News, Sentiment, and Inflation Expectations: Insights From Social Media Data and Experiments

Sergii Drobot*

January 29, 2025

Abstract

News media serves as a primary information source for most people, with social media taking on a growing role in the way news is distributed and interacted. This paper investigates how households use information conveyed through media form their inflation forecasts. Leveraging microdata, social media news data, and machine learning techniques, I show that households dynamically update the news topics they focus on when forming inflation expectations. However, the impact of news media on expectations is time varying — at times, it accounts for a significant portion of the variation in expectations, while at other times, its influence diminishes. Using social media reactions as a proxy for sentiment generated by the news, I demonstrate that sentiment plays a critical role in shaping expectations and can predict the direction of forecast adjustments. A novel information provision experiment, incorporating a sentiment elicitation method, further confirms the importance of sentiment in the expectations formation process. Moreover, I demonstrate that identical policy-related information, when framed with different tones, evokes distinct sentiment that lead to asymmetric effects on forecast revisions.

Keywords: survey, social media, expectations, machine learning
JEL Classification: E31, E52, E71

*Indiana University Bloomington. Email: sdrobot@iu.edu. I sincerely thank my advisors, Daniela Puzzello and Rupal Kamdar, for their invaluable guidance and insightful feedback. I also thank the participants of the Macro Brown Bag at Indiana University for their thoughtful feedback and constructive suggestions. The Indiana University IRB: 17972.

1 Introduction

This paper seeks to assess how households use information from the media to form their expectations about future inflation. In economic theory, information plays a crucial role in shaping individuals' expectations, which in turn drive macroeconomic outcomes. The mechanism is straightforward: households form expectations about future economic conditions by processing the information available, which then influences their consumption, savings, and investment decisions. However, information is far from simple or uniform. It varies widely in form and quality and evolves over time, making it a complex phenomenon that requires careful study, particularly in the context of expectation formation. Despite the centrality of information in expectation formation, little is known about how households process diverse and evolving media signals. This study addresses this gap by combining micro-level expectations data, social media analysis, machine learning techniques and a randomized controlled trial (RCT) to investigate this crucial issue.

Recent technological advancements have made information even more accessible and widespread than ever before. The internet, social media, and mobile devices have lowered the cost and increased the speed of accessing information, transforming the landscape of expectation formation. As a result, the quantity of information has also increased, but not necessarily its quality. Yet, this increased availability has introduced new challenges: the abundance of information has made it more difficult for individuals to process and interpret it effectively.

Understanding how households navigate this complex information environment is essential to grasping the mechanisms that shape economic expectations. In this landscape of information overload and a growing number of alternatives, the news media still plays a central role as the primary source through which households receive updates on economic, political, and social developments (Happer and Philo, 2013). What is even more important in the context of expectation formation is that media coverage not only informs the public about current events but also conveys signals about possible future trends. For example, during the COVID-19 pandemic, the media highlighted economic threats such as job losses, supply shortages, and long-term health challenges. Similarly, coverage of the war in Ukraine raised concerns about potential sanctions on Russia, leading to energy crises and rising energy prices. In another instance, the U.S. election coverage revealed insights into candidates' political and economic policies, influencing expectations about the economy's future direction.

While economic experts and policymakers have the tools to sift through this complex array of signals and form forecasts aligned with economic theory, the average household may struggle to process such a vast and often ambiguous volume of information. This paper explores how typical households without formal economic training digest the information presented by the media and form inflation expectations. By understanding this dynamic, we can gain insights into the broader mechanisms of expectation formation and their implications for economic stability and policy-making.

A unique feature of this paper is that it leverages the strengths of empirical analysis based on machine learning and experimental economics methods to explore the research question. The empirical analysis enables me to examine the dynamic relationship between news content and inflation expectations over time, capturing evolving patterns in public attention and sentiment. Specifically, I construct a daily time series of news topics from social media posts using the Latent Dirichlet Allocation (LDA) model, combine this with daily inflation expectations derived from microdata, and analyze the relationships using rolling-LASSO regressions.

The results highlight the dynamic nature of public attention, showing that the topics households focus on when forming inflation expectations shift in response to major events. Additionally, the explanatory power of these topics fluctuates over time, reflecting the complex and context-dependent influence of media on expectations. To better understand the role of sentiment, I incorporate social media reactions — such as "Love," "Angry," and "Sad" — as proxies for the emotional responses generated by news topics. This enables the construction of a panel dataset linking media sentiment to changes in inflation expectations. The analysis reveals that households frequently base their expectations not only on the factual content of news but also on the sentiment it generates. For instance, news associated with negative sentiment often leads to upward revisions in inflation expectations, suggesting that sentiment plays a significant role in shaping economic perceptions.

To confirm these findings and establish a causal relationship, I conduct a survey experiment using Randomized Control Trial (RCT) methodology. One treatment uses non-economic news that dominates the media and generates negative reactions, providing additional evidence supporting the insights from the empirical analysis. Furthermore, economic treatments based on the same information about a Federal Reserve interest rate cut — presented with different tones — demonstrate that media framing can influence households' emotional responses to economic information and, consequently, their interpretation of the information. These emotional responses significantly affect how households update their inflation expectations, but prior beliefs also play a critical role. For example, households with high prior inflation expectations who respond negatively to the information tend to revise their expectations upward, while those who react positively revise their expectations downward. Conversely, the relationship is reversed for households expecting deflation.

The paper speaks directly to the growing literature that highlights the role of media in the expectations formation process. It built on the work of Carroll (2003), Doms and Morin (2004), Pfajfar and Santoro (2013), Lamla and Lein (2014), and Dräger and Lamla (2017), but similar to Larsen et al. (2021) I use modern textual analysis and machine learning techniques to explore the research question. What distinguishes this study is the use of social media data, specifically users' reactions, which provides a more accurate measure of how individuals respond to the news. Unlike the traditional approach that focuses solely on the count and the tone of the news itself, this study incorporates user-generated responses to propose a sentiment-based mechanism that captures the direction of the effect of news topics on expectations. Additionally, by leveraging high-frequency data, this paper explores the dynamics of the expectations formation process. This approach offers a more granular view of how the significance of different news topics in shaping inflation expectations, and the role of media more broadly, evolves over time. In doing so, the study offers a fresh perspective on the evolving relationship between media, public sentiment, and economic expectations. It is worth noting that while some studies have used Twitter data to analyze central bank communication (Korhonen and Newby, 2019; Gorodnichenko et al., 2024), they focus exclusively on central banking topics. In contrast, this study examines a broader range of general media topics. Additionally, by using Facebook data, which offers a richer variety of user reactions compared to Twitter, this study broadens the scope of analysis and provides deeper insights into how individuals respond to news.

Similar to Kamdar and Ray (2024), this paper underscores the critical role that sentiment plays in the expectations formation process of households. A key contribution is the positioning of news media as a sentiment-generating agent. In addition to the empirical analysis, I provide support for the mechanisms driving my findings with a survey experiment. This experiment follows the framework of the widely used information provision experiments (Binder and Rodrigue, 2018; Armona et al., 2019; Coibion et al., 2023a), but with a unique emphasis on capturing the sentiment generated by the information provided — a focus that, to my knowledge, is novel in the existing literature. This approach not only extends the empirical understanding of sentiment's role in expectation formation but also offers a methodological contribution to survey-based studies on inflation expectations. By capturing sentiment directly, my experiment advances survey techniques, allowing for more nuanced insights into how households emotionally process economic information and how these sentiment drives inflation expectations.

Finally, more broadly, this paper contributes to the literature that explores the determinants of inflation expectations and deviations from the Full Information Rational Expectations (FIRE) assumption proposed by Muth (1961). It engages with theories such as rational inattention (Sims, 2003), information stickiness (Mankiw and Reis, 2002), experience (Malmendier and Nagel, 2016), and heuristics (Bordalo et al., 2016), among others.

The paper proceeds as follows. Section 2 focuses on the empirical analysis, introducing the data, examining the dynamics of news topics, and demonstrating the role of news-generated sentiment in the expectations formation process. Section 3 is dedicated to a survey experiment that provides additional causal evidence and new facts. Then, section 4 concludes.

2 Empirical Analysis

I employ textual analysis and machine learning techniques on high-frequency data to uncover new evidence regarding the role of media in shaping household inflation expectations. The analysis reveals three key findings: (1) the list of news topics influencing household inflation expectations changes over time, (2) the magnitude of the effect of news media on expectation formation is dynamic, and (3) the sentiment generated by news stories help explain the direction in which inflation expectations move.

2.1 News and Expectations Data

The data for this study comes from social media, collected through CrowdTangle, a tool that tracks and aggregates public content from social media platforms. Given the increasing influence of social networks in the dissemination of information, social media has become a vital instrument for shaping public opinion (Newman et al., 2023; Tandoc Jr et al., 2020; Hermida et al., 2012). Unlike traditional media, social media enables users to interact with news content through comments, shares, and reactions, providing valuable insights into audience sentiment and news popularity. This additional information makes social media an excellent substitute for the conventional news datasets typically used in this area of research.

To leverage the unique characteristics of social media data, I collected news posts from the official Facebook pages of 64 major U.S. news outlets, including ABC News, Bloomberg, CNN, Fox News, and The New York Times. The data spans from November 2013 to January 2024, covering approximately 7.9 million posts.

