Index



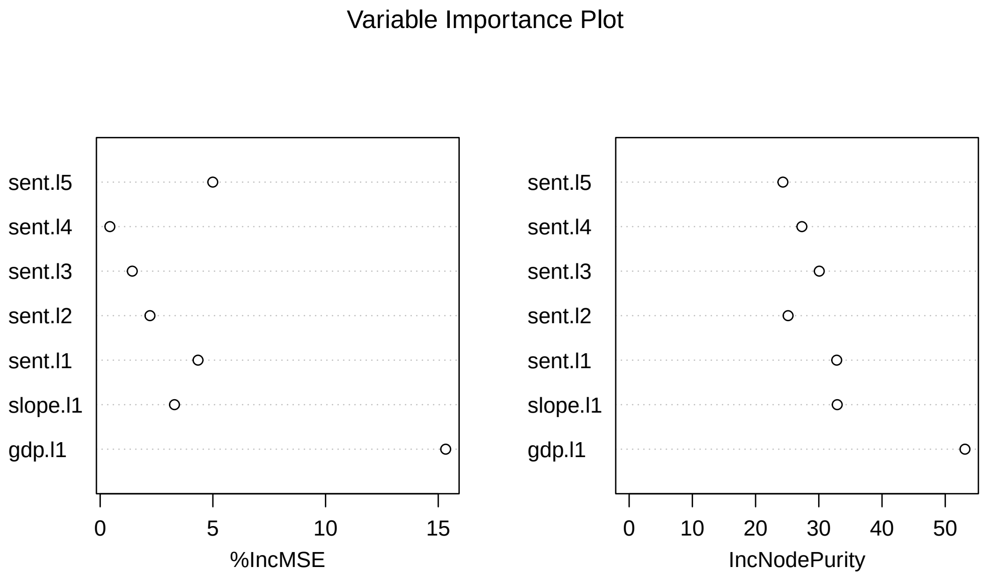
(a) News-Based Economic Sentiment and OECD Consumer Confidence Index



(b) News-Based Economic Sentiment and OECD Business Confidence Index

The figure shows the national news-based economic sentiment (NBES) OECD confidence indices. Panel (a) displays NBES and OECD Consumer Confidence Index, panel (b) plots NBES, OECD Business Confidence Index, and OECD Composite Lending Index. Each panel relies on the data sample available for all the corresponding time series.

Figure A5: Economic sentiment and GDP growth, Random Forest estimation



The figure shows variable importance in a nonlinear model predicting GDP growth, using the lagegd value of GDP growth, slope of the yield curve, and five lags of news-based economic sentiment and predictive variables. The left panel demonstrates the mean increase in accuracy (per variable), while the right panel shows variable-specific mean decrease in Mean Squred Error. The model is esimated on 500 trees, yielding the pseudo value of R-squared of 5.77%.

