



EUROPEAN CENTRAL BANK
EUROSYSTEM

Working Paper Series

Giovanni Trebbi Inflation narratives and expectations

No 3158

Disclaimer: This paper should not be reported as representing the views of the European Central Bank (ECB). The views expressed are those of the authors and do not necessarily reflect those of the ECB.

Abstract

I study how demand-supply narrative disagreement between general and specialized newspapers can explain households' absolute gap in inflation expectations with experts. I measure inflation narratives via a Causality Extraction algorithm that can identify causal relationships between events in a text and, hence, extract the perceived triggers of inflation. Causal relations can explain why narratives affect people's beliefs and cannot be captured by dictionary methods, topic models, and word embeddings. I then classify inflation narratives into demand and supply narratives based on their focus on demand and supply triggers. I measure narrative disagreement between general and specialized newspapers from their attention difference on demand and supply narratives. The absolute expectation gap widens when narrative disagreement increases, especially for non-college-educated and older households. Unlike the narratives of specialized newspapers, the narratives of general newspapers incorrectly align with experts' demand-supply views.

JEL Classification: C53, D1, D8, E3

Keywords: News Media, Causality Extraction, Natural Language Processing

Non-Technical Summary

Central banks attach great importance to anchoring private-sector inflation expectations, as unanchored expectations weaken their credibility and hinder the pursuit of price stability. However, expectations are not homogeneously anchored across private agents: while professionals' expectations are typically well anchored, households' expectations often diverge from them. In theory, greater newspaper coverage of inflation should help narrow this expectation gap by reducing the costs households face in acquiring and processing information. However, empirical evidence suggests the opposite. A potential reason is that general-interest newspapers, which households are more likely to read, may transmit professionals' views less precisely than specialized outlets. As a result, disagreement among newspapers might signal when the expectation gap widens.

This paper investigates whether differences in how general-audience and specialised newspapers describe the causes of inflation help account for movements in the expectation gap. The analysis focuses on the demand–supply narrative disagreement, i.e., the extent to which these two types of newspapers differ in attributing inflation to demand-side factors (e.g., strong consumer spending) versus supply-side factors (e.g., energy price increases). This focus originates from recent survey evidence suggesting households and experts attribute the post-pandemic inflation waves to different economic drivers.

The study employs a novel text-based measure of demand and supply narratives, derived from over 180,000 U.S. newspaper articles on inflation published between 1991 and 2022. Articles are drawn from three general-audience newspapers (The New York Times, USA Today, and The Washington Post) and one specialized outlet (The Wall Street Journal). A text analysis algorithm called Causality Extraction identifies explicit causal statements about inflation and classifies them into demand- or supply-related categories. These media-based indicators are then compared with the monthly inflation expectations of U.S. households and professionals, as reported in the University of Michigan Survey of Consumers and the Survey of Professional Forecasters, respectively.

The results show that the inflation expectation gap between U.S. households and experts widens when narrative disagreement between general and specialised newspapers increases. The relationship is stronger for households without a college degree and for older individuals, i.e., groups more likely to read general-audience newspapers. The relationship also strengthens when the level and persistence of inflation rise, namely when the costs of being uninformed about inflation increase. Moreover, the analysis finds that while general newspapers convey demand–supply stories that are consistent with households' views of the economy, these stories are inconsistent with both experts' economic views and with macroe-

conomic data.

For monetary policymakers, these results imply that efforts to bridge the gap between households' and experts' expectations cannot rely solely on increasing media coverage. Instead, attention should be paid to ensuring that clear and consistent explanations of the inflation drivers reach a broad audience through multiple channels, including general-audience outlets. To track whether these explanations are homogeneously absorbed across private agents, this paper introduces simple measures of demand and supply narratives that serve as real-time proxies for households' and professionals' economic views.

1 Introduction

It is well-known that there is a (sometimes) large and volatile gap between the inflation expectations of experts and households (see figure 1).¹ This paper suggests that their absolute expectation gap widens with demand-supply narrative disagreement between general and specialized newspapers. To explain how the absolute expectation gap varies over time, Carroll (2003) proposes a model in which households form expectations from newspapers that report experts' views. Its main prediction, in turn, is a negative relationship between the absolute expectation gap and inflation press coverage. Unlike Carroll (2003),² Pfajfar and Santoro (2013) find evidence of a positive relationship, which I confirm using data up to 2022 (see table 5). A plausible rationale is that households read general newspapers reporting experts' views less accurately than specialized ones, as newspapers differ in their audiences (Pew Research Center, 2012) and reporting styles (Nimark & Pitschner, 2019). Therefore, disagreement between general and specialized newspapers may explain why the absolute expectation gap widens with inflation press coverage.

Andre, Haaland, Roth, Wiederholt, and Wohlfart (2024) suggest that disagreement about the triggers of inflation, i.e., inflation narratives, may lead to differences in economic expectations. Their survey evidence from late 2021 to early 2022 shows that households attribute inflation to supply factors (e.g., "labor shortage") more often than to demand factors (e.g., "fiscal stimulus") compared to experts. In addition, the early-2021 survey evidence by Andre, Pizzinelli, Roth, and Wohlfart (2022) shows that their narrative disagreement explains differences in their expected inflation responses to macroeconomic shocks. However, the historic context of these surveys raises the question of whether narrative disagreement accounts for the absolute expectation gap over time. While both surveys occur during increases in the absolute expectation gap and inflation, Figure 1 shows that there are times when only the absolute expectation gap widens. Do these periods also exhibit greater narrative disagreement? The lack of time-series data on households' and experts' narratives limits our ability to answer this question. However, narratives from general and specialized newspapers may serve as their proxies. Therefore, this study investigates whether the absolute expectation gap widens with narrative disagreement between general and specialized newspapers.

This paper introduces a news-based measure of demand and supply narratives to answer this ques-

¹For a detailed review of the literature on households' inflation expectations, see Weber, D'Acunto, Gorodnichenko, and Coibion (2022).

²Using inflation expectations from the Michigan Survey of Consumers (MSC) and the Survey of Professional Forecasters (SPF) to measure households' and experts' expectations, as well as inflation articles from the New York Times and the Washington Post.

tion. This measure is derived from over 180,000 inflation articles³ published between 1991 and 2022 by three general newspapers (New York Times, USA Today, and Washington Post)⁴ and a specialized one (Wall Street Journal).⁵ Demand and supply narratives are derived via Causality Extraction (CE), a Natural Language Processing (NLP) tool designed to extract causal relations from text and introduced to Economics and Finance by Baele, De Jong, and Trebbi (2023). In a nutshell, my CE algorithm 1) finds within a sentence causal relations expressed via explicit causal keywords (e.g., *because*, *trigger*), 2) checks that inflation is the mentioned effect, and 3) extracts inflation narratives as the corresponding causes. I then use a dictionary method to classify all inflation narratives into the categories of demand and supply narratives by Andre et al. (2024). My narrative measures for general and specialized newspapers, $NetDemand^G$ and $NetDemand^S$, are the differences in the monthly volume of articles with demand and supply narratives. Finally, narrative disagreement $NetDemand^{G-S}$ is the difference in $NetDemand$ between general and specialized newspapers.

My central testable hypothesis is that the absolute expectation gap widens with narrative disagreement. My regressor of interest is the absolute value of $NetDemand^{G-S}$, which captures the quantity of narrative disagreement between newspapers. The results confirm my central testable hypothesis at both the aggregate and individual levels, while controlling for the level and volatility of inflation, as well as households' demographics and perceptions of news about inflation. In addition, the absolute expectation gap widens with $|NetDemand^{G-S}|$ also after controlling for disagreement about two alternative dimensions: a) whether inflation is increasing or decreasing, and b) whether articles talk about realized or future inflation episodes. Digging deeper, I assess how the results vary across household demographics. Bryan and Venkatu (2001) and D'Acunto, Hoang, Paloviita, and Weber (2019) find that age and education can explain expectation differences across households. Building on their findings, I show that the positive association between the absolute expectation gap and $|NetDemand^{G-S}|$ is stronger for individuals without a college degree and for older individuals. These results are not surprising, as the news readership of college-educated households is closer to that of experts (Pew Research Center, 2012), and older people are more likely to read newspapers (Pew Research Center, 2023).

A crucial assumption underlying my results is that general newspapers report expert opinions less accurately than specialized newspapers. To test this assumption, I examine whether the narratives of general and specialized newspapers capture the demand and supply views of individual households and experts

³Inflation articles mention at least one of the inflation expressions "inflation," "cpi," "consumer price," "ppi," or "producer price."

⁴I treat the inflation articles published by these three newspapers as if a single general newspaper published them.

⁵These newspapers are the top four newspapers by daily circulation in the U.S., and three appear among the top 20 inflation news outlets consulted in the survey by Andre et al. (2024).

differently. Following Dräger, Lamla, and Pfajfar (2016), I measure individual demand and supply views from the product of expectations regarding future changes in inflation and unemployment. This product is positive/negative depending on whether supply/demand views are dominant. Therefore, it should decline with both $NetDemand^G$ and $NetDemand^S$ for general and specialized narratives to align correctly with individual demand and supply views. The results show this is the case for households but not for experts. In particular, experts' demand and supply views are correctly aligned with the narratives of specialized newspapers but incorrectly aligned with those of general newspapers. Therefore, the narratives of general newspapers differ in how they capture the views of households and experts. Interestingly, the narratives of general newspapers also incorrectly align with the joint dynamics of realized inflation and unemployment, as $NetDemand^G$ rises with the product of realized inflation and unemployment. In contrast, for specialized newspapers, $NetDemand^S$ correctly declines with the product of realized inflation and unemployment. Therefore, these results suggest that general newspapers communicate incorrect narratives to households.

I confirm the robustness of my findings with numerous checks. First, the absolute expectation gap might widen differently based on the specific type of demand or supply narrative underlying newspaper disagreement. I show that it widens the most when newspapers disagree about the importance of monetary policy narratives. Second, the absolute expectation gap might move differently based on the sign of $NetDemand^{G-S}$. That is, based on whether general or specialized newspapers publish relatively more demand narratives. I find that the absolute expectation gap widens regardless of the sign of $NetDemand^{G-S}$, though the sign can explain differences in the magnitude of this association. Third, my three general newspapers differ in their partisanship, implying that narrative disagreement might be a proxy for partisan disagreement. Measuring narrative disagreement between individual general newspapers and the Wall Street Journal leaves the results virtually unchanged, thereby alleviating potential partisanship concerns. Fourth, changes in the level and persistence of inflation might influence the incentives to gather information and, hence, affect the relationship between the absolute expectation gap and $|NetDemand^{G-S}|$. In addition, narrative disagreement may serve as a proxy for forecast disagreement among experts. Including interactions with these variables reveals that the relationship becomes stronger when inflation rises but is unaffected by expert disagreement. These findings contribute to the work of Andre et al. (2022) and Andre et al. (2024) by demonstrating that the relationship between the absolute expectation gap and narrative disagreement varies over time. Finally, I test whether households' expectations also deteriorate in the presence of narrative disagreement. Households' forecast errors also widen with narrative disagreement, though only when not controlling for macroeconomic variables.

These findings have significant implications for policymakers, as increased reporting on inflation may bridge the gap between households' and experts' expectations only when there is minimal narrative disagreement in the media landscape. Understanding the fluctuations in the expectation gap is critical for central banks, which strive to anchor inflation expectations (Powell, 2020). Studying when households' expectations align with those of experts is especially important, as experts' expectations are more reactive to central bank communication (Blinder, Ehrmann, Fratzscher, De Haan, & Jansen, 2008). With its focus on the role of media in the expectation formation process, this question is also timely given recent calls to broaden central bank communication to the public via newspapers (Blinder, Ehrmann, De Haan, & Jansen, 2024). In particular, these findings suggest that if central bank communication aims to reduce the dispersion of inflation forecasts among different groups of individuals, it must disseminate its inflation narratives across a broad range of channels.

This paper adds to a rich body of empirical research on the formation of inflation expectations. This literature highlights the roles of experiences (Malmendier & Nagel, 2016), cognitive abilities (D'Acunto et al., 2019; D'acunto, Hoang, Paloviita, & Weber, 2023), grocery prices (Cavallo, Cruces, & Perez-Truglia, 2017), gas prices (Coibion & Gorodnichenko, 2015), and monetary policy communication (Coibion, Gorodnichenko, & Weber, 2022). Using the narratives of general and specialized newspapers as proxies for those of households and experts, I show that the relationship between disagreement in narratives and expectations, as documented by Andre et al. (2022) and Andre et al. (2024), varies over time. In particular, it strengthens with the level of inflation, thus at the times of their surveys, as well as with its persistence, which is also high in other periods (see figure A9).

By analyzing media's role in the expectation formation process, this paper speaks more directly to the work by Pfajfar and Santoro (2013), Ehrmann, Pfajfar, and Santoro (2018), Larsen, Thorsrud, and Zhulanova (2021), and Mazumder (2021). While also building on the epidemiological model by Carroll (2003), this work makes an important contribution to the use of textual data. The earlier literature employs dictionary-based methods or topic models, which do not capture causal relations in text and, therefore, cannot be used to measure narratives. Based on recent evidence highlighting the role of inflation narratives in the expectation formation process (Andre et al., 2022; Andre et al., 2024), I adopt causality extraction to measure them from newspapers.

By analyzing individual expectations about the future comovement of inflation and unemployment, this work also speaks to work by Fendel, Lis, and Rülke (2011), Dräger et al. (2016), and Geiger and Scharler (2021). These studies examine whether households' and experts' expectations align with the Phillips curve, i.e., whether they anticipate inflation and unemployment to move in opposite directions.

All three studies show that expectations are revised in a way consistent with the Phillips curve. However, Dräger et al. (2016) show that this consistency deteriorates with press coverage about inflation from the New York Times and Washington Post. In contrast, I show that households' expectations become more consistent with the Phillips curve when these newspapers publish more demand narratives.

A final section connects this paper to the media bias literature exemplified by Gentzkow and Shapiro (2006) and discusses the potential reverse causality concern affecting narrative disagreement. In fact, narrative disagreement may stem from the non-random selection of narratives by general and specialized newspapers to cater to households' and experts' different priors. My results are inconclusive regarding whether narrative disagreement is indeed endogenous because there is no systematic narrative disagreement in the direction of demand or supply narratives (see table 4), but general newspapers' supply narratives are misaligned with macroeconomic dynamics (see table A12). Unfortunately, the use of aggregate news measures prevents me from identifying the causal effect of narrative disagreement on the absolute expectation gap. In addition, existing methods for improving identification (Chahrour, Shapiro, & Wilson, 2024) cannot be applied here due to the lack of time series data on the narratives of households and experts. Therefore, while this paper provides at best evidence of a relationship between the absolute expectation gap and narrative disagreement, much work remains to be done to understand how media narratives shape the beliefs and expectations of their readers.

2 Hypotheses Development

The research question of this project is whether the absolute expectation gap widens when narrative disagreement between general and specialized newspapers increases. This research question yields seven testable hypotheses.

Carroll (2003) predicts that the absolute expectation gap narrows with inflation press coverage. While he finds supporting evidence, Pfajfar and Santoro (2013) provide contrary evidence, possibly due to their different sample: they use data up to 2011, while Carroll (2003) uses data up to 2000. This sample difference highlights the need for testing the relationship proposed by Carroll (2003) with a larger sample. Additionally, the existence of two groups of newspapers (general and specialized) allows for exploring whether the results vary based on the type of newspaper reporting on inflation. Therefore, my first hypothesis is:

The absolute expectation gap narrows with inflation press coverage by general
or specialized newspapers. (H1)

Andre et al. (2024) show that households and experts disagree in their attention to demand and supply narratives. Andre et al. (2022) show, in turn, that narrative heterogeneity might explain inflation expectation heterogeneity across households and experts. Since households and experts differ in their readership of general and specialized newspapers (Pew Research Center, 2012), narrative disagreement between these newspapers might capture narrative disagreement between households and experts. Therefore, the absolute expectation gap may widen with disagreement in newspaper narratives. Narrative disagreement is observed when general newspapers publish relatively more demand or supply narratives than specialized newspapers. This could be the case when general and specialized newspapers respectively attribute a rise in inflation to a surge in energy costs (a supply narrative) and looser fiscal policy (a demand narrative). Therefore, my second hypothesis is:

The absolute expectation gap widens with narrative disagreement between general and specialized newspapers. (H2)

Figure 3 displays the cross-sectional interquartile ranges of MSC and SPF inflation expectations, revealing a significant dispersion across households, which is substantially larger than the dispersion among experts. Existing evidence attributes this cross-sectional dispersion to demographic characteristics, namely sex (D’Acunto, Malmendier, & Weber, 2021), cognitive abilities (D’Acunto et al., 2019), socioeconomic status (Bruine de Bruin et al., 2010), and age (Bryan & Venkatu, 2001). In particular, inflation expectations are higher for women than men and decrease with income, education, and age. These demographics might also explain whether the relationship between the absolute expectation gap and narrative disagreement differs across households. For instance, the relationship might be weaker for highly educated and rich households because they are more likely to be readers of specialized newspapers (Pew Research Center, 2012). In contrast, it should be stronger for older individuals because they pay more attention to newspapers (Pew Research Center, 2023). However, it should not change based on sex because there are no significant differences in newspaper consumption between men and women (Pew Research Center, 2023). Therefore, my third hypothesis is:

Household demographics moderate the relationship between the absolute expectation gap and narrative disagreement. In particular, the relationship strengthens with age, weakens with income and education, and does not change with sex. (H3)

The second hypothesis assumes that narrative disagreement between newspapers captures narrative disagreement between households and experts. For this to happen, the narratives of general and spe-

cialized newspapers should align differently with those of households and experts. Otherwise, narrative disagreement between newspapers would not translate into narrative disagreement between households and experts, as they treat both newspapers equally. The narratives of households and experts can be proxied by their expectations on the future co-movement of inflation and unemployment. The rationale for this proxy is that demand and supply narratives prescribe different signs for this comovement. In particular, demand narratives prescribe a negative comovement (e.g., "looser monetary policy"), while supply narratives prescribe a positive one (e.g., "increasing energy prices"). Thus, newspapers' demand and supply narratives may influence how individuals expect inflation and unemployment to behave. For example, individuals might anticipate that inflation and unemployment will move in opposite directions to a greater extent when newspapers publish relatively more demand narratives. Importantly, this relationship should vary across newspapers and types of individuals to justify the assumption that narrative disagreement between newspapers captures narrative disagreement between households and experts. For example, experts' expectations may align more closely with the narratives of specialized newspapers than those of general newspapers. If this were the case, one policy lesson for central bankers would be to increase their communication through general newspapers, so that their narratives align more closely with those of experts. Therefore, my fourth and fifth hypotheses are:

Households and experts expect inflation and unemployment to move in opposite directions by a larger/smaller degree when newspapers publish relatively more demand/supply narratives. (H4)

Households' expectations are more strongly correlated with the narratives of general newspapers than specialized newspapers, while the opposite holds for experts' expectations. (H5)

3 Data

3.1 Inflation News

The source of news articles is Factiva, a comprehensive online database of news articles. I download all news articles that mention at least one of the keywords from the inflation dictionary compiled by Baker, Bloom, Davis, and Kost (2021): "inflation", "consumer price", "producer price", "cpi", and "ppi." I call these keywords "inflation expressions." My sample includes all days between 1991 and 2022. To focus on news about U.S. inflation, I download only articles that mention "United States" in their Factiva

regional identifier. I filter out short news articles with fewer than 200 words (3.87% of the corpus), as dictionary-based methods are typically noisy for brief texts (Shapiro, Sudhof, & Wilson, 2022).

The news sources considered are the New York Times (NYT), USA Today (USAT), the Washington Post (WaPo), and the Wall Street Journal (WSJ). These are the top four U.S. newspapers by daily circulation⁶. WSJ, NYT, and WaPo also appear among the top 20 inflation news outlets consulted in the survey by Andre et al. (2024). On the other hand, USAT is the news source from which Coibion et al. (2022) sample articles to measure press coverage of central bank communications. Therefore, these sources are likely to represent the inflation narratives of the general U.S. population. The final corpus comprises 157,130 inflation articles, 33,588 from NYT, 8,065 from USAT, 22,503 from WaPo, and 92,974 from WSJ.

I separate my four sources into two groups to measure narrative disagreement between general and specialized newspapers. On the one hand, I classify WSJ as a specialized newspaper. WSJ recognizes itself as *"the best way for marketers to reach the business leaders, active investors, and affluent consumers"*⁷ and its articles have a long history of applications in the Finance literature to measure investor beliefs, namely sentiment.⁸ On the other hand, I classify the other three sources as general newspapers. USAT is the national newspaper whose readership demographics align most closely with the general public (Pew Research Center, 2012). At the same time, Carroll (2003), Pfajfar and Santoro (2013), and Ehrmann et al. (2018) use the New York Times and Washington Post as the sources of the inflation articles households turn to in the model by Carroll (2003). Therefore, these three sources, combined, provide me with a source of newspaper articles directed at the general U.S. household. To motivate my distinction between general and specialized newspapers, I formally test whether their narratives capture the expectations of households and experts differently (more details in sections 2 and 5.3.4).

