Inflation Narratives and Expectations

Giovanni Trebbi*

Click here for latest version

Abstract

I study how disagreement in demand and supply narratives between newspaper articles read by households and professionals 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. 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. General newspapers' demand and supply narratives correctly predict only households' demand-supply expectations, while specialized newspapers' narratives predict the demand-supply expectations of both experts and households. I then use general and specialized newspapers' demand and supply narratives to measure narrative disagreement. Households' absolute inflation expectation gap widens when narrative disagreement increases, especially for non-college-educated and older households.

*I am deeply grateful to my supervisors, Lieven Baele and Frank de Jong, for all their support during this project. I thank Anna Mogilevskaja (discussant), Constantin Charles, Christian Wagner, Mariachiara Tedde (discussant), Katerina Nikalexi (discussant), Evgenia Passari, Philip Schnorpfeil (discussant), Peter Schotman (discussant), and Frank Smets, for their valuable discussion and 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, and the CEPR Paris Symposium 2024, and seminar participants at Tilburg University, the European Central Bank, and Ghent University for their helpful feedback. The author is from Tilburg University. A previous draft of this paper circulated under the title *Does Demand-Supply Narrative Disagreement Help Explain Households' Inflation Expectation Gap?*. Contact: g.trebbi@tilburguniversity.edu

Narratives are accounts of the sources of events (Akerlof and Snower, 2016); being at their core, causality explains why narratives affect people's beliefs (Eliaz and Spiegler, 2020), potentially in ways inconsistent with rational theory (Charles and Kendall, 2023). Andre, Haaland, Roth, and Wohlfart (2023) provide the first evidence of narrative disagreement between households and academic experts. The authors find that experts' narratives of the surges in inflation in late 2021 and 2022 are more balanced between demand (e.g., "increases in government spending") and supply (e.g., "shortage of workers") triggers that the narratives of households. Moreover, they show that heterogeneity in households' narratives accounts for a substantial share of the total variation in inflation expectations. However, the relationship between narrative disagreement and expectation differences is left unexplored.

The rich literature on households' inflation expectations[1] shows us that, over the past thirty years, households' inflation expectation gap with experts has been, on average, positive and has varied considerably over time (see figure 1[2]). To explain this time-series variation, Carroll (2003) develops a sticky information model à la Mankiw and Reis (2002) in which households learn about inflation from newspapers, which transmit experts' views. His model implies a negative relation between the absolute expectation gap and the intensity of inflation press coverage. He finds evidence supporting his prediction,[3] whereas Pfajfar and Santoro (2013) find evidence that the relation is positive,[4] which I confirm using data up to 2022 (see table 5). The authors conclude that newspapers might not transmit experts' views to households as assumed. This might be the case if households mostly learn about inflation from general newspapers that reflect experts' opinions less accurately than specialized newspapers do, as suggested by evidence from Mazumder (2021). Therefore, narrative disagreement between general and specialized newspapers might capture the narratives on which households and experts disagree (Andre et al., 2023). My research question is whether the absolute expectation gap widens with narrative disagreement between general and specialized newspapers.

[1]For a detailed review of this literature, see Weber, D'Acunto, Gorodnichenko, and Coibion (2022).

[2]Based on survey data from the Michigan Survey of Consumers (MSC) and the SPF to measure households' and experts' expectations.

[3]Using inflation articles from the New York Times and the Washington Post.

[4]Contrary to the evidence by Carroll (2003), Pfajfar and Santoro (2013) study both mean and individual household inflation expectations and finds a positive relationship in both cases.

To answer this question, I collect more than 180,000 newspaper articles that talk about inflation via explicit inflation expressions [5] and are published between 1991 and 2022 by three general newspapers (New York Times, USA Today, and Washington Post)[6] and one specialized newspaper (Wall Street Journal).[7] This general-specialized newspaper separation follows from statistics on newspaper readership (Pew Research Center, 2012) and recent research on how the absolute expectation gap is differently related to news reporting about the Fed by general and specialized newspapers I consider (Mazumder, 2021).

