# Inflation expectations and the news: A historical study using sentiment analysis and word embedding

Dissertation

In part completion of the MSc in Applied Social Data Science

Department of Methodology

London School of Economics and Political Science

Candidate: Michel Dilmanian

Supervisor: Pablo Barberá

Word count: 9,569

8 August, 2019

**Abstract**

Developments in the United States economy in the late 1970s and early 1980s have become one of the most widely studied events in post-war macroeconomic history. The period known as the Volcker disinflation has given rise to a substantial body of research examining the determinants of, and changes in, public inflation perceptions. Most of these studies suggest that the news media plays an important role in delivering and interpreting inflation-related information for the public. However, existing research has been restricted to newspaper data starting in the 1980s, and has typically focused on standard automated dictionary methods to construct variables of media coverage, making it difficult to measure the media's representation of any one economic quantity in particular. Using a rare corpus of *New York Times* economic news articles dating back to 1947, I present evidence that the relationship between inflation-related news coverage and inflation expectations has weakened in the decades after the Volcker disinflation. I also use word embedding, an emerging technique in natural language processing, to show that semantic associations of the word "inflation" have shifted quite dramatically in the periods before and after the Volcker disinflation. The latter analysis highlights some of the challenges associated with producing robust time-series of semantic shifts when relying on relatively small corpora. With this challenge in mind, I conclude by reviewing recent research that outlines several promising strategies for using word embedding to produce temporal variables.

# Contents

# 1 Introduction

Developments in the United States economy in the late 1970s and early 1980s have become one of the most important events studied by macroeconomists in post-war history (Goodfriend and King, 2005). Many economists see the appointment of Paul Volcker to the Chair of the Federal Reserve in 1979 as the start of a regime change: his tenure is believed to have ended the 1970s period of "stagflation"—that particularly toxic combination of low economic growth and high inflation that puts economic policy-making into a dilemma—thus transitioning the US from a period of high and volatile inflation to low and stable inflation (Bordo et al., 2007). This transition is such a prominent event of economic history that it is often referred to simply as the Volcker disinflation.

Most economists attribute Volcker's success in fighting inflation in large part to the decline in the public's inflation expectations during his tenure (Bernanke, 2007; Yellen, 2015). Indeed, expected inflation is a key variable in the analysis of business cycles and monetary policy, and has therefore prompted a large body of research investigating its roots. Most economists agree that households formulate assessments of future inflation not only by observing the prices of goods around them, but also by gathering information from the news (Coibion, Gorodnichenko, et al., 2018). By this account, the media acts both as a source of pure statistics and as an interpreter of professional opinions regarding the future paths of such statistics.

While the relationship between the media and inflation expectations has been examined in some detail, less attention has been paid to how this relationship may have changed over time, especially since the economic turmoil of the 1970s. This is partly a result of data availability. Full textual data from news articles is not easily collected and its historical availability is limited, meaning that so far (and to the best of the author's knowledge), such research has been applied only to the period beginning roughly in 1980.

There are also methodological restrictions in the existing literature. The vast majority of studies examining media-effects on inflation expectations focuses on two media variables: the volume of coverage and its tone (or sentiment). The latter is typically calculated based on automated dictionary methods, which count the number of words associated with certain categories. While these methods can be useful, recent advances in neural-network modelling have made possible the use of much more sophisticated natural language processing techniques at manageable computational costs. My paper explores one such technique, word embedding, as a means of measuring semantic changes in very specific concepts (in this case, inflation) over time.

The limited availability of textual data and the methodological restriction in the literature present two interesting questions. First, has the relationship between the news media and inflation expectations changed in the decades after the Volcker disinflation? Second, can word embedding be used as an alternative to dictionary methods when measuring this relationship?

Using a corpus of *New York Times* print stories about the economy going back to 1947 and a simple model of inflation expectations, I present evidence that the relationship between inflation-related news coverage and inflation expectations has weakened in the decades after the Volcker disinflation. My analysis also reveals the challenges of distinguishing between the tone of news with respect to inflation, and the tone of economic news more generally, when using dictionary-based sentiment analysis. To overcome this problem, I propose the use of word embedding to track the semantic relationship between "inflation" and positive/negative words over time. While a granular temporal application of this strategy appears to require meaningfully more data, I show that semantic associations of the word "inflation" have shifted quite dramatically in the period before and after the Volcker disinflation.

The remainder of this paper is structured as follows. I begin with an overview of the literature covering effects of the news media on economic perceptions. I then provide historical and theoretical context for my analysis, including a brief overview of the Volcker disinflation, the determinants of inflation expectations, and the role of the news media. Next, I outline the data and methods used for analysis, followed by its results. I conclude the paper by identifying potential solutions to the challenges faced in my analysis and describing areas for further research.

## 2 Literature review

The relationship between news coverage and public perceptions of the economy has been examined in a substantial body of research, much of it in the context of voting behavior. The focus of this literature has been on two variables: the tone or content of coverage, and its reporting volume. Importantly, a number of studies have found that these variables affect economic perceptions even after controlling for measured economic conditions. Doms and Morin (2004) show that consumer sentiment in the United States is affected both by the media's pure transmission of economic data and professional opinions, as well through the tone and volume of reporting. Their work suggests that volume of coverage also operates through a second channel, leading consumers to update their expectations about the economy more frequently when volume is higher. Using a custom survey instrument, Ansolabehere, Meredith, and Snowberg (2014) find that assessments of the national unemployment rate are less responsive to a respondent's personal employment situation when he or she has watched television news—i.e. sociotropic assessments are updated rationally in response to the news. Blood and Phillips (1995) go a step further and suggest that headline recession news has a causal impact on consumer sentiment. To form the causal argument, the authors use vector autoregressions that control for coincident and leading indicators of actual economic activity.

It also appears that media effects are stronger during recessions. Hollanders and Vliegenthart (2011) find that negative news is associated with lower consumer confidence in Netherlands between 1990 and 2009, with a stronger effect during the financial crisis. Similarly, using a three-wave panel survey, Boomgaarden et al. (2011) find a strong impact of news coverage on changes in economic perceptions during the 2008-2009 financial crisis (also in the Netherlands).

But not all research finds support for the media-effect hypothesis. De Boef and Kellstedt (2004) analysis of New York Times stories published between 1981 and 2001 yields no explanatory power of media tone on consumer confidence, though it does find an effect on economic approval ratings. Hopkins, Kim and Kim (2017) question the causal direction of much of the aforementioned research finding media effects. Using Granger causality tests, they find evidence that economic perceptions typically lead, rather than follow, national media coverage. S. N. Soroka, Stecula, and Wlezien (2015) argue that the effect goes in both directions: news coverage both affects, and is affected by, public economic perceptions.

The link between media and economic perceptions has also been examined with respect to one specific economic variable (the focus of this paper): inflation. Here, the body of work is noticeably smaller, but the main findings are comparable: both the reporting intensity and the tone (or content) of the news affect inflation expectations of households. Early strides were made in 2003, when Carroll, 2003 constructed a model in which households form inflation expectations by incorporating projections of professional forecasters via the news. The study's main finding is in line with results from much of the work on economic perceptions: reporting intensity affects how well consumers absorb projections of professional forecasters from the media. Carroll also provides empirical evidence for what is referred to as "inattention", namely that households only periodically incorporate information from the news into their expectations. Inattention creates "stickiness" in aggregate household expectations, and this has meaningful consequences for macroeconomic outcomes.

Around the same time as Carroll's work was published, Mankiw, Reis, and

Wolfers (2003) showed that there is substantial disagreement among economic agents' inflation expectations, and that the amount of disagreement moves over time with other economic aggregates. Their findings highlighted the need for models of information acquisition and processing and steered subsequent research towards determining the causes of such disagreement. For example, Lamla and Maag (2012) leverage a German media database to find that disagreement in household inflation forecasts depends on both the heterogeneity of news stories and reporting intensity—especially when inflation is rising. Using similar data, Lamla and Lein (2014) investigate the role of quantity, tone and content of news in the accuracy of consumers' inflation expectations. The authors find that the quantity of news improves accuracy of inflation expectations, but that this effect is conditional on the tone of the news (badly-toned news reduces accuracy while neutral news improves accuracy). Their conclusions also corroborate the aforementioned work by Carroll (2003).

There are, however, caveats to the idea that the media affects consumers' perceptions of prices. Ansolabehere, Meredith, and Snowberg (2014) found no systematic differences in assessments of gasoline prices between those that do and those that do not watch national television news. The theory is that the price of gasoline is directly observable, therefore assessments of gas prices should be less sensitive to what is seen or heard in the news. Another question concerns the validity of revised forecasts: Pfajfar and Santoro (2013) find only weak support for the idea that individuals revise their inflation expectations based on the news. More importantly, the authors conclude that any revised expectations move away from, rather than towards, those of professional forecasters. Here again, the content of the news is important: the counterintuitive effect is stronger for unfavorable news (news about higher prices) than for favorable news (news about lower prices).

The literature reviewed above reveals several gaps. First, none of the studies about inflation expectations reach back further in time than 1980, the year in which

Lexis-Nexis coverage of the New York Times and Washington Post begins. This restriction misses a major period in US economic history—the stagflation of the 1970s—and denies us the opportunity to draw conclusions about changes in media effects before and after this period. Second, the studies rely on traditional dictionary methods to produce measures of media tone, while recent developments in computational methods enable the use of a much more sophisticated set of models. This paper aims to begin to fill both of these gaps. In the sections that follow, I first provide a contextual basis for my analysis, followed by an overview of the methods and the data to be used.