Table 1 provides some initial insights into what the data looks like. Inflation articles are published almost daily, and incidence does not vary across general and specialized newspapers. Specialized newspapers publish almost 45% more inflation articles per month than general newspapers, but publish shorter and similarly complex inflation articles.⁹ The Jaccard index, which measures the frequency with which

⁶Source: PressGazette (2022). Note: This ranking is based on average Monday-Friday circulation figures for the six months to March 2022.

⁷Source: <https://classifieds.wsj.com/products/#::~:~:text=The%20Wall%20Street%20Journal%20is,the%20world's%20most%20in%EF%AC%82uential%20audience.>

⁸Bybee, Kelly, Manela, and Xiu (2023); Dougal, Engelberg, Garcia, and Parsons (2012); Garcia, Hu, and Rohrer (2023); Manela and Moreira (2017); Tetlock (2007); Tetlock, Saar-Tsechansky, and Macskassy (2008)

⁹Complexity is measured using the Flesch-Kincaid index, which is equal to $0.39 * (\text{number of words}/\text{number of sentences}) + 11.8 * (\text{number of syllables}/\text{number of words}) - 15.59$. Its uses in Economics include Smales and Apergis (2017) and Hayo, Henseler, Rapp, and Zahner (2022).

two sources publish inflation articles on the same day out of all their publication days, indicates that this is the case 83% of the time. Finally, figure 2 shows the evolution of inflation press coverage intensity. Following Carroll (2003), I measure inflation press coverage intensity with the monthly volume of inflation articles scaled by its maximum in any month. I do this separately for general and specialized newspapers, obtaining the measures $News^G$ and $News^S$. Press coverage of inflation reached its highest levels for both general and specialized newspapers in the second half of 2022, with a notable increase in November for general newspapers and in July for specialized newspapers. As a reference, June 2022 is the month with the highest annual CPI inflation rate for my sample. However, the inflation press coverage in specialized newspapers is more volatile and exceeds half its maximum multiple times before 2022. In particular, this happens between the last quarter of 2010 and the first quarter of 2011, in January 2014, and between the last quarter of 2016 and the first quarter of 2017. All three of these periods also coincide with months of accelerating inflation.

As my objective is to extract inflation narratives at the sentence level, I separate each article into sentences using the sentence separator by spaCy¹⁰. In doing so, I obtain 793,333 sentences that contain at least one inflation expression (inflation sentences). I call these sentences "inflation sentences."

3.2 Inflation Expectations

I measure households' inflation expectations using the Survey of Consumer Attitudes and Behavior conducted by the Survey Research Center at the University of Michigan. Participants in the Surveys of Consumers (henceforth, MSC) are asked two questions about expected changes in prices:

1. *"During the next 12 months, do you think that prices in general will go up, or go down, or stay where they are now?"*
2. *"By what percent do you expect prices to go up, on average, during the next 12 months?"*

Following Weber et al. (2022), I discard observations if the respondent expects inflation to be less than -2 percent or more than +15 percent.¹¹ Throughout the paper, I use both households' average and individual expectations. In the latter case, I also make use of several household-level attributes used in the previous literature, namely gender, age, income, education, marital status, and residence region in the United

¹⁰See <https://spacy.io/api/sentencizer> for more details and code. Since the sentence separator fails to recognize lists or tables as separate from their adjacent sentences, I select only sentences with at most 70 words, as suggested by Core NLP.

¹¹Adopting a less restrictive truncation that retains observations only if the respondent expects inflation to be between -5 and +30 percent leads to nearly unchanged results.

States.¹²

Concerning experts, the analysis uses both aggregate and forecaster-level data from the Survey of Professional Forecasters (SPF). Currently conducted by the Federal Reserve Bank of Philadelphia, the SPF collects and summarizes forecasts from leading private forecasting firms. The survey questionnaire is distributed once a quarter and asks participants for quarter-by-quarter forecasts that span the current and next five quarters.¹³ Throughout the analysis, I employ mean and individual-level nowcasts and one-year-ahead forecasts of CPI inflation.

3.3 Inflation News Perceptions

The study also employs a measure of households' perceptions of new information about prices. This measure is intended to complement the news-based variables $News^G$ and $News^S$ in capturing households' attention to news on inflation, as households might learn about inflation from other sources (e.g., grocery prices, D'Acunto, Malmendier, Ospina, and Weber 2021). The use of perceived inflation news can also be motivated by theories of rational inattention, where agents have limited information-processing capacity and, therefore, cannot absorb all available information (Dräger et al., 2016). Such a variable is directly available from the MSC, where respondents are asked whether they have heard of any changes in business conditions during the previous few months. In the case of an affirmative response, the respondents can report two types of news they have heard about, including either higher or lower prices. Therefore, answers to this question allow me to construct two variables, $News_t^P$ and $News_{i,t}^P$. $News_t^P$ is the percentage of MSC respondents who report having heard of recent price changes, while $News_{i,t}^P$ is a dummy indicating if the MSC i -th respondent reports having heard of recent price changes.

3.4 Unemployment Expectations

The study also employs a qualitative measure of households' unemployment expectations, which derives from the answers to the following question in the MS: *"How about people out of work during the coming 12 months — do you think that there will be more unemployment than now, about the same, or less?"* Therefore, I construct the variable UNEMP, which takes values 1, 0, and -1 if the respondent expects the

¹²Household income is grouped into quintiles and age is measured in integers, while education is split into six groups: "Grade 0–8, no high school diploma," "Grade 9–12, no high school diploma," "Grade 0–12, with high school diploma," "4 yrs. of college, no degree," "3 yrs. of college, with degree," and "4 yrs. of college, with degree." Marital status is given as "Married/with a partner," "Divorced," "Widowed," or "Never married," while the region of residence is grouped into "West," "North Central," "Northeast," or "South."

¹³To obtain a monthly estimate of the SPF, I follow Ehrmann et al. (2018) and linearly interpolate the data. Replacing missing monthly values with the last available forecast yields virtually unchanged results.

unemployment rate to increase, stay the same, or decrease. As for experts, I use individual-level nowcasts and one-year-ahead unemployment forecasts from the SPF.

4 Methodology

Section 4.1 discusses the causality extraction method used in this paper. Section 4.2 explains how inflation narratives are extracted from articles. Section 4.3 describes how inflation narratives are categorized into demand and supply categories. Section 4.4 illustrates how to distinguish hawkish narratives from dovish ones, whereas section 4.5 explains how I discriminate observed narratives from expected ones.

4.1 Causality Extraction

Causal relationships in text can be defined as relations between two occurrences or nouns X and Y such that X is described as the "cause" of Y, the "effect." In this study, Y is an inflation expression, and X is an inflation driver. Causality extraction (CE) aims to extract these causal relationships from the text. I use the CE method introduced by Baele et al. (2023) to extract the drivers of inflation from sentences that mention inflation keywords. I refer to these drivers as inflation narratives. This CE method extracts causal relations that are based on predefined causal keywords (e.g., "*because*", "*caused*") and relates cause and effect within the same sentence. Baele et al. (2023) choose these causal keywords in three steps.

First, they select the types of causal relations that their method can capture. As their method identifies explicit causal relations, they follow Khoo, Kornfilt, Oddy, and Myaeng (1998) and focus on four types of explicit causal relationships:

1. Conditionals (i.e., "*if ... then ...*").
2. Resultative constructions (e.g., "*A **tight labor market** keeps **inflation** high.*");¹⁴
3. Causal links (e.g., "*so*", "*because of*", "*that's why*"); and
4. Causal verbs (e.g., "*triggers*").

Second, they assign explicit causal keywords to each type of causal relation. These keywords are used to identify causal relations from text. Baele et al. (2023) define conditionals as those described by "*if-then*" constructions, so they identify them via the use of the keyword "*if*". Then, **resultative constructions**

¹⁴Resultative constructions are sentences in which the object of a verb is followed by a phrase describing the state of the object as a result of the action denoted by the verb. Baele et al. (2023) focus on resultative constructions in which the resultative phrase is an adjective.

are identified from the appearance of the grammatical pattern subject-verb-object-adjective in which the verb is in active form¹⁵. Next, the keywords of **causal links** come from the list of non-adverbial links by Altenberg (1984). Finally, keywords for **causal verbs** are the transitive verbs used in the causal sentences listed in the Penn Discourse Treebank (PDTB) dataset (Prasad et al., 2008) and the causal verbs identified by Girju (2003).

Finally, they define the cause-and-effect order implied by each causal keyword. **Conditionals** consist of two sentence clauses, one describing the effect and the other describing the cause. The latter clause always starts with the causal keyword "*if*" and is also the first subordinate of the former clause. Then, the definition of **resultative constructions** implies that the cause appears in the subject position and the effect in the object position. Next, the authors assign a cause-effect order to each **causal link** keyword based on the direction of causality that Altenberg (1984) assigns to his non-adverbial links. Finally, for **causal verb**, the authors exploit a subcategorization feature of the PDTB dataset that indicates whether cause and effect, respectively, appear before and after a verb or vice versa.

Tables A1 and A2 in Appendix A show the causal keywords identified for causal verbs and links, along with their cause-effect order.

4.2 Inflation Narratives

The CE method used in this study extracts inflation narratives in three steps. First, it selects all sentences mentioning both an inflation expression and a causal keyword. Second, it identifies causal relations whose effects mention inflation keywords. Third, it extracts the text of the cause from each identified causal relation and adds it to the list of inflation narratives. I refer the reader to Baele et al. (2023) for a detailed discussion of how to extract causal relations.

4.3 Demand and Supply Narratives

The output of the previous section is a long, unstructured list of inflation narratives. I classify them into the demand and supply narratives used by Andre et al. (2024). Their demand narratives include consumer spending/sentiment, government spending, and monetary policy, while their supply narratives include supply chain, labor, and energy. The authors also have residual narrative categories for narratives that cannot be classified into either demand or supply narratives. I do not consider these residual narrative categories in my classifications.

¹⁵The authors do not impose any causality requirement on the verbs used in this type of construction.

I use a dictionary method for my classification, whereby a narrative is assigned to a category whenever the narrative's text contains any keyword from the category's dictionary. The extensive literature on text analysis applications in Economics and Finance provides dictionaries for all my categories. On the demand side, I utilize the dictionaries "Spending/Deficit/Debt," "Monetary Policy," and "Consumer Spending and Sentiment" from Baker et al. (2021). On the supply side, I borrow from Baker et al. (2021) the dictionaries "Labor Markets" and "Labor Disputes" for the Labor narrative category, and the dictionary "Commodity Markets" for the Energy category. In addition, the supply-chain dictionary includes all the top one-hundred supply-chain risk bigrams compiled by Ersahin, Giannetti, and Huang (2024). Finally, I add a (limited) number of keywords to each dictionary; these words come from my raw list of narratives and are narrowly related to their respective narrative category. For instance, I add the words "energy", "electricity", "fuel", and "gasoline" to the commodity-markets dictionary. Table 2 provides full transparency on all manually added keywords.

Insert table 2 here.

For each article, I take the difference in the number of mentions of demand and supply keywords in its narratives. I classify an article as a demand/supply article if the difference between the two is positive/negative. Therefore, I leave an article unclassified if it has no narrative, or the difference in the number of mentions of demand and supply keywords in its narratives is zero. For an article j published in month t by newspaper n , this step yields two dummies, $Demand_{j,t}^n$ and $Supply_{j,t}^n$, taking the value one if the article is classified as a demand or supply article, respectively.

I then compute each newspaper's monthly demand and supply narrative indicators $Demand_t^n$ and $Supply_t^n$ from the volume of its demand and supply articles. To measure each newspaper's relative attention to demand and supply articles, I use $Demand_t^n$ and $Supply_t^n$ to compute the newspaper-specific indicator $NetDemand_t^n$ by dividing their difference by its maximum absolute value across all months. This scaling choice ensures that $NetDemand_t^n$ ranges between -1 and 1, whereby positive/negative values represent months when newspapers publish predominantly demand/supply articles.

An alternative scaling choice would be to divide the difference between $Demand_t^n$ and $Supply_t^n$ by their sum or the maximum sum across all months. However, this choice might give a small weight to months with a large number of demand or supply articles, even though there is a predominance of demand or supply articles. Thus, scaling by the maximum absolute difference in the volume of demand and supply articles across all months ensures that $NetDemand_t^n$ correctly measures a newspaper's relative attention

to demand and supply articles.¹⁶

Finally, I compute demand-supply narrative disagreement between general and specialized newspapers, $NetDemand_t^{G-S}$, as the difference between $NetDemand_t^G$ and $NetDemand_t^S$. This measure ranges between -2 and 2, whereby positive/negative values represent months when general newspapers publish relatively more demand/supply articles than specialized newspapers.

I also construct a monthly measure of narrative disagreement for each narrative type in a similar fashion:

$$\begin{aligned}
ConsSpendSent_t^{G-S} &= \frac{\sum_{j=1}^{N_t^G} ConsSpendSent_{j,t}^G}{\max_t \sum_{j=1}^{N_t^G} ConsSpendSent_{j,t}^G} - \frac{\sum_{j=1}^{N_t^S} ConsSpendSent_{j,t}^S}{\max_t \sum_{j=1}^{N_t^S} ConsSpendSent_{j,t}^S} \\
MonPol_t^{G-S} &= \frac{\sum_{j=1}^{N_t^G} MonPol_{j,t}^G}{\max_t \sum_{j=1}^{N_t^G} MonPol_{j,t}^G} - \frac{\sum_{j=1}^{N_t^S} MonPol_{j,t}^S}{\max_t \sum_{j=1}^{N_t^S} MonPol_{j,t}^S} \\
SpendDefDebt_t^{G-S} &= \frac{\sum_{j=1}^{N_t^G} SpendDefDebt_{j,t}^G}{\max_t \sum_{j=1}^{N_t^G} SpendDefDebt_{j,t}^G} - \frac{\sum_{j=1}^{N_t^S} SpendDefDebt_{j,t}^S}{\max_t \sum_{j=1}^{N_t^S} SpendDefDebt_{j,t}^S} \\
ComEne_t^{G-S} &= \frac{\sum_{j=1}^{N_t^G} ComEne_{j,t}^G}{\max_t \sum_{j=1}^{N_t^G} ComEne_{j,t}^G} - \frac{\sum_{j=1}^{N_t^S} ComEne_{j,t}^S}{\max_t \sum_{j=1}^{N_t^S} ComEne_{j,t}^S} \\
Labor_t^{G-S} &= \frac{\sum_{j=1}^{N_t^G} Labor_{j,t}^G}{\max_t \sum_{j=1}^{N_t^G} Labor_{j,t}^G} - \frac{\sum_{j=1}^{N_t^S} Labor_{j,t}^S}{\max_t \sum_{j=1}^{N_t^S} Labor_{j,t}^S} \\
SupplyChain_t^{G-S} &= \frac{\sum_{j=1}^{N_t^G} SupplyChain_{j,t}^G}{\max_t \sum_{j=1}^{N_t^G} SupplyChain_{j,t}^G} - \frac{\sum_{j=1}^{N_t^S} SupplyChain_{j,t}^S}{\max_t \sum_{j=1}^{N_t^S} SupplyChain_{j,t}^S}
\end{aligned}$$

Where $ConsSpendSent_{j,t}^G$, $MonPol_{j,t}^G$, $SpendDefDebt_{j,t}^G$, $ComEne_{j,t}^G$, $Labor_{j,t}^G$, and $SupplyChain_{j,t}^G$ are dummy variables taking value one when article j published in month t by general newspapers has inflation narratives that mention terms from the dictionaries of the narrative categories consumer spending/sentiment, monetary policy, spending/deficit/debt, commodities/energy, labor, and supply chain, respectively. The specialized newspapers' variables $ConsSpendSent_{j,t}^S$, $MonPol_{j,t}^S$, $SpendDefDebt_{j,t}^S$, $ComEne_{j,t}^S$, $Labor_{j,t}^S$, and $SupplyChain_{j,t}^S$ are measured similarly.

¹⁶Replacing the current scaling choice with the two alternatives leaves the main results virtually unchanged.

4.4 Hawkishness of Demand and Supply Narratives

The output of the previous section is a classification of inflation narratives into demand and supply narratives. A missing element of this classification is whether an inflation narrative is about increasing or decreasing inflation. I adopt the dictionary by Apel, Blix Grimaldi, and Hull (2022) to discern whether narratives mention inflation as increasing or decreasing. This dictionary measures the extent to which a central bank text or speech is predominantly hawkish or dovish. In particular, it consists of two lists of adjectives and verbs called modifiers: one list for hawkish modifiers (e.g., "*accelerating*") and another for dovish modifiers (e.g., "*decelerating*"). The authors first count the hawkish and dovish modifiers mentioned within seven words from the word "inflation". Then, they compute a net hawkishness score by subtracting the count of dovish modifiers from the count of hawkish modifiers, scaled by their sum. This net hawkishness score is positive/negative when inflation is described as accelerating/decelerating, commanding a hawkish/dovish policy response.

As a refinement of the approach adopted by Apel et al. (2022), I use a dependency parser¹⁷ to precisely identify when a modifier is used in conjunction with an inflation expression. A dependency parser analyzes the grammatical structure of a sentence,¹⁸ so it can verify whether adjective modifiers refer to inflation. For instance, in the sentence "*Looser monetary policy might lead to high inflation and low unemployment*", a narrative is identified and marked as hawkish because the hawkish modifier "high" refers to inflation. However, the dictionary method by Apel et al. (2022) would incorrectly mark it as neither hawkish nor dovish because the hawkish and dovish modifiers "high" and "low" cancel each other out. Similarly, a dependency parser can ascertain whether a verbal modifier has an inflation keyword as its subject or object and, hence, is directly related to it rather than simply appearing in the same sentence. For instance, in the sentence "*Inflation is decelerating because of higher interest rates.*", a narrative is identified and marked as dovish because the dovish modifier "decelerating" has an inflation keyword as its subject. However, the dictionary method by Apel et al. (2022) would incorrectly mark it as neither hawkish nor dovish because the hawkish and dovish modifiers "higher" and "decelerating" cancel each other out.

For each article, I take the difference in the number of hawkish and dovish narratives. I classify an article as hawkish/dovish based on whether this difference is positive/negative. Therefore, I leave an article unclassified if it has no narrative, or the difference in the number of hawkish and dovish narratives is zero. For an article j published in month t by newspaper n , this step yields two dummies, $Hawkish_{j,t}^n$

¹⁷I use the Python implementation of the spaCy dependency parser.

¹⁸Appendix B provides a detailed description of dependency parsing.

and $Dovish_{j,t}^n$, taking the value one if the article is hawkish and dovish, respectively.

I then compute each newspaper's monthly hawkish and dovish narrative indicators $Hawkish_t^n$ and $Dovish_t^n$ from the volume of its hawkish and dovish articles. To measure each newspaper's relative attention to hawkish and dovish articles, I use $Hawkish_t^n$ and $Dovish_t^n$ to compute the newspaper-specific indicator $NetHawkish_t^n$ by dividing their difference by its maximum absolute value across all months. This scaling choice ensures that $NetHawkish_t^n$ ranges between -1 and 1, whereby positive/negative values represent months when newspapers publish predominantly hawkish/dovish articles.¹⁹

An alternative scaling choice would be to divide the difference between $Hawkish_t^n$ and $Dovish_t^n$ by their sum or their maximum sum across all months. However, this choice might give a small weight to months with a large number of hawkish or dovish articles, even though there is a predominance of hawkish or dovish articles. Thus, scaling by the maximum absolute difference in the volume of hawkish and dovish articles across all months ensures that $NetHawkish_t^n$ correctly measures a newspaper's relative attention to hawkish and dovish articles.

Finally, I compute hawkish-dovish narrative disagreement between general and specialized newspapers $NetHawkish_t^{G-S}$ as the difference between $NetHawkish_t^G$ and $NetHawkish_t^S$. This measure ranges between -2 and 2, whereby positive/negative values represent months when general newspapers publish relatively more hawkish/dovish articles than specialized newspapers.

4.5 Observed vs. Expected Inflation Narratives

A key question is whether inflation narratives focus on past/present inflation episodes rather than future/potential ones. To discriminate between these two cases, I follow Baele et al. (2023) and formulate three non-exclusive conditions under which an inflation narrative is classified as expected; if no condition is met, then a narrative is classified as observed. First, inflation narratives extracted using conditionals are identified as expected. Conditionals state the conditions under which inflation episodes occur and naturally refer to inflation episodes that have not happened yet. Second, inflation narratives extracted from causal relations mentioning modal verbs²⁰ are identified as expected. One of the functions of modal verbs is to express possibility, so they are natural candidates to verify whether an inflation narrative is expected. Finally, inflation narratives extracted from causal relations that mention the verb "to expect" or any of its synonyms²¹ are identified as expected.

¹⁹Replacing the current scaling choice with the two alternatives leaves the main results virtually unchanged.

²⁰The list of the modal verbs I use comes from here.

²¹Synonyms of the verb "to expect" are from the thesaurus by Merriam-Webster.