I measure newspapers' demand and supply narratives and their disagreement in three steps. First, I use a Causality Extraction (CE) method based on Baele, De Jong, and Trebbi (2023) to identify inflation narratives as described in Andre et al. (2023). In a nutshell, the CE method finds within a sentence causal relations expressed via explicit causal keywords (e.g., *because*, *trigger*), checks that inflation expressions are mentioned as the effects, and extracts inflation narratives as the corresponding causes. Second, I use a dictionary method to classify inflation narratives into the demand and supply narrative categories specified in Andre et al. (2023). My general and specialized measures of demand and supply narratives, $NetDemand^{G}$ and $NetDemand^{S}$, are the scaled monthly differences of articles with demand or supply narratives from general and specialized newspapers, respectively. Finally, narrative disagreement $NetDemand^{G-S}$ is the difference in $NetDemand$ between general and specialized newspapers.

Before relating narrative disagreement to the absolute expectation gap, I test if different newspapers' demand and supply narratives correctly capture the demand and supply views of households and experts differently. I measure households' and experts' demand and supply views with the product in their expectations of future changes in inflation and unemployment. This product is positive if supply views dominate and negative if demand views dominate. Any $NetDemand$ measure takes positive values when the respective newspaper publishes relatively more demand than supply narratives and vice versa for negative values. Therefore, there should be a negative relation between

[5]They need to mention the inflation expressions "inflation," "cpi," "consumer price," "ppi," or "producer price."

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

[7]These 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. (2023).

the inflation-unemployment expectation product and $NetDemand$. The results show that general newspapers' demand and supply narratives align correctly only with households' demand and supply views. In contrast, specialized newspapers' demand and supply narratives align correctly with households' and experts' demand and supply views. This finding suggests that the narratives of general and specialized newspapers capture the expectations of different audiences. Interestingly, general newspapers' demand and supply narratives also incorrectly align with the joint dynamics of realized inflation and unemployment, as their $NetDemand$ measure increases when inflation and unemployment are jointly rising. In contrast, specialized newspapers' demand and supply narratives correctly align with the joint dynamics of realized inflation and unemployment. This result suggests that general newspapers might communicate incorrect narratives to households.

My primary hypothesis is that the absolute expectation gap widens when narrative disagreement increases. My regressor of interest is the absolute value of narrative disagreement to abstract from whether general newspapers publish more demand or supply narratives than specialized newspapers. I explore the relevance of the sign of narrative disagreement in later analysis. The results confirm my primary hypothesis by showing that the absolute expectation gap positively relates to narrative disagreement between newspapers. This finding holds both at the aggregate and individual levels and when controlling for the level and volatility of inflation, as well as for households' demographics and perceptions of news about inflation. In addition, it holds also after controlling for narrative disagreement about whether inflation is increasing or decreasing and whether articles talk about realized or future inflation episodes. Digging deeper, I assess how this positive association changes across household demographics. D'Acunto, Hoang, Paloviita, and Weber (2019) find education can explain cross-sectional differences in inflation expectations across households. Building on this finding, I show that the positive association between my newspaper narrative disagreement measure and the absolute expectation gap is stronger for individuals without a college degree. This result is perhaps to be expected, as the news readership of college-educated households is closer to that of experts (Pew Research Center, 2012). On the other hand, this positive association is at its highest when narrative disagreement revolves around the importance of monetary policy narratives.

I confirm the robustness of my findings with respect to the individual types of demand and supply narratives on which newspapers disagree, the sign of narrative disagreement and the level of inflation, as well as across different general newspapers and when introducing disagreement among experts, and verify if the positive association between absolute expectation gap and narrative disagreement translates into a positive association between households' forecast errors and narrative disagreement between newspapers. First, narrative disagreement about most individual types of demand and supply narratives is positively related to the absolute expectation gap. The positive association is at its highest when narrative disagreement revolves around the importance of monetary policy narratives. Second, the positive association between narrative disagreement and absolute expectation gap remains positive irrespective of the sign of narrative disagreement, though the sign can explain differences in the magnitude of this association, and become stronger as inflation increases. Third, there is also a positive association between narrative disagreement and absolute expectation gap when individual general newspapers are used to measure narrative disagreement. Fourth, expert disagreement is positively related to narrative disagreement, but narrative disagreement remains positively related to the aggregate absolute expectation gap even after controlling for expert disagreement. Finally, households' forecast errors also become narrower when narrative disagreement between newspapers declines, though only when not controlling for macroeconomic variables.

These findings have exciting implications for policymakers because increased reporting about inflation might bridge the gap between households' and experts' expectations only when there is low narrative disagreement in the media landscape. In particular, these findings suggest that if central bank communication is aimed at lowering the dispersion of inflation forecasts among different groups of individuals, it needs to direct its inflation narratives across a broad range of channels.