Table A1: Newspaper coverage

	Observations	Mean	P10	P25	Median	P75	P90
Panel A: National							
National	170	1,153.05	339.00	421.00	1,365.00	1,722.00	2,116.00
Panel B: State							
Alabama	149	8.27	2.00	4.00	8.00	12.00	15.00
Alaska	122	5.74	3.00	4.00	5.00	7.00	10.00
Arizona	158	16.58	4.00	8.00	16.00	25.00	30.00
Arkansas	145	8.34	2.00	4.00	7.00	12.00	16.00
California	169	35.14	8.00	14.00	32.00	48.00	70.00
Colorado	138	7.46	2.00	2.00	6.00	10.00	19.00
Connecticut	149	9.67	2.00	4.00	10.00	14.00	17.00
Delaware	74	4.97	2.00	4.00	5.00	6.00	7.00
District of Columbia	107	6.98	3.00	4.00	7.00	9.00	11.00
Florida	152	12.80	3.00	6.00	10.00	20.00	26.00
Georgia	128	36.98	19.00	22.00	29.50	52.00	69.00
Idaho	117	20.48	2.00	11.00	17.00	22.00	51.00
Illinois	170	79.14	13.00	36.00	85.00	115.00	138.50
Indiana	168	84.00	13.00	30.00	52.00	122.00	229.00
Iowa	170	116.01	13.00	23.00	141.50	189.00	229.00
Kansas	166	23.59	6.00	14.00	18.50	38.00	42.00
Kentucky	167	11.72	2.00	4.00	7.00	11.00	36.00
Louisiana	134	15.70	6.00	8.00	15.00	22.00	27.00
Maine	153	7.10	3.00	4.00	7.00	9.00	12.00
Maryland	170	20.26	8.00	12.00	21.00	26.00	32.00
Massachusetts	165	14.59	2.00	4.00	15.00	21.00	31.00
Michigan	170	23.96	6.00	9.00	29.00	37.00	42.50
Minnesota	132	21.00	6.00	10.50	16.50	30.00	45.00
Mississippi	163	14.14	2.00	4.00	12.00	18.00	35.00
Missouri	169	65.83	6.00	9.00	61.00	106.00	146.00
Montana	151	22.59	2.00	4.00	21.00	36.00	47.00
Nebraska	123	23.79	3.00	17.00	23.00	30.00	44.00
Nevada	115	8.71	4.00	5.00	8.00	12.00	14.00
New Hampshire	118	3.14	1.00	2.00	2.00	4.00	4.00
New Jersey	130	23.13	4.00	16.00	22.00	28.00	45.00
New Mexico	168	21.98	2.00	14.00	20.00	28.00	46.00
New York	170	55.89	4.50	13.00	77.00	84.00	91.50
North Carolina	170	24.61	8.00	10.00	26.50	37.00	43.50
North Dakota	91	12.11	2.00	3.00	4.00	20.00	33.00
Ohio	170	63.21	19.50	32.00	70.50	94.00	99.00
Oklahoma	110	17.25	3.00	8.00	15.00	24.00	35.50
Oregon	130	14.25	2.00	3.00	6.00	22.00	44.50
Pennsylvania	170	64.38	30.00	38.00	69.00	89.00	95.00
Rhode Island	143	3.13	2.00	2.00	3.00	4.00	5.00
South Carolina	170	4.49	2.00	3.00	4.00	6.00	7.50
South Dakota	113	7.09	2.00	4.00	4.00	10.00	16.00
Tennessee	128	3.45	1.00	2.00	3.00	6.00	6.00
Texas	170	117.88	7.00	29.00	67.50	199.00	307.00
Utah	127	11.04	2.00	5.00	12.00	16.00	20.00
Vermont	96	8.06	2.00	4.50	8.00	11.00	13.00
Virginia	170	15.57	4.00	7.00	13.50	24.00	29.50
Washington	166	8.72	2.00	4.00	6.00	11.00	20.00
West Virginia	169	11.41	3.00	4.00	7.00	16.00	28.00
Wisconsin	169	34.67	10.00	20.00	35.00	44.00	62.00
Wyoming	110	9.27	2.00	4.00	5.50	10.00	26.00

The table shows the distribution of the newspaper coverage across years at national and state levels.

Table A2: News sentiment measures: Summary statistics

	Observations	Mean	StDev	P10	P25	Median	P75	P90
Econ Sentiment	680	0.00	1.00	-1.26	-0.62	0.04	0.63	1.18
Econ Sentiment(Unresidualize)	680	-0.00	1.00	-1.27	-0.41	0.20	0.65	1.04
Non-econ Sentiment	680	0.00	1.00	-1.70	-0.27	0.39	0.65	0.84
Δ Econ Sentiment	679	-0.00	0.46	-0.51	-0.27	0.00	0.28	0.50
Δ Econ Sentiment(Unresidualize)	679	-0.00	0.30	-0.35	-0.18	-0.01	0.17	0.36
Δ Non-econ Sentiment	679	-0.01	0.17	-0.22	-0.12	0.00	0.10	0.20

The table shows the summary statistics for the economic sentiment, unresidualized economic sentiment, and non-economic sentiment news measure. The table also presents the summary statistics of the changes in each of these variables.