For each article, I take the difference in the number of observed and expected narratives. I classify an article as observed/expected if this difference is positive/negative. Therefore, I leave an article unclassified if it has no narrative or the difference in the number of observed and expected narratives is zero. For an article j published in month t by newspaper n , this step yields two dummies, $Observed_{j,t}^n$ and $Expected_{j,t}^n$, taking the value one if the article is classified as observed and expected, respectively.

I then compute each newspaper's monthly observed and expected narrative indicators $Observed_t^n$ and $Expected_t^n$ from the volume of its observed and expected articles. To measure each newspaper's relative attention to observed and expected articles, I use $Observed_t^n$ and $Expected_t^n$ to compute the newspaper-specific indicator $NetObserved_t^n$ by dividing their difference by its maximum absolute value across all months. This scaling choice ensures that $NetObserved_t^n$ ranges between -1 and 1, where positive/negative values represent months when newspapers publish predominantly observed/expected articles, respectively.

An alternative scaling choice would be to divide the difference between $Observed_t^n$ and $Expected_t^n$ by their sum or their maximum sum across all months. However, this choice might give a small weight to months with a large number of observed or expected articles, even though there is a predominance of observed or expected articles. Thus, scaling by the maximum absolute difference in the volume of observed and expected articles across all months ensures that $NetObserved_t^n$ correctly measures a newspaper's relative attention to observed and expected articles.²²

Finally, I compute observed-expected narrative disagreement between general and specialized newspapers, $NetObserved_t^{G-S}$, as the difference between $NetObserved_t^G$ and $NetObserved_t^S$. This measure ranges between -2 and 2, whereby positive/negative values represent months when general newspapers publish relatively more observed/expected articles than specialized newspapers.

5 Results

This section discusses the main empirical findings. Section 5.1 describes the output of the CE algorithm, section 5.2 discusses the results from the classification of my inflation narratives, section 5.3 presents the model used to test the hypotheses from section 2 and the test results, and section 5.4 concludes with a battery of tests to verify the robustness of my results.

²²Replacing the current scaling choice with the two alternatives leaves the main results virtually unchanged.

5.1 Inflation Narratives

Applying the CE method outlined in section 4.2 to the 793,333 sentences containing an inflation expression yields 39,510 inflation narratives. Inflation narratives appear in 29,135 (causal) inflation articles, published on approximately 77% of publication days of inflation articles (8,880 out of 11,555). Of all 39,510 inflation narratives, 5,565 are extracted via conditionals, 3,638 and 30,269 via causal verbs and links, respectively, and 38 via resultative constructions.

Baele et al. (2023) validate their CE algorithm by manually inspecting a subset of causal relations to verify that they mention their flight-to-safety expressions as the effect. As their context differs from mine, my CE algorithm might perform differently from theirs when extracting inflation narratives. Therefore, I repeat their validation exercise within this study. For each causal keyword, I select two sets of twenty inflation sentences that mention that causal keyword: the first one with inflation narratives, the second without (e.g., an inflation expression can be the cause or can be used neither as the cause nor the effect). For instance, I inspect twenty inflation sentences with inflation narratives found via the causal keyword *"because"* and twenty inflation sentences where *"because"* is mentioned, but no inflation narrative is found. Therefore, I exclude all causal keywords used in fewer than twenty (causal) inflation sentences. My manual annotations and evaluations are described in section C of the Appendix.

Table A3 shows the results of the manual annotations and reveals that most causal keywords achieve an F-score above 70%. For comparison, Yang, Han, and Poon (2022) show that studies using comparable CE methods generally achieve F-scores ranging between 54 and 71%. As some causal keywords achieve F-scores below 54%, I drop all inflation narratives that are extracted based on them. Therefore, I retain all narratives extracted via resultative constructions, four out of nine causal link keywords, and seventeen out of twenty-two causal verb keywords. As a consequence, the number of inflation narratives drops to 5,204. The retained narratives appear in 4,896 causal inflation articles, 1,582 of which are from general newspapers, and 3,314 from specialized newspapers. The retained causal inflation articles are published on about 27% of publication days of inflation articles (3,086 of 11,555).

Table 3 and figure 4 reproduce table 1 and figure 2 for the retained causal inflation articles. The frequency and volume of causal articles published by specialized newspapers are almost twice as high as for general newspapers. Additionally, their publication timing is significantly less aligned, with a Jaccard index of 13%. However, there are no significant differences in terms of length and complexity. Looking at their causal press coverage over time, we can see that specialized newspapers' causal inflation press coverage is more volatile than that of general newspapers and reaches beyond half its maximum multiple

times. In addition, causal inflation press coverage spikes more frequently than press coverage, namely in June 2009, the first quarter of 2011, and July 2022. Interestingly, these spikes coincide with spikes in the expectation gap.

5.2 Narrative Features and Disagreement

I now describe the results of classifying my inflation narratives along all three narrative dimensions described in sections 4.3 to 4.5. Table 4 provides some summary statistics on *NetDemand*, *NetHawkish*, and *NetObserved*, as well as measures of disagreement at the level of different demand and supply narrative categories. In addition, figures 5 to 8 show how these measures evolve.

The first block of table 4 shows that the averages of $NetDemand^G$ and $NetDemand^S$ are both negative and significantly different from zero. Therefore, general and specialized newspapers publish more supply than demand narratives. Moreover, the median of $NetDemand^S$ is negative while the first quartile of $NetDemand^G$ equals zero. This means that specialized newspapers publish relatively more supply narratives more often than general newspapers, which suggests there might be systematic narrative disagreement between them. However, a two-sided t-test reveals that the $NetDemand^{G-S}$ is statistically different from zero only at a significance level of 10%. Therefore, there is only weak evidence of systematic demand-supply narrative disagreement between general and specialized newspapers.

Nonetheless, the top panel of figure 6 documents multiple periods in which general and specialized newspapers disagree in their demand and narratives. In particular, demand-supply narrative disagreement peaks in April-May 2006, March 2018, and February 2022. Delving deeper into the types of narratives behind these spikes, Figure 8 shows specialized newspapers published relatively more commodities/energy narratives in April-May 2006 and February 2022. This seems to be an exception, as the bottom panel of table 4 indicates that the mean of $ComEne^{G-S}$ is not statistically different from zero. On the other hand, figure 7 shows specialized newspapers published relatively more government spending/deficit/debt narratives in March 2018. This does not seem to be an exception, as the bottom panel of table 4 indicates that the mean of $SpendDefDebt^{G-S}$ is positive and statistically different from zero.

Moving to other narrative dimensions, the second block of table 4 indicates that the averages of $NetHawkish^G$ and $NetHawkish^S$ are both positive and significantly different from zero. This means that the narratives of both general and specialized newspapers are predominantly hawkish in nature. This is also showcased in the central panel of figure 5 and aligns with existing evidence that the media pays more attention to rising inflation (Lamla & Lein, 2014). Moreover, the median of $NetHawkish^S$ is positive while the median of $NetHawkish^G$ equals zero. This means that specialized newspapers pub-

lish relatively more hawkish narratives more often than general newspapers, suggesting another form of systematic narrative disagreement between them. However, a two-sided t-test reveals the null of no hawkish-dovish narrative disagreement cannot be rejected.

Concluding with observed and expected narratives, the third block of table 4 indicates that the averages of $NetObserved^G$ and $NetObserved^S$ are both positive and significantly different from zero. Therefore, the narratives of both general and specialized newspapers predominantly focus on realized inflation episodes rather than expected ones. This is also showcased in the third panel of figure 5. Moreover, the median of $NetObserved^S$ is positive while the third quartile of $NetObserved^G$ equals zero. This means that specialized newspapers publish relatively more observed narratives more often than general newspapers, suggesting another form of systematic narrative disagreement. However, a two-sided t-test reveals that the null of observed-expected narrative disagreement cannot be rejected.

Overall, this section shows that general and specialized newspapers predominantly publish narratives attributing realized inflationary episodes to supply factors. There is also no systematic disagreement along my three main narrative dimensions.

5.3 Narratives and Expectations

5.3.1 Inflation Press Coverage and Expectations

The first testable hypothesis from section 2 states that the absolute expectation gap narrows with inflation press coverage. To test this hypothesis, I estimate the following models based on Carroll (2003) and Pfajfar and Santoro (2013):

$$GAP_t = \alpha_1 + \alpha_2 * News_{t-1}^G + \alpha_3 * News_{t-1}^S + \alpha_4 * News_t^P + \alpha_5 * \pi_{t-1} + \alpha_6 * \sigma_{\pi,t-1}^2 + \epsilon_t \quad (1)$$

$$GAP_{i,t} = \alpha_1 + \alpha_2 * News_{t-1}^G + \alpha_3 * News_{t-1}^S + \alpha_4 * News_{i,t}^P + \alpha_5 * \pi_{t-1} + \alpha_6 * \sigma_{\pi,t-1}^2 + x_{i,t}\alpha_7 + \epsilon_{i,t} \quad (2)$$

Where $GAP_t = |\pi_{t,t+12}^{MSC} - \pi_{t,t+12}^{SPF}|$ and $GAP_{i,t} = |\pi_{i,t,t+12}^{MSC} - \pi_{i,t,t+12}^{SPF}|$ respectively represent the aggregate and individual measures of the absolute expectation gap. $\pi_{t,t+12}^{MSC}$ and $\pi_{i,t,t+12}^{MSC}$ respectively represent the MSC mean and individual inflation forecasts, while $\pi_{t,t+12}^{SPF}$ is the SPF mean inflation forecast. Following Carroll (2003) and Pfajfar and Santoro (2013), I measure general and specialized newspapers' inflation press coverage $News^G$ and $News^S$ by dividing the monthly volumes of inflation articles published by

general and specialized newspapers by their maxima in any month. They appear with a lag to ensure I use articles that households and experts had access to before their interview. To test the first hypothesis, I test whether α_2 and α_3 are positive in equations 1 and 2. For this purpose, I also estimate modified versions of equations 1 and 2 where the news-based measures appear one by one.

Equations 1 and 2 include several control variables. $News_t^P$ and $News_{i,t}^P$ account for households' perceptions of news about inflation. π_{t-1} is the last observed value of CPI inflation, while $\sigma_{\pi,t-1}^2$ is a measure of inflation volatility built as the sum of squared inflation changes over the previous six months (Dräger et al., 2016). Survey respondents do not observe data on inflation for their interview month because of the publication lag in the numbers for inflation, which is why I use lags of inflation and its volatility. x_i is a vector of socioeconomic characteristics for MSC households (namely, gender, age, income, education, marital status, and location in the United States).²³ For equation 1, standard errors are computed with the Huber–White sandwich estimator. For equation 2, standard errors are clustered at the individual level. The results are shown in table 5.

The results indicate that only the individual absolute expectation gap changes with inflation press coverage. However, it rises with the inflation press coverage of either general or specialized newspapers, contrary to the theoretical prediction by Carroll (2003). When both news variables are included, the results indicate a positive correlation with general newspapers and a negative correlation with specialized newspapers. However, the difference in the magnitude of the slope coefficients and the adjusted R-squared changes from one press coverage measure to another indicate that the largest explanatory power resides with general newspapers' press coverage. Overall, there is only limited evidence to support the hypothesis that the absolute expectation gap becomes narrower with inflation press coverage.

The distinction between causal and non-causal inflation articles from Section 5.1 suggests that their relationship with the absolute expectation gap may differ. If households care about inflation narratives, they might only pay attention to causal inflation articles because they mention inflation narratives. To verify this, I replace $News^G$ and $News^S$ with $CausalNews^G$ and $CausalNews^S$ in equations 1 and

²³Household income is grouped into quintiles and age is measured in integers, while education is split into six groups: "Grade 0–8, no high school diploma," "Grade 9–12, no high school diploma," "Grade 0–12, with high school diploma," "4 yrs. of college, no degree," "3 yrs. of college, with degree," and "4 yrs. of college, with degree." Marital status is given as "Married/with a partner," "Divorced," "Widowed," or "Never married." Finally, the region of residence is grouped into "North Central" ("Midwest" in the Survey Information page online), "Northeast," "South," or "West." Region "Midwest" consists of Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin. Region "Northeast" consists of Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. Region "South" consists of Alabama, Arkansas, Delaware, the District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia. Region "West" consists of Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

2. $CausalNews^G$ and $CausalNews^S$ are constructed similarly to $News^G$ and $News^S$ by using causal inflation articles rather all inflation articles. These measures are also scaled by their maxima in any month. The results in table 6 indicate that the absolute expectation gap widens with causal inflation press coverage. In particular, the aggregate absolute expectation gap widens only with specialized newspapers' causal press coverage. In contrast, the individual absolute expectation gap widens with both newspapers' causal press coverage, particularly general newspapers.

Overall, the evidence rejects the first testable hypothesis and aligns with Pfajfar and Santoro (2013) in concluding that the absolute expectation gap does not narrow with inflation press coverage.

5.3.2 Narratives Disagreement and Differences in Inflation Expectations

The second testable hypothesis from section 2 states that the absolute expectation gap widens with demand-supply narrative disagreement between general and specialized newspapers. To test this hypothesis, I estimate the following modified versions of equations 1 and 2:

$$GAP_t = \alpha_1 + \alpha_2 * |NetDemand_{t-1}^{G-S}| + \alpha_3 * News_t^P + \alpha_4 * \pi_{t-1} + \alpha_5 * \sigma_{\pi,t-1}^2 + \epsilon_t \quad (3)$$

$$GAP_{i,t} = \alpha_1 + \alpha_2 * |NetDemand_{t-1}^{G-S}| + \alpha_3 * News_{i,t}^P + \alpha_4 * \pi_{t-1} + \alpha_5 * \sigma_{\pi,t-1}^2 + x_{i,t}\alpha_6 + \epsilon_{i,t} \quad (4)$$

$|NetDemand_{t-1}^{G-S}|$ measures disagreement about demand and supply narratives between newspapers, and all other variables are defined as before. To test the second hypothesis, I test whether α_2 is positive in equations 3 and 4. As there are other narrative dimensions along which newspapers might disagree, I also control for $|NetHawkish_{t-1}^{G-S}|$ and $|NetObserved_{t-1}^{G-S}|$. $|NetHawkish_{t-1}^{G-S}|$ measures disagreement about hawkish and dovish narratives between newspapers. $|NetObserved_{t-1}^{G-S}|$ measures disagreement about observed and expected narratives between newspapers. I add these disagreement measures first, one by one, then together. For equation 3, standard errors are computed with the Huber–White sandwich estimator. For equation 4, standard errors are clustered at the individual level. The results are shown in table 7.

Both the aggregate and individual absolute expectation gaps widen with demand-supply narrative disagreement. The aggregate absolute expectation gap widens statistically significantly only with demand-supply narrative disagreement, though only at the 10% level. In contrast, the individual absolute expectation gap widens statistically significantly with all three measures and at levels of at least 10%. Therefore, multiple dimensions of narrative disagreement matter for the absolute expectation gap between house-

holds and experts.

Overall, the results confirm the second testable hypothesis that the absolute expectation gap widens with demand-supply narrative disagreement between general and specialized newspapers.

5.3.3 Narratives and Expectations across Individual Characteristics

The third testable hypothesis from section 2 states that household demographics moderate the relationship between the absolute expectation gap and demand-supply narrative disagreement between general and specialized newspapers. To test this hypothesis, I estimate a modified version of equation 4 by sequentially interacting the variable $|NetDemand_t^{G-S}|$ with some of the consumer characteristics represented in x_i , namely:

- $FEMALE_{i,t}$, which is a dummy taking the value one when the respondent is a woman;
- $AGE_{i,t}$, which measures the age of the respondent in integers (minus 40);
- $INC1_{i,t}$, $INC2_{i,t}$, $INC4_{i,t}$, and $INC5_{i,t}$, which are dummies respectively taking value one when the income of the respondent belongs to the first, second, fourth, and fifth quintiles of the cross-sectional MSC income distribution;
- $EDUC1_{i,t}$, $EDUC2_{i,t}$, $EDUC4_{i,t}$, $EDUC5_{i,t}$, and $EDUC6_{i,t}$, which are dummies respectively taking value one when the respondent's education respectively belongs to the group "Grade 0–8, no high school diploma," "Grade 9–12, no high school diploma," "4 yrs. of college, no degree," "3 yrs. of college, with degree," and "4 yrs. of college, with degree".

The interaction term between these characteristics and $|NetDemand_{t-1}^{G-S}|$ reflects how the relationship between narrative disagreement and the individual absolute expectation gap changes across individual MSC respondents vis-à-vis the benchmark one.²⁴ Table 8 shows the results.

The relationship between the absolute expectation gap and narrative disagreement varies with households' age and education; specifically, it strengthens with age and weakens with college education. The first result aligns with what Ehrmann et al. (2018) show for inflation press coverage intensity. Additionally, it can be reconciled with evidence from readership studies that older people are more likely to read newspapers (Pew Research Center, 2023). Moreover, the second result is novel and can be explained by

²⁴Married, male, forty years old, with a high school diploma (EDUC3), having an income in the middle quintile of the distribution (INC3), and living in the North Center of the country.

the higher likelihood of college-educated households reading specialized newspapers (Pew Research Center, 2012). Minor results are the absence of a clear pattern in the coefficients of the interaction terms with the income and sex dummies. In particular, the relationship between the absolute expectation gap and narrative disagreement weakens for individuals moving from the middle-income quintile to the second- and fifth-income quintiles. On the other hand, the relationship does not change with sex. This last result aligns with what Ehrmann et al. (2018) show for inflation press coverage intensity and with readership evidence that men and women do not differ in their news readership (Pew Research Center, 2023).

Overall, the evidence broadly confirms the third testable hypothesis, indicating that the relationship between the absolute expectation gap and narrative disagreement strengthens with age, weakens with education, and remains unchanged with sex.

5.3.4 Newspaper Demand and Supply Narratives and Individual Demand and Supply Expectations

The fourth hypothesis from Section 2 states that households and experts expect inflation and unemployment to move in opposite directions to a greater degree when newspapers publish relatively more demand narratives. The fifth hypothesis posits that the expectations of households are more strongly correlated with the narratives of general newspapers than with those of specialized newspapers, whereas the opposite holds for experts' expectations. To test these hypotheses, I estimate the following two models:

$$\Delta\pi_{i,t,t+12}^{MSC} * \Delta u_{i,t,t+12}^{MSC} = \alpha_1 + \alpha_2 * NetDemand_{t-1}^G + \alpha_3 * NetDemand_{t-1}^S + \alpha_4 * \pi_{t-1} + \alpha_5 * \sigma_{\pi,t-1}^2 + \alpha_6 * News_{i,t}^P + x_{i,t}\alpha_7 + \epsilon_{i,t}^{MSC} \quad (5)$$

$$\Delta\pi_{i,t,t+4}^{SPF} * \Delta u_{i,t,t+4}^{SPF} = \alpha_1 + \alpha_2 * NetDemand_{t-1}^G + \alpha_3 * NetDemand_{t-1}^S + \alpha_4 * \pi_{t-1} + \alpha_5 * \sigma_{\pi,t-1}^2 + \epsilon_{i,t}^{SPF} \quad (6)$$

$\Delta\pi_{i,t,t+12}^{MSC}$ and $\Delta\pi_{i,t,t+4}^{SPF}$ are individual households' and expert' expected changes in one-year-ahead inflation. $\Delta u_{i,t,t+12}^{MSC}$ and $\Delta u_{i,t,t+4}^{SPF}$ are individual households' and experts' expected changes in unemployment in the following year. I follow Dräger et al. (2016) in measuring expected changes in inflation and unemployment. $\Delta\pi_{i,t,t+12}^{MSC}$ is the difference between one-year-ahead inflation expectation ($\pi_{i,t,t+12}^{MSC}$) and the average inflation over the previous twelve months ($\tilde{\pi}_t$). $\Delta u_{i,t,t+12}^{MSC}$ is an indicator taking value 1, 0, -1 when the individual household expects the unemployment rate to increase, stay the same, and decrease in the following year, respectively. $\Delta\pi_{i,t,t+4}^{SPF}$ is the differences between the SPF respondent's one-year-ahead

expectation of inflation and its nowcast. $\Delta u_{i,t,t+4}^{SPF}$ is constructed similarly to $\Delta u_{i,t,t+12}^{MSC}$ using the SPF respondent's one-year-ahead expectation of unemployment and its nowcast. All other variables are defined as before. As the MSC is conducted monthly and the SPF is run quarterly, equations 5 and 6 are estimated at monthly and quarterly frequencies, respectively. Therefore, $NetDemand^G$ and $NetDemand^S$ are computed using quarterly numbers in equation 6. Standard errors are clustered at the individual level as respondents in the MSC and SPF can be reinterviewed.

$NetDemand^G$ ($NetDemand^S$) is positive/negative when general (specialized) newspapers publish relatively more demand/supply articles. Demand/supply articles predominantly contain demand/supply narratives, which describe a negative/positive relationship between changes in inflation and unemployment. Therefore, to test the fourth hypothesis, I test whether α_2 and α_3 are negative in equations 5 and 6. In addition, to test the fifth hypothesis, I test whether the magnitude of α_2 is lower than that of α_3 in equation 6, and vice versa in 5. For this purpose, I also estimate modified versions of equations 5 and 6 where the narrative measures appear one by one. The results are shown in table 9.