This paper builds on and adds to four closely related studies. First, Andre, Pizzinelli, Roth, and Wohlfart (2022) show that the expectation gap relates to narrative disagreement between households and experts. By using time-series data, I confirm that their experimental evidence extends to periods when differences in narratives and inflation expectations are less prominent (Weber et al., 2023).

Second, Larsen, Thorsrud, and Zhulanova (2021) study which news topics best predict households' inflation expectations. My contribution is to relate inflation expectations to inflation narratives, which are captured by the CE method I employ and not by news topics. Third, Dräger, Lamla, and Pfajfar (2016) study how news about inflation, unemployment, or monetary policy help households make predictions more in line with the Phillips curve. My narrative measures differ from theirs because they causally relate inflation to unemployment, monetary policy, and other factors. Doing so, this study shows that households make predictions more aligned with the Phillips curve when the volume of demand narratives increases. Fourth, I add to Mazumder (2021) by showing that the choice of newspapers matters also for the relation between narratives and expectations.

1 Data

1.1 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 et al. (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 200 words or less (3.87% of the corpus), as dictionary-based methods are typically noisy for brief texts (Shapiro, Sudhof, and 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 circulation[8]. WSJ, NYT, and WaPo also appear among the top 20 inflation news outlets consulted in the survey by Andre et al. (2023). On the other hand, USAT is the news source from which Coibion, Gorodnichenko, and Weber (2022) sample articles to measure press coverage of central

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

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"*[9] and its articles have a long history of applications in the Finance literature to measure investor beliefs, namely sentiment.[10] 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 to 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 4.3.5).

Table 1 offers some first 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. [11] The Jaccard index, which expresses how often 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

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

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

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

volume of inflation articles scaled by its maximum in any month. I do so separately for general and specialized newspapers and obtain the measures $News^G$ and $News^S$. Press coverage about inflation is at its highest for general and specialized newspapers in the second half of 2022, particularly in November for general newspapers and July for specialized newspapers. As a reference, June 2022 is the month with the highest annual CPI inflation rate in my sample. However, the inflation press coverage of specialized newspapers is more volatile and goes above 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 these three 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 spaCy[12]. In doing so, I obtain 793,333 sentences that contain at least one inflation expression (inflation sentences). I call these sentences "inflation sentences".

1.2 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.[13] Throughout the paper, I use both mean and individual

[12] See `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.

[13] Adopting 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.

households' 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 States.[14]

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[15] and asks participants for quarter-by-quarter forecasts that span the current and next five quarters. Throughout the analysis, I employ mean and individual-level nowcasts and one-year-ahead forecasts of CPI inflation.

1.3 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 et al. 2021a). 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 give two types of news they have heard about, among them being either higher or lower prices. Therefore, answers to this question allow me to construct two variables, $News^P_t$ and $News^P_{i,t}$. $News^P_t$ is the percentage of MSC respondents who report having heard of recent price changes, while $News^P_{i,t}$ is a dummy indicating if the MSC i-th respondent reports having

[14]Household 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."

[15]To obtain a monthly estimate of the SPF, I follow Ehrmann, Pfajfar, and Santoro (2018) and linearly interpolate the data. Replacing missing monthly values using the last available forecast leaves the results virtually unchanged.

heard of recent price changes.

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

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.

The key prediction of the model designed by Carroll (2003) is that the absolute expectation gap becomes narrower when newspapers publish more articles about inflation. The author finds evidence supporting his prediction, whereas Pfajfar and Santoro (2013) find evidence showing the contrary. An essential difference between the two studies is the sample used, as Carroll (2003) uses data between 1981 and 2000, while Pfajfar and Santoro (2013) use data until 2011. Differences in the sample used and, correspondingly, in the results obtained call for a test of the relationship posited by Carroll (2003) using a more extended sample. The presence of two groups of newspapers (general and specialized) further allows me to test if the results depend on who reports about inflation. Therefore, my first hypothesis is:

> The absolute expectation gap narrows when inflation press coverage by general or specialized newspapers becomes more intense. (H1)

A direct but untested implication of the prediction by Carroll (2003) is that households become more informed about inflation when newspapers publish more articles about inflation. While newspapers are one source from which households might learn about inflation, there might be other sources (e.g., grocery prices, D'Acunto et al. 2021a). In this respect, there might be no relationship between the absolute expectation gap and inflation press coverage if households do not read newspapers. The presence of two groups of newspapers (general and specialized) further allows me to test if the results depend on who reports about inflation. Therefore, my second hypothesis is:

> Households are more likely to pay attention to news about inflation when inflation press coverage by general or specialized newspapers becomes more intense. (H2)

Andre et al. (2023) show the inflation narratives of households and experts are different. In particular, households focus more on supply narratives relative to experts. The authors further show that households with different narratives also have different inflation expectations. This and the evidence by Andre et al. (2022) suggest that the inflation expectations of households and expert differ when their inflation narratives differ. If the narratives of general and specialized newspapers capture the expectations of households and experts differently, then their narrative disagreement might be related to the absolute expectation gap in two ways. First, the systematic component of the absolute expectation gap might be related to the systematic narrative disagreement between newspapers. Second, the absolute expectation gap might widen when narrative disagreement increases. Narrative disagreement is observed when general newspapers publish relatively more demand or supply narratives than specialized newspapers. For instance, there might be disagreement when specialized newspapers attribute a rise in inflation to looser fiscal policy (a demand narrative) and general newspapers attribute it to a surge in energy costs (a supply narrative). Therefore, my third and fourth hypotheses are:

> There is systematic narrative disagreement, i.e., general newspapers publish relatively more demand or supply narratives than specialized newspapers. (H3)

The absolute expectation gap widens when general newspapers publish relatively more demand or supply narratives than specialized newspapers, i.e. when narrative disagreement increases. (H4)

Figure 3 shows the cross-sectional interquartile ranges of MSC and SPF inflation expectations and indicates a large dispersion across households, far larger than across experts. Existing evidence suggests some individual characteristics help explain this cross-sectional dispersion, namely sex (D'Acunto, Malmendier, and Weber, 2021b), cognitive abilities (D'Acunto, Hoang, Paloviita, and Weber, 2019), socioeconomic status (Bruine de Bruin, Vanderklaauw, Downs, Fischhoff, Topa, and Armantier, 2010), and age (Bryan and Venkatu, 2001). In particular, inflation expectations are higher for women than men and decrease with income, education, and age. On the one hand, highly educated and rich households are more likely to be readers of specialized newspapers (Pew Research Center, 2012), so the relationship between narrative disagreement and their absolute expectation gap should be weaker. On the other hand, older individuals pay more attention to newspapers (Pew Research Center, 2023), so the relationship between narrative disagreement and their absolute expectation gap should be stronger. Concerning sex, there is no newspaper consumption is similar across men and women (Pew Research Center, 2023), so the relationship between narrative disagreement and their absolute expectation gap should not change based on sex. Therefore, my fifth hypothesis is:

Household characteristics moderate the correlation between narrative disagreement and the absolute expectation gap. In particular, the correlation increases with age and decreases with income and education, while it does not change with sex. (H5)

The fourth hypothesis argues that the absolute expectation gap widens with increasing disagreement between newspapers over their demand and supply narratives. The reason for this positive relationship might be that the demand and supply narratives capture differences in demand and supply views between households and experts. Supply narratives describe events that move inflation and unemployment in the same direction, while the opposite holds for demand narratives. For instance, "increasing energy prices" (a supply narrative) is expected to raise both inflation and unemployment. In contrast "looser monetary policy" (a demand narrative) is expected to increase inflation and lower unemployment. Therefore, if newspapers' narratives capture individuals' expectations about future macroeconomic outcomes, their demand and supply narratives might capture how individuals expect inflation and unemployment to move. For instance, individuals might expect inflation and unemployment to move in opposite directions to a larger extent when newspapers publish relatively more demand narratives and vice versa when newspapers publish relatively more supply narratives. Crucially, to help explain why the absolute expectation gap widens with narrative disagreement between newspapers, the relationship between newspapers' demand-supply narratives and individuals' demand-supply views should change across newspapers and types of individuals. For instance, the expectations of experts might align more closely with the narratives of general newspapers than of specialized newspapers. Therefore, my sixth and seventh hypotheses are:

> Households and experts expect that inflation and unemployment will move in opposite directions by a larger degree when newspapers publish relatively more demand narratives. The opposite holds when they publish relatively more supply narratives. (H6)

> Households' expectations are more strongly correlated with the narratives of general newspapers than specialized newspapers, while the opposite holds for experts' expectations. (H7)

3 Methodology

Section 3.1 discusses the causality extraction method used in this paper. Section 3.2 explains how inflation narratives are extracted from articles. Section 3.3 describes how inflation narratives are categorized into demand and supply categories. Section 3.4 illustrates how to distinguish hawkish narratives from dovish ones, whereas section 3.5 explains how I discriminate observed narratives from expected ones.