# 3 Theory and context

## 3.1 The Volcker disinflation

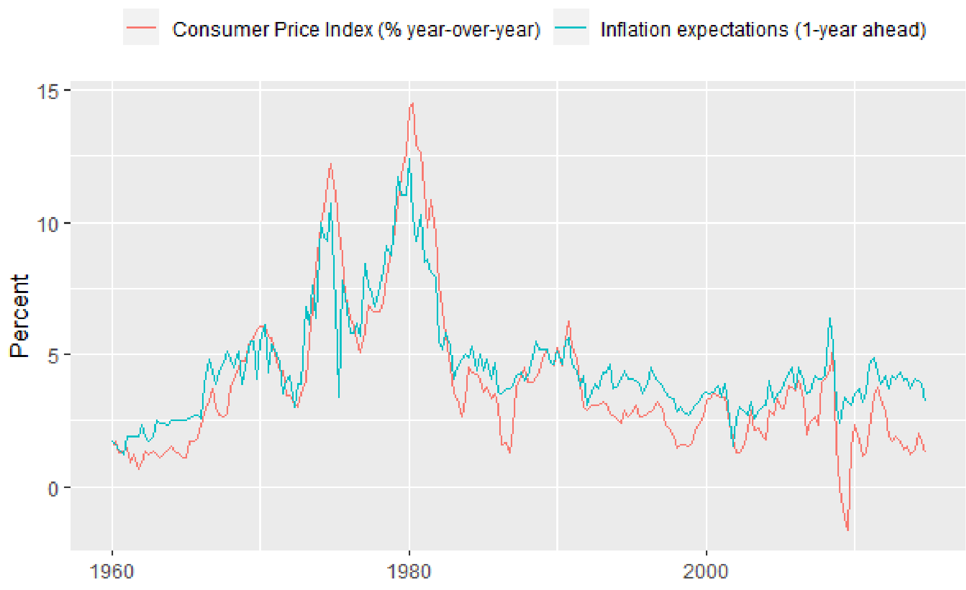
The conventional narrative of the Volcker disinflation is as follows.[1] In the early 1970s, the oil crisis caused a surge in US consumer price inflation that Federal Reserve policymakers were either unable or unwilling to control, given political pressures and the trade-off between lowering inflation and promoting economic growth. Inflation rose again in the run-up to and during the oil crisis caused by the Iranian Revolution of 1979. By the time Paul Volcker was appointed to the chair of the Federal Reserve in August 1979, consumer price inflation was running close to 12 percent year-over-year, compared to less than 3 percent in the early 1970s. Volcker publicly announced a break from past policies and began raising the federal funds rate sharply to combat inflation. This policy contributed to a brief recession in 1980, during which the Federal Reserve reversed its policy in a move that appears to have been motivated by political pressure (Bordo et al., 2007). However, by late 1980, the Volcker Fed resumed its tight monetary policy, raising the federal funds rate to an unprecedented 20 percent and allowing it to fall only commensurately with infla-

[1] See Bordo et al. (2007) for a more detailed account.

tion. As a result, consumer price inflation fell from almost 13 percent in late 1980 to under 4 percent by late 1982, at the cost of a deep recession in 1981-82. It is this last campaign of restrictive monetary policy that reduced inflation sustainably and convinced the public of the Federal Reserve's credibility and political independence. In the 30 years since, inflation in the US and most advanced economies has fallen further and has remained remarkably subdued.

Most of the academic literature attributes Volcker's success in fighting inflation in large part to the decline in the public's inflation expectations (Bernanke, 2007; Yellen, 2015). As William Silber (Silber, 2012) notes in his biography of Volcker, the Federal Reserve under his chairmanship succeeded in beating the "inflationary virus" that had "wormed its way into people's brains" (p. 133). One of the consequences of this is that household expectations have become more "anchored", meaning that they should be less sensitive to changes in observed prices or rhetoric in the news because individuals trust the central bank to manage inflation adequately. The raw data certainly corroborate this idea: surveys of inflation expectations, as well as actual inflation, exhibit a significant structural downtrend starting in the 1980s (see Figure 1).

Figure 1: Actual inflation and inflation expectations



More generally, the role of public expectations in macroeconomic theory and policy cannot be understated—anticipated inflation is a key variable in the most widely used macroeconomic models (Williamson, 2018; Mankiw, 2016). The centrality of expectations in economic models calls for a theory of how individual citizens form their forecasts of future price growth. While the exact mechanisms of the expectations-formation process remain poorly understood, there is general support for a number of factors. I outline these factors next.

## 3.2 Inflation expectations

The economics literature has identified four broadly accepted determinants of household inflation expectations. The four determinants, as discussed in Coibion, Gorodnichenko, et al. (2018), are observed inflation (what the authors call "shopping experience"), priors and perceptions, the media and knowledge about monetary policy. I discuss each of these factors, as well as how I intend to control for them in the first part of my analysis, below.

1. <u>Observed inflation</u>: As Coibion et al. point out, "the strongest predictor of a household's inflation forecast is typically what they believe inflation has been over the recent past" (p. 103). Most commonly, consumers form opinions about past inflation by observing prices of food at the supermarket or gasoline at the pump. In my model, I control for this determinant by including the percent change in the Consumer Price Index (as measured by the Bureau of Labor Statistics) over the previous year—this includes both food and energy prices but also other consumer goods, weighted according to their proportion in a typical consumer basket. While a number of researchers have documented disproportionately strong effects of food (Clark and Davig, 2008) and gasoline (Wong, 2015; Coibion and Gorodnichenko, 2015) on US households' inflation expectations, I do not attempt to model this explicitly, given that the focus of my analysis is the news media.

2. Priors and perceptions: Observed inflation, however, can vary across households. For some time now, studies have found that consumers have heterogeneous beliefs about what inflation has been in the past, despite the availability of official statistics (Jonung, 1981). Such "disagreement" tends to be higher in low-inflation countries, where misestimating price-changes is less consequential for the typical consumer's saving and consumption decisions (Cavallo, Cruces, and Perez-Truglia, 2017). Instead of attempting to control for this variable in my analysis, I rely on the Consumer Price Index as a proxy. The rationale here is that the variable's omission would only be problematic if it varied over time, and by explicitly splitting the sample period based on the Volcker disinflation, I largely avoid this complication.

3. Media: As discussed in more detail in the following section, the news media are a major source of information for most households and should be a candidate for determining inflation expectations. In essence, the news media acts as a mediator of a complex set of economic circumstances. It does so in two ways. First, it transmits economic news and data releases that many consumers otherwise would not have been exposed to. Second, it interprets the reported information by relaying the views of professional forecasters, government officials, political pundits and the like. To investigate the media's impact on inflation expectations in my model, I calculate the volume and tone of articles in a corpus of *New York Times* articles, which is presented below.

4. Knowledge about monetary policy: This refers to the general public's awareness of the central bank's objectives and/or credibility in pre-empting excessively high inflation. Households living in a country with a less reliable central bank are likely to have more volatile inflation expectations, while a credible and transparent central bank fosters stability of inflation expectations. As Coibion et al. explain, the literature has typically found that monetary policy

is typically poorly understood in countries with low and stable inflation environments. Since there is no cross-country comparison in my analysis, I do not include any regressors that attempt to measure it in my regressions. Now, of course the US might be considered a high-inflation country before 1983 and a low-inflation country from 1983 onwards. However, since the sample is explicitly split into these two periods, it is reasonable to believe that knowledge of monetary policy is not changing meaningfully throughout either one of the periods, and this suggests that the risk of omitted-variable bias is low.

## 3.3 The role of the media

The US economy is large and complex. National economic fluctuations are the result of millions of heterogeneous households making microeconomic decisions with imperfect information. In making these decisions, households must form expectations about both local and national economic variables and assess how the developments in these variables will affect their own economic lives. And individual households only directly observe a minute portion of the full set of economic realities at any point in time.

Inflation can be a particularly esoteric concept. There exist a number of theories that attempt to model its behavior, and a large number of variables are relevant in assessing the predictions of these theories; these include economic growth, unemployment, wage growth, and fiscal and monetary policies. Constantly monitoring a plethora of individual economic indicators would be an inefficient task that is unlikely to pass the cost-benefit test for most individuals. Even so, there is no shortage of disagreement over the path of future inflation among professional economists, who follow these indicators profusely (Mankiw, Reis, and Wolfers, 2003).

So, besides observing the price of food at the supermarket and the cost of gasoline at the pump, how might consumers form expectations about the future path of inflation? The good news is that individuals need not have a firm grasp of economic theories in order to formulate opinions about current and future economic performance (MacKuen, Erikson and Stimson, 1992).

This is where the news media can play a role. As De Boef and Kellstedt (2004) note, "the national media is the quintessential mediator" (p. 640). Various parts of the news media consistently reach a large and diverse segment of the population, distilling complex events into more digestible stories. The media also transmits projections from professional forecasters, helping consumers form their own opinions about the economy and inflation. As the authors highlight, there are two main steps through which the media can influence economic perceptions. The first is purely informational: the media transmits data and information that consumers would otherwise not learn about. In the case of news about inflation, a number of news items are important: general economic growth, unemployment, wage growth, fiscal and budgetary decisions of the government, and central bank actions, to name a few. The second step follows naturally from the first: almost by definition, the media interprets the information it transmits. This can include relaying opinions of economic forecasters, political pundits and government officials; but it is also inherent in any editorial process, as subtle differences in word choice and semantics will alter the way a particular news story is perceived.