The first three columns indicate that households expect inflation and unemployment to move in the same direction to a greater extent when both newspapers publish relatively more supply articles. Therefore, the evidence suggests that the narratives of both general and specialized newspapers align correctly with the expectations of households. A one-sided t-test shows that the null hypothesis of equal slope coefficients cannot be rejected (t-test = -0.195). Therefore, the narratives of both newspapers align equally well with the expectations of households. On the other hand, the last three columns show that experts expect inflation and unemployment to move in the same direction to a larger degree when specialized newspapers publish relatively more supply articles, but to a smaller degree when general newspapers do so. Therefore, the evidence suggests that only the narratives of specialized newspapers correctly align with experts' expectations.

All in all, the evidence partially confirms the fourth and fifth testable hypotheses by indicating that households' expectations are similarly aligned with the narratives of both general and specialized newspapers, whereas experts' expectations are correctly aligned with the narratives of specialized newspapers only.

5.4 Robustness checks

5.4.1 Narratives and Expectations across Narrative Types

My analysis so far abstracts from the underlying types of demand and supply narratives. However, these types might also matter. For instance, when general newspapers publish narratives about energy prices (a supply narrative), the absolute expectation gap might widen differently based on whether specialized newspapers publish narratives about loose government spending or monetary policy (both demand narratives). Alternatively, if both general and specialized newspapers publish supply narratives, the absolute expectation gap might widen differently based on whether the same or different supply narratives are published (e.g., higher energy prices vs. labor market tightness). Consequently, the absolute expectation gap might widen differently based on how general and specialized newspapers allocate their attention across different demand and supply narratives. To verify this, I estimate the following modified versions of equations 3 and 4:

$$\begin{aligned} GAP_t = & \alpha_1 + \alpha_2 * |MonPol_{t-1}^{G-S}| + \alpha_3 * |SpendDefDebt_{t-1}^{G-S}| \\ & + \alpha_4 * |ComEne_{t-1}^{G-S}| + \alpha_5 * |Labor_{t-1}^{G-S}| \\ & + \alpha_6 * News_t^P + \alpha_7 * \pi_{t-1} + \alpha_8 * \sigma_{\pi,t-1}^2 + \epsilon_t \end{aligned} \quad (7)$$

$$\begin{aligned} GAP_{i,t} = & \alpha_1 + \alpha_2 * |MonPol_{t-1}^{G-S}| + \alpha_3 * |SpendDefDebt_{t-1}^{G-S}| \\ & + \alpha_4 * |ComEne_{t-1}^{G-S}| + \alpha_5 * |Labor_{t-1}^{G-S}| \\ & + \alpha_6 * News_{i,t}^P + \alpha_7 * \pi_{t-1} + \alpha_8 * \sigma_{\pi,t-1}^2 + x_{i,t}\alpha_9 + \epsilon_{i,t} \end{aligned} \quad (8)$$

$MonPol^S$, $SpendDefDebt^S$, $ComEne^S$, and $Labor^S$ represent narrative disagreement about monetary policy, government spending/deficit/debt, commodities/energy, and labor, respectively.²⁵ All other variables are defined as before. I also estimate modified versions of equations 7 and 8 in which the measures of disagreement about narrative types appear individually. For equation 7, standard errors are computed with the Huber–White sandwich estimator. For equation 8, standard errors are clustered at the individual level. Table A4 in Appendix D shows the results.

The aggregate absolute expectation gap widens only with disagreement around monetary policy narratives. In contrast, the individual absolute expectation gap widens with narrative disagreement about most narrative types. In particular, it widens the most with narrative disagreement around monetary pol-

²⁵I exclude narrative disagreement about Consumer Spending/Sentiment and Supply Chain because there are fewer than 20 months in which both general and specialized newspapers publish narratives from these categories.

icy (t-test = 4.352). This result is particularly interesting, given the recent survey evidence by Stantcheva (2024), which suggests that higher-income and college-educated individuals are more likely to entertain monetary policy narratives. In addition, the individual absolute expectation gap similarly widens with narrative disagreement around government spending/deficit/debt and commodities/energy (t-test = 0.363). Surprisingly, both aggregate and individual absolute expectation gaps narrow with disagreement about labor narratives. More work is needed to understand why this is the case.

All in all, the evidence suggests that the absolute expectation gap widens differently depending on how general and specialized newspapers allocate their attention across various demand and supply narratives.

5.4.2 Asymmetric Narratives Disagreement and Differences in Inflation Expectations

My hypotheses so far have been tested using the absolute value of demand-supply narrative disagreement, as these hypotheses revolve around the magnitude of narrative disagreement, not its sign. An important question is whether the sign also matters. That is, whether the results change based on whether general or specialized newspapers publish relatively more demand narratives than the other. To answer this question, I estimate the following modified versions of equations 3 and 4:

$$\begin{aligned} GAP_t = & \alpha_1 + \alpha_2 * |NetDemand_{t-1}^{G-S}| * 1_{NetDemand_{t-1}^{G-S} > 0} \\ & + \alpha_3 * |NetDemand_{t-1}^{G-S}| * 1_{NetDemand_{t-1}^{G-S} < 0} \\ & + \alpha_4 * News_t^P + \alpha_5 * \pi_{t-1} + \alpha_6 * \sigma_{\pi,t-1}^2 + \epsilon_t \end{aligned} \quad (9)$$

$$\begin{aligned} GAP_{i,t} = & \alpha_1 + \alpha_2 * |NetDemand_{t-1}^{G-S}| * 1_{NetDemand_{t-1}^{G-S} > 0} \\ & + \alpha_3 * |NetDemand_{t-1}^{G-S}| * 1_{NetDemand_{t-1}^{G-S} < 0} \\ & + \alpha_4 * News_{i,t}^P + \alpha_5 * \pi_{t-1} + \alpha_6 * \sigma_{\pi,t-1}^2 + x_{i,t} \alpha_7 + \epsilon_{i,t} \end{aligned} \quad (10)$$

$1_{NetDemand_{t-1}^{G-S} > 0}$ ($1_{NetDemand_{t-1}^{G-S} < 0}$) is a dummy variable that takes the value one only when $NetDemand_{t-1}^{G-S}$ is positive (negative), i.e., when general newspapers publish relatively more demand (supply) narratives than specialized newspapers. All other variables are defined as before. For equation 9, standard errors are computed with the Huber–White sandwich estimator. For equation 10, standard errors are clustered at the individual level. Table A5 in Appendix E shows the results.

The absolute expectation gap widens with narrative disagreement, irrespective of its sign. In fact, neither α_2 nor α_3 are negative. However, their magnitudes differ across dependent variables. In particular, the aggregate absolute expectation gap widens only when general newspapers publish relatively more de-

mand narratives than specialized newspapers. In contrast, the individual absolute expectation gap widens only when general newspapers publish relatively more supply narratives than specialized newspapers.

Overall, the absolute expectation gap always widens with narrative disagreement, though differently based on its sign.

5.4.3 Between-General-Newspaper Heterogeneity

So far, the analysis treats the New York Times, USA Today, and Washington Post as one general newspaper. A key question is whether the results differ depending on which general newspaper disagrees with specialized newspapers. This might be the case if the readership demographics of the three general newspapers differ (Pew Research Center, 2012). Additionally, disagreements between general and specialized newspapers may reveal partisan bias. In particular, the New York Times and Washington Post are typically considered liberal, USA Today is considered moderate, and the Wall Street Journal is considered conservative (Gallup/Knight, 2020). Therefore, if narrative disagreement proxies for partisan disagreement, the absolute expectation gap should widen only when there is narrative disagreement between liberal and conservative newspapers.

To answer this question, I construct three general-newspaper-specific measures of demand and supply narratives based on the methods described in sections 4.2 and 4.3. I call these measures $NetDemand_t^{NYT}$, $NetDemand_t^{USAT}$, and $NetDemand_t^{WaPo}$ and use them to derive three distinct measures of demand-supply narrative disagreement as follows:

$$NetDemand_t^{NYT-S} = NetDemand_t^{NYT} - NetDemand_t^S \quad (11)$$

$$NetDemand_t^{USAT-S} = NetDemand_t^{USAT} - NetDemand_t^S \quad (12)$$

$$NetDemand_t^{WaPo-S} = NetDemand_t^{WaPo} - NetDemand_t^S \quad (13)$$

Therefore, $NetDemand_t^{NYT-S}$ measures demand-supply narrative disagreement between the New York Times and specialized newspapers, $NetDemand_t^{USAT-S}$ measures demand-supply narrative disagreement between USA Today and specialized newspapers, and $NetDemand_t^{WaPo-S}$ measures demand-supply narrative disagreement between the Washington Post and specialized newspapers.

Finally, I estimate modified versions of the models specified at equations 3 and 4 in which the general newspapers' narrative measure ($NetDemand^{G-S}$) is replaced by each of the three general newspaper-specific narrative measures ($NetDemand^{NYT-S}$, $NetDemand^{USAT-S}$, and $NetDemand^{WaPo-S}$).

The results in Table A6 in Appendix F indicate that the results from the general newspapers' narrative measure broadly extend to those obtained from each of the three general newspaper-specific narrative measures. Therefore, this finding suggests that the aggregate general measure does not hide differences between its components and does not proxy for partisan disagreement.

Overall, the results of treating the three general newspapers independently align with those from aggregating them into one general newspaper.

5.4.4 Incentives to Gather Information about Inflation

Cavallo et al. (2017) show that households in high-inflation environments are more informed about inflation than those in low-inflation settings. Therefore, incentives for gathering information are crucial. Given the volatility of inflation in my sample, the evidence from Cavallo et al. (2017) suggests that the absolute expectation gap may widen with narrative disagreement, particularly at higher levels of inflation. For instance, households may read more newspaper articles when inflation is high; as a result, the relationship between the absolute expectation gap and narrative disagreement may strengthen. However, incentives may also depend on the persistence of inflation, which indicates how long a particular level of inflation is likely to last. For example, if inflation is high and persistent, its shocks create lasting effects, encouraging households to pay closer attention to newspapers. Conversely, if inflation is high but transient, its shocks might quickly fade, leading households to pay less attention to inflation.

I begin my analysis of the role of incentives in gathering information by examining how the relationship between narrative disagreement and the absolute expectation gap varies with the level of inflation. For this purpose, I estimate modified versions of equations 3 and 4 in which I interact narrative disagreement with the most recently observed level of inflation. The results are shown in Table A7 in Appendix G. The aggregate absolute expectation gap widens with narrative disagreement, irrespective of the level of inflation. In contrast, the individual absolute expectation gap widens with narrative disagreement only when the level of inflation is above its mean. Therefore, narrative disagreement especially matters for the absolute expectation gap when inflation is high.

Next, I continue my analysis by studying how the relationship between narrative disagreement and the absolute expectation gap changes based on the persistence of inflation. To answer this question, I measure inflation persistence using rolling-sample estimates of the first-order autocorrelation coefficient for inflation. This is a simple measure of inflation persistence previously used by Pivetta and Reis (2007) and obtained from data from $t-120$ to $t-1$, a ten-year window. Figure A9 in Appendix G shows the time-series evolution of this measure of inflation persistence. For most of my sample period, the serial correlation

is relatively low, rising above 0.5 only between 2016 and 2018, as well as from the second half of 2021. Notable sudden changes are the fall from 0.3 to zero in the second half of 2000, the jump from 0.25 to 0.45 around the end of 2008, and the fall from 0.5 to 0.25 around the end of 2018. In addition, inflation persistence rises steadily between 2002 and 2008 and after 2018, while it remains stable between 0.45 and 0.5 between 2009 and 2018.

Finally, I estimate modified versions of equations 3 and 4 in which I interact narrative disagreement with the inflation persistence measure, which I standardize to ease interpretation. The regression results are shown in Table A8 in Appendix G. The aggregate absolute expectation gap widens with narrative disagreement, irrespective of the level of inflation persistence. In contrast, the individual absolute expectation gap widens even further with narrative disagreement when the level of inflation persistence exceeds its mean. Therefore, narrative disagreement especially matters for the absolute expectation gap when inflation persistence is high.

Overall, the relationship between the absolute expectation gap and narrative disagreement strengthens as incentives to collect information increase.

5.4.5 Expert Disagreement

The evidence from section 5.3.3 highlights heterogeneity in the relationship between the absolute expectation gap and narrative disagreement across different age and education levels. Therefore, narrative disagreement also matters for between-household disagreement in inflation expectations. A crucial question is whether narrative disagreement is also linked to expert disagreement. If so, expert disagreement could affect the relationship between the absolute expectation gap and narrative disagreement. For instance, this could be the case if general and specialized newspapers publish the narratives of different sets of experts.

To answer these questions, I first regress the interquartile range of the SPF point forecasts on $|NetDemand^{G-S}|$ while controlling for the most recent level and volatility of inflation.²⁶ The results in table A9 in Appendix H indicate that narrative disagreement is positively related to expert disagreement. Therefore, expert disagreement might subsume the predictive power of narrative disagreement toward the absolute expectation gap. To verify this, I estimate modified versions of equations 3 and 4 in which I interact narrative disagreement with expert disagreement.²⁷ The results are shown in table A10 in Appendix H. The aggregate absolute expectation gap widens with narrative disagreement, irrespective of the level of expert

²⁶I use quarterly observations.

²⁷This measure is monthly and interpolated using quarterly values as done to measure the aggregate and individual absolute expectation gaps.

disagreement. In contrast, when controlling for expert disagreement, the individual absolute expectation gap no longer widens with narrative disagreement. Therefore, there is only weak evidence that narrative disagreement proxies for expert disagreement.

Overall, the results show that expert disagreement widens with narrative disagreement but does not subsume its predictive power toward the absolute expectation gap, at least at the aggregate level.

5.4.6 Forecast Errors

The results so far indicate that the absolute expectation gap widens with narrative disagreement. An important question is whether households make worse inflation predictions when narrative disagreement increases. If so, households' forecasts would be both closer to those of experts and more accurate when narrative disagreement declines. To verify this, I estimate the following models:

$$FE_t = \alpha_1 + \alpha_2 * |NetDemand_{t-1}^{G-S}| + \alpha_3 * News_t^P + \alpha_4 * \pi_{t-1} + \alpha_5 * \sigma_{\pi,t-1}^2 + \epsilon_t \quad (14)$$

$$FE_{i,t} = \alpha_1 + \alpha_2 * |NetDemand_{t-1}^{G-S}| + \alpha_3 * News_{i,t}^P + \alpha_4 * \pi_{t-1} + \alpha_5 * \sigma_{\pi,t-1}^2 + x_{i,t}\alpha_6 + \epsilon_{i,t} \quad (15)$$

Where $FE_t = |\pi_{t,t+12}^{MSC} - \pi_{t+12}|$ represents the absolute difference between the aggregate MSC forecast and the CPI inflation (at the forecast horizon), and $FE_{i,t} = |\pi_{i,t,t+12}^{MSC} - \pi_{t+12}|$ represents its individual-level counterpart. All other variables are defined as before. For equation 14, standard errors are computed with the Huber–White sandwich estimator. For equation 15, standard errors are clustered at the individual level.

The results in Table A11 in Appendix I indicate that only individual forecast errors widen with narrative disagreement. However, this is no longer the case when I control for the most recent level and volatility of inflation. Therefore, there is only weak evidence that households make better forecasts when narrative disagreement is low.

5.4.7 Narratives and Macroeconomic Dynamics

A key question is whether demand and supply narratives proxy for unobservable demand and supply dynamics, as captured by the comovement of inflation and unemployment. Assessing this relationship is particularly important due to the evidence from section 5.3.4 that the narratives of general newspapers

align incorrectly with experts' expectations. To this end, I estimate the following models:

$$Demand_t^j = \alpha_1 + \alpha_2 * \pi_t + \alpha_3 * u_t + \alpha_4 * \pi_t * u_t + \alpha_5 * Demand_{t-1}^j + \epsilon_t \quad (16)$$

$$Supply_t^j = \alpha_1 + \alpha_2 * \pi_t + \alpha_3 * u_t + \alpha_4 * \pi_t * u_t + \alpha_5 * Supply_{t-1}^j + \epsilon_t \quad (17)$$

$$NetDemand_t^j = \alpha_1 + \alpha_2 * \pi_t + \alpha_3 * u_t + \alpha_4 * \pi_t * u_t + \alpha_5 * NetDemand_{t-1}^j + \epsilon_t \quad (18)$$

Where $j \in G, S$, meaning I estimate the model separately for general and specialized narratives, $Demand_t^j$ and $Supply_t^j$ are the newspaper-level measures of demand and supply narratives from section 4.3, and u_t is the seasonally adjusted civilian unemployment rate.²⁸

The content of demand articles predominantly describes a negative relationship between changes in inflation and unemployment, while the opposite holds for supply articles. Therefore, α_4 should be non-positive in equation 16 and non-negative in equation 17. In addition, since $NetDemand^j$ is the scaled difference between $Demand^j$ and $Supply^j$, α_4 should be non-positive in equation 18. Table A12 in Appendix J shows the results.

Both general and specialized newspapers publish fewer demand narratives when inflation and unemployment move in the same direction. Thus, the demand narratives from both types of newspapers accurately reflect their prescribed inflation-unemployment relationship. However, general newspapers also reduce the publication of supply narratives under the same conditions. In contrast, the movement of inflation and unemployment does not impact the supply narratives published by specialized newspapers. As a result, the supply narratives from general newspapers do not align with their prescribed inflation-unemployment relationship. Likewise, general newspapers mistakenly publish relatively more demand narratives when inflation and unemployment move in the same direction. The opposite is true for specialized newspapers.

Overall, the evidence indicates that the narratives of specialized newspapers are more correctly aligned with macroeconomic dynamics than those of general newspapers.

6 Discussion and Relationship to the News Media Literature

A potential criticism of this paper is that narrative disagreement is endogenous, raising concerns about reverse causality. Specifically, newspapers may disagree more in their inflation narratives as the expecta-

²⁸Barnichon and Shapiro (2024) show that the ratio of job vacancies to unemployed workers captures inflation dynamics better than the traditionally used unemployment rate. Replacing the latter with the former produces virtually identical results.

tion gap widens. This concern could be valid, as the media bias model by Gentzkow and Shapiro (2006) suggests that narrative disagreement may arise from incentives to cater to households' and experts' differing narrative priors. If households and experts hold different narrative priors, newspapers are motivated to segment the market (i.e., disagree) and choose narratives that resonate with their readers. The premise of this argument is reasonable, as both Andre et al. (2024) and Dräger et al. (2016) demonstrate that households and experts often disagree in their narratives. In particular, Dräger et al. (2016) show that experts are more likely to form expectations based on the Phillips curve (Phillips, 1958; Samuelson & Solow, 1960), i.e., in line with demand narratives. Furthermore, the media bias model indicates that the challenging ex-post verifiability of inflation narratives enhances disagreement incentives. In particular, this difficulty arises from the description of macroeconomic shocks, the nature of which is uncertain in the short term (Lane, 2024). Otherwise, readers would avoid general newspapers that publish imprecise narratives, thereby discouraging disagreement. Overall, newspapers' selection of narratives may be non-random because it could stem from differences of opinion between households and experts.

The potential concern of reverse causality raises the question of which media activities might lead to endogenous narrative disagreement. Ahern and Peress (2023) identify two media activities typically undertaken by newspapers: news selection and creation.²⁹ News selection could generate endogenous disagreement between newspapers because it aligns with the model proposed by Carroll (2003), which assumes that newspapers relay experts' views to households. In particular, the incentive to cater to households' priors might generate narrative disagreement through general newspapers' selection of only a subset of the narratives considered by experts. In addition, this selection may become less representative when there is increased disagreement among the experts themselves, as suggested by the results in Section 5.4.5. On the other hand, news creation could be justified by newspapers' efforts to enhance their readers' understanding of the news. For instance, Guest (2021) demonstrates that WSJ earnings articles enhance price discovery and increase trading volume during S&P 500 earnings announcements, owing to their longer and more readable nature. Thus, news creation may lead to disagreement among newspapers because general newspapers tend to publish less complex narratives, which are more prevalent across households (Andre et al., 2024). However, both tables 1 and 3 indicate that inflation and causal inflation articles are similarly lengthy and complex across general and specialized newspapers. Overall, news selection is the media activity most likely to lead to endogenous narrative disagreement.

Do general and specialized newspapers behave in a way that could confirm the potential reverse

²⁹I do not consider news propagation because Ahern and Peress (2023) attribute this activity to search platforms, social media websites, and newswires. The newspapers studied in my paper do not fit into any of these categories.

causality concern? For this concern to be valid, the existing evidence of different demand and supply perspectives between households and experts requires that general and specialized newspapers should align accordingly. Specifically, general newspapers should systematically report relatively more supply narratives than specialized newspapers, implying that *NetDemand* should be negative on average. This is not the case, as table 4 shows that *NetDemand* is zero on average, and figure 6 indicates that there are periods of both positive and negative *NetDemand* with similarly short and long spells. Nevertheless, the evidence from section 5.4.7 suggests that general newspapers inaccurately publish supply narratives when inflation and unemployment move in opposite directions. Therefore, the overall evidence is inconclusive regarding whether narrative disagreement is endogenous or not.