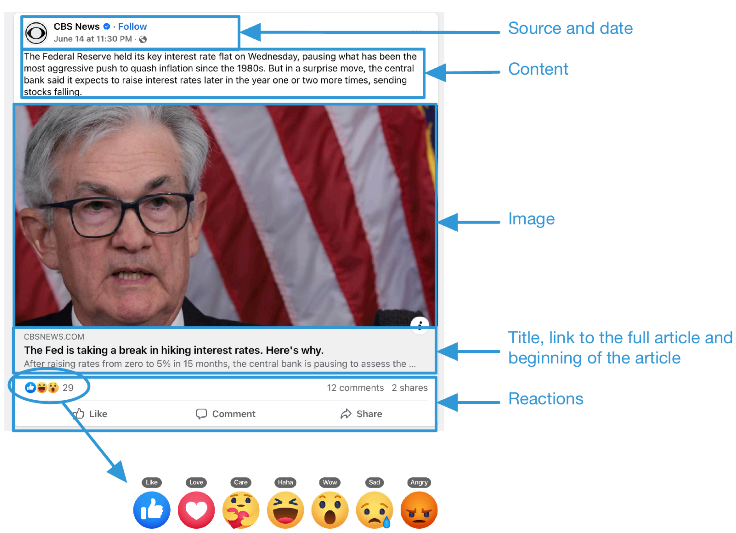


Figure 1: Facebook Post Example

The choice of Facebook as the source of news data was made with careful consideration. As Newman et al. (2023) highlighted, an increasing number of people are using social media as their primary source of news, with Facebook playing a central role in this trend. Mosquera et al. (2020) further demonstrated through a field experiment that news consumption decreased significantly among those who abstained from using Facebook for a week. Despite its decline in popularity, Facebook remains the most widely used platform for news consumption in the U.S. Major media companies regularly post their content on Facebook, typically offering brief summaries along with links to full articles. This format makes news more accessible and often free, increasing the likelihood of reaching a broader audience.

Importantly, Facebook offers a unique mechanism for capturing user feedback through its reaction buttons — "Like," "Love," "Care," "Haha," "Wow," "Sad," and "Angry" — which provide rich insights into users' emotional responses to news content and the extent of its dissemination. Figure 1 shows an example of a Facebook post.

I processed the social media news posts for empirical analysis using standard Natural Language Processing (NLP) techniques. Using the Latent Dirichlet Allocation (LDA) model, the posts were algorithmically grouped into 80 distinct topics based on the dominant topic identified in the text. LDA operates on the premise that documents (or posts, in this case) are random mixtures of topics, with each topic consisting of a distribution of words. For example, the topic labeled "Pandemic" was defined by terms such as "covid19," "pandemic," "vaccine," "case," "test," "virus," "health," and so on. Each post was then assigned the topic that best described its content. These topics were further transformed into tone-adjusted time series based on their frequencies. Detailed information on data cleaning and transformation can be found in Appendix A.

The time series of news topics is based on both the frequency and tone of the news, similar to previous literature (e.g., Larsen et al., 2021), rather than on social media metrics such as reactions. There are several reasons for this. First, using frequency and tone makes the results comparable to earlier studies. Second, Facebook reaction buttons were introduced only in 2016, which would significantly shorten the time series available for analysis. However, I do leverage social media metrics to construct a measure of the sentiment generated by the news and build a panel dataset, demonstrating that these sentiment can help predict the direction of changes in inflation expectations.

To measure median household inflation expectations, I use micro-data from the Federal Reserve Bank of New York's Survey of Consumer Expectations (SCE). Although the University of Michigan's Survey of Consumers (MSC) offers a longer time series, the SCE has distinct advantages for this study. With a sample size of approximately 1,300 respondents per month — more than twice the MSC's sample size — the SCE offers greater granularity and allows for the measurement of daily inflation expectations, providing deeper insights into the dynamics of expectation formation. Additionally, as Armantier et al. (2017) noted, the SCE specifically asks respondents about "inflation" rather than "price changes," reducing ambiguity and ensuring a clearer understanding of the concept among respondents. Finally, respondents participate in the SCE survey up to twelve times, making it highly valuable for constructing the panel dataset discussed later.

2.2 Dynamic Attention Toward News

In the following discussion, I extend the analysis of Larsen et al. (2021) by relaxing the assumption that the list of news topics influencing household inflation expectations is static over time. Instead of relying on monthly data, I utilize rolling LASSO (Least Absolute Shrinkage and Selection Operator) regressions on daily data with a rolling window of 90 days. This approach allows for a more dynamic understanding of how the relevance of different news topics evolves over time.

LASSO is particularly well-suited for this analysis because it not only identifies which variables (in this case, news topics) are most important, but it also shrinks the coefficients of less important variables to zero, effectively reducing the number of relevant topics. The model specification is as follows:

$$F_t \pi_{t+365} = \beta_0 + \sum_{n=1}^{N} \beta_n \cdot T_{n,t} + \epsilon_t \tag{1}$$

In this model, $F_t \pi_{t+365}$ represents households' one year ahead (365 days) median forecast of inflation at time t, and N is the number of news topics $T_{n,t}$. While LASSO does not explain the direction of the effect — since all variables are normalized — it highlights the news topics that correlate with the median inflation expectations of households. For each LASSO regression, I record which topics are selected, and I present the results in Figure 2, showing how the list of news topics that people are paying attention to evolves over time.

It is notable that topics related to politics are frequently selected, aligning with recent literature that highlights the role of political affiliation in shaping inflation expectations (Kamdar and Ray, 2023; Mian et al., 2023; Binder et al., 2024; Huseynov and Murad, 2024). Furthermore, the topics identified by LASSO reflect major global events such as the war in Ukraine and the COVID-19 pandemic, underscoring the relevance of media discourse in capturing public attention.[1] These results suggest that media coverage of significant events and politically salient issues likely plays a role in influencing public expectations over time.

[1] It should be noted that the topic labels were assigned based on word clusters derived from the data, providing a systematic, albeit approximate, representation of the underlying themes.

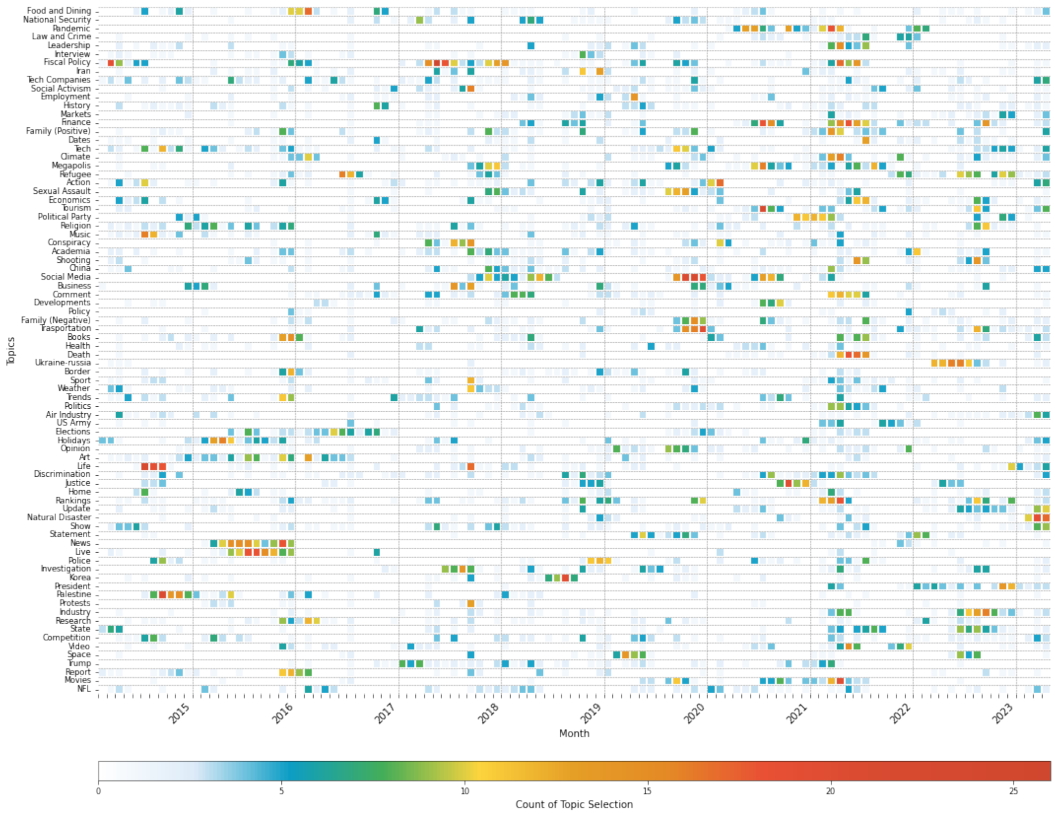


Figure 2: Inflation Expectations and Attention to the News Topics

Notes: The heatmap presents the results of rolling LASSO regressions using a 90-day window. The dependent variable is the daily median inflation expectations, while the independent variables consist of 80 tone-adjusted news topics. For each regression, I record the topics with non-zero coefficients and assign them to the latest day within the regression window. The heatmap displays the count of times each topic was selected by LASSO per month.

Additionally, the results indicate that certain topics attract public attention only temporarily, while others play a sustained role in shaping household inflation expectations over extended periods. This dynamic underscores the constantly evolving nature of the information environment, as public focus shifts between topics in response to changing circumstances. Notably, the coefficient of determination, R^2, illustrated in Figure 3, exhibits significant volatility. This variability suggests that while news media can occasionally be a dominant factor influencing inflation expectations, at other times, its explanatory power diminishes substantially, and other determinants play a more significant role.