In the literature, these steps are typically modelled using volume (step 1) and tone (step 2) of news articles. I follow the same logic in the first part of my analysis. In the second part, I explore measuring semantic shifts using word embedding, a cutting-edge language modelling technique.

# 4 Data

## 4.1 New York Times corpus

### 4.1.1 Data collection

The corpus for this project was provided by the authors of Barberá et al. (2019), who collected 68,000 *New York Times* stories about the US economy published between 1947 and 2014 from the ProQuest Archive of Historical Newspapers and used optical character recognition (OCR) software to convert the collected PDF files into machine-readable text. The authors also collected similar data directly as plain text files from ProQuest Newsstand for the period 1980-2014 and subsequently merged the two datasets into one. The corpus contains the following seven variables: a unique article ID, the full text of the article, the number of words in the article, text of the article headline, newspaper section, page number and publish date. The collection method yielded only articles from the print edition of the paper. Checks were made to confirm that no articles are duplicated.

The *New York Times* stories were collected using a Boolean regular expression search (keyword search). As the authors note in their 2019 paper (Barberá, Boydstun, Linn, Nagler & Mcmahon, 2019), the goal of identifying articles on the US economy was pursued as follows:

> To generate a sample of economic news stories using a keyword search, we downloaded all articles from the *New York Times* with any of the following terms: employment, unemployment, ination, consumer price index, GDP, gross domestic product, interest rates, household income, per capita income, stock market, federal reserve, consumer sentiment, recession, economic crisis, economic recovery, globalization, outsourcing, trade decit, consumer spending, full employment, average wage, federal decit, budget decit, gas price, price of gas, deation, existing home sales, new home sales, productivity, retail trade gures, wholesale prices AND United States. We used a filter to remove any articles that mentioned a country name, country capital, nationality or continent name that did NOT also mention U.S., U.S.A. or United States in the headline or first 1000 characters of the article (Schrodt 2011). (p. 7)

The selected corpus presents both advantages and challenges. It is clearly suited to the goal of examining inflation perceptions over a long period of time—given that it contains articles from only one newspaper, the corpus provides reliability and consistency for time-series analysis. The *New York Times* is arguably the most established paper in the US, with journalistic practices that have remained largely stable over time. Moreover, it is unlikely the corpus would have any important blind-spots, given the paper's ambition and ability to cover everything that matters.

By the same token, being restricted to one news outlet means that the results of the textual analysis will likely be more reflective of progressive segments of the US population (readers of the *New York Times*) rather than the median individual. But this bias is likely mitigated by the central position of the newspaper in the US (and global) media landscape. For example, in their study on agenda-setting effects of *New York times* content, Winter and Eyal (2016) find evidence supporting the idea of the paper as a "national media indicator" (p. 381). Neuman (1990) similarly describes the *New York Times* as a "recognized source of record" that "serves as a model for editors and reporters of other newspapers as well as the wire services" (see also (Crouse, 1973)). The paper has also been found to have an impact on television news coverage (Brown, 1971).

While it is impossible to rule out any demographic bias from using the *New York Times*, I propose that any bias is outweighed by the benefit of having consistency over a 70-year sample period. The focus of this analysis is how changes in inflation expectations are affected by changes in economic news coverage over time. The assumption here is that variability in volume and tone of economic news articles in the *New York Times* over time broadly mirrors that of most other outlets in the US, despite any ideological gap that may exist at any point in time.

### 4.1.2 Descriptive statistics

To focus the analysis on inflation perceptions rather than overall economic perceptions, I subset the original corpus to include only articles that refer to inflation.

I do this using regular expressions functions that select only articles that include one or more of the following terms: “inflation”, “price index”, “deflator”. Roughly one third of the original corpus, or about 23,500 articles, match this description. The terms “price index” and “deflator” are included to catch articles that are more technical in nature. The former is intended to identify articles about the Consumer Price Index or the Personal Consumption Expenditures Price Index, the two most widely followed national indicators of price inflation. The latter aims to identify articles about the Implicit GDP Price Deflator and its variants, published by the Bureau of Economic Analysis in its GDP accounts—these are similar to the consumer price indices but include goods exported to other countries while excluding those imported from other countries. In practice, the choice of including the latter two terms bears little consequence for the analysis: of the 23,469 articles in the filtered corpus, only 1,048 (or 4 percent) contain the terms “price index” or “deflator” without also mentioning “inflation”.

The terms “price” and “prices” could also have been used to filter for articles related to inflation. This would have resulted in over half of the raw corpus (about 35,500 articles) being included in the analysis. I opt for the more restrictive filter since the term “price” can be found in any number of (very microeconomic) contexts that would be unlikely to prompt the typical reader to think about broader (macroeconomic) inflation. It would also have included articles about prices of securities, such as individual stocks, that are largely unrelated to inflation.

Summary statistics for the original corpus and the filtered corpus are shown in Table 1. The filtered corpus will be the basis for this analysis. It contains almost 23,500 unique articles, with a mean length of 893 words. Article length is relatively variable, with standard deviation of over 600 words, but this value is inflated by a severely right-skewed distribution. The vast majority of articles in the filtered corpus (over 90 percent) are between 100 and 1,500 words in length. Filtering the original corpus does not meaningfully skew basic article-level characteristics: the statistics above are broadly similar in the original corpus.

Table 1: Descriptive statistics for *New York Times* corpora

| | **Original corpus** | **Filtered corpus** |
|---|---|---|
| Total number of articles | 68,001 | 23,469 |
| Total number of words | 52,259,224 | 20,959,204 |
| **Article length (# of words)** | | |
| Mean | 769 | 893 |
| Std. dev. | 562 | 622 |
| Minimum | 2 | 8 |
| Maximum | 11,106 | 11,106 |
| **Articles per quarter** | | |
| Mean | 250 | 86 |
| Std. dev. | 97 | 58 |
| Minimum | 86 | 7 |
| Maximum | 573 | 375 |
| **Words per quarter** | | |
| Mean | 192,130 | 77,056 |
| Std. dev. | 79,152 | 53,914 |
| Minimum | 76,630 | 4,195 |
| Maximum | 539,585 | 332,705 |

The corpus contains a fairly large number of articles per calendar quarter, which is the unit of analysis for this paper. The mean quarter contains about 86 articles about inflation, which provides about 77,000 words for textual analysis. The minimum article count is 7 (roughly 4,200 words) and a maximum of 375 over 330,000 words). Most of the low article counts, however, occur before 1960, which is the start of the sample period for the main portion of this analysis. Post-1960, the minimum quarterly article count is 13, amounting to 9,057 total words available for analysis.

## 4.2 Variables and modelling strategy

The first goal of my paper is to examine whether US inflation expectations have become less responsive to news coverage following the Volcker disinflation. To do this, I regress inflation expectations of US households on actual (prior) inflation, a dummy for energy crises, and the volume and tone of inflation-related news coverage.

Most of the publicly available data used in these regressions is drawn from University of Michigan's Survey of Consumers, henceforth the Michigan Survey. This is a monthly survey of US households containing roughly 50 core questions about consumer attitudes and expectations regarding both sociotropic and ecotropic economic variables. Each month, at least 500 households are interviewed via telephone using a rotating panel design.

The following variables are used in the regression analysis. The response variable is the mean expected 12-month-ahead percent change in consumer prices (i.e., 12-month inflation expectations) from the Michigan Survey.[2] Past observed inflation is proxied using percent change in the Consumer Price Index over the 12 months leading up to the survey period.[3] To control for the unusual spikes in inflation expectations seen during the two energy crises of the 1970s, I include a series from the Michigan Survey that asks whether respondents have recently heard news about an energy crisis; although it is not strictly binary, this series acts as a dummy variable of sorts. Lastly, to investigate media effects, I calculate the volume and tone of inflation-related news articles from the *New York Times* corpus and include these as independent variables. The resulting regression model is as follows:

$$\pi_t^e = \beta_{1t} CPI_t + \beta_{2t} crisis_t + \beta_{3t} volume_t + \beta_{4t} tone_t,$$

where $\pi^e$ is expected inflation over the next year from the Michigan Survey, $CPI$ is the percent change in the Consumer Price Index over the previous year (including the current period), $crisis$ refers to the index of news about energy crises from the Michigan Survey, and $volume$ and $tone$ are calculated based on the *New York Times* corpus. The subscript $t$ denotes the calendar quarter of each observation.

[2]The survey also asks about expected inflation over the next 5 to 10 years, but this measure is not available as far back as the 12-month metric.

[3]See https://www.bls.gov/cpi/home.htm. Specifically, I use the Consumer Price Index for all Urban Consumers: All Items from the Bureau of Labor Statistics. I use the series that is not seasonally adjusted, since that is the rate that consumers can observe in practice. In any case, the seasonal factors for this series are negligible, particularly when it comes to year-over-year changes, making this choice inconsequential.

The model is estimated via OLS on quarterly data between 1960 and 2014. I use quarterly, rather than monthly, aggregation based on availability of the inflation expectations variable. Using the monthly series would have restricted my analysis to begin in 1978, omitting the majority of the stagflationary period and reducing statistical power for the pre-1983 sample. I standardize the variables before estimation and therefore exclude an intercept from the model.