Nonetheless, could we do anything to assuage the concern about reverse causality? Unfortunately, little guidance comes from studies on the media's role in the expectation formation process Carroll (2003); Larsen et al. (2021), as they face the same criticism due to their predominant use of aggregate time-series news data. Chahrour et al. (2024) and Andre et al. (2024) are two recent exceptions. Chahrour et al. (2024) exploit the MSC's panel structure in which some respondents are reinterviewed after six months. This feature allows the use of both time and fixed effects to isolate the causal impact of the responses to the survey question about inflation news consumption. Unfortunately, this approach cannot be used with inflation narratives, as no existing survey collects household-level data on narrative consumption. On the other hand, Andre et al. (2024) conduct an experiment on how news consumption affects households' narratives. While their results indicate that news consumption brings households' narratives closer to those of experts, the WSJ is the most frequently consulted news source by their respondents. If the WSJ were the primary source of inflation narratives for households, the expectation gap should narrow with increased WSJ inflation narrative press coverage. However, this is not the case empirically (see table 6). Overall, more work is needed to understand how media narratives shape people's beliefs and expectations.

7 Conclusion

This paper examines whether disagreement between newspapers' demand-supply narratives can explain the absolute gap in inflation expectations held by households compared to experts. I measure narrative disagreement by applying causality extraction (CE) and dictionary-based algorithms to more than 180,000 articles on inflation, published by three major general newspapers and a specialized one between 1991 and 2022. CE is designed to extract causal relations and, hence, can be used to construct measures of inflation narratives that describe the triggers of inflation. After applying CE, I categorize the extracted

inflation narratives into demand and supply narratives using existing dictionaries of demand and supply factors. Finally, I measure narrative disagreement based on attention differences to demand and supply narratives between general and specialized newspapers.

Using household and expert expectations from the MSC and the SPF, I find that the absolute expectation gap widens with narrative disagreement, regardless of the direction in which it is increasing. Across different households, the absolute expectation gap widens with narrative disagreement relatively more for older individuals and less for college-educated individuals. These findings are relevant and intuitive, as non-college-educated households are less likely to be readers of specialized newspapers, and older individuals are more likely to read newspapers. Importantly, the positive relationship between the absolute expectation gap and narrative disagreement strengthens as inflation rises and becomes more persistent, i.e., when incentives to be informed about inflation rise. In contrast, it is unaffected by expert disagreement and the type of general newspaper used to measure narrative disagreement, thereby removing potential concerns of partisan disagreement. In addition, the predictive power of demand-supply narrative disagreement is not subsumed by narrative disagreement regarding whether inflation is increasing or decreasing, or whether articles discuss realized or future inflation episodes. Finally, I provide suggestive evidence of households' forecast errors widening with narrative disagreement.

To explain the results, I connect the narratives of general and specialized newspapers to the demand and supply views of households and experts. I proxy for these views with the product of individual expected changes in inflation and unemployment, which should be positive/negative under predominantly supply/demand views. Unlike the narratives of specialized newspapers, the narratives of general newspapers align correctly only with households' views. This finding suggests that the narratives of general and specialized newspapers capture the opinions of different audiences. Interestingly, the narratives of general newspapers also incorrectly align with the joint dynamics of realized inflation and unemployment, while the narratives of specialized newspapers do not. Overall, these findings further suggest that general newspapers may convey narratives that are incorrect and do not accurately reflect the views of experts.

These results have interesting implications for policymakers and the media. They suggest that the absolute expectation gap shrinks with inflation press coverage when media disagreement is low. One policy lesson for central banks is to communicate their narratives through various channels to lower the dispersion of inflation forecasts across different demographics.

References

- Ahern, K. R., & Peress, J. (2023). *The role of media in financial decision-making*. Edward Elgar Publishing.
- Altenberg, B. (1984). Causal linking in spoken and written english. *Studia linguistica*, 38(1), 20–69.
- Andre, P., Haaland, I., Roth, C., Wiederholt, M., & Wohlfart, J. (2024). *Narratives about the macroeconomy* (Tech. Rep.). SAFE Working Paper.
- Andre, P., Pizzinelli, C., Roth, C., & Wohlfart, J. (2022). Subjective models of the macroeconomy: Evidence from experts and representative samples. *The Review of Economic Studies*, 89(6), 2958–2991.
- Apel, M., Blix Grimaldi, M., & Hull, I. (2022). How much information do monetary policy committees disclose? evidence from the fomc’s minutes and transcripts. *Journal of Money, Credit and Banking*, 54(5), 1459–1490.
- Baele, L., De Jong, F., & Trebbi, G. (2023). What triggers flights to safety? Available at SSRN 4503287.
- Baker, S. R., Bloom, N., Davis, S. J., & Kost, K. (2021). Policy news and stock market volatility. *University of Chicago, Becker Friedman Institute for Economics Working Paper*(2019-53).
- Barnichon, R., & Shapiro, A. H. (2024). How much has the cooling economy reduced inflation? *FRBSF Economic Letter*, 2024(30), 1–5.
- Blinder, A. S., Ehrmann, M., De Haan, J., & Jansen, D.-J. (2024). Central bank communication with the general public: Promise or false hope? *Journal of Economic Literature*, 62(2), 425–457.
- Blinder, A. S., Ehrmann, M., Fratzscher, M., De Haan, J., & Jansen, D.-J. (2008). Central bank communication and monetary policy: A survey of theory and evidence. *Journal of Economic Literature*, 46(4), 910–945.
- Bruine de Bruin, W., Vanderklaauw, W., Downs, J. S., Fischhoff, B., Topa, G., & Armantier, O. (2010). Expectations of inflation: The role of demographic variables, expectation formation, and financial literacy. *Journal of Consumer Affairs*, 44(2), 381–402.
- Bryan, M. F., & Venkatu, G. (2001). The demographics of inflation opinion surveys. *Economic Commentary*(10/15/2001).
- Bybee, L., Kelly, B. T., Manela, A., & Xiu, D. (2023). Business news and business cycles. *Journal of Finance, Forthcoming*.
- Carroll, C. D. (2003). Macroeconomic expectations of households and professional forecasters. *The Quarterly Journal of Economics*, 118(1), 269–298.

- Cavallo, A., Cruces, G., & Perez-Truglia, R. (2017). Inflation expectations, learning, and supermarket prices: Evidence from survey experiments. *American Economic Journal: Macroeconomics*, 9(3), 1–35.
- Chahrour, R., Shapiro, A. H., & Wilson, D. J. (2024). News selection and household inflation expectations..
- Coibion, O., & Gorodnichenko, Y. (2015). Is the phillips curve alive and well after all? inflation expectations and the missing disinflation. *American Economic Journal: Macroeconomics*, 7(1), 197–232.
- Coibion, O., Gorodnichenko, Y., & Weber, M. (2022). Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy*, 130(6), 000–000.
- Dougal, C., Engelberg, J., Garcia, D., & Parsons, C. A. (2012). Journalists and the stock market. *The Review of Financial Studies*, 25(3), 639–679.
- Dräger, L., Lamla, M. J., & Pfajfar, D. (2016). Are survey expectations theory-consistent? the role of central bank communication and news. *European Economic Review*, 85, 84–111.
- D’Acunto, F., Hoang, D., Paloviita, M., & Weber, M. (2019). Cognitive abilities and inflation expectations. In *Aea papers and proceedings* (Vol. 109, pp. 562–566).
- D’acunto, F., Hoang, D., Paloviita, M., & Weber, M. (2023). Iq, expectations, and choice. *The Review of Economic Studies*, 90(5), 2292–2325.
- D’Acunto, F., Malmendier, U., Ospina, J., & Weber, M. (2021). Exposure to grocery prices and inflation expectations. *Journal of Political Economy*, 129(5), 1615–1639.
- D’Acunto, F., Malmendier, U., & Weber, M. (2021). Gender roles produce divergent economic expectations. *Proceedings of the National Academy of Sciences*, 118(21), e2008534118.
- Ehrmann, M., Pfajfar, D., & Santoro, E. (2018). Consumers’ attitudes and their inflation expectations. *47th issue (February 2017) of the International Journal of Central Banking*.
- Ersahin, N., Giannetti, M., & Huang, R. (2024). Supply chain risk: Changes in supplier composition and vertical integration. *Journal of International Economics*, 147, 103854.
- Fendel, R., Lis, E. M., & Rülke, J.-C. (2011). Do professional forecasters believe in the phillips curve? evidence from the g7 countries. *Journal of Forecasting*, 30(2), 268–287.
- Gallup/Knight. (2020). American views 2020: Trust, media and democracy. *Knight Foundation*.
- Garcia, D., Hu, X., & Rohrer, M. (2023). The colour of finance words. *Journal of Financial Economics*, 147(3), 525–549.
- Geiger, M., & Scharler, J. (2021). How do people interpret macroeconomic shocks? evidence from u.s. survey data. *Journal of Money, Credit and Banking*, 53(4), 813–843.

- Gentzkow, M., & Shapiro, J. M. (2006). Media bias and reputation. *Journal of Political Economy*, 114(2), 280–316.
- Girju, R. (2003). Automatic detection of causal relations for question answering. In *Proceedings of the acl 2003 workshop on multilingual summarization and question answering* (pp. 76–83).
- Guest, N. M. (2021). The information role of the media in earnings news. *Journal of Accounting Research*, 59(3), 1021–1076.
- Hayo, B., Henseler, K., Rapp, M. S., & Zahner, J. (2022). Complexity of ecb communication and financial market trading. *Journal of International Money and Finance*, 128, 102709.
- Khoo, C. S., Kornfilt, J., Oddy, R. N., & Myaeng, S. H. (1998). Automatic extraction of cause-effect information from newspaper text without knowledge-based inferencing. *Literary and Linguistic Computing*, 13(4), 177–186.
- Lamla, M. J., & Lein, S. M. (2014). The role of media for consumers' inflation expectation formation. *Journal of Economic Behavior & Organization*, 106, 62–77.
- Lane, P. (2024). The 2021-2022 inflation surges and the monetary policy response through the lens of macroeconomic models. *SUERF Policy Note*, 364.
- Larsen, V. H., Thorsrud, L. A., & Zhulanova, J. (2021). News-driven inflation expectations and information rigidities. *Journal of Monetary Economics*, 117, 507–520.
- Malmendier, U., & Nagel, S. (2016). Learning from inflation experiences. *The Quarterly Journal of Economics*, 131(1), 53–87.
- Manela, A., & Moreira, A. (2017). News implied volatility and disaster concerns. *Journal of Financial Economics*, 123(1), 137–162.
- Mazumder, S. (2021). The reaction of inflation forecasts to news about the fed. *Economic Modelling*, 94, 256–264.
- Nimark, K. P., & Pitschner, S. (2019). News media and delegated information choice. *Journal of Economic Theory*, 181, 160–196.
- Pew Research Center. (2012). *In changing news landscape, even television is vulnerable*. Retrieved 2012-09-27, from <https://www.pewresearch.org/politics/2012/09/27/section-4-demographics-and-political-views-of-news-audiences/#attitudes-about-the-news>
- Pew Research Center. (2023). *News platform fact sheet*. Retrieved 2023-11-15, from <https://www.pewresearch.org/journalism/fact-sheet/news-platform-fact-sheet/>
- Pfajfar, D., & Santoro, E. (2013). News on inflation and the epidemiology of inflation expectations.

- Journal of Money, Credit and Banking*, 45(6), 1045–1067.
- Phillips, A. W. (1958). The relation between unemployment and the rate of change of money wage rates in the united kingdom, 1861-1957. *Economica*, 25(100), 283–299.
- Pivetta, F., & Reis, R. (2007). The persistence of inflation in the united states. *Journal of Economic Dynamics and Control*, 31(4), 1326–1358.
- Powell, J. H. (2020). New economic challenges and the fed’s monetary policy review. In *Speech at economic policy symposium sponsored by the federal reserve bank of kansas city, jackson hole, wyoming*.
- Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A. K., & Webber, B. L. (2008). The penn discourse treebank 2.0. In *Lrec*.
- Samuelson, P. A., & Solow, R. M. (1960). Analytical aspects of anti-inflation policy. *The American Economic Review*, 50(2), 177–194.
- Shapiro, A. H., Sudhof, M., & Wilson, D. J. (2022). Measuring news sentiment. *Journal of Econometrics*, 228(2), 221–243.
- Smales, L., & Apergis, N. (2017). Understanding the impact of monetary policy announcements: The importance of language and surprises. *Journal of Banking & Finance*, 80, 33–50.
- Stantcheva, S. (2024). *Why do we dislike inflation?* (Tech. Rep.). National Bureau of Economic Research.
- Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3), 1139–1168.
- Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms’ fundamentals. *The Journal of Finance*, 63(3), 1437–1467.
- Weber, M., D’Acunto, F., Gorodnichenko, Y., & Coibion, O. (2022, August). The subjective inflation expectations of households and firms: Measurement, determinants, and implications. *Journal of Economic Perspectives*, 36(3), 157-84.
- Yang, J., Han, S. C., & Poon, J. (2022). A survey on extraction of causal relations from natural language text. *Knowledge and Information Systems*, 1–26.

8 Tables

Table 1: Inflation articles - Summary statistics

		Mean	SD	Q1	Median	Q3
	Source					
# publication days per month	All	30	1	30	30	31
	General	28	2	28	30	30
	Specialized	26	4	22	27	30
Monthly volume	All	408	224	300	377	454
	General	166	141	101	141	188
	Specialized	241	99	186	233	280
Word count per article	All	1062	1663	572	840	1188
	General	1236	2384	634	909	1336
	Specialized	941	848	539	794	1099
Flesch–Kincaid per article	All	10.1	2.0	8.8	10.1	11.3
	General	10.6	2.1	9.3	10.6	11.8
	Specialized	9.7	1.9	8.6	9.7	10.9

Note: This table reports statistics on the monthly volume of inflation articles, their length, and the monthly number of publication days of inflation articles. Inflation articles contain at least one inflation expression: "inflation", "deflation", "consumer price", "producer price", "cpi", and "ppi." General newspapers include the NYT, USAT, and WaPo, whereas the WSJ is a specialized newspaper. The sample includes all days between 1991 and 2022.

Table 2: Dictionaries of categories of inflation narratives

Demand/Supply	Inflation narrative	Dictionary
Demand	Consumer Spending/Sentiment	"Consumer Spending and Sentiment" dictionary from Baker et al. (2021)
Demand	Monetary Policy	"Monetary Policy" dictionary from Baker et al. (2021) + { <i>"central-bank," "dallas fed," "easing," "easing of rate," "ecb," "fed assistance," "fed credibility," "fed easing," "fed expansion," "fed official," "fed rate," "fed response," "fed's bond-buying," "ffr," "hard landing," "high interest rate," "higher interest rate," "higher rate," "keep interest rate," "keep rate," "low interest rate," "low rate," "low rates," "lower base rate," "lower interest rate," "lower rate," "m1," "monetary," "monetary easing," "money growth," "money printing," "money-creation," "money-printing," "money-printing," "natural rate," "negative rates," "paul volcker," "printing money," "qe2," "raised interest rate," "rate cut," "rate increase," "rate reduction," "rise in interest rate," "rising interest rate," "slashing of short-term rate," "soaring interest rate," "soft landing," "volcker," "volckerism"</i> }
Demand	Spending/Deficit/Debt	"Spending/Deficit/Debt" dictionary from Baker et al. (2021) + { <i>"budget," "budgetary," "debt buildup," "debt burden," "deficit," "excessive debt," "federal fund," "federal spending," "government debt," "government support," "growth package," "recovery plan," "relief package," "rescue package," "social spending"</i> }
Supply	Commodities/Energy	"Commodity Markets" dictionary from Baker et al. (2021) + { <i>"commodity," "crop," "crude," "diesel," "electric," "electricity," "energy," "fuel," "gasoline," "grain," "lumber," "opec," "petroleum," "soybean"</i> }
Supply	Labor	"Labor Markets" and "Labor Disputes" dictionaries from Baker et al. (2021) + { <i>"collective bargaining agreement," "job creation," "job market," "jobless," "pay," "pay raise," "paycheck," "union," "work force," "workforce," "worker"</i> }
Supply	Supply chain	Top 100 supply-chain risk bigrams from Ersahin et al. (2024)

Table 3: Causal inflation articles - Summary statistics

		Mean	SD	Q1	Median	Q3
	Source					
# publication days per month	All	8	4	5	7	10
	General	3	3	1	3	4
	Specialized	5	3	3	5	7
Monthly volume	All	12	10	6	10	14
	General	4	5	1	3	5
	Specialized	8	6	4	7	11
Word count per article	All	1201	1133	726	969	1350
	General	1251	1150	764	1033	1504
	Specialized	1136	1109	694	898	1183
Flesch–Kincaid per article	All	10.2	1.7	9.2	10.2	11.3
	General	10.7	1.8	9.5	10.7	11.8
	Specialized	10.0	1.6	9.0	10.0	11.0

Note: This table reports statistics on the monthly volume of causal inflation articles, their length, and the monthly number of publication days of causal inflation articles. Inflation articles contain at least one inflation expression: "inflation", "deflation", "consumer price", "producer price", "cpi", and "ppi." General newspapers include the NYT, USAT, and WaPo, whereas the WSJ is a specialized newspaper. The sample includes all days between 1991 and 2022.

Table 4: Narratives and their disagreement - Summary statistics

	Count	Mean	Std	Min	25%	50%	75%	Max
<i>NetDemand</i> ^G	384.0	-0.073***	0.227	-1.000	-0.125	0.000	0.000	0.875
<i>NetDemand</i> ^S	384.0	-0.049***	0.166	-1.000	-0.143	-0.048	0.000	0.714
<i>NetDemand</i> ^{G-S}	384.0	-0.023*	0.242	-0.952	-0.125	0.000	0.109	0.780
<i>NetHawkish</i> ^G	384.0	0.064***	0.163	-0.500	0.000	0.000	0.100	1.000
<i>NetHawkish</i> ^S	384.0	0.081***	0.174	-0.350	0.000	0.050	0.150	1.000
<i>NetHawkish</i> ^{G-S}	384.0	-0.017	0.202	-0.750	-0.100	0.000	0.100	0.850
<i>NetObserved</i> ^G	384.0	0.045***	0.165	-0.500	-0.056	0.000	0.111	1.000
<i>NetObserved</i> ^S	384.0	0.048***	0.225	-0.632	-0.105	0.053	0.158	1.000
<i>NetObserved</i> ^{G-S}	384.0	-0.003	0.256	-0.868	-0.158	0.000	0.115	1.091
<i>ConsSpendSent</i> ^{G-S}	384.0	0.005	0.144	-1.000	0.000	0.000	0.000	1.000
<i>MonPol</i> ^{G-S}	384.0	-0.031***	0.161	-0.846	-0.077	0.000	0.000	0.701
<i>SpendDefDebt</i> ^{G-S}	384.0	0.029***	0.162	-0.364	0.000	0.000	0.000	1.000
<i>ComEne</i> ^{G-S}	384.0	-0.004	0.143	-0.636	-0.052	0.000	0.043	0.810
<i>Labor</i> ^{G-S}	384.0	-0.030***	0.138	-0.545	-0.082	0.000	0.000	0.727
<i>SupplyChain</i> ^{G-S}	384.0	0.002	0.106	-1.000	0.000	0.000	0.000	1.000

Note: This table shows summary statistics for *NetDemand*, *NetHawkish*, and *NetObserved* measures, as well as in the individual demand-supply narrative types over time between general and specialized newspapers. *NetDemand* and individual demand-supply narrative factors are measured as described in section 4.3. *NetHawkish* is measured as described in section 4.4. *NetObserved* is measured as described in section 4.5. General newspapers include the NYT, USAT, and WaPo, whereas the WSJ is a specialized newspaper. * $p < 10\%$; ** $p < 5\%$; *** $p < 1\%$ refer to two-sided tests for the null of the variable being equal to zero. The sample includes all months between January 1991 and December 2022.

Table 5: Inflation press coverage and expectation gap

	GAP_t			$GAP_{i,t}$		
$News_{t-1}^G$	0.452 (0.530)		0.981 (0.725)	1.244*** (0.068)		1.693*** (0.090)
$News_{t-1}^S$		-0.187 (0.306)	-0.624 (0.448)		0.343*** (0.053)	-0.544*** (0.070)
Demographics control	-	-	-	Yes	Yes	Yes
Past inflation control	Yes	Yes	Yes	Yes	Yes	Yes
Past inflation volatility control	Yes	Yes	Yes	Yes	Yes	Yes
Heard of inflation news control	Yes	Yes	Yes	Yes	Yes	Yes
Adj-R2	39.34	39.17	39.75	4.20	3.99	4.24
N	384	384	384	162453	162453	162453

Note: This table shows the results obtained from estimating the models specified in equations 1 and 2, as well as modified versions of them where the news measures appear one by one. General newspapers include the NYT, USAT, and WaPo, whereas the WSJ is a specialized newspaper. Demographic controls include gender, age, income, education, marital status, and location in the United States. For equation 1, standard errors are computed with the Huber–White sandwich estimator. For equation 2, standard errors are clustered at the individual level. The sample includes all months between 1991 and 2022.