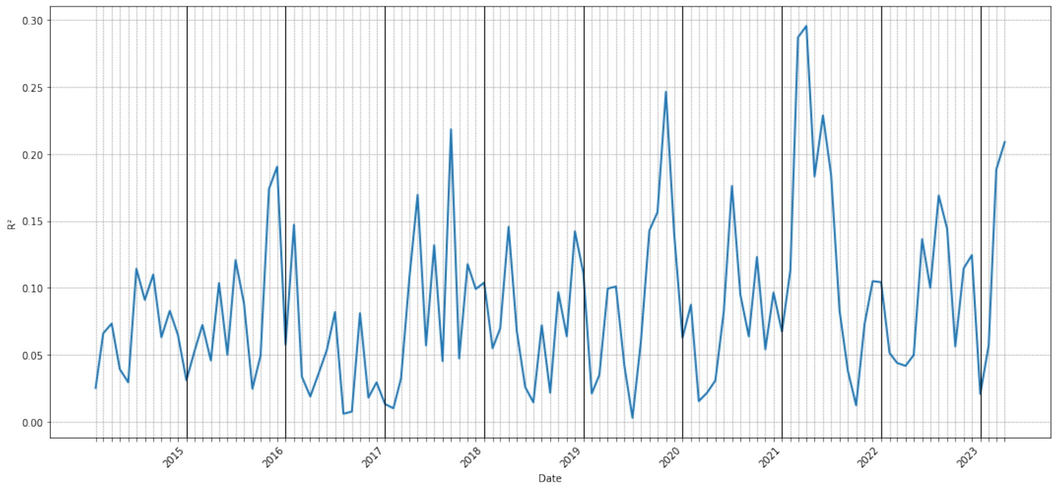


Figure 3: Rolling LASSO Coefficient of Determination

Notes: For each rolling LASSO regression, the coefficient of determination (R^2) is calculated and assigned to the last day within the regression window. The chart displays the monthly average of these R^2 values.

2.3 Sentiment-Driven Inflation Expectations

Next, I seek to uncover the mechanism through which media-reported news shapes household beliefs about future inflation. Specifically, I aim to understand how households transform complex and, in many cases, seemingly economically unrelated information into inflation forecasts. My hypothesis is that most households rely on a sentiment-based model — meaning that inflation expectations are driven by the sentiment (which reflects attitudes toward or feelings about current and future developments) generated by news, alongside other factors.

To test this hypothesis, I construct a panel dataset using microdata from the SCE, leveraging the fact that each respondent is surveyed multiple times. The dependent variable is the percent points change in the respondent's one-year-ahead inflation expectations compared to their previous survey response. The key independent variable is the measure of media sentiment circulating between the respondent's survey participation dates.

To construct the individual specific sentiment measure, ξ_i, I calculate the total number of media reactions (e.g., "Love," "Sad," "Angry") that occurred between the two survey dates for each respondent i. The sentiment measure is defined by the following formula:

$$\xi_i = \underbrace{\frac{N_i}{100,000,000}}_{\text{Spread}} \cdot \underbrace{\frac{n_i^{\text{love}} - n_i^{\text{sad}} - n_i^{\text{angry}}}{n_i^{\text{love}} + n_i^{\text{sad}} + n_i^{\text{angry}}}}_{\text{Sentiment}}, \tag{2}$$

where N_i represents the total number of interactions with Facebook posts (including reactions, comments, shares, and reposts), and n_i^j denotes the total number of each reaction type $j \in \{$"Love," "Sad," and "Angry"$\}$. The first term represents the spread of information, while the second term captures the balance of sentiment. I exclude the "Haha" and "Wow" reactions from the baseline sentiment balance, as they can convey both positive and negative sentiment, as confirmed in the subsequent survey experiment. Additionally, the "Care" reaction is excluded because it was introduced later, during the COVID-19 pandemic. The measure is individual-specific because it is based on the survey dates unique to each respondent i, but it reflects the aggregate number of media reactions observed during that period. However, it does not represent reactions made by respondent i.

To examine the updating behavior, I regress the individual revision of inflation expectations between two surveys, $\Delta\pi_i = \pi_i' - \pi_i$, on the measure of sentiment, ξ_i. Specifically, I estimate the following regression,

$$\Delta\pi_i = \alpha + \beta \cdot \xi_i + \epsilon_i \tag{3}$$

This approach allows me to quantify the overall sentiment generated by the media during the period between surveys and assess its influence on changes in individuals' inflation expectations. By focusing on sentiment rather than the specific details of news stories, I aim to identify the emotional cues that households use to adjust their forecasts for future inflation.

Table 1: Forecast Revision and Sentiment

	I	II	III	IV	V
ξ	-0.078***	-0.093***	-0.330***	-0.826*	-0.056
	(0.021)	(0.035)	(0.093)	(0.481)	(0.157)
Estimator	Huber	MixedLM	MixedLM	Huber	Huber
Controls	Yes	No	No	Yes	Yes
Obs.	62,365	62,391	9,501	2,679	2,679
Sample	Full	Full	FB users	Full	Full
Topics	All	All	All	LASSO	All
Period	Jan.18-Apr.23	Jan.18-Apr.23	Jan.18-Apr.23	Apr.20-Jun.20	Apr.20-Jun.20

Notes: The table presents various model specifications where the dependent variable is the change in expected inflation between two survey dates, and the independent variable is the sentiment measure (ξ_i) generated by news published on social media. Column (I) employs a Huber regression to control for outliers and influential variables, and also includes controls for demographic characteristics of survey participants (gender, age, race, employment, region, education, income, and numerical literacy). Column (II) is a mixed linear model (MixedLM) that accounts for both fixed effects and individual-specific random effects. Column (III) is a MixedLM applied to a sample of survey participants who resemble typical Facebook news consumers. Columns (IV) and (V) are both Huber regressions covering the period at the onset of the COVID-19 pandemic. In (IV), ξ_i is based on topics selected through LASSO, whereas in (V), it includes all topics. In all specifications, I filter out outliers by retaining only observations where the change in inflation expectations is within three standard deviations, based on the Z-score. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1 presents the results of multiple regression models. Across all specifications, the slope coefficient is negative, indicating that when news generates predominantly negative sentiment (i.e., accumulates negative reactions), inflation expectations are revised upward. For example, according to the model (I), a 1-unit increase in negative sentiment is associated with a 0.078 percentage point increase in inflation expectations, holding other factors constant. This finding supports my hypothesis based on a heuristic expectation formation model, where individuals project negative sentiment from the news onto their expectations for future inflation, which they also perceive unfavorably (Shiller, 1997; Stantcheva, 2024).

The effect of sentiment is stronger when accounting for respondent-specific characteristics using a mixed linear model (MixedLM) in column (II) than when using a Huber regression with demographic controls, which also adjusts for outliers and influential observations in column (I). The effect becomes even stronger when the analysis is restricted to respondents who match the demographic profile of typical U.S. Facebook news consumers — primarily white women aged 30-49, according to Pew Research Center.[2] This finding suggests that the source of news consumption can shape perceptions of future inflation.

As a case study, I examine the onset of the COVID-19 pandemic (April-June 2020) and focus exclusively on the news topics selected by LASSO in the prior analysis. As shown in column (IV) of Table 1, the impact of sentiment is significantly stronger during this period. The beginning of COVID-19 is especially noteworthy, as it coincides with a puzzling divergence between the inflation expectations of households and professional forecasters (see Figure B1). Given the strong and statistically significant effect of the sentiment measure, this divergence likely stems from households' reliance on news-driven sentiment, while professional forecasters continued to ground their expectations in observed inflation and the Phillips curve, anticipating an increase in unemployment. To confirm this result is not coincidental, I run the same regression over the same time period without filtering news topics. The results, presented in column (V), are not statistically significant, indicating that not all news topics within a given time shaped household expectations.

3 Information Treatment Experiment

A limitation of the empirical analysis is that it only provides correlational evidence. To establish causality and explore these relationships more deeply, I collect data in a controlled environment, such as a survey experiment. Specifically, to study the causal effect of news on inflation expectations, I conduct an information provision experiment using a Randomized Control Trial (RCT), a methodology that has recently gained popularity in macroeconomics. In the RCT, respondents are randomly divided into control and treatment groups.

[2]Pew Research Social Media and News Fact Sheet (September 2024).

Since only the treatment groups receive information interventions, any statistically significant differences between the groups can be attributed to the effect of the information.

The primary objective of the survey is to demonstrate that the news media plays an important role in shaping household inflation expectations by generating sentiment through the information they convey. Based on the insights from the empirical analysis, I formulate the following key hypotheses to test in the experiment:

> **Hypothesis 1. Sentiment-Driven Expectations.** Sentiment from general media news, even without direct economic information, can influence inflation expectations by shaping beliefs about economic conditions.

> **Hypothesis 2. Sentiment Shape Perception of Economic Data.** Sentiment generated by economic news influence how individuals perceive the economic data it presents.

3.1 Survey Design

To maintain consistency with empirical analysis, many survey questions are based on the SCE instrument.[3] To elicit inflation expectations before the information treatment, I use a probabilistic question, where respondents assign probabilities to a range of inflation and deflation bins. They are then randomly assigned to one of five groups: a control group with no information, a *SentimentOnly* group, and three economic treatment groups (categorized as *Negative*, *Neutral*, and *Positive*).