# 5 Methods

## 5.1 Sentiment analysis

The first part of my analysis aims to extend existing research linking media variables to sociotropic economic perceptions, instead examining the media's impact on inflation expectations in particular. The media variables that have been consistently featured in the literature are volume (i.e. reporting intensity) and tone. This makes intuitive sense: as the volume of news on a certain economic aggregate increases, the average reader is more likely to take note of the information contained in the news. The ultimate effect on the reader, however, should still be dependent on the tone of the reporting: it is unclear a priori whether a greater volume of news on inflation must necessarily affect expectations if the tone of the reporting is strictly neutral.

Researchers interested in measuring the tone of a document have traditionally relied on automated dictionary methods, likely because this appears to be the most straightforward and intuitive technique (Grimmer/Steward). Dictionary methods count the occurrence of pre-defined lists of words (dictionaries) associated with certain categories. In this paper, I focus on categories relating to tone or sentiment. More specifically, the measure of tone I use counts the occurrence of words in the "positive" and "negative" categories of the Lexicoder Sentiment Dictionary (Mohammad and Turney, 2010), each with around 3,000 entries. I choose the Lexicoder dictionary for three reasons. First, while the its word patterns are not subject-specific, there are demonstrated applications of its use in analyzing both economic and noneconomic news coverage. In particular, it has been used to analyze the impact of media tone on sociotropic economic perceptions (Young and S. Soroka, 2012). Second, it has readily available preprocessing scripts aimed at improving the dictionary's performance, freely available on the Lexicoder website (I discuss these scripts in the section on preprocessing below). Third, in addition to positive and negative words, it enables the user to identify negations: positive words preceded by a negation and negative words preceded by a negation. As Young and Soroka have shown, incorporating negations into the calculation of tone can improve performance.

The calculation of tone is straightforward. The quantity of interest is net tone, the difference between the number of positive and negative word patterns after accounting for negations. To adjust for varying document sizes, this quantity is divided by the total number of words in the document. More formally, I calculate net tone as follows:

$$tone_i = \frac{c_i^p - c_i^n + c_i^{nn} - c_i^{np}}{N_i},$$

where $c_i^p$, $c_i^n$, $c_i^{nn}$ and $c_i^{np}$ refer to the count of positive, negative, negative-negative and negative-positive word patterns, $N_i$ refers to the count of all words in the document and $i$ denotes the subscript for the document (in this case a calendar quarter).

## 5.2 Word embedding

The analytical focus of the second portion of this paper is word embedding. Word embedding is a statistical technique that extracts semantic meanings of words by considering the context in which they are found in large text corpora. The central idea underlying the theory of embedding is that "you shall know a word by the company it keeps", famously stated by linguist J. R. Firth in the late 1950s (Firth, 1957). Instead of counting frequencies, word embedding represents words as a position in latent multidimensional space. The result is that the angle between words with similar meanings is smaller in vector space, while words with opposing meanings have a larger angle between them (they are orthogonal in the most "dis-similar" case). Resulting vectors are typically used as features in natural language processing and machine-learning applications.

The idea that the meaning of a word is derived from its context has been around for decades; it is a key insight in the distributional hypothesis in linguistics (Harris, 1954). However, word embedding has undergone something of a renaissance since 2013, when scientists at Google Inc. developed a new algorithm to generate embedding vectors using neural networks (Mikolov et al., 2013). The speed with which embedding vectors can now be estimated has given rise to a growing, but still small, body of academic research employing word embedding in fields as diverse as medicine, social science and linguistics (examples of this literature are provided below).

The novelty of the implementation method warrants a brief explanation of how it functions. Word vectors are typically learned using one of two types of neural network architectures: continuous bag-of-words (CBOW) or Skip-gram. The models work in similar ways but have different advantages. CBOW aims to predict a given word based on its context without considering the order of the context words. Skip-gram, meanwhile, uses the given word to predict its context and weighs closer words more heavily that distant words. The latter model is more computationally expensive but typically produces more accurate vectors for infrequent words. Given the importance of relatively infrequent positive and negative words for this analysis, I use the Skip-gram implementation throughout this paper.

There is scant evidence of the use of word embedding to examine the media's portrayal of inflation, but results from applications in fields outside of economics are encouraging. Word embedding has been used extensively for sentiment analysis, for example in Tweets (Tang, Wei, et al., 2014) and user-generated reviews of products and films (Tang, Qin, and Liu, 2015). In medicine, the technique has been used to obtain mentions of adverse drug reactions in social media (Nikfarjam et al., 2015). An important motivator of this project is a study by Caliskan, Bryson, and Narayanan (2017). Using word embeddings and texts from the internet, the authors show that text corpora "contain recoverable and accurate imprints of our historic biases". For example, the study replicates findings related to gender (female names are more closely associated with family than careers), race (pleasantness associations of European-American vs. African-American names) and other areas of human bias.

My analysis aims to test whether word embedding can uncover changes in the way the media discusses inflation over time. If so, it might be used as an alternative to dictionary methods to help analyze the determinants of inflation expectations in future research. Given the shortage of prior work that uses word embedding temporally to examine inflation perceptions in the news, this part of my analysis is exploratory and descriptive in nature.

My methodology for testing word embedding is as follows. I first convert the raw *New York times* corpus into a time series by concatenating the texts of all articles based on the year and calendar quarter in which they were published. This yields a time series with 272 year-quarter observations, each of which is a chunk of text. Preprocessing steps for the textual data are outlined in the next section. After preprocessing, I separately train the `word2vec` model on the text contained in each of the 272 quarters. I use the first 100 dimensions (`word2vec`'s default) and train the model through multiple iterations to stabilize the word vectors (`iter = 20`). For the window of text considered around each word, I use the default size of 10. For each quarter, this procedure produces a matrix of 100-dimensional vectors, one for each unique word contained in the corpus in that quarter.

Once the model is trained, I compute the cosine similarity between vector-representations of "inflation" and various negative and positive words in each of these 272 quarters. I use the same positive and negative words as in the sentiment analysis described above, namely those from the Lexicoder Sentiment Dictionary (LSD). Cosine similarity is based on the angle between two vectors of an inner product space and is defined as follows (Han, Kamber, and Pei, 2012):

$$similarity(\vec{x}, \vec{y}) = cos(\theta) = \frac{\vec{x} \cdot \vec{y}}{\|x\| \, \|y\|},$$

where $\vec{x}$ and $\vec{y}$ are vectors, $\theta$ measures the angle between them, $\|x\|$ is the Euclidean norm of vector $\vec{x}$, and the numerator represents the inner product of the two vectors in question. This metric has two main advantages. First, always ranges between -1 and 1, with a value of -1 indicating perfectly opposite vectors, 1 meaning the vectors are directionally identical and 0 representing orthogonal (unrelated) vectors. Second, it is robust to the magnitudes of the vectors being measured, reflecting only their directions (in the way that Euclidean distance is not).

I then calculate the mean of pairwise cosine similarities between "inflation" and negative words, and "inflation" and positive words. The final result in each case is a quarterly time-series that I refer to as the word embedding index. Higher values in the negative word embedding index suggest a greater semantic similarity between "inflation" and negative words, while higher values in the positive word embedding index indicate a greater semantic similarity between "inflation" and positive words.

## 5.3 Preprocessing

Running sentiment analysis and training the word embeddings model required different sets of preprocessing steps. To prepare the text for sentiment analysis, I used preprocessing modules written specifically for the Lexicoder Sentiment Dictionary by Luxon (2017) with the main goal of disambiguating and contextualizing commonly occurring dictionary words. [4] Specifically, I applied seven of the modules provided. The first removes contractions (e.g. "won't" becomes "will not"). The second excludes a number of dictionary words that would otherwise be counted as positive or negative, based on surrounding punctuation (e.g. "Well," at the beginning of the sentence should not be considered a positive word). The third and fourth remove punctuation from acronyms and abbreviations to ensure consistency across the corpus. Since identification of dictionary words would be impaired when words end in punctuation, I use a fifth module to separate punctuation from words. The sixth simplifies phrases that use negations (e.g. "not very" becomes "not") to improve their identification in the dictionary. Finally, the seventh removes another set of variations of dictionary words that should not be counted (e.g. "well" within the phrase "may very well" is removed).

Preprocessing for word embeddings is less involved, since the model estimates vector representations based on the lexical context of each word. Therefore, only three steps were undertaken to prepare the text: I first removed all punctuation, then removed excess whitespace characters (including newline and tab characters), and finally converted all words to lower case.

For both the sentiment analysis and word embeddings approaches, the preprocessed text was aggregated by calendar quarter. In other words, articles published in the same calendar quarter were concatenated to produce one continuous string of words for each quarter.

Finally, some examples in the literature (e.g. Soroka et al. 2015) remove articles with less than 100 words prior to analysis. Results in this paper are based on analysis that does not exclude such articles since they comprise only 0.3 percent of observations in the corpus. The alternative results do not change meaningfully if they are excluded and are available upon request. Moreover, a random selection of some of these articles suggests that they include stories with relevant information

[4] See also Young and Soroka (2012) for a description of preprocessing modules for the Lexicoder dictionary.

that could reasonably affect the reader's inflation expectations.