Table 6: Causal inflation press coverage and expectation gap

	GAP_t			$GAP_{i,t}$		
$CausalNews_{t-1}^G$	0.412 (0.453)		0.136 (0.482)	1.133*** (0.067)		1.086*** (0.073)
$CausalNews_{t-1}^S$		0.531** (0.206)	0.502** (0.238)		0.367*** (0.044)	0.085* (0.047)
Demographics control	-	-	-	Yes	Yes	Yes
Past inflation control	Yes	Yes	Yes	Yes	Yes	Yes
Past inflation volatility control	Yes	Yes	Yes	Yes	Yes	Yes
Heard of inflation news control	Yes	Yes	Yes	Yes	Yes	Yes
Adj-R2	39.31	40.00	39.86	4.16	4.00	4.16
N	384	384	384	162453	162453	162453

Note: This table shows the results obtained from estimating modified versions of the models specified in equations 1 and 2, where measures of causal inflation press coverage intensity replace measures of inflation press coverage intensity. General newspapers include the NYT, USAT, and WaPo, whereas the WSJ is a specialized newspaper. Demographic controls include gender, age, income, education, marital status, and location in the United States. For the modified version of equation 1, standard errors are computed with the Huber–White sandwich estimator. Standard errors are clustered at the individual level for the modified version of equation 2. The sample includes all months between 1991 and 2022.

Table 7: Narrative disagreement and expectation gap

(a) Aggregate

		GAP_t			
$ NetDemand_{t-1}^{G-S} $	0.376*	0.371*	0.376*	0.373*	
	(0.200)	(0.201)	(0.206)	(0.206)	
$ NetHawkish_{t-1}^{G-S} $		0.024		0.027	
		(0.216)		(0.227)	
$ NetObserved_{t-1}^{G-S} $			-0.004	-0.010	
			(0.193)	(0.202)	
Heard of inflation news control	Yes	Yes	Yes	Yes	
Past inflation control	Yes	Yes	Yes	Yes	
Past inflation volatility control	Yes	Yes	Yes	Yes	
Adj-R2	39.81	39.65	39.65	39.49	
N	384	384	384	384	

(b) Individual

	$GAP_{i,t}$			
$ NetDemand_{t-1}^{G-S} $	0.133*** (0.037)	0.097** (0.038)	0.090** (0.038)	0.071* (0.038)
$ NetHawkish_{t-1}^{G-S} $		0.179*** (0.043)		0.124*** (0.045)
$ NetObserved_{t-1}^{G-S} $			0.202*** (0.038)	0.174*** (0.040)
Demographic controls	Yes	Yes	Yes	Yes
Heard of inflation news control	Yes	Yes	Yes	Yes
Past inflation control	Yes	Yes	Yes	Yes
Past inflation volatility control	Yes	Yes	Yes	Yes
Adj-R2	3.97	3.98	3.99	3.99
N	162453	162453	162453	162453

Note: This table shows the slope estimates obtained from estimating the models at equations 3 and 4. GAP_t is the absolute difference in one-year-ahead mean inflation expectations between households and experts. $GAP_{i,t}$ is the absolute difference in one-year-ahead inflation expectations between individual households and the expert consensus. $NetDemand^{G-S}$ is the demand-supply narrative disagreement measure constructed as described in section 4.3. $NetHawkish^{G-S}$ is the hawkish-dovish narrative disagreement measure and is constructed as described in section 4.4. $NetObserved^{G-S}$ is the observed-expected narrative disagreement measure and is constructed as described in section 4.4. For equation 3, standard errors are computed with the Huber–White sandwich estimator. For equation 4, standard errors are clustered at the individual level. The sample includes all months between 1991 and 2022.

Table 8: Narrative disagreement and expectation gap across demographics

	$GAP_{i,t}$			
$ NetDemand_{t-1}^{G-S} $	0.144*** (0.047)	0.078* (0.043)	0.314*** (0.083)	0.320*** (0.083)
$ NetDemand_{t-1}^{G-S} * FEMALE_{i,t}$	-0.022 (0.073)			
$ NetDemand_{t-1}^{G-S} * AGE_{i,t}$		0.005** (0.002)		
$ NetDemand_{t-1}^{G-S} * EDUC1_{i,t}$			-0.087 (0.333)	
$ NetDemand_{t-1}^{G-S} * EDUC2_{i,t}$			0.128 (0.269)	
$ NetDemand_{t-1}^{G-S} * EDUC4_{i,t}$			-0.133 (0.111)	
$ NetDemand_{t-1}^{G-S} * EDUC5_{i,t}$			-0.360*** (0.106)	
$ NetDemand_{t-1}^{G-S} * EDUC6_{i,t}$			-0.267** (0.109)	
$ NetDemand_{t-1}^{G-S} * INC1_{i,t}$				-0.128 (0.141)
$ NetDemand_{t-1}^{G-S} * INC2_{i,t}$				-0.343*** (0.122)
$ NetDemand_{t-1}^{G-S} * INC4_{i,t}$				-0.183* (0.110)
$ NetDemand_{t-1}^{G-S} * INC5_{i,t}$				-0.272*** (0.102)
Demographic controls	Yes	Yes	Yes	Yes
Heard of inflation news control	Yes	Yes	Yes	Yes
Past inflation control	Yes	Yes	Yes	Yes
Past inflation volatility control	Yes	Yes	Yes	Yes
Adj-R2	3.97	3.97	3.98	3.97
N	162453	162453	162453	162453

Note: This table shows the results obtained from estimating a modified version of 4. The model differs because I sequentially interact the variable $|NetDemand_{t-1}^{G-S}|$ with several consumer characteristics represented in x_i . $FEMALE_{i,t}$ is a dummy taking the value of one when the respondent is a woman. $AGE_{i,t}$ measures the respondent's age in integers. $INC1_{i,t}$, $INC2_{i,t}$, $INC4_{i,t}$, and $INC5_{i,t}$ are dummies taking value one when the income of the respondent belongs to the first, second, fourth, and fifth quintiles of the cross-sectional MSC income distribution, respectively. $EDUC1_{i,t}$, $EDUC2_{i,t}$, $EDUC4_{i,t}$, $EDUC5_{i,t}$, and $EDUC6_{i,t}$ are dummies taking value one when the respondent's education respectively belongs to the group "Grade 0–8, no high school diploma," "Grade 9–12, no high school diploma," "4 yrs. of college, no degree," "3 yrs. of college, with degree," and "4 yrs. of college, with degree." Standard errors are clustered at the individual level, as some respondents in the MSC are reinterviewed. The sample includes all months between 1991 and 2022.

Table 9: Newspaper and individual narratives

	$\Delta\pi_{i,t,t+12}^{MSC} * \Delta u_{i,t,t+12}^{MSC}$			$\Delta\pi_{i,t,t+4}^{SPF} * \Delta u_{i,t,t+4}^{SPF}$		
$NetDemand_{t-1}^G$	-0.149*** (0.033)		-0.131*** (0.033)	0.282*** (0.040)		0.405*** (0.049)
$NetDemand_{t-1}^S$		-0.157*** (0.044)	-0.119*** (0.044)		-0.226** (0.036)	-0.345*** (0.044)
Demographic controls	Yes	Yes	Yes	-	-	-
Heard of inflation news control	Yes	Yes	Yes	-	-	-
Past inflation control	Yes	Yes	Yes	Yes	Yes	Yes
Past inflation volatility control	Yes	Yes	Yes	Yes	Yes	Yes
Adj-R2	0.50	0.49	0.51	8.07	7.99	8.60
N	161044	161044	161044	4614	4614	4614

Note: This table shows the results obtained from estimating the models specified in equations 5 and 6, as well as modified versions of them where the narrative measures appear one by one. $\pi_{i,t,t+12}^{MSC}$ measures the MSC's household expected change in inflation, whereas $\Delta u_{i,t,t+12}^{MSC}$ measures the MSC's household expected change in unemployment. $\Delta\pi_{i,t,t+4}^{SPF}$ and $\Delta u_{i,t,t+4}^{SPF}$ respectively measure the SPF respondent's expected change in inflation and unemployment. $NetDemand^G$ and $NetDemand^S$ are the NetDemand measures for general and specialized newspapers, respectively. They are constructed as described in section 4.3. General newspapers include the NYT, USAT, and WaPo, whereas the WSJ is a specialized newspaper. Demographic controls include gender, age, income, education, marital status, and location in the United States. The model with results in the first three columns is estimated at the monthly frequency, whereas the one in the other columns is estimated at the quarterly frequency. Standard errors are clustered at the individual level, as some respondents in the MSC and SPF are reinterviewed. The sample includes all months and quarters between 1991 and 2022.

9 Figures

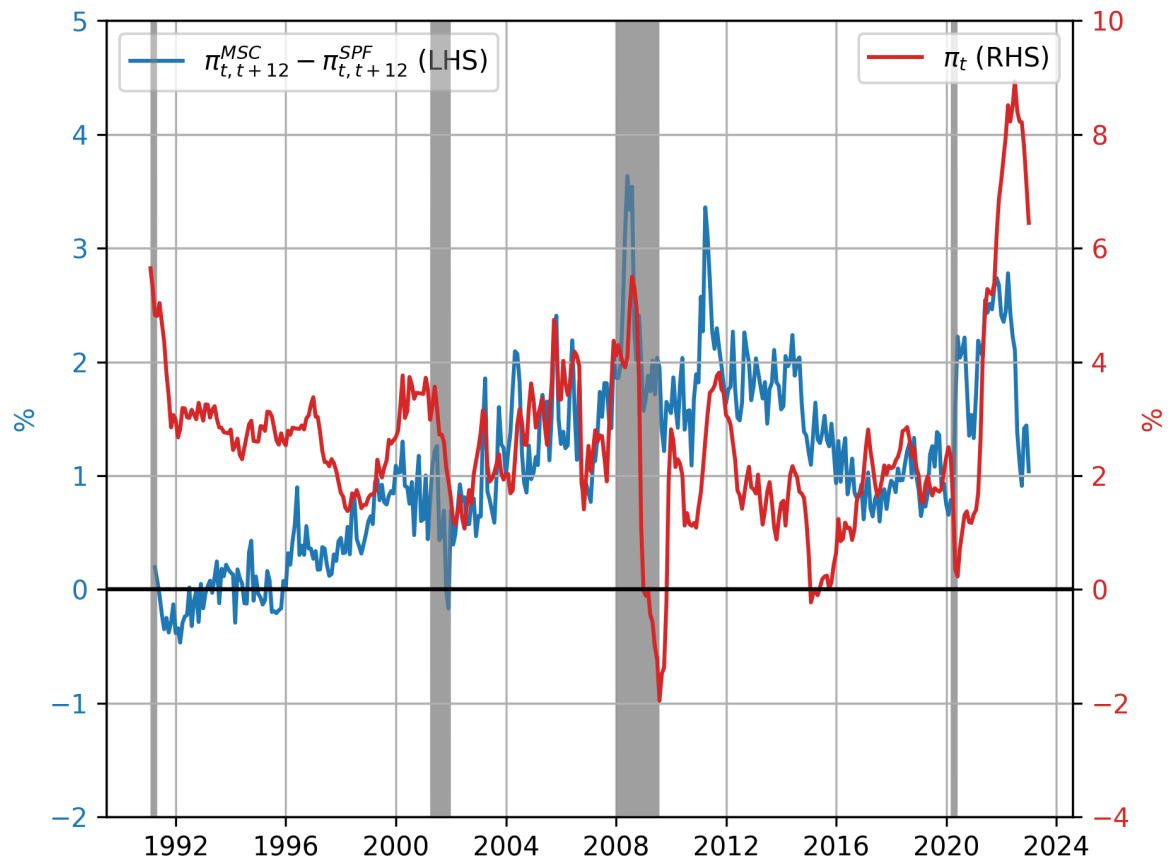


Figure 1: Expectation gap and inflation

Note: The blue line (left axis) shows the difference between the average one-year ahead inflation expectations from the Michigan Survey of Consumers (MSC) and the Survey of Professional Forecasters. The red line (right axis) shows the monthly year-on-year CPI inflation rate. Shaded areas represent NBER recession periods. The sample includes all months between January 1991 and December 2022.

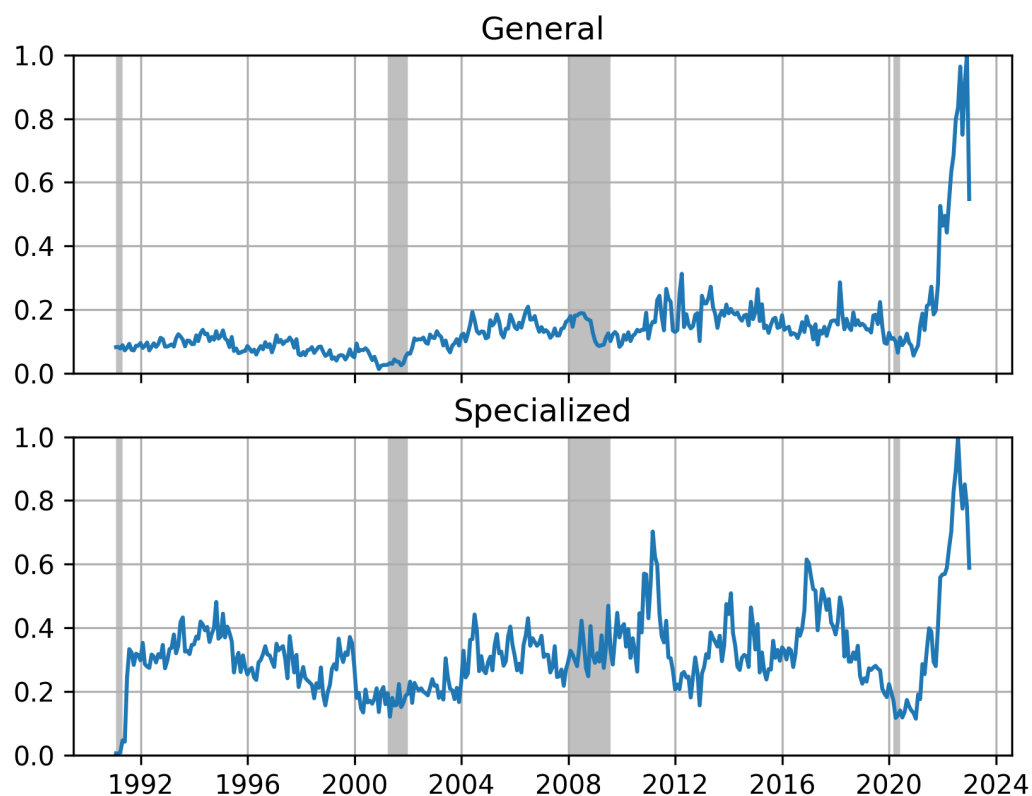


Figure 2: Inflation press coverage intensity

Note: This figure shows the monthly volume of inflation articles scaled by its maximum in any month separately for general and specialized newspapers. General newspapers include the NYT, USAT, and WaPo, whereas the WSJ is a specialized newspaper. Shaded areas represent NBER recession periods. The sample includes all months between January 1991 and December 2022.

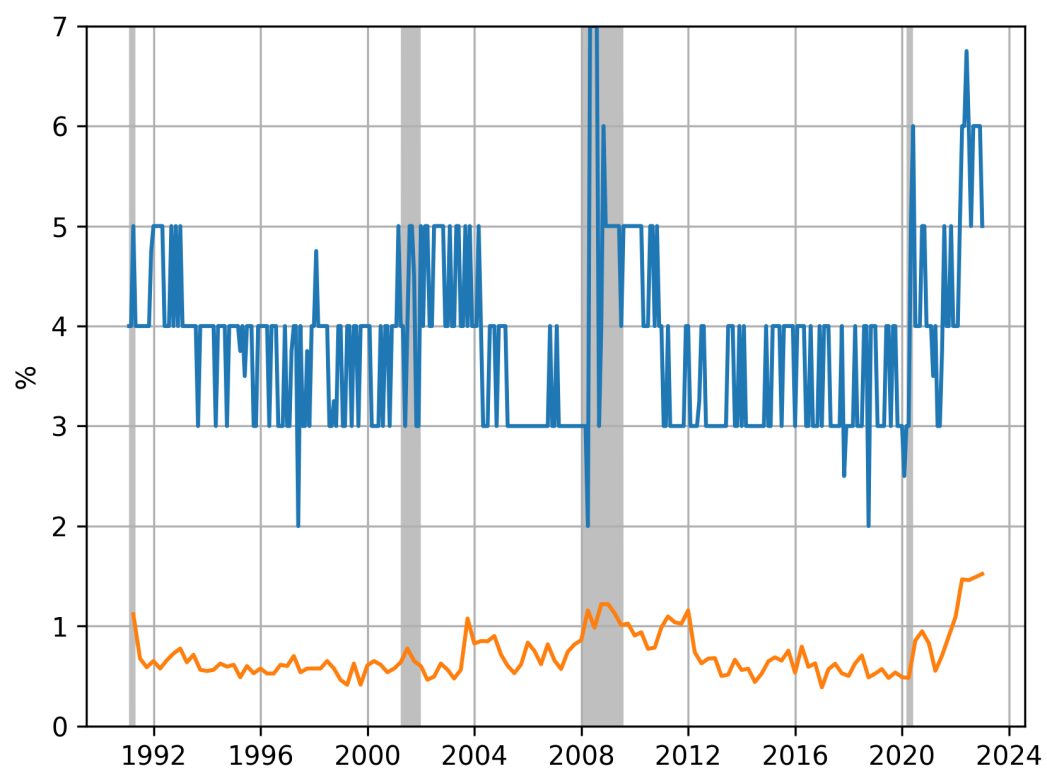


Figure 3: Inflation expectations disagreement

Note: This figure shows the cross-sectional interquartile ranges of the MSC (blue) and SPF (orange) inflation forecasts. Shaded areas represent NBER recession periods. The sample includes all months between January 1991 and December 2022.

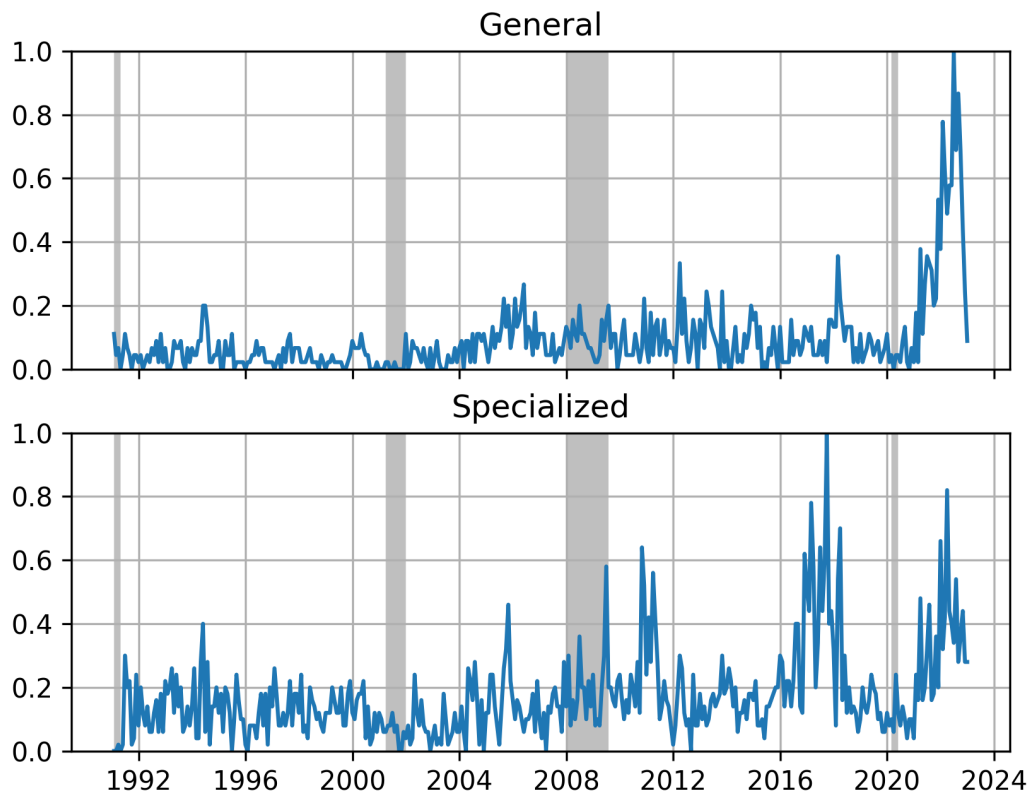


Figure 4: Causal inflation press coverage intensity

Note: This figure shows the monthly volume of causal inflation articles scaled by its maximum in any month separately for general and specialized newspapers. General newspapers include the NYT, USAT, and WaPo, whereas the WSJ is a specialized newspaper. Shaded areas represent NBER recession periods. The sample includes all months between January 1991 and December 2022.