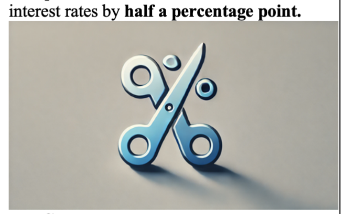
POSITIVE	**NEGATIVE**	**NEUTRAL**
In September, the Federal Reserve reduced interest rates by **half a percentage point**, a decision widely praised by market participants as a crucial step in response to current conditions after the rate increases during the Covid pandemic. This decisive action is seen by many as a positive, **providing significant relief for millions of US families and fostering greater confidence**.  **Fed Cuts Rates: A Timely and Appropriate Response to Current Conditions**	In September, the Federal Reserve reduced interest rates by a **mere half a percentage point**, a move dismissed by some market participants as inadequate in response to current conditions, especially considering the sharp rate hikes during the Covid pandemic. Despite this minimal adjustment, the decision is seen by many as too little, too late, **leaving the everyday lives of millions of US families struggling with uncertainty and offering no meaningful relief.**  **Fed Finally Cuts Rates: A Weak and Ineffective Response to Current Challenges**	In September, the Federal Reserve reduced interest rates by **half a percentage point.**  **Fed Cuts Rates**

Figure 4: Economic Treatments

Notes: Green and red indicate the differences between treatments. All treatments are designed to mimic the format of a Facebook post.

[3] Link to the instrument is here.

The *SentimentOnly* treatment incorporates five real Facebook posts collected from the official page of the New York Times that have recently triggered negative reactions. I focus on negative news, as previous studies, such as Nguyen and Claus (2013), have shown that consumer sentiment is more responsive to bad news. This treatment assesses whether the sentiment generated by these posts can move inflation expectations.

All economic treatments present the same news about a Federal Reserve interest rate cut that was recent at the time the survey was conducted, providing an opportunity to examine how indirect yet closely related economic signals in the media influence inflation expectations. Research by Coibion et al. (2023b) demonstrates that information about interest rates can prompt households to revise their inflation expectations, making this a particularly relevant and impactful choice for the information treatment. Importantly, at the time of the experiment, the recent shift toward monetary easing by the Fed offered a timely opportunity to examine how interest rate announcements influence inflation expectations and how these communications can be made more effective.

Each economic treatment is presented in a Facebook-style post with identical information about the rate cut but varying in tone — either positive, neutral, or negative. Crucially, none of the treatments include additional economic data or imply future economic developments. This design ensures that any observed effects on expectations are due to tone of the news and sentiment generated by it rather than content differences. To ensure comparability across treatments, I use a pre-trained Word2Vec model based on Google News to calculate the Cosine Similarity between the *Positive* and *Negative* treatments, yielding a score of 0.92 (where 1 indicates identical texts). Sentiment analysis indicates a polarity score of -0.12 for the *Negative* treatment and 0.18 for the *Positive* treatment, reflecting a clear difference in tone.

A novel feature of this survey experiment is that respondents react directly to the treatment, akin to engaging with an actual Facebook post. Respondents are asked to select a reaction that best reflects their feelings after reading the news. This approach not only verifies whether the treatment generated the intended sentiment but also allows for comparisons among respondents who reacted differently within the same treatment group, addressing concerns about varying interpretations of the treatments.

After each information treatment, all respondents, including those in the control group who receive no information, are asked again about their inflation forecast, but this time in the form of a point estimate question. This change in question format avoids asking the same question twice, helps measure forecast revision after the treatment while reducing the risk of demand effects, following current best practices in the literature (Coibion et al., 2023a).

To account for respondents' prior knowledge of interest rates and their understanding of the relationship between interest rates and inflation, additional questions are included after gathering their updated beliefs. Furthermore, data on demographic characteristics, news consumption, numerical literacy, and other data is collected.

3.2 Survey Data

The survey, conducted on November 1-2, 2024 via the Prolific platform, collected responses from 986 individuals representative of the U.S. population. The demographic composition of the sample is relatively balanced across treatment groups, though Democrats are notably overrepresented compared to Republicans. The average perceived inflation over the past 12 months was higher than the actual inflation rate, reflecting both the lasting impact of elevated prices during COVID-19 and households' general tendency to overestimate inflation (D'Acunto et al., 2023).

To compare prior and posterior inflation expectations, I estimate the implied mean for responses to the density forecast question, following Engelberg et al. (2009). Consistent with Armantier et al. (2017), I first fit a generalized beta distribution to the responses and then calculate the mean based on the estimated parameters. However, if a respondent assigns probability to an open interval, assumptions about boundary values are necessary. This, combined with the differing question formats, leads to noticeably different standard deviations for prior and posterior expectations — a common outcome in this type of analysis (see, e.g., Coibion et al., 2023b). Additionally, the higher variability in posterior expectations across treatment groups compared to control group serves as a preliminary indication that the information provision had an impact on expectations. The descriptive statistics are presented in Table C1.

3.3 Results

Given that the *SentimentOnly* treatment and the economic treatments have noticeably different designs — the former based on five real Facebook posts not implicitly related to economics, and the latter based on information about an interest rate cut delivered in different ways — I analyze them separately.

3.3.1 *SentimenOnly* Treatment

To begin, I examine the effect of the *SentimentOnly* treatment on households' forecast revision, $\Delta\pi_i$, which is the difference between the implied mean of the posterior density forecast and the prior point forecast. I capture the effect of treatment using the following baseline model specification:

$$\Delta\pi_i = \alpha + \beta \cdot T_i^{(SentimentOnly)} + \boldsymbol{W}_i \cdot \boldsymbol{\phi} + \epsilon_i \tag{4}$$

where $T_i^{(SentimentOnly)}$ is a binary variable equal to 1 if respondent i was exposed to the *SentimentOnly* information treatment, and $\boldsymbol{W}_i$ is a vector of demographic characteristic. This treatment is distinct in that it conveys emotional, non-economic stimuli, making it a test of whether purely sentimental reactions — independent of economic content — can influence inflation expectations.

Selecting news with strong potential to influence households' expectations posed a challenge, as emotionally charged information spreads rapidly and often dominates media consumption. Furthermore, differences in the phrasing of the prior and posterior survey questions, differences across respondents' understanding of inflation, and the possibility of mean reversion in their answers, required meticulous data processing and careful consideration in choosing the appropriate estimation method.

To estimate the model, I employ and compare three methods: Ordinary Least Squares (OLS), Weighted Least Squares (WLS), and Huber robust regression (Huber, 1979). The last one is prevailing in the literature. Table 2 presents the results. Notably, the slope coefficient β is consistently positive across all estimation methods, supporting the hypothesis that negative sentiment generated by the treatment increases inflation expectations. However, the coefficient achieves statistical significance only under WLS estimation.

Table 2: Effect of SentimentOnly Treatment on Forecast Revision

	I	II	III
const	12.332***	8.275***	2.086
	(4.514)	(1.196)	(1.774)
SentimentOnly	1.291	0.967***	0.166
	(1.058)	(0.233)	(0.416)
Estimator	OLS	WLS	Huber
R-squared	0.040	0.235	0.007
Obs.	394	394	394

Notes: The table presents the estimation results for equation (4) using different estimation methods. Controls include gender, age, age-squared, education level, and income. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

WLS's improved explanatory power, as reflected in its higher R^2, and the statistical significance of its results suggest that it is better suited to capturing subtle variance patterns that might not be fully addressed by OLS or Huber robust regression. This advantage likely arises from WLS's ability to account for variability in responses to emotionally charged stimuli, like *SentimentOnly* treatment, by assigning smaller weights to observations with higher variance. For instance, some individuals may react to the same stimulus with heightened emotional intensity, resulting in greater variability in their responses. In contrast, others who are less sensitive or less familiar with such stimuli may exhibit more consistent and moderate responses, leading to lower variance. By addressing these systematic differences, WLS effectively accounts for the structured heteroscedasticity inherent in the data.

3.3.2 Economic Treatments

Next, I analyze the effects of the economic treatments. Here there are several nuances that must be carefully considered. First, it is crucial to ensure that each treatment induces the intended sentiment. Figure C1 demonstrates that around 70% of respondents exposed to the *Negative* treatment selected reactions associated with negative sentiment. For the *Neutral* and *Positive* treatments, most respondents (60-70%) reported positive sentiment, with a similar balance of positive and negative reactions across the two treatments. Interestingly, the fact that some respondents in the *Negative* treatment group selected positive reactions, and vice versa, highlights the subjective nature of treatments' interpretation.

Another important aspect is understanding the connections respondents make between inflation and interest rates. The survey reveals that only 10% of respondents believe that lower interest rates lead to higher inflation, while 22% think higher interest rates cause inflation to rise. These views may stem from the 41% of respondents who subscribe to the belief that higher inflation leads to higher interest rates. Additionally, most respondents perceive the probability of a rate change within the next 12 months to be low, and fewer than 20% believe such a change would significantly impact inflation. Moreover, half of the respondents were already aware of the rate cut news on which the information treatment was based before participating in the survey. This suggests that the impact of the economic information itself may be limited, necessitating an approach that emphasizes respondents' attentiveness to the information provided. Hence, I adopt the framework proposed by Coibion et al. (2022), with a detailed explanation provided by Weber et al., 2023, to estimate the following specification:

$$\pi_i' = \alpha \cdot \pi_i + \sum_{k=2}^{4} \beta_k \cdot T_i^{(k)} + \sum_{k=2}^{4} \gamma_k \cdot T_i^{(k)} \cdot \pi_i + \boldsymbol{W}_i \cdot \boldsymbol{\phi} + \epsilon_i \tag{5}$$

where π_i is the prior inflation forecast of respondent i, $T_i^{(k)}$ is an indicator variable for individual i and treatment $k \in \{Control, Negative, Neutral, Positive\}$, and, similar to (4), W_i is a vector of demographic characteristic.