# 6 Results and discussion

## 6.1 Sentiment analysis

The first step in exploring the effects of media coverage on inflation expectations was constructing the variables based on the tone and volume of inflation coverage in the *New York Times* corpus. Figure 2 plots the quarterly count of inflation-related news coverage between the first quarter of 1960 and the last quarter of 2014. It is immediately visible that this measure might be strongly correlated to inflation expectations. The plot exhibits two trends: an ascent from the beginning of the sample period until the early 1980s, and a downward trend from the early 1980s until the end of the sample period. The peak of the series[5] occurs in the first quarter of 1980. This makes sense: there is broad agreement that Volcker's inflation-fighting actions began bearing fruit in the early 1980s, and this would be consistent with the news media drifting away from the topic of inflation once the worst had passed.

The same observations cannot be made for the tone of coverage, illustrated in Figure 3 (higher values indicate more negative tone). To begin with, the series is quite noisy, and trends in the 1970s and 80s do not seem to systematically reflect the inflation narrative during that period. However, more general economic events are represented: tone plummets during the recessions of the early 1980s, early 1990s and 2008 and swings wildly around the period of the September 11, 2011 attacks.

Pairwise correlations will help contextualize broader results. Table 2 shows Pearson's correlation coefficients between the variables considered in my model. The variables considered are 1-year inflation expectations from the Michigan Survey, year-over-year growth in the Consumer Price Index (CPI), net news heard about

[5]The spike in the second half of 1974 may reflect news stories about the imminent end of the crisis, as US diplomacy in the near east had progressed

Figure 2: Volume of news (by calendar quarter)

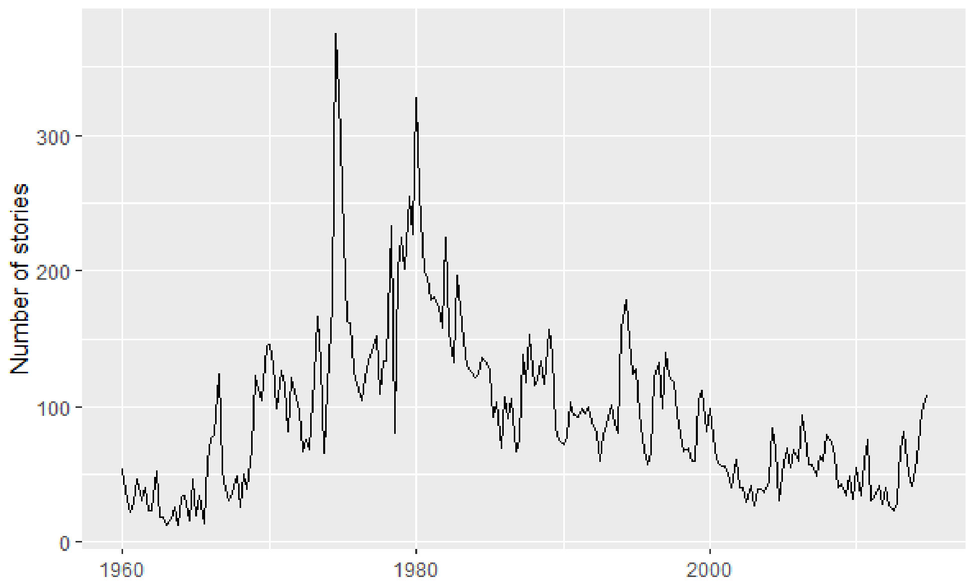
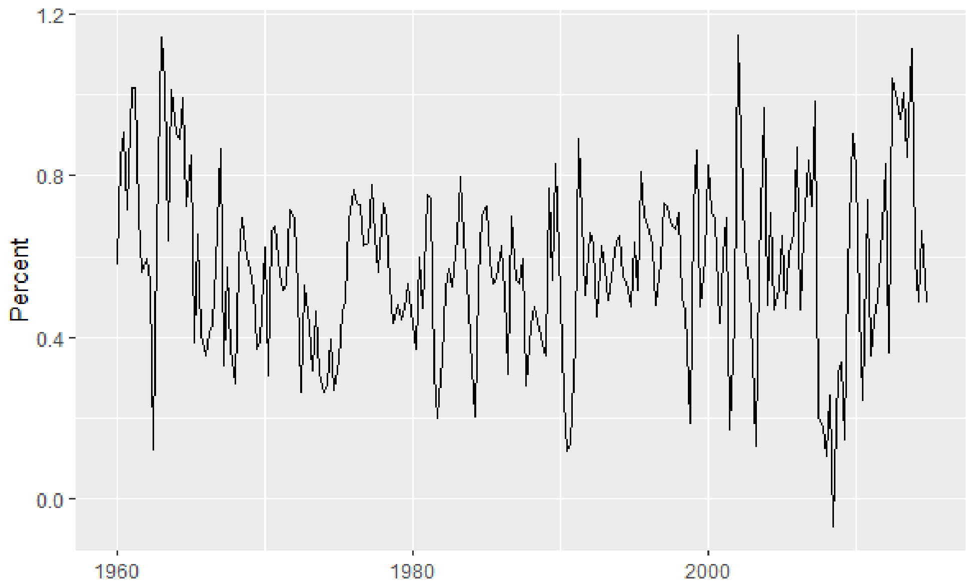



Figure 3: Net tone of news (negative minus positive; detrended)



a crisis from the Michigan Survey and the two media measures mentioned above. Results are reported for two periods: 1960-1982, and 1983-2014, with the split made to coincide with the the end of the Volcker disinflation. Inflation expectations are highly correlated with actual prior inflation and the volume of news coverage ($p > 0.8$) up until Volcker's appointment to the chairmanship of the Federal Reserve. They are also negatively correlated with the tone of news coverage, and while this correlation is somewhat lower in magnitude ($r = -0.6$), it is statistically different from zero at the 5 percent significance level. From the beginning of 1983 onwards, the correlations weaken across the board, but particularly so for tone. This is in line with intuition: if Volcker truly managed to "break the back of inflation" and anchor households' outlooks for future price changes, inflation expectations should be more stable and less responsive to news coverage or temporary changes in observed consumer prices.

Table 2: Pairwise correlations of variables in levels

*Statistically significant at the 5% level

| **1960-1982** | Infl. Exp. | Crisis news | CPI inflation | Volume |
|---|---|---|---|---|
| Infl. Exp. | 1.00 | | | |
| Crisis news | 0.56* | 1.00 | | |
| CPI inflation | 0.90* | 0.50* | 1.00 | |
| Volume | 0.80* | 0.32* | 0.87* | 1.00 |
| Tone | -0.59* | -0.35* | -0.62* | -0.55* |

| **1983-2014** | Infl. Exp. | Crisis news | CPI inflation | Volume |
|---|---|---|---|---|
| Infl. Exp. | 1.00 | | | |
| Crisis news | 0.13 | 1.00 | | |
| CPI inflation | 0.70* | 0.14 | 1.00 | |
| Volume | 0.42* | -0.06 | 0.41* | 1.00 |
| Tone | -0.04 | -0.26* | -0.01 | 0.23* |

Since some of the variables appear to be nonstationary in level terms, I also tabulate correlations for variables in first differences (which are stationary) in Table 3.[6] A similar picture emerges here, though with some notable changes. Up until the

[6]Stationarity was tested using Dickey-Fuller and Augmented Dickey-Fuller procedures.

start of the Volcker era, changes in inflation expectations are positively correlated with changes in actual inflation and the volume of news coverage ($p = 0.35$ and $p = 0.22$, respectively), but the relationship with news tone disappears. From 1983 onwards, changes in inflation expectations have a statistically meaningful relationship only with actual inflation; changes in both media volume and tone appear to lose their effect.

Table 3: Pairwise correlations of variables in first differences

*Statistically significant at the 5% level

| **1960-1982** | Infl. Exp. | Crisis news | CPI inflation | Volume |
|---|---|---|---|---|
| Infl. Exp. | 1.00 | | | |
| Crisis news | 0.45* | 1.00 | | |
| CPI inflation | 0.35* | 0.24* | 1.00 | |
| Volume | 0.22* | -0.06 | 0.28* | 1.00 |
| Tone | -0.05 | -0.07 | -0.12 | -0.05 |

| **1983-2014** | Infl. Exp. | Crisis news | CPI inflation | Volume |
|---|---|---|---|---|
| Infl. Exp. | 1.00 | | | |
| Crisis news | 0.18* | 1.00 | | |
| CPI inflation | 0.44* | 0.11 | 1.00 | |
| Volume | 0.06 | -0.05 | 0.13 | 1.00 |
| Tone | -0.04 | -0.19* | -0.08 | -0.05 |

To control for other variables that may affect inflation expectations, I include the two news variables in a model regressing inflation expectations on observed inflation and net news heard about an oil crisis. The regressors were chosen based on four broadly accepted determinants of inflation expectations discussed in Coibion, Gorodnichenko, et al. (2018) and in section 3 of this paper.

Tables 4 and 5 summarize results from four variations of this model. I run the model in both levels and differences, in each case for both the 1960-82 and 1983-2014 sample period. Since some of the existing literature has found interaction effects between volume and tone of coverage (Lamla and Lein, 2014; Lamla and Maag, 2012), I include the product of the two measures in addition to each measure in isolation. I do so in separate regressions given collinearity between these regressors.