Table A3: National economic and non-economic sentiment and GDP

	$\Delta ln(GDP_t)$			
	I	II	III	IV
$\Delta ln(GDPPerCapita_{t-1})$	0.356***	0.332***	0.357***	0.334***
	(5.79)	(5.35)	(5.74)	(5.35)
$Term\ Spread_{t-1}$	0.072*	0.062	0.072**	0.063*
	(1.97)	(1.64)	(1.97)	(1.65)
$\Delta Sent(Notresidualized)_{t-1}$		0.356**		0.366**
		(2.52)		(2.57)
$\Delta Non-EconSent_{t-1}$			-0.083	-0.180
			(-0.32)	(-0.67)
Observations	290	290	290	290
Adjusted R-squared	0.13	0.15	0.13	0.15

The table reports a linear predictive model for GDP per capita. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

Table A4: News-Based Economic Sentiment and GDP Growth, 1947Q1–2019Q4

	$\Delta ln(GDP_t)$				
	I	II	III	IV	V
$\Delta ln(GDP_{t-1})$	0.212***	0.152	0.163***		0.140***
	[0.010]	[0.030]	[0.010]		[0.010]
$Term\ Spread_{t-1}$	0.073*		0.018		0.027*
	[0.059]		[0.327]		[0.089]
$\Delta Sent_{t-1}$		0.030	0.019	0.097***	0.051**
		[0.307]	[0.277]	[0.010]	[0.020]
$\Delta Sent_{t-2}$				0.010	-0.001
				[0.505]	[0.693]
$\Delta Sent_{t-3}$				0.039	0.019
				[0.149]	[0.297]
$\Delta Sent_{t-4}$				0.033	0.018
				[0.139]	[0.277]
$\Delta Sent_{t-5}$				0.013	0.026
				[0.406]	[0.109]
Observations	290	290	290	290	290
Pseudo R-squared	-0.050	-0.057	0.018	-0.031	0.043

The table reports variable importance permutation-based p-values for a nonlinear predictive model for GDP per capita, estimated via Random Forest. p-valued based on the permutation approach of Altmann, Tolosi, Sander, and Lengauer (2010) are reported in squared bracked. Significance levels are indicated by * (10%), ** (5%), and *** (1%). Estimation is done via the **ranger** package in **R**.

Table A5: News-Based Economic Sentiment and GDP Growth, 1947Q1–2019Q4: With twelve lags of GDP growth