Theoretically, α should be equal to 1, as in the absence of any new information, respondents would not update their beliefs. However, as shown in Table 3, $\alpha < 1$ across all estimation methods. This discrepancy likely reflects the methodological nuances of belief elicitation discussed earlier. Thus, α serves as a benchmark for comparison. Bayesian updating implies that $\gamma_k \in [-1, 0]$, as posterior expectations should represent a weighted average of prior beliefs and the information signal. For instance, equation (5) implies that in the presence of an information treatment:

$$\pi_i' = (\alpha + \gamma_k) \cdot \pi_i + \beta_k$$

When $\alpha + \gamma_k = 0$, households place full weight on the signal (β_k), disregarding their prior beliefs. Conversely, when $\alpha + \gamma_k = 1$, $\pi_i' = \pi_i$, indicating that households do not update their beliefs. However, both WLS and Huber robust regression results show $\gamma_k > 0$, suggesting that households place more weight on their prior beliefs than Bayesian updating would predict, warranting further investigation.

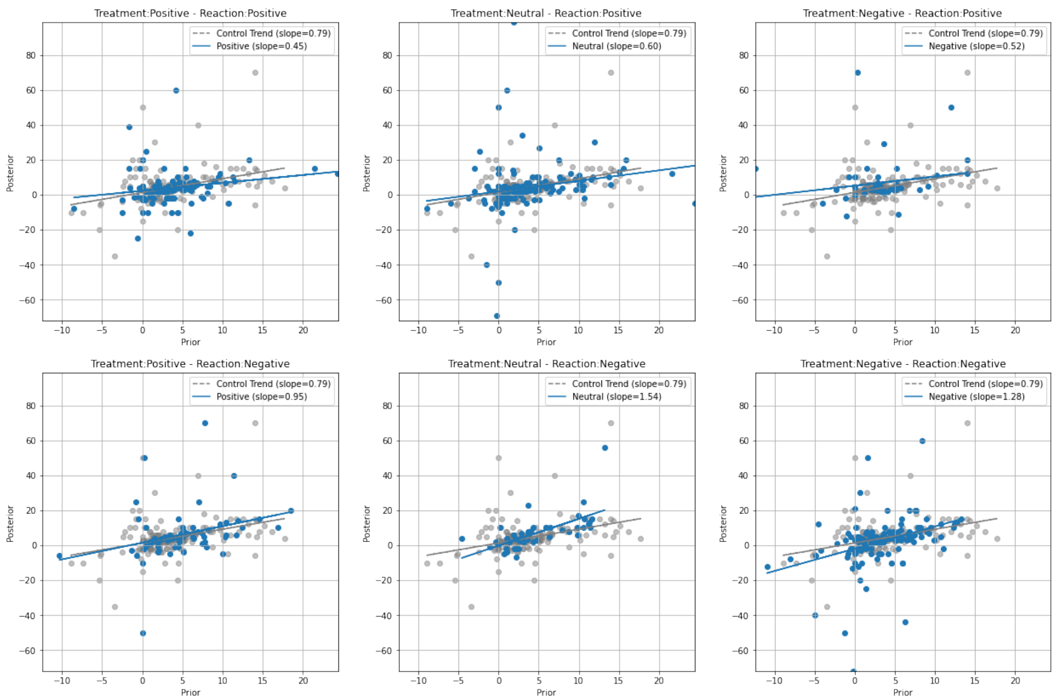


Figure 5: Treatment and Reaction Effects on Forecast Revision

Notes: This figure illustrates the relationship between prior and posterior inflation beliefs across different treatment groups (Positive, Neutral, Negative) and the Control group. Each subplot displays regression trend lines for both the Control group (gray line) and treatment-specific subgroups (blue line), organized by different reaction types. The Control group serves as a baseline, enabling comparisons of treatment effects on inflation forecast revisions.

Forecast revisions across different groups and reactions, as shown in Figure 5, reveal distinct patterns. Respondents' reactions to treatments significantly influence both the weight assigned to new information and the direction of belief adjustments. Those who exhibit negative reactions to treatments tend to discount their prior beliefs and revise expectations upward relative to control group, suggesting that negative sentiment amplify pessimism, leading to higher inflation forecasts. In contrast, respondents with positive reactions display a stronger anchoring effect, assigning greater weight to prior beliefs and adjusting their expectations slightly downward. This implies that positive sentiment generated by the treatments reinforced confidence in existing beliefs, reducing the extent of revisions.

These findings suggest that the sentiment generated by the treatments must be explicitly accounted for in the analysis. To capture this effect, I extend equation (5) to the following specification:

$$\begin{aligned}\pi_i' = \alpha \cdot \pi_i + \sum_{k=2}^{4} \beta_k \cdot T_i^{(k)} + \sum_{k=2}^{4} \gamma_k \cdot T_i^{(k)} \cdot \pi_i + \sum_{k=2}^{4} \rho_k \cdot T_i^{(k)} \cdot R_i^{(-)} + \\ + \sum_{k=2}^{4} \tau_k \cdot T_i^{(k)} \cdot R_i^{(-)} \cdot \pi_i + \boldsymbol{W}_i \cdot \boldsymbol{\phi} + \epsilon_i\end{aligned} \tag{6}$$

where $R_i^{(-)}$ is indicator variable that equal to 1 if respondent i selected a negative reaction after the information treatment. By incorporating sentiment-based interaction terms, equation (6) allows for a more nuanced understanding of how emotional reactions shape belief updating processes.

Comparing the regression results of models (5) and (6) yields several intriguing findings. First, incorporating sentiment into the model causes γ_k to fall within the theoretically expected range of $[-1, 0]$. However, for the *Negative* treatment, this result is not statistically significant, while for other treatments, it is significant only under WLS estimation. Second, $\tau_k > 0$, and this result is robust across all treatment groups and estimation methods. This indicates that negative sentiment generated by the information treatment were a key driver of deviations from Bayesian updating behavior described by the model (5).

Additionally, a strongly robust finding is that $\rho < 0$ for the Negative treatment. Interestingly, this coefficient is also negative for other treatments, suggesting it captures how respondents who reacted negatively interpreted the shared information across all treatments — the Federal Reserve's interest rate cut. According to Binetti et al. (2024), there is a prevalent belief that interest rate cuts can help reduce inflation. Thus, it appears that negative sentiment influenced respondents' attentiveness to the economic information in the treatment. However, a lack of formal economic knowledge led to a misinterpretation of the treatment's implications.

Table 3: Effect of Economic Treatments and Prior Inflation Expectations on Posterior Inflation Expectations

	I	II	III	IV
const	6.494***	3.496***	1.500***	3.775***
	(0.763)	(1.339)	(0.546)	(1.359)
π	0.766***	0.703***	0.780***	0.699***
	(0.058)	(0.071)	(0.037)	(0.072)
$T^{(Negative)}$	-1.145***	-0.628	2.354***	1.394
	(0.339)	(0.555)	(0.747)	(0.850)
$T^{(Neutral)}$	0.055	-0.349	0.418	-0.249
	(0.348)	(0.557)	(0.303)	(0.597)
$T^{(Positive)}$	0.474	-0.174	0.733**	0.055
	(0.354)	(0.570)	(0.349)	(0.652)
$T^{(Negative)} \times \pi$	0.256***	0.192*	-0.186	-0.131
	(0.073)	(0.105)	(0.123)	(0.160)
$T^{(Neutral)} \times \pi$	0.077	0.109	-0.162***	-0.022
	(0.071)	(0.097)	(0.058)	(0.104)
$T^{(Positive)} \times \pi$	-0.042	0.024	-0.283***	-0.080
	(0.070)	(0.100)	(0.068)	(0.118)
$T^{(Negative)} \times R^{(-)}$			-5.148***	-2.658***
			(0.776)	(0.888)
$T^{(Neutral)} \times R^{(-)}$			-1.400**	-0.887
			(0.612)	(1.038)
$T^{(Positive)} \times R^{(-)}$			-0.572	-0.506
			(0.461)	(0.889)
$T^{(Negative)} \times R^{(-)} \times \pi$			0.622***	0.407**
			(0.124)	(0.170)
$T^{(Neutral)} \times R^{(-)} \times \pi$			0.738***	0.537***
			(0.141)	(0.168)
$T^{(Positive)} \times R^{(-)} \times \pi$			0.450***	0.218
			(0.085)	(0.145)
Estimator	WLS	Huber	WLS	Huber
R-squared	0.699	0.090	0.726	0.104
Obs.	791	791	791	791

Notes: Columns I-II present the estimation results of (5). Columns III-IV present the estimation results of (6). Standard errors are in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Although ρ is negative, the net effect of the treatment on the posterior forecast can still be positive if the prior forecast is sufficiently high, as shown in the following relationship:

$$\pi_i' = (\alpha + \gamma + \tau) \cdot \pi_i + (\beta + \rho)$$

Given the complexity of the model (6), Figure 6 provides a visual representation of the model's predictions across different percentiles of prior expectations. It clearly shows that increasing pessimism about future inflation (higher prior) combined with negative reactions to the treatments leads to upward revisions in inflation expectations. Conversely, positive reactions result in downward revisions, reflecting a mitigating effect on inflation concerns. Interestingly, positive treatments and positive reactions to other treatments lead to upward revisions in inflation expectations for relatively optimistic households (with lower prior). A plausible explanation is that this upward movement for lower percentiles may reflect an improvement, as expectations shift closer to the inflation target, reducing extreme optimism.