Regression results introduce a few caveats to the simple correlations observed above. In level terms and after controlling for other variables, news volume remains a meaningful covariate in explaining inflation expectations both before and after the Volcker era. News tone, however, only appears to be relevant before 1983. Although this dovetails neatly with the qualitative story about the Volcker era, uncertainty about which variables are stationary leaves open the possibility that these relationships are spurious. I therefore run the same regressions on the variables in first differences. Unsurprisingly, differencing the variables weakens measured relationships. News volume is still associated positively with inflation expectations until the Volcker disinflation, although only at the 10 percent significance level ($p = 0.08$); afterwards, it loses its effect. Tone seems to be irrelevant in either time period.

Table 4: Regression results: variables in levels

*, **, ***: 10, 5, 1% signif. level

| | Dependent var: inflation expectations | | | | | |
|---|---|---|---|---|---|---|
| | 1960-1982 | | | 1983-2014 | | |
| CPI | 0.732*** | 0.882*** | 0.872*** | 0.559*** | 0.620*** | |
| | (0.083) | (0.043) | (0.042) | (0.056) | (0.051) | |
| Crisis news | 0.151*** | 0.136*** | 0.147*** | 0.239** | 0.207* | 0.608*** |
| | (0.048) | (0.038) | (0.037) | (0.107) | (0.109) | (0.053) |
| Volume | 0.186** | | | 0.106** | | 0.217** |
| | (0.082) | | | (0.042) | | (0.109) |
| Tone | | -0.085* | | | 0.013 | |
| | | (0.050) | | | (0.026) | |
| Tone*volume | | | 0.127*** | | | -0.041 |
| | | | (0.043) | | | (0.038) |
| Observations | 92 | 92 | 92 | 128 | 128 | 128 |
| Adj. R-squared | 0.881 | 0.878 | 0.885 | 0.696 | 0.681 | 0.684 |
| Resid. std. err. | 0.496 | 0.502 | 0.487 | 0.262 | 0.268 | 0.267 |
| D.f. | 89 | 89 | 89 | 125 | 125 | 125 |

Looking beyond the news, observed inflation is clearly the most consistent factor in all of these regressions: a one-standard-deviation move in year-over-year consumer price inflation is associated with a change in inflation expectations of anywhere

Table 5: Regression results: variables in first differences

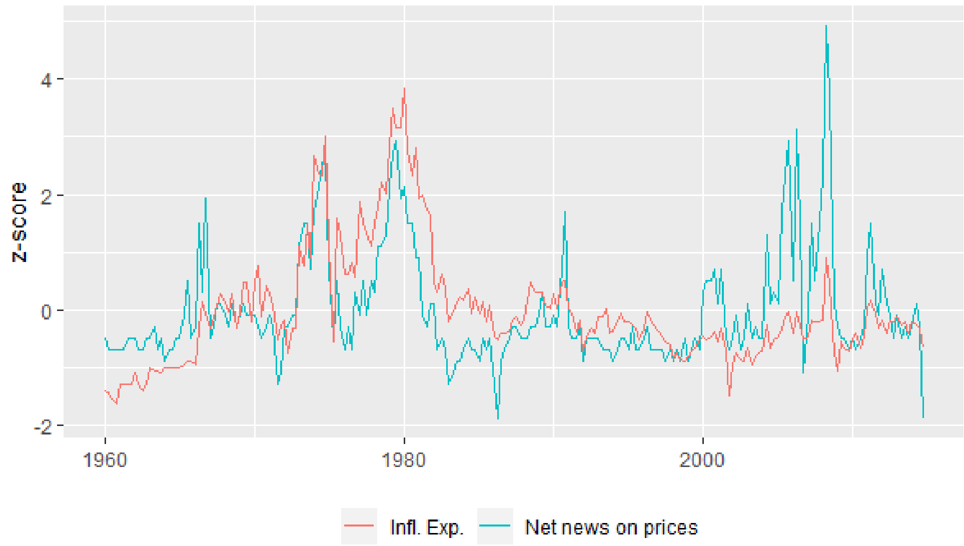
*, **, ***: 10, 5, 1% signif. level

| | Dependent var: inflation expectations | | | | | |
|---|---|---|---|---|---|---|
| | 1960-1982 | | | 1983-2014 | | |
| CPI | 0.351*** | 0.407*** | 0.397*** | 0.282*** | 0.284*** | 0.285*** |
| | (0.122) | (0.122) | (0.121) | (0.053) | (0.053) | (0.053) |
| Crisis news | 0.335*** | 0.323*** | 0.328*** | 0.432* | 0.442* | 0.419 |
| | (0.080) | (0.081) | (0.082) | (0.260) | (0.264) | (0.261) |
| Volume | 0.167* | | | 0.012 | | |
| | (0.093) | | | (0.079) | | |
| Tone | | -0.005 | | | 0.012 | |
| | | (0.144) | | | (0.048) | |
| Tone*volume | | | 0.087 | | | -0.03 |
| | | | (0.162) | | | -0.09 |
| Observations | 91 | 91 | 91 | 127 | 127 | 127 |
| Adj. R-squared | 0.269 | 0.242 | 0.245 | 0.196 | 0.197 | 0.197 |
| Resid. std. err. | 1.155 | 1.176 | 1.174 | 0.572 | 0.572 | 0.572 |
| D.f. | 88 | 88 | 88 | 125 | 125 | 125 |

between 0.3 and 0.9 standard deviations in the same direction, depending on the period and form of the variables. The variable also behaves in line with the idea that inflation expectations became more anchored after the early 1980s; its effect clearly diminishes after 1983. In level terms, its coefficient falls from between 0.7 and 0.9 to around 0.6; in differences, from 0.4 to 0.3.

The lack of explanatory power of newspaper tone in the regressions above has several plausible explanations. One explanation is that the tone of the news truly does not matter in the formation of household inflation expectations—this cannot be ruled out. Another is that tone does matter, but is being mismeasured by the method used here. There is some support for this explanation in the Michigan survey. Indeed, two questions early on in the survey's questionnaire inquire about news that the respondent has heard: the question "During the last few months, have you heard of any favorable or unfavorable changes in business conditions?" is followed up by "What did you hear?" The latter gives rise to two series that track the number of respondents who report having heard (unfavorable) news about "higher prices" and (favorable) news about "lower prices". The net result of this index (amount "higher" minus "lower"), plotted in Figure 4, tracks relatively well with actual inflation expectations from the survey. This suggests that, on average, respondents in the Michigan survey do pay attention to the news as a source of information about prospective inflation.

Figure 4: Net news on prices from the Michigan Survey



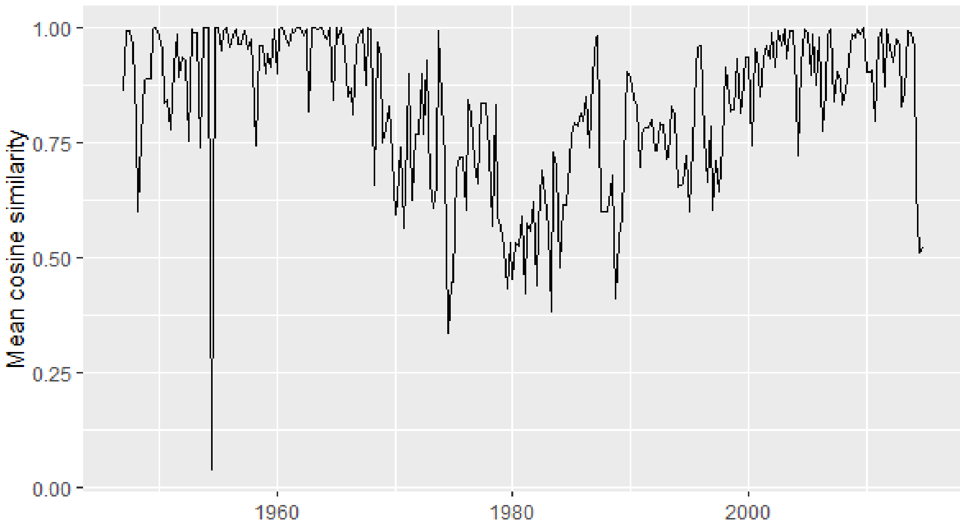

It therefore seems that sentiment analysis is unable to capture the tone of the news specifically with respect to inflation. Instead, it appears to be capturing attitudes about the general economy, which often coincide with inflation worries but do not do so in systematic ways: one can be worried about inflation when the economy is growing too quickly, but equally worried about inflation when an oil-price shock causes a recession (as in the "stagflation" of the 1970s). Using a more restrictive filter could mitigate this problem in theory, but would require a substantially larger corpus: filtering only for articles that contain the word "inflation" in the headline produced only about 6,000 stories. We would also be excluding important comments about inflation that are contained in articles about the Federal Reserve or general economic conditions.

The in ability of the dictionary-based measure of tone to capture inflation dynamics leads us to word embedding, which I will discuss next.