	$\Delta ln(GDP_t)$					
	I	II	III	IV	V	VI
$\Delta ln(GDP_{t-1})$	0.330***	0.331***	0.306***	0.307***	0.305***	0.303***
	(5.12)	(4.93)	(4.71)	(4.57)	(4.26)	(4.18)
$\Delta ln(GDP_{t-2})$	0.146**	0.152**	0.148**	0.154**	0.137**	0.148**
	(2.17)	(2.30)	(2.22)	(2.34)	(2.16)	(2.27)
$\Delta ln(GDP_{t-3})$	-0.066	-0.052	-0.043	-0.032	-0.071	-0.069
	(-1.06)	(-0.80)	(-0.70)	(-0.49)	(-1.14)	(-1.11)
$\Delta ln(GDP_{t-4})$	-0.027	-0.017	-0.036	-0.027	-0.046	-0.060
	(-0.38)	(-0.25)	(-0.53)	(-0.40)	(-0.67)	(-0.88)
$\Delta ln(GDP_{t-5})$	-0.098	-0.090	-0.100	-0.093	-0.072	-0.068
	(-1.46)	(-1.30)	(-1.49)	(-1.34)	(-1.02)	(-0.96)
$\Delta ln(GDP_{t-6})$	0.074	0.092	0.078	0.094	0.108	0.117*
	(1.07)	(1.35)	(1.15)	(1.40)	(1.56)	(1.79)
$\Delta ln(GDP_{t-7})$	-0.072	-0.056	-0.075	-0.061	-0.068	-0.061
	(-1.27)	(-1.01)	(-1.35)	(-1.10)	(-1.30)	(-1.18)
$\Delta ln(GDP_{t-8})$	-0.056	-0.045	-0.058	-0.047	-0.048	-0.065
	(-0.76)	(-0.62)	(-0.79)	(-0.66)	(-0.65)	(-0.90)
$\Delta ln(GDP_{t-9})$	0.065	0.070	0.082	0.085	0.088	0.082
	(1.04)	(1.14)	(1.30)	(1.38)	(1.49)	(1.41)
$\Delta ln(GDP_{t-10})$	0.085	0.099	0.080	0.092	0.085	0.097
	(1.36)	(1.59)	(1.31)	(1.53)	(1.39)	(1.54)
$\Delta ln(GDP_{t-11})$	-0.006	0.018	0.001	0.022	0.038	0.045
	(-0.08)	(0.23)	(0.01)	(0.28)	(0.51)	(0.61)
$\Delta ln(GDP_{t-12})$	-0.167**	-0.156**	-0.172**	-0.162**	-0.147**	-0.164**
	(-2.32)	(-2.20)	(-2.54)	(-2.41)	(-2.27)	(-2.49)
$TermSpread_{t-1}$		0.073**		0.066*	0.063*	0.064*
		(1.98)		(1.78)	(1.67)	(1.69)
$\Delta Sent_{t-1}$			0.219***	0.209***	0.252***	0.185**
			(2.98)	(2.85)	(3.00)	(1.98)
$\Delta Sent_{t-2}$					0.010	0.023
					(0.11)	(0.25)
$\Delta Sent_{t-3}$					0.246**	0.243**
					(2.19)	(2.17)
$\Delta Sent_{t-4}$					0.235**	0.237**
					(2.20)	(2.23)
$\Delta Sent_{t-5}$					0.047	0.012
					(0.50)	(0.13)
$\Delta Sent_{t-1} \times Q4Dummy$						0.113
						(0.67)
Q4 Dummy						-0.210*
						(-1.79)
F statistic ($\Delta Sent$ (t-2 to t-n))					2.539	2.823
P-value ($\Delta Sent$ (t-2 to t-n))					0.040	0.026
Observations	279	279	279	279	279	279
Adjusted R-squared	0.18	0.19	0.20	0.20	0.22	0.22

The table reports a linear predictive model for GDP per capita after controlling for twelve lags of GDP Growth. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

Table A6: National Economic Sentiment and GDP Growth, 1947Q1–2019Q4: With DP Ratio

	$\Delta ln(GDP_t)$			
	I	II	III	IV
$\Delta ln(GDP_{t-1})$	0.357***	0.336***	0.328***	0.331***
	(5.80)	(5.49)	(5.19)	(5.16)
Term Spread$_{t-1}$	0.076**	0.067*	0.064*	0.064*
	(2.12)	(1.82)	(1.72)	(1.71)
$DPRatio_{t-1}$	0.054	0.047	0.043	0.042
	(0.42)	(0.37)	(0.33)	(0.32)
$\Delta Sent_{t-1}$		0.241***	0.287***	0.251**
		(2.60)	(2.73)	(2.22)
$\Delta Sent_{t-2}$			-0.017	-0.007
			(-0.17)	(-0.07)
$\Delta Sent_{t-3}$			0.224**	0.223**
			(2.04)	(2.01)
$\Delta Sent_{t-4}$			0.152	0.151
			(1.35)	(1.34)
$\Delta Sent_{t-5}$			0.007	-0.017
			(0.08)	(-0.18)
$\Delta Sent_{t-1} \times Q4Dummy$				0.050
				(0.26)
Q4 Dummy				-0.133
				(-1.06)
F statistic ($\Delta Sent$ (t-2 to t-n))			2.495	2.429
P-value ($\Delta Sent$ (t-2 to t-n))			0.043	0.048
Observations	290	290	290	290
Adjusted R-squared	0.13	0.15	0.16	0.16

The table reports a linear predictive model for GDP per capita after controlling for the lagged dividend-to-price ratio. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

Table A7: State-Aggregated National Economic Sentiment and GDP Growth, 1947Q1–2019Q4