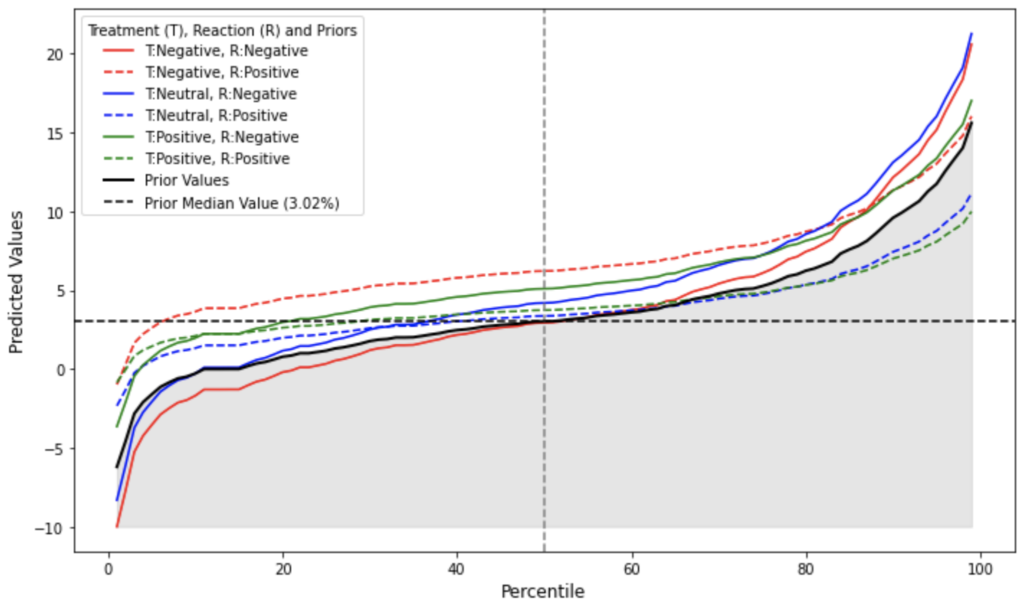


Figure 6: Model Predictions

Notes: The figure shows model (III) predictions from Table 3 across different percentiles of the prior inflation forecast.

4 Discussion

This study offers new insights into how news media shape households' inflation expectations by leveraging Facebook data to analyze emotional responses to news, captured through users' reactions and interactions. Unlike traditional datasets, which often focus solely on the content and tone of news articles, Facebook data provides a direct measure of user reactions. These interactions offer a valuable window into the emotional mechanisms driving expectation formation, allowing us to link sentiment to forecast revisions in ways that traditional approaches cannot.

Facebook news data helps to reveal several important lessons. First, the dynamic nature of topic attention demonstrates that households are not passive consumers of information. Instead, they actively shift focus to different economic and social issues depending on media coverage. This highlights the volatility of media influence on expectations and underscores the need to study expectation formation as a process embedded in evolving information environments.

Second, sentiment analysis based on Facebook reactions highlights the significant role of media-generated sentiment in shaping inflation expectations. Negative reactions associated with negative sentiment, for example, often push inflation expectations upward. The survey experiment — designed to mimic Facebook's interactive dynamics — further validated the importance of sentiment-driven mechanisms. By integrating sentiment elicitation directly into the experimental design, I showed that emotional responses to news are not peripheral but central to how households process and update their economic beliefs. This underscores the role of media not only as disseminators of information but also as generators of sentiment that influence public perceptions.

Furthermore, the findings reveal that identical policy-related news, when framed with different tones, can evoke distinct sentiments, leading to asymmetrical effects on forecast revisions that depend on prior expectations. Specifically, households with relatively high prior inflation expectations who react negatively to the information tend to revise their forecasts upward, while those who react positively adjust their forecasts downward. Conversely, households with relatively low prior expectations exhibit the opposite pattern, with positive reactions resulting in upward adjustments. These findings underscore the importance of considering both the tone used to deliver news and the public's prior expectations in designing effective policy communications.

This paper highlights the need for future research to consider the dual role of media as both an information source and a sentiment-shaping agent. By examining the implications of media-driven sentiment, policymakers and researchers can better understand and anticipate public economic responses in an era of increasingly dynamic and emotionally charged information flows.

References

Armantier, O., Topa, G., Van der Klaauw, W., and Zafar, B. (2017). An overview of the survey of consumer expectations. *Economic Policy Review*, (23-2):51–72.

Armona, L., Fuster, A., and Zafar, B. (2019). Home price expectations and behaviour: Evidence from a randomized information experiment. *The Review of Economic Studies*, 86(4):1371–1410.

Binder, C. and Rodrigue, A. (2018). Household informedness and long-run inflation expectations: Experimental evidence. *Southern Economic Journal*, 85(2):580–598.

Binder, C. C., Kamdar, R., and Ryngaert, J. M. (2024). Partisan expectations and covid-era inflation. *Journal of Monetary Economics*, page 103649.

Binetti, A., Nuzzi, F., and Stantcheva, S. (2024). People's understanding of inflation. Technical report, National Bureau of Economic Research.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022.

Bordalo, P., Coffman, K., Gennaioli, N., and Shleifer, A. (2016). Stereotypes. *The Quarterly Journal of Economics*, 131(4):1753–1794.

Carroll, C. D. (2003). Macroeconomic expectations of households and professional forecasters. *the Quarterly Journal of economics*, 118(1):269–298.

Coibion, O., Georgarakos, D., Gorodnichenko, Y., and Van Rooij, M. (2023a). How does consumption respond to news about inflation? field evidence from a randomized control trial. *American Economic Journal: Macroeconomics*, 15(3):109–152.

Coibion, O., Georgarakos, D., Gorodnichenko, Y., and Weber, M. (2023b). Forward guidance and household expectations. *Journal of the European Economic Association*, 21(5):2131–2171.

Coibion, O., Gorodnichenko, Y., and Weber, M. (2022). Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy*, 130(6):1537–1584.

D'Acunto, F., Malmendier, U., and Weber, M. (2023). What do the data tell us about inflation expectations? In *Handbook of economic expectations*, pages 133–161. Elsevier.

Doms, M. E. and Morin, N. J. (2004). Consumer sentiment, the economy, and the news media. *FRB of San Francisco Working Paper*, (2004-09).

Dräger, L. and Lamla, M. J. (2017). Imperfect information and consumer inflation expectations: Evidence from microdata. *Oxford Bulletin of Economics and Statistics*, 79(6):933–968.

Engelberg, J., Manski, C. F., and Williams, J. (2009). Comparing the point predictions and subjective probability distributions of professional forecasters. *Journal of Business & Economic Statistics*, 27(1):30–41.

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

Gorodnichenko, Y., Pham, T., and Talavera, O. (2024). Central bank communication on social media: What, to whom, and how? *Journal of Econometrics*, page 105869.

Happer, C. and Philo, G. (2013). The role of the media in the construction of public belief and social change. *Journal of social and political psychology*, 1(1):321–336.

Hermida, A., Fletcher, F., Korell, D., and Logan, D. (2012). Share, like, recommend: Decoding the social media news consumer. *Journalism studies*, 13(5-6):815–824.

Huber, P. J. (1979). Robust smoothing. *Robustness in statistics*, pages 33–47.

Huseynov, S. and Murad, Z. (2024). Seeing through different lenses: Partisanship and updating of inflation expectations. *Available at SSRN*.

Kamdar, R. and Ray, W. (2023). Polarized expectations, polarized consumption. *Polarized Consumption (September 19, 2023)*.

Kamdar, R. and Ray, W. (2024). Attention-driven sentiment and the business cycle. Technical report, Center for Applied Economics and Policy Research, Department of Economics . . . .

Korhonen, I. and Newby, E. (2019). Mastering central bank communication challenges via twitter. Technical report, BoF Economics Review.

Lamla, M. J. and Lein, S. M. (2014). The role of media for consumers' inflation expectation formation. *Journal of Economic Behavior & Organization*, 106:62–77.

Larsen, V. H., Thorsrud, L. A., and Zhulanova, J. (2021). News-driven inflation expectations and information rigidities. *Journal of Monetary Economics*, 117:507–520.

Malmendier, U. and Nagel, S. (2016). Learning from inflation experiences. *The Quarterly Journal of Economics*, 131(1):53–87.

Mankiw, N. G. and Reis, R. (2002). Sticky information versus sticky prices: a proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics*, 117(4):1295–1328.

McCallum, A. K. (2002). Mallet: A machine learning for language toolkit. http://mallet.cs.umass.edu.

Mian, A., Sufi, A., and Khoshkhou, N. (2023). Partisan bias, economic expectations, and household spending. *Review of Economics and Statistics*, 105(3):493–510.

Mosquera, R., Odunowo, M., McNamara, T., Guo, X., and Petrie, R. (2020). The economic effects of facebook. *Experimental Economics*, 23:575–602.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica: journal of the Econometric Society*, pages 315–335.

Newman, N., Fletcher, R., Eddy, K., T., R. C., and Nielsen, R. K. (2023). Digital news report 2023.

Nguyen, V. H. and Claus, E. (2013). Good news, bad news, consumer sentiment and consumption behavior. *Journal of Economic Psychology*, 39:426–438.

Pfajfar, D. and Santoro, E. (2013). News on inflation and the epidemiology of inflation expectations. *Journal of Money, Credit and Banking*, 45(6):1045–1067.

Shiller, R. J. (1997). Why do people dislike inflation? In *Reducing inflation: Motivation and strategy*, pages 13–70. University of Chicago Press.

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690.

Stantcheva, S. (2024). Why do we dislike inflation? Technical report, National Bureau of Economic Research.