## 6.2 Word embedding

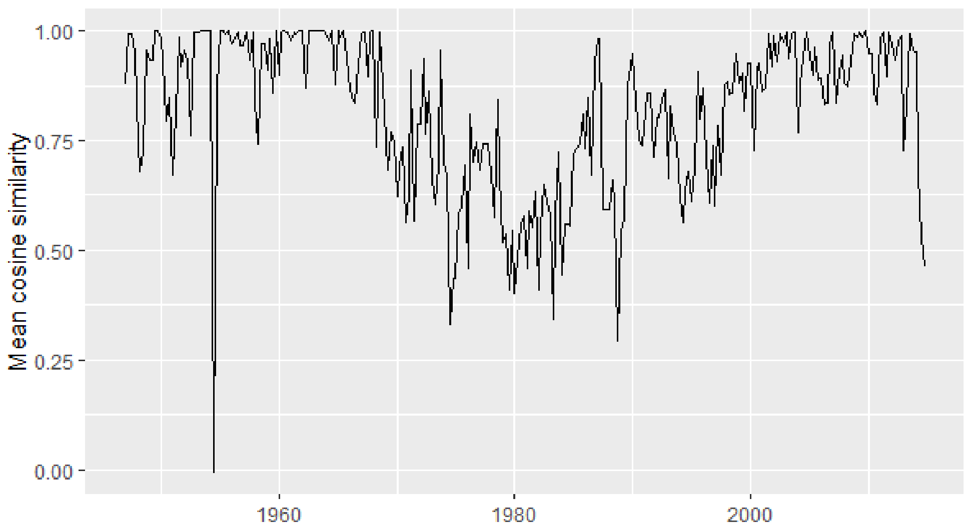
Applying word embedding in a temporal fashion proved to be a difficult task with the dataset at hand. Figure 5 plots the positive word embedding index, i.e. the quarterly cosine similarity between the word "inflation" and positive words from the Lexicoder dictionary. At first glance, the series looks promising: inflation has a very close semantic proximity (cosine similarity) to positive words up until the late 1960s, and then enters a steep decline that culminates around 1980. The series then begins increasing again, climbing back to extremely high similarity metrics by the early 2000s.

Figure 5: Positive word embedding index




The shape of this series fits neatly into the history of inflation in the US—an initial increase with the first oil crisis in 1973, further acceleration into the second oil crisis of 1979, and a moderation after the sharp rise in interest rates from the Volcker-led Federal Reserve caused the 1980 and 1981-82 double-dip recessions. Unfortunately, this correlation is quasi-spurious. Figure 6 plots the negative word embedding index, i.e. the quarterly cosine similarity between "inflation" and negative words. This index is almost identical to the former (indeed, plotting the difference between the two yields a white-noise-like series). This result is plainly incorrect: the similarity between “inflation” and negative words should be roughly the inverse of the similarity between “inflation” and positive words. Clearly, this invalidates results from both the indices.

Figure 6: Negative word embedding index




Extensive testing suggests that the odd initial results are a matter of wordcount. That is, the size of the corpus for much of the sample period is simply too small when divided into three-month chunks. This should not be too surprising: a large share of observations have wordcounts below 20,000, particularly before the 1970s. This has the effect that the model does not train properly, producing inaccurate word vectors. The litmus test for this problem is to examine the closest words for a chosen target phrase using the `distance` function in the `word2vec` package (available in the `R` implementation of this package). When the closest words have limited relation to the target phrase and the corresponding cosine similarities are surprisingly high, the model is likely to have trained poorly.

To illustrate this problem with an example, Figure 6 shows the output of the `distance` function for the target word “inflation”, using the model trained on the first quarter of 1962. Some of the words produced by the function seem reasonably related to “inflation”, such as “budget”, “nation’s” and “industrial”. However, the

Table 6: Embedding output, 1962 Q1

| word | cosine similarity |
|---|---|
| budget | 0.9969 |
| who | 0.9964 |
| provide | 0.9964 |
| action | 0.9957 |
| nation's | 0.9956 |
| before | 0.9955 |
| balanced | 0.9954 |
| law | 0.9954 |
| d | 0.9953 |
| industrial | 0.9952 |

cosine similarities for all of these words are essentially equal to 1—an unreasonably close relationship for these words. Moreover, very common words such as "who" and "provide" are also included in the list, with virtually the same similarity scores. Similar results are found using the trained vectors from other early periods in the dataset.

The U-shaped trend in the word embedding indices between 1970 and the early 2000s therefore simply reflects the wordcount in the *New York Times* corpus over the same period. As more words become available for analysis, the model trains more effectively and no longer produces vectors that are spuriously close to each other. This mechanically lowers the average cosine similarity between "inflation" and dictionary words.

An example of what appears to be the result of properly trained word embeddings is shown in Figure 7. This is the output of the `distance` function for the first quarter of 1980—around the peak of both volume of news stories and wordcount. Here, the words seem to be much more closely related to the concept of inflation, and the similarity scores are at more realistic levels of between 0.4 and 0.5. The second and third words, "unemployment" and "expectations", both refer to key concepts in inflation forecasting: the unemployment rate is strongly linked to inflation in

Table 7: Embedding output, 1980 Q1

| word | cosine similarity |
|---|---|
| inflationary | 0.5523 |
| unemployment | 0.5066 |
| expectations | 0.4562 |
| recession | 0.4217 |
| rises | 0.4089 |
| our | 0.4012 |
| economy | 0.3899 |
| public's | 0.3886 |
| prevailing | 0.3871 |
| risk | 0.3804 |

the Phillips curve model, as are expectations of future inflation in some variants of the Phillips curve. The fourth word, "recession", is an interesting result. The US was just entering a brief recession that lasted for the first half of 1980—under Paul Volcker, the Federal Reserve was in the process of severely raising interest rates in order to combat the inflation that had taken hold in the 1970s. But official determination of business-cycle dates only happens after the fact, when revised versions of the economic data have been released. The semantic relationship between "inflation" and "recession" during this period suggests that the media reflected an awareness that recession was imminent, and accurately linked this assessment to excessively high inflation.

To avoid the wordcount issue, I increase the training samples to the largest size that permits a meaningful analysis. That is, I demarcate a split between the fourth quarter of 1982 and the first quarter of 1983 as before, but in this case I use the full history of the *New York Times* dataset starting in 1947. This yields training samples with about 10.6 million and 10.3 million words, respectively, and amounts to a case study of semantic changes in the media's portrayal of the news before and after the era of stagflation in the US.

Figure 8 tabulates the words with the closest semantic proximity to "inflation" using the trained word vectors from these two periods. A glance at the table confirms that the model trained reasonably well on these data: most of the words are related to economics and prices. Moreover, the cosine similarities lie between 0.4 and 0.7, more realistic than the suspiciously high results seen previously. The results also fit the economic narrative I have stressed before. Prior to 1983, "recession" and "stagflation" are among the most closely related words to the concept of inflation; in the latter period, both words are farther down in similarity. By contrast, "deflation" shows up in the top three words in the latter period (and again as "deflationary" farther down) while it sits at eleventh place prior to 1983.

Some words are related to inflation exclusively in that period. Those include "depression" and "instability" before 1983 and "productivity" and "innovation" afterwards. The latter observation might be explained by the productivity boom in the late 1990s and its subsequent decline—one reason that economists place much importance on productivity is that it enables an economy to grow more quickly without generating inflationary pressures.

While the results above seem to be in line with intuition, I next set out to test them more quantitatively. To do this, I combine the output of the word embedding model with the dictionary methods used earlier in order to measure the semantic proximity of "inflation" to a range of words. This process follows several steps. First, I extract the matrix of trained word vectors for the two time periods, each of which contain around 35,000 entries. I then loop through each of the words in a given dictionary category and compute the cosine similarity between "inflation" and that word; the final cosine similarity for that category is the mean of individual scores. I iterate this process for four dictionary categories: positive and negative sentiment from the Lexicoder dictionary (roughly 3,000 words in each category) as well as fear and trust from the NRC Word-Emotion Association Lexicon, sometimes called EmoLex (Mohammad and Turney, 2013; Mohammad and Turney, 2010). EmoLex contains word-associations for two sentiments (negative and positive) and eight ba-

Table 8: Embedding output, 1947-1982 and 1983-2014

| **1947-1982** | | **1983-2014** | |
|---|---|---|---|
| word | cosine similarity | word | cosine similarity |
| unemployment | 0.6301 | inflationary | 0.6855 |
| inflationary | 0.6231 | unemployment | 0.5862 |
| recession | 0.5477 | deflation | 0.5807 |
| joblessness | 0.5449 | joblessness | 0.4885 |
| protectionism | 0.5340 | growth | 0.4833 |
| stagflation | 0.5230 | wages | 0.4649 |
| growth | 0.5187 | productivity | 0.4552 |
| recessionary | 0.4974 | innation | 0.4430 |
| jobless | 0.4915 | velocity | 0.4327 |
| deflation | 0.4641 | recession | 0.4287 |
| overheating | 0.4624 | appreciably | 0.4283 |
| depression | 0.4549 | interest | 0.4279 |
| disinflation | 0.4435 | volatility | 0.4203 |
| expansion | 0.4409 | stagflation | 0.4183 |
| wind | 0.4405 | overheating | 0.4157 |
| economy | 0.4311 | turmoil | 0.3991 |
| recovery | 0.4299 | deflationary | 0.3890 |
| instability | 0.4272 | affluence | 0.3873 |
| stagnation | 0.4181 | employment | 0.3847 |
| | | wageprice | 0.3786 |
| | | noninflationary | 0.3778 |
| | | inequality | 0.3775 |
| | | slowing | 0.3768 |
| | | prosperity | 0.3759 |
| | | moneysupply | 0.3663 |

sic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust), manually annotated via crowdsourcing through Amazon's Mechanical Turk. The fear category contains about 1,500 words, including "abandon", "despair", "difficult", "evil", "fear", "harmful", "helpless", "menacing" and "unstable". The trust category contains roughly 1,200 words, such as "achieve", "assure", "confident", "efficient", "remedy", "succeed" and "trust".