	$\Delta ln(GDP_t)$								
	I	II	III	IV	V	VI	VII	VIII	IX
$\Delta ln(GDP_{t-1})$	0.355***	0.356***		0.335***	0.337***			0.332***	0.334***
	(5.91)	(5.79)		(5.66)	(5.57)			(5.37)	(5.33)
$TermSpread_{t-1}$		0.072*			0.064*			0.056	0.057
		(1.97)			(1.72)			(1.51)	(1.53)
$\Delta Sent_{t-1}$			0.378***	0.303***	0.291***	0.496***	0.451***	0.354***	0.305***
			(3.92)	(3.12)	(3.00)	(4.33)	(3.78)	(3.25)	(2.69)
$\Delta Sent_{t-2}$						0.181	0.186	-0.033	-0.029
						(1.64)	(1.64)	(-0.32)	(-0.28)
$\Delta Sent_{t-3}$						0.379***	0.376***	0.265**	0.262**
						(2.75)	(2.70)	(2.14)	(2.10)
$\Delta Sent_{t-4}$						0.349**	0.348**	0.203	0.202
						(2.48)	(2.45)	(1.57)	(1.56)
$\Delta Sent_{t-5}$						0.134	0.111	0.030	0.002
						(1.17)	(0.96)	(0.29)	(0.02)
Q4 Dummy							-0.110		-0.139
							(-0.98)		(-1.17)
$\Delta Sent_{t-1} \times Q4Dummy$							0.107		0.098
							(0.45)		(0.45)
F statistic ($\Delta Sent$ (t-2 to t-n))						2.569	2.602	3.283	3.230
P-value ($\Delta Sent$ (t-2 to t-n))						0.038	0.036	0.012	0.013
Observations	290	290	290	290	290	290	290	290	290
Adjusted R-squared	0.12	0.13	0.04	0.15	0.15	0.07	0.07	0.17	0.17

The table reports a linear predictive model for GDP per capita. State-aggregated national economic sentiment is calculated as the population-weighted average of the state-level sentiments. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

Table A8: Economic and non-economic sentiment during 167 years of economic history

	$\Delta ln(GDP_t)$					
	I	II	III	IV	V	VI
$\Delta ln(GDP_{t-1})$	0.147		0.094	-0.184	0.216	0.122
	(1.29)		(0.85)	(-1.45)	(1.48)	(0.82)
$\Delta Sent(Notresidualize)_{t-1}$		3.100***	2.633***	2.381	3.305*	1.721**
		(3.23)	(3.36)	(0.95)	(1.69)	(2.54)
$\Delta Non-EconSent_{t-1}$		-3.557**	-3.854**	-4.022	-3.722	0.178
		(-2.10)	(-2.30)	(-1.67)	(-1.04)	(0.10)
Period	Full	Full	Full	1850-1914	1915-1980	1981-2017
Observations	167	167	167	63	66	38
Adjusted R-squared	0.02	0.05	0.06	0.02	0.10	0.24

The table reports a linear predictive model for GDP per capita. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. Significance levels are indicated by * (10%), ** (5%), and *** (1%).

Table A9: National economic sentiment and durable consumption

	$\Delta ln(Cons^{dur}_{t-1})$				
	I	II	III	IV	V
$\Delta ln(Cons^{dur}_{t-1})$	-0.102			-0.113	-0.118
	(-1.16)			(-1.28)	(-1.39)
$Term\ Spread_{t-1}$	0.452***			0.435***	0.429***
	(2.80)			(2.69)	(2.67)
$\Delta Sent_{t-1}$		0.549*	0.645	0.555	0.658
		(1.67)	(1.63)	(1.58)	(1.48)
$\Delta Sent_{t-2}$			0.129		0.178
			(0.39)		(0.49)
$\Delta Sent_{t-3}$			0.595		0.581
			(1.25)		(1.10)
$\Delta Sent_{t-4}$			0.295		0.354
			(0.65)		(0.80)
$\Delta Sent_{t-5}$			0.293		0.333
			(0.93)		(0.99)
F statistic (all $\Delta Sent$)			0.605		0.495
P-value (all $\Delta Sent$)			0.659		0.739
Observations	290	290	290	290	290
Adjusted R-squared	0.03	0.00	-0.00	0.03	0.03

The table reports a linear predictive model for durable consumption. t-statistics based on Newey and West (1987) standard errors are reported in parentheses. Significance levels are indicated by * (10%), ** (5%), and *** (1%).