Tandoc Jr, E. C., Lim, D., and Ling, R. (2020). Diffusion of disinformation: How social media users respond to fake news and why. *Journalism*, 21(3):381–398.

Weber, M., Candia, B., Afrouzi, H., Ropele, T., Lluberas, R., Frache, S., Meyer, B. H., Kumar, S., Gorodnichenko, Y., Georgarakos, D., et al. (2023). Tell me something i don't already know: Learning in low and high-inflation settings. Technical report, National Bureau of Economic Research.

Appendix A: Data and Methods

A1. News Data

The news data consists of Facebook posts from the official pages of 64 major U.S. media outlets, covering a wide range of topics. The historical data was collected using CrowdTangle, a platform provided by Meta Platforms, Inc., that aggregates publicly available content from social media (Facebook, Twitter, Instagram, and Reddit). CrowdTangle tracks verified accounts and profiles, excluding personal information and private or restricted accounts.

A2. Data Cleaning

To focus on textual news, all posts containing only videos or photos were removed. The remaining posts, classified as "Link" by CrowdTangle, included a title, a description, a link to the news provider's official page, and occasionally, an image or video. To ensure all posts were in English, the Python library `langdetect` was used to identify and exclude non-English content. The textual components of each post (title, description, and link text) were combined into a single text entry.

Standard Natural Language Processing (NLP) steps were applied for data preparation, following Gentzkow et al. (2019). Text was converted to lowercase, and punctuation, accents, and stop-words (e.g., "is," "the," "and") were removed. N-gram analysis eliminated repetitive and irrelevant phrases (e.g., media names). Bigrams (two-word phrases) and trigrams (three-word phrases) were created to preserve meaning. Lemmatization reduced words to their root forms (e.g., "running," "runs," and "ran" became "run"). Only informative parts of speech (nouns, adjectives, verbs, adverbs) were retained. Short tokens and any remaining stop-words were removed.

A dictionary was created to map each unique word to a numeric ID, enabling conversion of posts into a "bag-of-words" format: a list of tuples where each tuple contains a word ID and its frequency in the post. Words appearing in fewer than ten posts or more than 50% of posts were removed for additional cleaning. The final dataset contained 7.9 million posts spanning November 2013 to January 2024.

A3. Topic Modeling

Topics were extracted using the Latent Dirichlet Allocation (LDA) model (Blei et al. (2003)), a widely used technique in NLP and economics. LDA treats each document as a mixture of topics and each topic as a collection of words, assuming Dirichlet distributions for both. For each post, LDA identifies a mixture of topics, assigning probabilities to words within each topic. The final output includes word distributions for topics and topic distributions for documents.

The LDA model was implemented using MALLET (Machine Learning for Language Toolkit, McCallum (2002)). The dictionary and "bag-of-words" representation were used as inputs. Following Larsen et al. (2021), the number of topics was set to 80. Labels for the topics were assigned based on the associated word lists and a random sample of posts linked to each topic.

A4. Topic Time Series

The trained LDA model was applied to the entire dataset to create topic time series. Each post was assigned a dominant topic based on the highest probability. For each day, the frequency of each dominant topic was calculated. Sentiment analysis was performed daily using the Harvard IV-4 dictionary, scoring texts for positive and negative sentiment. For days with multiple posts on a topic, the post contributing most to the topic (as determined by LDA) was used for sentiment scoring.

Sentiment scores were calculated as the difference between the number of positive and negative words, divided by their total. These scores were used to adjust daily topic frequencies. Finally, the daily time series were standardized, and monthly averages were computed for use in LASSO analysis.

Appendix B: Empirical Part

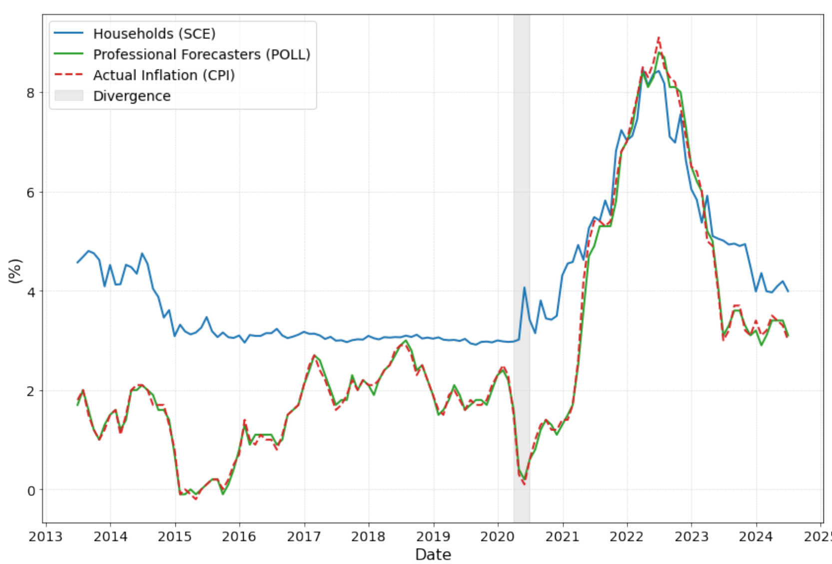


Figure B1: Inflation Expectations of Households and Professionals

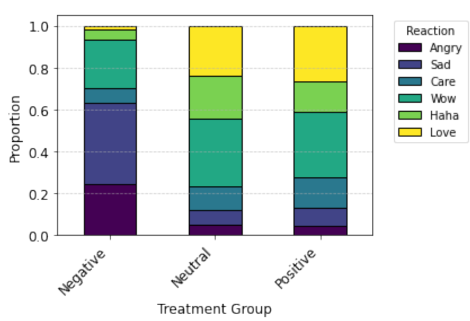
Notes: The blue line represents the median one-year-ahead inflation expectations from the SCE. The green line shows the median inflation expectations of professional forecasters from the Reuters Poll. The red dashed line indicates the actual inflation (core CPI).

Appendix C: RCT Experiment

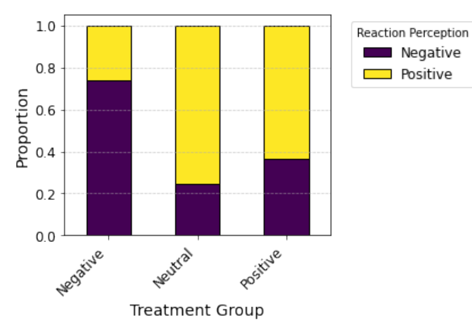
Table C1: Descriptive Statistics

Variable	Total (# obs = 986)	Control (# obs = 199)	Negative (# obs = 196)	Neutral (# obs = 196)	Positive (# obs = 200)	SentimentOnly (# obs = 195)
Demographics						
Age	44.57 (15.91)	44.48 (16.22)	45.75 (15.39)	43.83 (16.16)	44.07 (15.20)	44.72 (16.63)
Female	50.61%	50.25%	55.61%	48.47%	49.00%	49.74%
White	68.15%	64.32%	63.27%	70.92%	73.50%	68.72%
Democrat	49.09%	45.73%	54.59%	49.48%	44.72%	52.33%
Republican	36.41%	37.19%	36.22%	38.14%	36.68%	34.72%
Employed (part/full)	68.66%	71.86%	69.90%	65.82%	72.00%	63.59%
College	51.42%	49.75%	54.08%	53.57%	52.00%	47.69%
Less than 30K	28.70%	32.66%	28.57%	29.59%	22.50%	30.26%
30-100K	54.26%	50.75%	56.12%	53.57%	60.50%	50.26%
Inflation						
Prior Expectations	3.66 (4.23)	3.62 (4.18)	3.34 (3.92)	3.95 (4.54)	4.16 (4.25)	3.21 (4.22)
Posterior Expectations	4.46 (12.03)	4.16 (9.16)	3.53 (13.35)	4.84 (13.63)	4.73 (10.95)	5.03 (12.60)
Personal Inflation	11.93 (22.47)	10.53 (13.01)	11.56 (16.61)	12.87 (15.78)	10.71 (14.13)	14.06 (40.39)
Perceived Inflation	7.36 (14.39)	7.39 (10.47)	6.40 (10.68)	7.96 (11.06)	6.62 (12.43)	8.44 (23.23)

Notes: Values represent percentages/shares (indicated by %) or means with standard deviations in parentheses. Groups correspond to different treatments in the study. Sample sizes for each group are provided in parentheses in the header row.



(a) Reaction



(b) Reaction Perception

Figure C1: Analysis of respondents' reactions and associated sentiment.

Notes: Figure (a) illustrates the share of respondents in each treatment group who selected one of the available reactions. Figure (b) depicts the perception (positive or negative) associated with the selected reactions.