Table 9: Embedding output, 1947-1982 and 1983-2014
Note: For simplicity, negative cosine similarities are reported in absolute values.

| | Mean cosine similarity | |
|---|---|---|
| Dictionary category | 1947-82 | 1983-2014 |
| Negative (LSD) | 0.047 | 0.008 |
| Positive (LSD) | 0.003 | 0.028 |
| *Ratio (neg/pos)* | *16.4* | *0.3* |
| Fear (EmoLex) | 0.063 | 0.005 |
| Trust (EmoLex) | 0.015 | 0.044 |
| *Ratio (fear/trust)* | *4.2* | *0.1* |

Results are shown in Table 9 and are aligned with the more informal discussion above. Negative words have a closer semantic relationship with "inflation" than positive words before 1983 by a ratio of about 16. By contrast, after 1983, negative words have very weak semantic relationship with "inflation", one that is only about 70 percent weaker than the cosine similarity with positive words (ratio of 0.3). Results using emotions from the EmoLex dictionary are similar. Prior to 1983, inflation has a higher semantic proximity to words associated with fear than with trust, by a ratio of 4.2. For the period after 1983, that ratio falls to 0.1.

# 7 Conclusion

Using a corpus of *New York Times* articles about the US economy from 1947 through 2014, I showed that 1-year inflation expectations are less sensitive to the volume and tone of newspaper coverage in the decades following the Volcker disinflation compared to the period leading up to it. Results were obtained using OLS regressions of inflation expectations on past inflation, an indicator of oil crises, and the two media variables, and hold for specifications in both levels and first-differences.

I also used word embedding, an emerging technique in natural language processing, to demonstrate semantic shifts around the concept of inflation in *New York Times* coverage of the economy before and after the Volcker disinflation. Between 1947 and 1982, "inflation" was much more closely associated with negative sentiment than with positive sentiment, and more closely related to fear than to trust. For the period between 1983 and 2014, these semantic relationships are reversed. Semantic proximity to positive and negative sentiment was measured using those categories in the Lexicoder Sentiment Dictionary; proximity to fear and trust was measured using the NRC Word-Emotion Association Lexicon.

The main challenge I encountered in my analysis was producing a robust time-series of semantic shifts in the meaning of "inflation". Evidently, either the size or the lexical diversity of the *New York Times* corpus is insufficient for training word vectorizations over three-month periods. This is an area ripe for further research. Indeed, a (very) recently published paper by Emma Rodman (Rodman, 2019) tackles exactly this problem.[7] Rodman outlines several strategies that can facilitate the use of word embeddings to draw time-series conclusions using small corpora. These strategies include bootstrap resampling of the text, initializing embeddings using pre-trained vectors from prior periods, and using selective stemming to mitigate the problem of language instability over long periods of time. Applying these techniques to the study of historical inflation perceptions seems very promising, and I hope that this paper motivates further research in this domain.

[7] Rodman's paper was published two weeks before the submission deadline for this paper.

# References


Ansolabehere, Stephen, Marc Meredith, and Erik Snowberg (2014). "Mecro-economic voting: Local information and micro-perceptions of the macro-economy". In: *Economics and Politics*.

Barberá, Pablo et al. (2019). "Automated Text Classification of News Articles: A Practical Guide".

Bernanke, Ben S. (2007). "Inflation Expectations and Inflation Forecasting". In: *Monetary Economics Workshop of the NBER Summer Institute*. Cambridge, Massachusetts, pp. 1–10.

Blood, Deborah J. and Peter C.B. Phillips (1995). "Recession headline news, consumer sentiment, the state of the economy and presidential popularity: A time series analysis 1989-1993". In: *International Journal of Public Opinion Research*.

Boomgaarden, Hajo G et al. (2011). "Covering the crisis: Media coverage of the economic crisis and citizens' economic expectations". In: *Acta Politica* 46.4, pp. 353–379.

Bordo, Michael D. et al. (2007). "Three Great American Disinflations". Cambridge, MA.

Brown, Les (1971). *Television*. New York: Harcourt Brace Jovanovich, Inc.

Caliskan, Aylin, Joanna J. Bryson, and Arvind Narayanan (2017). "Semantics derived automatically from language corpora contain human-like biases". In: *Science* 356.6334, pp. 183–186.

Carroll, Christopher D. (2003). *Macroeconomic expectations of households and professional forecasters*.

Cavallo, Alberto, Guillermo Cruces, and Ricardo Perez-Truglia (2017). "Inflation expectations, learning, and supermarket prices: Evidence from survey experiments". In: *American Economic Journal: Macroeconomics* 9.3.

Clark, Todd E and Troy Davig (2008). "An Empirical Assessment of the Relationships Among Inflation and Short- and Long-Term Expectations". In: *Federal Reserve Bank of Kansas City Research Working Paper*.

Coibion, Olivier and Yuriy Gorodnichenko (2015). "Is the phillips curve alive and well after all? Inflation expectations and the missing disinflation". In: *American Economic Journal: Macroeconomics* 7.1, pp. 197–232.

Coibion, Olivier, Yuriy Gorodnichenko, et al. (2018). "Inflation Expectations – a Policy Tool?" In: *ECB Forum on Central Banking: Price and wage-setting in advanced economies*. Sintra, Portugal, pp. 93–151.

Crouse, Timothy (1973). *The Boys on the Bus*. New York: Random House.

De Boef, Suzanna and Paul M. Kellstedt (2004). "The political (and economic) origins of consumer confidence". In: *American Journal of Political Science*.

Doms, Mark E. and Norman J. Morin (2004). "Consumer Sentiment, the Economy, and the News Media". In: *SSRN Electronic Journal*.

Firth, John R (1957). "A synopsis of linguistic theory, 1930-1955". In: *Studies in Linguistic Analysis*, pp. 1–32.

Goodfriend, Marvin and Robert G. King (2005). "The incredible Volcker disinflation". In: *Journal of Monetary Economics* 52.5, pp. 981–1015.

Han, Jiawei, Micheline Kamber, and Jian Pei (2012). *Data Mining: Concepts and Techniques*. 3rd. URL: `https://www.sciencedirect.com/book/9780123814791/data-mining-concepts-and-techniques#book-info`.

Harris, Zellig S. (1954). "Distributional Structure". In: *WORD* 10.2-3, pp. 146–162.

Hollanders, David and Rens Vliegenthart (2011). "The influence of negative newspaper coverage on consumer confidence: The Dutch case". In: *Journal of Economic Psychology*.

Jonung, Lars (1981). "Perceived and Expected Rates of Inflation in Sweden". In: *American Economic Review* 71.5, pp. 961–968.

Lamla, Michael J. and Sarah M. Lein (2014). "The role of media for consumers' inflation expectation formation". In: *Journal of Economic Behavior and Organization*.

Lamla, Michael J. and Thomas Maag (2012). "The Role of Media for Inflation Forecast Disagreement of Households and Professional Forecasters". In: *Journal of Money, Credit and Banking* 44.7, pp. 1325–1350.

Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers (2003). "Disagreement about Inflation Expectations". In: *NBER Macroeconomics Annual* 18, pp. 209–248.

Mikolov, Tomas et al. (2013). "Distributed Representations of Words and Phrases and their Compositionality". In: *Proceedings of NIPS*.

Mohammad, Saif M and Peter D Turney (2010). "Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon". In: *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, Los Angeles, California, June 2010*.

Neuman, W. Russell (1990). "The Threshold of Public Attention". In: *Public Opinion Quarterly* 54, pp. 159–76.

Nikfarjam, Azadeh et al. (2015). "Pharmacovigilance from social media: Mining adverse drug reaction mentions using sequence labeling with word embedding cluster features". In: *Journal of the American Medical Informatics Association*.

Pfajfar, Damjan and Emiliano Santoro (2013). "News on inflation and the epidemiology of inflation expectations". In: *Journal of Money, Credit and Banking*.

Rodman, Emma (2019). "A Timely Intervention: Tracking the Changing Meanings of Political Concepts with Word Vectors". In: *Political Analysis*, pp. 1–25.

Silber, William L. (2012). *Volcker: The Triumph of Persistence*. New York: Bloomsbury Press.

Soroka, Stuart N., Dominik A. Stecula, and Christopher Wlezien (2015). "It's (Change in) the (Future) Economy, Stupid: Economic Indicators, the Media, and Public Opinion". In: *American Journal of Political Science* 59.2, pp. 457–474.

Tang, Duyu, Bing Qin, and Ting Liu (2015). "Learning Semantic Representations of Users and Products for Document Level Sentiment Classification". In:

Tang, Duyu, Furu Wei, et al. (2014). "Learning Sentiment-Specific Word Embedding". In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL 2014).*

Winter, James P. and Chaim H. Eyal (2016). "Agenda-setting for the civil rights issue". In: *Agenda Setting: Readings on Media, Public Opinion, and Policymaking.*

Wong, Benjamin (2015). "Do Inflation Expectations Propagate the Inflationary Impact of Real Oil Price Shocks?: Evidence from the Michigan Survey". In: *Journal of Money, Credit and Banking* 47.8, pp. 1673–1689.

Yellen, Janet L. (2015). "Inflation Dynamics and Monetary Policy". In: *Philip Gamble Memorial Lecture, University of Massachusetts, Amherst, Amherst, Massachusetts.*

Young, Lori and Stuart Soroka (2012). "Affective News: The Automated Coding of Sentiment in Political Texts". In: *Political Communication* 29.2, pp. 205–231.