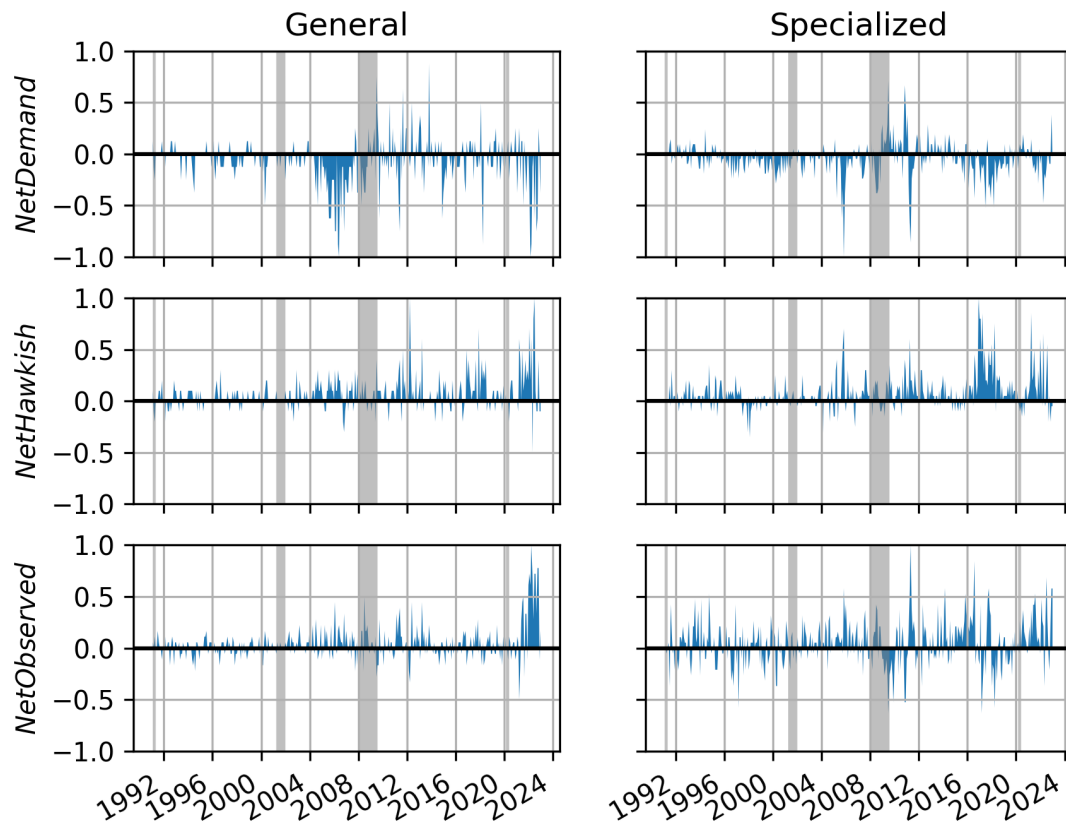


Figure 5: Demand-Supply, Hawkish-Dovish, and Observed-Expected narratives

Note: This figure shows the evolution of the *NetDemand*, *NetHawkish*, and *NetObserved* measures over time for general and specialized newspapers separately. *NetDemand* is measured as described in Section 4.3. *NetHawkish* is measured as described in section 4.4. *NetObserved* is measured as described in section 4.5. General newspapers include the NYT, USAT, and WaPo, whereas the WSJ is a specialized newspaper. Shaded areas represent NBER recession periods. The sample includes all months between January 1991 and December 2022.

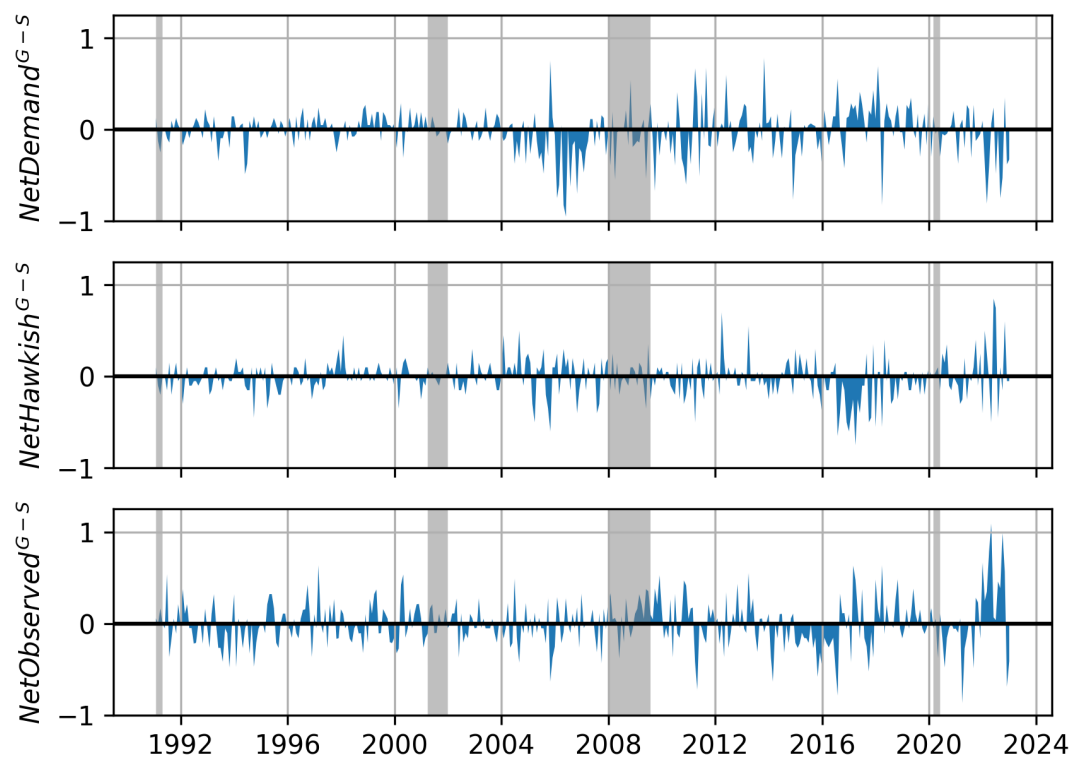


Figure 6: Disagreement about demand-supply, hawkish-dovish, and observed-expected narratives

Note: This figure shows the evolution of the difference in the *NetDemand*, *NetHawkish*, and *NetObserved* measures over time between general and specialized newspapers. *NetDemand* is measured as described in Section 4.3. *NetHawkish* is measured as described in section 4.4. *NetObserved* is measured as described in section 4.5. General newspapers include the NYT, USAT, and WaPo, whereas the WSJ is a specialized newspaper. Shaded areas represent NBER recession periods. The sample includes all months between January 1991 and December 2022.

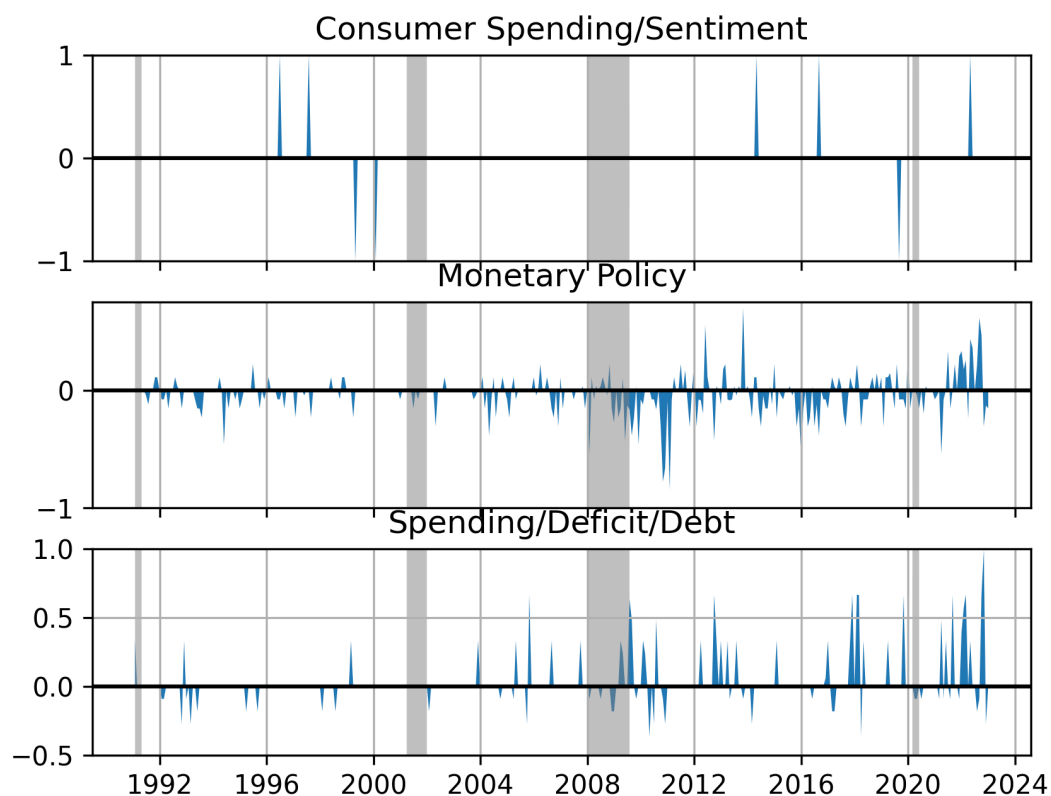


Figure 7: Disagreement about demand narratives

Note: This figure shows the evolution of the difference in the scaled monthly volume of individual demand narrative types over time between general and specialized newspapers. Individual demand narrative types are measured as described in section 4.3. General newspapers include the NYT, USAT, and WaPo, whereas the WSJ is a specialized newspaper. Shaded areas represent NBER recession periods. The sample includes all months between January 1991 and December 2022.

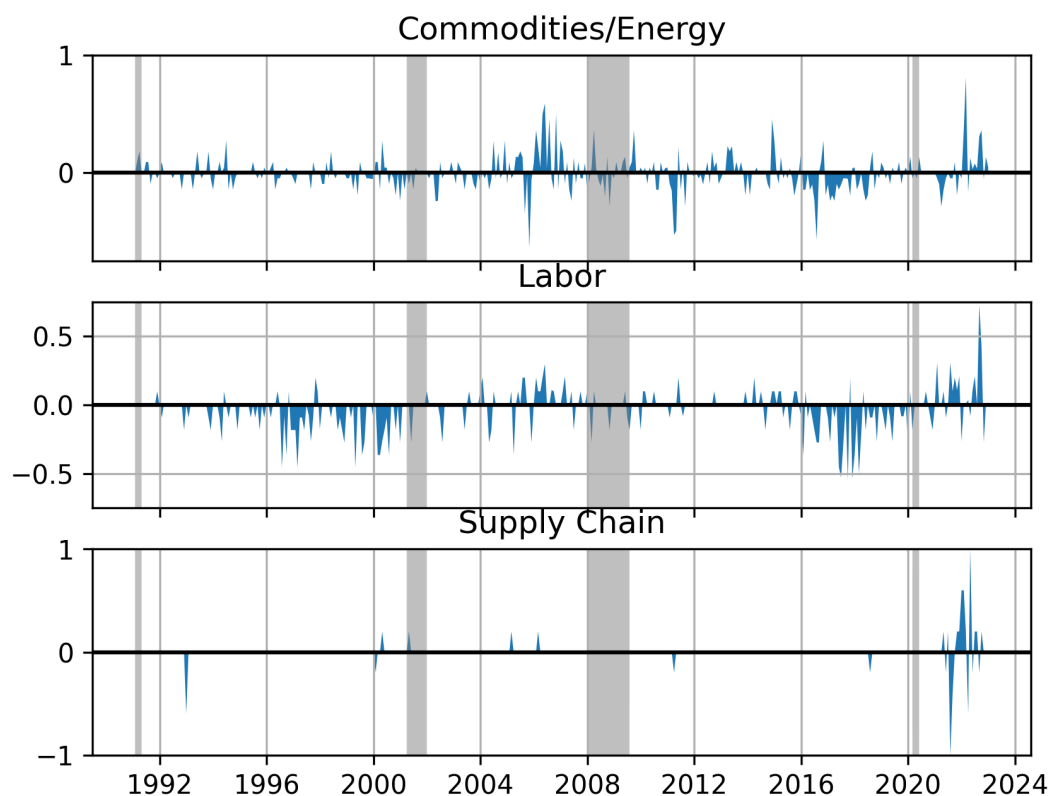


Figure 8: Disagreement about supply narratives

Note: This figure shows the evolution of the difference in the scaled monthly volume of individual supply narrative types over time between general and specialized newspapers. Individual supply narrative types are measured as described in section 4.3. General newspapers include the NYT, USAT, and WaPo, whereas the WSJ is a specialized newspaper. Shaded areas represent NBER recession periods. The sample includes all months between January 1991 and December 2022.

A Causal Verb and Link Keywords

Table A1: Benchmark causal verbs and links

Causal Verb	When
benefit	C-V-E (Active) or E-V-C (Passive)
boost	C-V-E (Active) or E-V-C (Passive)
break	C-V-E (Active) or E-V-C (Passive)
bring	C-V-E (Active) or E-V-C (Passive)
cause	C-V-E (Active) or E-V-C (Passive)
cinch	C-V-E (Active) or E-V-C (Passive)
compel	C-V-E (Active) or E-V-C (Passive)
consign	C-V-E (Active) or E-V-C (Passive)
double	C-V-E (Active) or E-V-C (Passive)
drive	C-V-E (Active) or E-V-C (Passive)
fuel	C-V-E (Active) or E-V-C (Passive)
give	C-V-E (Active) or E-V-C (Passive)
hurt	C-V-E (Active) or E-V-C (Passive)
impact	C-V-E (Active) or E-V-C (Passive)
increase	C-V-E (Active) or E-V-C (Passive)
persuade	C-V-E (Active) or E-V-C (Passive)
portend	C-V-E (Active) or E-V-C (Passive)
produce	C-V-E (Active) or E-V-C (Passive)
prompt	C-V-E (Active) or E-V-C (Passive)
push	C-V-E (Active) or E-V-C (Passive)
put	C-V-E (Active) or E-V-C (Passive)
remove	C-V-E (Active) or E-V-C (Passive)
require	C-V-E (Active) or E-V-C (Passive)
save	C-V-E (Active) or E-V-C (Passive)
vault	C-V-E (Active) or E-V-C (Passive)
attribute	E-V-C (Passive)
blame	E-V-C (Passive)
head	E-V-C (Passive)
link	E-V-C (Passive)

Note: This table lists the causal verbs extracted from the PDTB dataset. Column "When" shows how a causal relationship involving a causal verb can be expressed in an SVO pattern. In particular, C-V-E is used when the cause is in the subject position, the effect is in the object position, and the inverse holds for E-V-C. In addition, the word between in parentheses is the form in which the causal verb needs to be used.

Table A2: Causal links from Altenberg (1984) and causal verbs from Girju (2003)

Causal Link	When	Causal Verb	When
a consequence of	Before Cause	result	C-V-E (Active)
as a result of	Before Cause	give birth to	C-V-E (Active)
because	Before Cause	activate	C-V-E (Active) or E-V-C (Passive)
because of	Before Cause	actuate	C-V-E (Active) or E-V-C (Passive)
due to	Before Cause	arouse	C-V-E (Active) or E-V-C (Passive)
for	Before Cause	begin	C-V-E (Active) or E-V-C (Passive)
for the sake of	Before Cause	bring	C-V-E (Active) or E-V-C (Passive)
on account of	Before Cause	call	C-V-E (Active) or E-V-C (Passive)
on grounds of	Before Cause	cause	C-V-E (Active) or E-V-C (Passive)
on the grounds of	Before Cause	commence	C-V-E (Active) or E-V-C (Passive)
owing to	Before Cause	conduce	C-V-E (Active) or E-V-C (Passive)
since	Before Causes	contribute	C-V-E (Active) or E-V-C (Passive)
the conclusion of	Before Cause	create	C-V-E (Active) or E-V-C (Passive)
the consequence of	Before Cause	develop	C-V-E (Active) or E-V-C (Passive)
the result of	Before Cause	educate	C-V-E (Active) or E-V-C (Passive)
by reason of	Before Effect	effect	C-V-E (Active) or E-V-C (Passive)
so that	Before Effect	effectuate	C-V-E (Active) or E-V-C (Passive)
the reason for	Before Effect	elicit	C-V-E (Active) or E-V-C (Passive)
the reason why	Before Effect	entail	C-V-E (Active) or E-V-C (Passive)
why	Before Effect	evoke	C-V-E (Active) or E-V-C (Passive)
		fire	C-V-E (Active) or E-V-C (Passive)
		generate	C-V-E (Active) or E-V-C (Passive)
		implicate	C-V-E (Active) or E-V-C (Passive)
		induce	C-V-E (Active) or E-V-C (Passive)
		launch	C-V-E (Active) or E-V-C (Passive)
		lead	C-V-E (Active) or E-V-C (Passive)
		make	C-V-E (Active) or E-V-C (Passive)
		kick	C-V-E (Active) or E-V-C (Passive)
		kindle	C-V-E (Active) or E-V-C (Passive)
		originate	C-V-E (Active) or E-V-C (Passive)
		produce	C-V-E (Active) or E-V-C (Passive)
		provoke	C-V-E (Active) or E-V-C (Passive)
		set in motion	C-V-E (Active) or E-V-C (Passive)
		set off	C-V-E (Active) or E-V-C (Passive)
		set up	C-V-E (Active) or E-V-C (Passive)
		spark	C-V-E (Active) or E-V-C (Passive)
		start	C-V-E (Active) or E-V-C (Passive)
		stimulate	C-V-E (Active) or E-V-C (Passive)
		stir	C-V-E (Active) or E-V-C (Passive)
		trigger	C-V-E (Active) or E-V-C (Passive)
		unleash	C-V-E (Active) or E-V-C (Passive)
		stem	E-V-C (Active)
		derive	E-V-C (Active) or E-V-C (Passive)
		associate	E-V-C (Passive)
		link	E-V-C (Passive)
		relate	E-V-C (Passive)

Note: The left panel shows the causal links selected from the list published by Altenberg (1984). Column "When" shows whether a causal link precedes a phrase containing the cause or the effect. The right panel shows the causal verbs selected from the list published by Girju (2003). As for the causal verbs from table A1 in Appendix A, column "When" shows how a causal relationship involving a causal verb can be expressed in an SVO pattern. In particular, C-V-E is used when the cause is in the subject position, the effect is in the object position, and the inverse holds for E-V-C. In addition, the word between in parentheses is the form in which the causal verb needs to be used.

Table A3: Manual evaluation of causal relations extracted from causal links, causal verbs, conditionals, and resultative constructions

Causal relation	Causal connective	TP	TN	FP	FN	TPR	TNR	PPV	ACC	F1
Causal link	the result of	18	14	6	2	90%	70%	75%	80%	82%
Causal link	because of	17	15	5	3	85%	75%	77%	80%	81%
Causal link	due to	16	14	6	4	80%	70%	73%	75%	76%
Causal link	as a result of	15	13	7	5	75%	65%	68%	70%	71%
Causal link	why	8	17	3	12	40%	85%	73%	63%	52%
Causal link	because	5	17	3	15	25%	85%	63%	55%	36%
Causal link	so that	3	19	1	17	15%	95%	75%	55%	25%
Causal link	since	1	20	0	19	5%	100%	100%	53%	10%
Causal link	for	0	20	0	20	0%	100%	0%	50%	0%
Causal verb	prompt	20	17	3	0	100%	85%	87%	93%	93%
Causal verb	cause	20	16	4	0	100%	80%	83%	90%	91%
Causal verb	break	19	16	4	1	95%	80%	83%	88%	88%
Causal verb	increase	15	19	1	5	75%	95%	94%	85%	83%
Causal verb	create	19	13	7	1	95%	65%	73%	80%	83%
Causal verb	boost	18	14	6	2	90%	70%	75%	80%	82%
Causal verb	generate	20	11	9	0	100%	55%	69%	78%	82%
Causal verb	push	20	11	9	0	100%	55%	69%	78%	82%
Causal verb	produce	17	15	5	3	85%	75%	77%	80%	81%
Causal verb	spark	20	10	10	0	100%	50%	67%	75%	80%
Causal verb	bring	17	14	6	3	85%	70%	74%	78%	79%
Causal verb	drive	17	14	6	3	85%	70%	74%	78%	79%
Causal verb	stimulate	16	15	5	4	80%	75%	76%	78%	78%
Causal verb	stir	19	10	10	1	95%	50%	66%	73%	78%
Causal verb	trigger	18	10	10	2	90%	50%	64%	70%	75%
Causal verb	fuel	20	5	15	0	100%	25%	57%	63%	73%
Causal verb	unleash	19	6	14	1	95%	30%	58%	63%	72%
Causal verb	put	4	17	3	16	20%	85%	57%	53%	30%
Causal verb	give	2	19	1	18	10%	95%	67%	53%	17%
Causal verb	make	2	18	2	18	10%	90%	50%	50%	17%
Causal verb	call	0	20	0	20	0%	100%	0%	50%	0%
Causal verb	require	0	20	0	20	0%	100%	0%	50%	0%
Conditional	if	5	20	0	15	25%	100%	100%	63%	40%
Resultative construction	keep	18	20	0	2	90%	100%	100%	95%	95%

Note: This table shows the results of the manual evaluation conducted on causal relations extracted from the causal verbs and links selected from tables A1 and A2 in Appendix A, as well as conditionals and resultative constructions, and following the steps detailed in section C. I select causal verbs and links shown in tables A1 and A2 in Appendix A based on whether they appear in at least 20 causal inflation sentences. The verbs shown for resultative constructions are similarly selected. The third to sixth columns report the number of true positives, true negatives, false positives, and false negatives for each causal connective. The following columns report, respectively, the true positive rate, true negative rate, positive predicted value, accuracy, and F1 score, as computed in Section C.