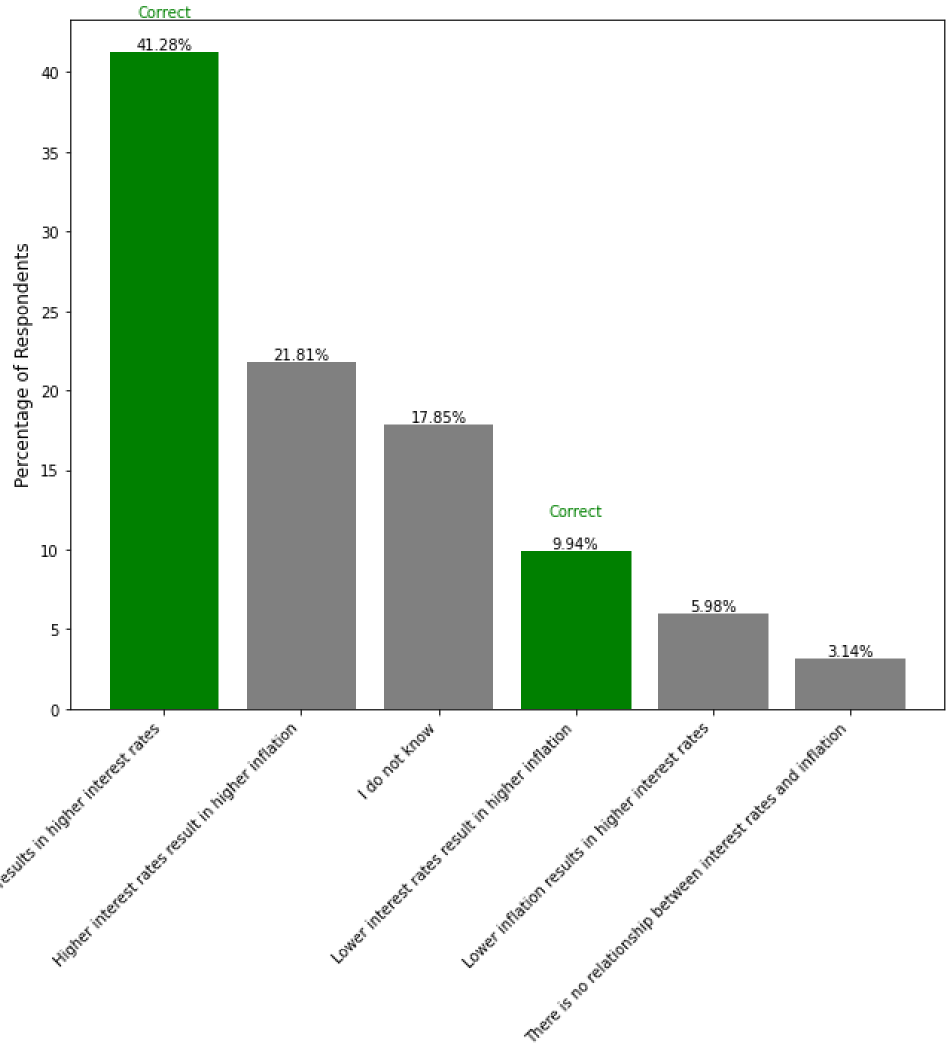
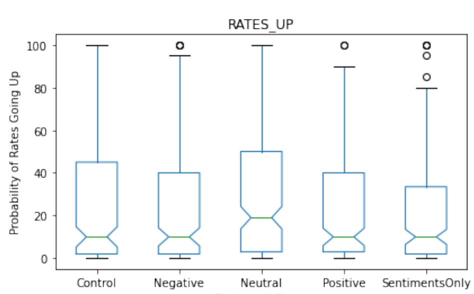
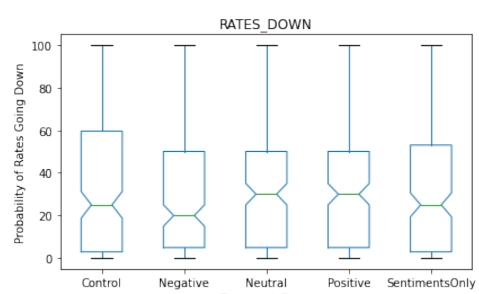


Figure C2: "In general, what do you think is the relationship between interest rates and inflation?"



(a) Rates Up



(b) Rates Down

Figure C3: "What do you think is the percent chance that the Federal Reserve will increase/decrease interest rates in the next 12 months?"

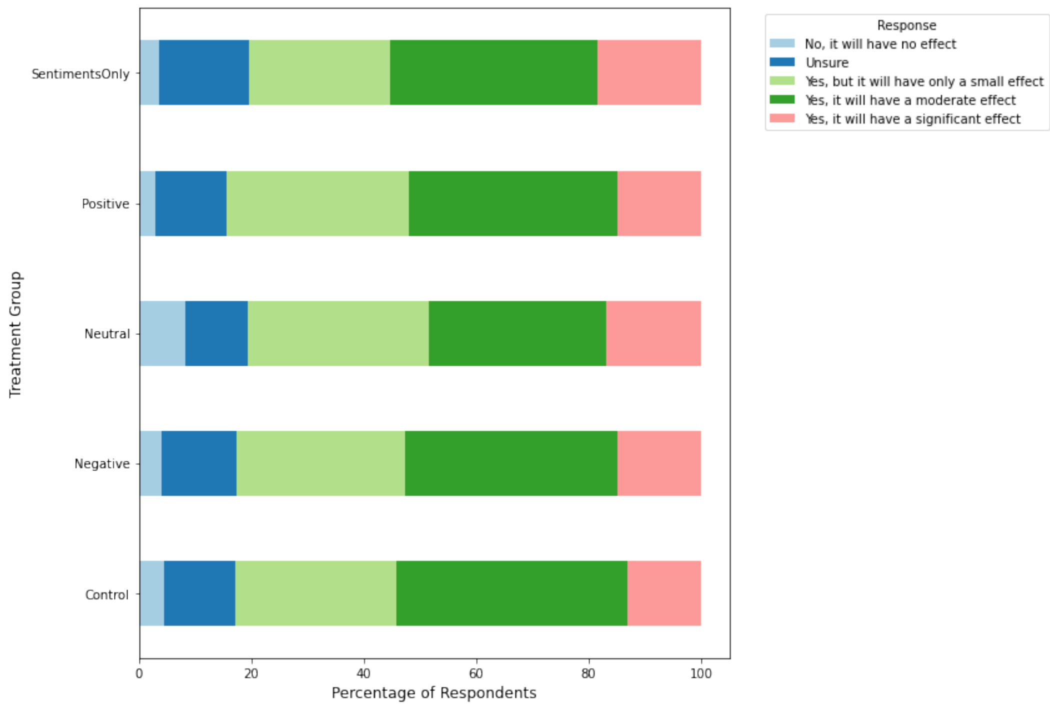


Figure C4: "Do you think the Federal Reserve's future change in interest rates will affect inflation in the next 12 months?"

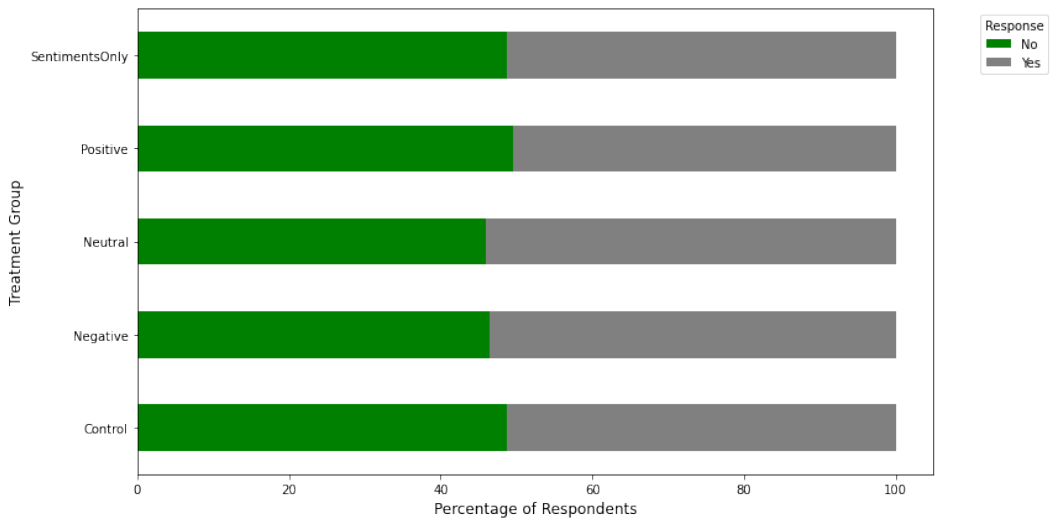


Figure C5: "Were you aware that the Federal Reserve cut interest rates in September, prior to taking this survey?"

Table C2: Comparison of Treatment Effects Across Demographic Groups with Prior Expectations Interaction

	Baseline	Gender	Age	Education	Low Income
	(1)	(2)	(3)	(4)	(5)
const	0.416***	0.401*	0.457***	0.364	0.409***
	(0.132)	(0.206)	(0.150)	(0.256)	(0.145)
$T^{SentimentOnly}$	1.080***	0.745**	1.089***	1.216***	0.759***
	(0.189)	(0.330)	(0.224)	(0.301)	(0.262)
Female		0.026			
		(0.268)			
$T^{SentimentOnly}$·Female		0.458			
		(0.405)			
Old			-0.180		
			(0.314)		
$T^{SentimentOnly}$·Old			0.026		
			(0.427)		
College				0.071	
				(0.299)	
$T^{SentimentOnly}$·College				-0.389	
				(0.429)	
LowIncome					0.038
					(0.341)
$T^{SentimentOnly}$·LowIncome					0.494
					(0.440)
Observations	394	394	394	394	394
R^2	0.077	0.083	0.078	0.079	0.085
Adjusted R^2	0.074	0.076	0.071	0.072	0.078
Residual Std. Error	2.284(df = 392)	2.282(df = 390)	2.288(df = 390)	2.286(df = 390)	2.279(df = 390)
F Statistic	32.596*** (df = 1.0; 392.0)	11.733*** (df = 3.0; 390.0)	11.031*** (df = 3.0; 390.0)	11.216*** (df = 3.0; 390.0)	12.137*** (df = 3.0; 390.0)

Note: *p<0.1; **p<0.05; ***p<0.01