B Dependency Parsing

Assigning a syntactic structure to a sentence is a core task in NLP called sentence parsing. There are currently two main approaches to sentence parsing: constituency parsing and dependency parsing. As dependency parsing is the de facto tool in RE and, hence, of CE, I adopt it and describe its details in this section.

Dependency parsing forms the syntactic structure of a sentence by identifying directed binary relations between words, known as dependencies. The concept of dependency is based on the idea that words in a sentence follow a hierarchical structure, establishing relations between headwords and dependent words. To give an example of dependencies, let's take the sentence example "*I ate pizza with Giuseppe yesterday*". Any dependency parser starts from the assumption that the finite verb used in each clause does not depend on any other word and is called the root of the clause. Then, a dependency parser would identify the word *ate* as the root of the sentence and the words *I*, *pizza*, and *yesterday* as its dependencies. In addition, it would identify two additional dependencies, one where the headword is *pizza* and *with* is the dependent word, and another one where the headword is *with* and *Giuseppe* is the dependent word. Intuitively, the dependency parser begins with the smallest independent sentence and expands it by sequentially adding words based on their importance to the sentence.

More formally, a dependency parser represents a sentence as a tree with a set of connected nodes corresponding to individual words. This tree is built such that each node has links (dependencies) through which it is connected to its child nodes (dependents), but is connected to only one parent node (head), except for the root node, which is connected to no parent node. Each node has exactly one path connected to the root node. In addition, each dependency comes with a label that defines the dependent's role towards its head. For instance, continuing with the previous sentence example, *I*, *pizza*, and *yesterday* are dependents of the word *eat*, and their dependency labels identify respectively as the subject, the object, and the adverbial modifier of *eat*. In addition, *with* is a dependent of *pizza*. Its dependency label defines *with* as the preposition of *pizza*, whereas *Giuseppe* is a dependent of *with*, and its dependency label defines *Giuseppe* as the object of the preposition *with*.

A natural question at this point is how a dependency parser works, namely, which actions it takes and with what objective.

Concerning the actions taken, a dependency parse does a series of bottom-up steps to connect each word with its head. In particular, it maintains two data structures: a buffer for the terms to be processed and a

stack for the currently processed terms. It takes two types of actions: shifting a word from the buffer to the stack and adding a left or right arc between the top two items on the stack, after which the dependent is "popped" from the stack. Importantly, no arc is formed between a head and a dependent until the dependent has been linked to all its dependents via left or right arcs, and these dependents have been popped from the stack. To illustrate how this works, let's take our sentence example and apply the set of operations just described:

1. Shift *I* to the stack.
2. Shift *eat* to the stack.
3. Add left arc from *eat* to *I* and pop *I*.
4. Shift *pizza* to the stack.
5. Shift *with* to the stack.
6. Shift *Giuseppe* to the stack.
7. Add right arc from *with* to *Giuseppe* and pop *Giuseppe*.
8. Add right arc from *pizza* to *with* and pop *with*.
9. Add right arc from *eat* to *pizza* and pop *pizza*.
10. Shift *yesterday* to the stack.
11. Add right arc from *eat* to *yesterday* and pop *yesterday*.
12. Pop *eat*.

As the process illustrates, *eat* could have connected to *pizza* in the fifth step, but *pizza* was the head of other dependencies, so other words were shifted to the stack.

Regarding the objective of the dependency parser, it is trained on a treebank, a corpus manually annotated with labeled dependencies. In particular, a dependency parser is typically trained to maximize the number of correctly identified dependencies. Concerning the dependency parser from spaCy, the training and testing data used come from the fifth release of OntoNotes³⁰, which is a large annotated corpus of various genres of text (news, conversational telephone speech, weblogs, Usenet newsgroups, broadcast, talk shows).

³⁰<https://catalog.ldc.upenn.edu/LDC2013T19>

C CE Evaluation

The set of inflation sentences to be tested is constructed by selecting all those containing any of the causal verbs or links listed in tables A1 and A2 in Appendix A and randomly drawing:

- For each causal verb:
 - 20 causal inflation sentences from those where the CE algorithm finds a causal relationship involving the chosen causal verb.
 - 20 inflation sentences from those where the chosen causal verb appears, but the CE algorithm finds no causal relationship with an inflation expression as the effect.
- For each causal link
 - 20 causal inflation sentences from those where the CE algorithm finds a causal relationship involving the chosen causal link.
 - 20 inflation sentences from those where the chosen causal link appears, but the CE algorithm finds no causal relationship with an inflation expression as the effect.

Additionally, 20 causal inflation sentences are drawn from those where the CE algorithm finds a causal relationship involving a conditional, and 20 more are drawn from those where a conditional appears, but the CE algorithm finds no causal relationship with an inflation expression as the effect. Similarly, for each verb used, 20 causal inflation sentences are drawn from those where the CE algorithm finds a causal relationship involving a resultative construction, and 20 more are drawn from those where a resultative construction appears, but the CE algorithm finds no causal relationship with an inflation expression as the effect.

The evaluation of each sentence is conducted separately for each type of causal relation found. I describe the steps to be followed when finding any of the types of causal relations in an inflation sentence:

- Causal links:
 1. Find the subordinate clause starting with the causal link.
 2. If this clause mentions (does not mention) the effect based on the causal link's prescribed cause-effect order, as from column 'When' in table A2 in Appendix A, check whether this clause mentions (does not mention) the inflation expression.

3. If so, find the main clause from which the previous subordinate clause depends.
 4. If the main clause does not mention (mentions) the inflation expression, evaluate whether there is a causal relationship based on the causal link and with the inflation expression as the effect.
- Causal verbs:
 1. Find the subject and object of the causal verb.
 2. Check whether the causal verb is used in active or passive form.
 3. If so, check whether the inflation expression is the subject or object of the causal verb.
 4. If so, based on whether the verb is used in the active or passive form and the prescribed cause-effect order from column 'When' in Tables A1 and A2 in Appendix A, check that the inflation expression appears in the position of the effect.
 5. If so, evaluate whether a causal relationship is based on the causal verb and with the inflation expression as the effect.
 - Conditionals:
 1. Find the subordinate clause starting with *if*.
 2. Check whether this clause does not mention the inflation expression. If so, find the main clause from which the previous subordinate clause depends. If the main clause mentions the inflation expression, evaluate whether there is a causal relationship based on the conditional and with the inflation expression as the effect.
 - Resultative construction:
 1. Check whether the verb used in the resultative construction is in active form.
 2. Find the subject and object of the verb.
 3. If so, check whether the inflation expression is the verb's object.
 4. If so, evaluate whether there is a causal relationship based on the verb and with the inflation expression as the effect.

Once the inflation sentences are annotated, I compare the manual annotation with the results from the CE algorithm and classify them as:

- True positives for all causal inflation sentences with a manually annotated inflation narrative whose text overlaps with the inflation narrative found by the CE algorithm.
- True negatives for all inflation sentences where neither the manual annotator nor the CE algorithm can find an inflation narrative.
- False positives for all causal inflation sentences where the CE algorithm finds an inflation narrative, but either no manually annotated inflation narrative is found, or the text of the manually annotated inflation narrative does not overlap with that of the inflation narrative found by the CE algorithm.
- False negatives for all inflation sentences with a manually annotated inflation narrative, but no inflation narrative found by the CE algorithm.

Finally, I evaluate the performance of the CE algorithm in terms of both accuracy and F-score, which are the most popular adopted metrics in CE and are computed as:

$$ACC = \frac{TP + TN}{TP + TN + FN + FP} \quad (19)$$

$$F1 = \frac{2 * TP}{2 * TP + FN + FP} \quad (20)$$

I also compute three more evaluation metrics, namely true positive rate, true negative rate, and positive predicted value, as follows:

$$TPR = \frac{TP}{TP + FN} \quad (21)$$

$$TNR = \frac{TN}{TN + FP} \quad (22)$$

$$PPV = \frac{TP}{TP + FP} \quad (23)$$

D Demand and Supply Narrative Types

Table A4: Narrative disagreement and expectation gap across narrative types

(a) Aggregate					
GAP_t					
$ MonPol_{t-1}^{G-S} $	0.589** (0.241)				0.598** (0.233)
$ SpendDefDebt_{t-1}^{G-S} $		0.210 (0.212)			0.140 (0.220)
$ ComEne_{t-1}^{G-S} $			0.312 (0.342)		0.347 (0.337)
$ Labor_{t-1}^{G-S} $				-0.597*** (0.227)	-0.652*** (0.240)
Heard of inflation news control	Yes	Yes	Yes	Yes	Yes
Past inflation control	Yes	Yes	Yes	Yes	Yes
Past inflation volatility control	Yes	Yes	Yes	Yes	Yes
Adj-R2	40.16	39.26	39.28	39.95	40.97
N	384	384	384	384	384

(b) Individual					
$GAP_{i,t}$					
$ MonPol_{t-1}^{G-S} $	0.591*** (0.048)				0.573*** (0.048)
$ SpendDefDebt_{t-1}^{G-S} $		0.276*** (0.042)			0.209*** (0.042)
$ ComEne_{t-1}^{G-S} $			0.252*** (0.060)		0.237*** (0.060)
$ Labor_{t-1}^{G-S} $				-0.510*** (0.049)	-0.582*** (0.049)
Demographic controls	Yes	Yes	Yes	Yes	Yes
Heard of inflation news control	Yes	Yes	Yes	Yes	Yes
Past inflation control	Yes	Yes	Yes	Yes	Yes
Past inflation volatility control	Yes	Yes	Yes	Yes	Yes
Adj-R2	4.06	3.99	3.97	4.02	4.16
N	162453	162453	162453	162453	162453

Note: These tables show the slope estimates obtained from estimating the models at equations 7 and 8, as well as modified versions of them where the narrative factor disagreement measures appear one by one. $|MonPol_{t-1}^{G-S}|$ measures disagreement between general and specialized newspapers on Monetary Policy as an inflation narrative, $|SpendDefDebt_{t-1}^{G-S}|$ measures disagreement between general and specialized newspapers on Spending/Deficit/Debt as an inflation narrative, $|ComEne_{t-1}^{G-S}|$ measures disagreement between general and specialized newspapers on Commodities/Energy as an inflation narrative, and $|Labor_{t-1}^{G-S}|$ measures disagreement between general and specialized newspapers on Labor as an inflation narrative. These individual demand-supply narrative variables are measured as described in section 4.3. In the bottom panel, standard errors are clustered at the individual level, as some respondents in the MSC are reinterviewed. The sample includes all months between 1991 and 2022.

E Asymmetric Narrative Disagreement

Table A5: Asymmetric narrative disagreement and expectation gap

	GAP_t	$GAP_{i,t}$
$ NetDemand_{t-1}^{G-S} * 1_{NetDemand_{t-1}^{G-S} > 0}$	0.688** (0.292)	0.057 (0.052)
$ NetDemand_{t-1}^{G-S} * 1_{NetDemand_{t-1}^{G-S} < 0}$	0.225 (0.212)	0.168*** (0.041)
Demographic controls	-	Yes
Heard of inflation news control	Yes	Yes
Past inflation control	Yes	Yes
Past inflation volatility control	Yes	Yes
Adj-R2	40.13	3.97
N	384	162453

F Between-General-Newspaper Heterogeneity

Table A6: Narrative disagreement and expectation gap across general newspapers

	GAP_t				$GAP_{i,t}$			
$ NetDemand_{t-1}^{NYT-S} $	0.297 (0.265)			-0.103 (0.275)	0.236*** (0.042)			0.041 (0.048)
$ NetDemand_{t-1}^{USAT-S} $		0.580*** (0.185)				0.395*** (0.038)		0.312*** (0.041)
$ NetDemand_{t-1}^{WaPo-S} $			0.778*** (0.229)	0.693*** (0.238)			0.364*** (0.041)	0.231*** (0.048)
Demographic controls	Yes	Yes	Yes	Yes	-	-	-	-
Heard of inflation news control	Yes	Yes	Yes	Yes	-	-	-	-
Past inflation control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Past inflation volatility control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj-R2	39.43	40.61	41.49	41.86	3.98	4.03	4.01	4.05
N	384	384	384	384	162453	162453	162453	162453

G Incentives to Gather Information about Inflation

Table A7: Narrative disagreement and expectation gap — inflation level

	GAP_t	$GAP_{i,t}$
$ NetDemand_{t-1}^{G-S} $	0.634* (0.336)	0.002 (0.064)
$ NetDemand_{t-1}^{G-S} * \pi_{t-1}$	-0.090 (0.102)	0.045** (0.019)
Demographic controls	-	Yes
Heard of inflation news control	Yes	Yes
Past inflation control	Yes	Yes
Past inflation volatility control	Yes	Yes
Adj-R2	39.83	3.97
N	384	162453

Table A8: Narrative disagreement and expectation gap — inflation persistence persistence

	GAP_t	$GAP_{i,t}$
$ NetDemand_{t-1}^{G-S} $	0.351* (0.197)	0.062* (0.037)
$ NetDemand_{t-1}^{G-S} * \rho_{1,t-1}(\pi)$	0.121 (0.178)	0.292*** (0.037)
Demographic controls	-	Yes
Heard of inflation news control	Yes	Yes
Past inflation control	Yes	Yes
Past inflation volatility control	Yes	Yes
Adj-R2	39.75	4.08
N	384	162453

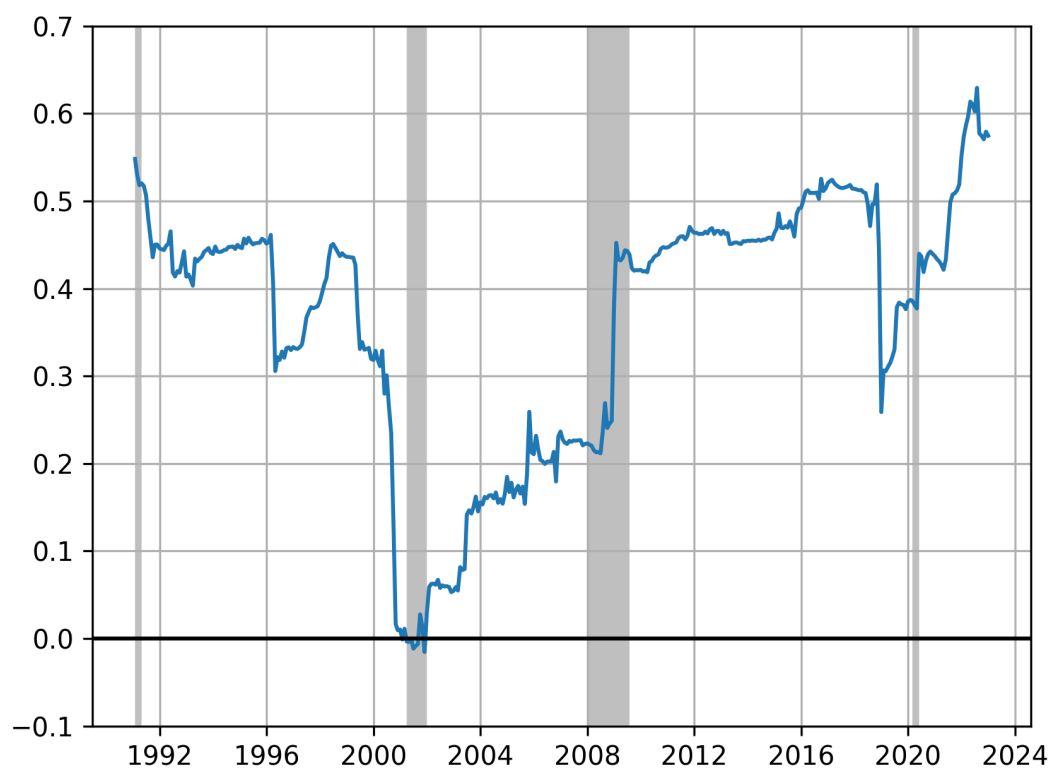


Figure A9: Inflation persistence

H Expert Disagreement

Table A9: Narrative and expert disagreement

	$var(\pi_{i,t,t+4}^{SPF})$
$ NetDemand_{t-1}^{G-S} $	0.270** (0.120)
Past inflation control	Yes
Past inflation volatility control	Yes
Adj-R2	40.63
N	127

Table A10: Narrative disagreement and expectation gap — expert disagreement

	GAP_t	$GAP_{i,t}$
$ NetDemand_{t-1}^{G-S} $	1.204* (0.660)	-0.073 (0.117)
$ NetDemand_{t-1}^{G-S} * var(\pi_{i,t,t+12}^{SPF})$	-1.152 (0.918)	0.080 (0.151)
Demographic controls	-	Yes
Heard of inflation news control	Yes	Yes
Past inflation control	Yes	Yes
Past inflation volatility control	Yes	Yes
Adj-R2	43.81	4.59
N	384	162453

I Forecast Errors

Table A11: Narrative disagreement and forecast errors

	FE_t		$FE_{i,t}$	
$ NetDemand_{t-1}^{G-S} $	-0.148 (0.404)	-0.205 (0.387)	0.154** (0.038)	-0.060 (0.038)
Demographic controls	Yes	Yes	-	-
Heard of inflation news control	Yes	Yes	-	-
Past inflation control	Yes	Yes	Yes	Yes
Past inflation volatility control	Yes	Yes	Yes	Yes
Adj-R2	12.89	15.47	2.43	3.36
N	384	384	162453	162453

J Macroeconomic Dynamics

Table A12: Narratives and macroeconomic dynamics

	$Demand_t^G$	$Supply_t^G$	$NetDemand_t^G$	$Demand_t^S$	$Supply_t^S$	$NetDemand_t^S$
π_t	0.763*** (0.104)	0.914*** (0.143)	-0.063*** (0.019)	0.933*** (0.180)	0.452* (0.231)	0.015 (0.013)
u_t	0.295*** (0.047)	0.185*** (0.061)	0.005 (0.009)	0.609*** (0.089)	-0.080 (0.109)	0.031*** (0.007)
$\pi_t * u_t$	-0.114*** (0.018)	-0.118*** (0.024)	0.006* (0.003)	-0.181*** (0.032)	-0.043 (0.041)	-0.005** (0.002)
Lag Dep. Var.	Yes	Yes	Yes	Yes	Yes	Yes
Adj-R2	20.69	39.60	12.22	19.04	20.36	21.07
N	383	383	383	383	383	383

Acknowledgements

I am deeply grateful to my supervisors, Lieven Baele and Frank de Jong, for all their support during this project. As this paper is the third chapter of my Ph.D. thesis, I want to thank my committee members, Stephan Hollander, Stefano Cassella, Marie Hoerova, and Christian Wagner, for taking the time to carefully read this paper and provide valuable feedback. I thank Anna Mogilevskaja, Mariachiara Tedde, Katerina Nikalexi, Philip Schnorpfeil, and Peter Schotman for their valuable discussions, and Constantin Charles, Evgenia Passari, and Frank Smets for their valuable suggestions. I also thank conference participants at the ifo Dresden Workshop on Macroeconomics and International Finance 2024, the 18th Belgian Financial Research Forum, the 7th Dauphine Finance PhD Workshop, the 30th Annual Meeting of the German Finance Association, the NETSPAR Pension Day 2024, the CEPR Paris Symposium 2024, and the 2025 ECB Forum on Central Banking, and seminar participants at Tilburg University, the European Central Bank, and Ghent University for their helpful feedback. A previous draft of this paper circulated under the title *Does Demand-Supply Narrative Disagreement Help Explain Households' Inflation Expectation Gap?*

Giovanni Trebbi

European Central Bank, Frankfurt am Main, Germany; email: giovanni.trebbi@ecb.europa.eu

© European Central Bank, 2025

Postal address 60640 Frankfurt am Main, Germany

Telephone +49 69 1344 0

Website www.ecb.europa.eu

All rights reserved. Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the ECB or the authors.

This paper can be downloaded without charge from www.ecb.europa.eu, from the [Social Science Research Network](#) electronic library or from [RePEc: Research Papers in Economics](#). Information on all of the papers published in the ECB Working Paper Series can be found on the [ECB's website](#).

PDF

ISBN 978-92-899-7534-6

ISSN 1725-2806

doi:10.2866/3606911

QB-01-25-280-EN-N