3.1 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 they can capture with their method. 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."*); [16]
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 causal relation type. 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** are identified from the appearance of the grammatical pattern subject-verb-object-adjective in which the verb is in active form[17]. 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

[16]Resultative 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.

[17]The authors do not impose any causality requirement on the verbs used in this type of constructions.

verbs used in the causal sentences listed in the Penn Discourse Treebank (PDTB) dataset (Prasad, Dinesh, Lee, Miltsakaki, Robaldo, Joshi, and Webber, 2008) and the causal verbs identified by Girju (2003).

Finally, they define the cause-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.

3.2 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.

3.3 Demand and Supply Narratives

The output of the previous section is a long, unstructured list of inflation narratives. I classify them into demand and supply narratives using a dictionary-based method. I first define what types of narratives can be classified as either demand or supply narratives. Then, I describe the dictionary-based method used for this classification.

I follow Andre et al. (2023) in their definition of which narratives can be described as either demand or supply narratives. The authors' survey elicits a group of households' and experts' perceived drivers (narratives) of the surge in U.S. inflation in late 2021 and 2022. 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.

My classification assigns each narrative to a category whenever words from the narrative's text appear in the dictionary associated with that specific category. I borrow dictionaries compiled by previous studies for most categories. On the demand side, I use the dictionaries "Spending/Deficit/Debt", "Monetary Policy", and "Consumer Spending and Sentiment" from Baker et al. (2021). On the supply side, I borrow dictionaries from Baker et al. (2021), namely their dictionaries "Labor Markets", "Labor Disputes", and "Commodity Markets". However, the supply-chain dictionary includes all the top one-hundred supply-chain risk bigrams compiled by Ersahin et al. (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.

Using these dictionaries, I count all the demand and supply terms across all the inflation narratives mentioned for each article and take the difference in their counts. I classify an article as a demand/supply article if this difference is positive/negative. Instead, if an article contains narratives without either demand or supply terms, or their count difference is zero, it is classified as neither a demand nor a supply article. For an article j published in month t by newspaper n, this step yields two dummies, $Demand^n_{j,t}$ and $Supply^n_{j,t}$, taking value one if the article is classified as a demand and supply article, respectively. Naturally, $Demand^n_{j,t}$ and $Supply^n_{j,t}$ cannot both take value one.

Next, I compute monthly newspaper-specific demand and supply narrative indicators as follows:

$$Demand_t^n = \sum_{j=1}^{N_t^n} Demand_{j,t}^n$$

$$Supply_t^n = \sum_{j=1}^{N_t^n} Supply_{j,t}^n$$

Where N_t^n is the number of causal articles published by newspaper n in month t. Finally, I compute a newspaper-specific net demand indicator as follows: [18]

$$NetDemand_t^n = \frac{Demand_t^n - Supply_t^n}{\max_t |Demand_t^n - Supply_t^n|}$$

$NetDemand_t^n$ can take any value between -1 and 1. Positive values represent months when newspapers publish more demand than supply articles; the opposite holds for negative values. Finally, I compute demand-supply narrative disagreement between general and specialized newspapers as:

$$NetDemand_t^{G-S} = NetDemand_t^G - NetDemand_t^S$$

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

$$ConsSpendSent_t^{G-S} = \frac{\sum_{j=1}^{N_t^G} ConsSpendSent_{j,t}^G}{\max_t \sum_{j=1}^{N_t^G} ConsSpendSent_{j,t}^G} - \frac{\sum_{j=1}^{N_t^S} ConsSpendSent_{j,t}^S}{\max_t \sum_{j=1}^{N_t^S} ConsSpendSent_{j,t}^S}$$

$$MonPol_t^{G-S} = \frac{\sum_{j=1}^{N_t^G} MonPol_{j,t}^G}{\max_t \sum_{j=1}^{N_t^G} MonPol_{j,t}^G} - \frac{\sum_{j=1}^{N_t^S} MonPol_{j,t}^S}{\max_t \sum_{j=1}^{N_t^S} MonPol_{j,t}^S}$$

$$SpendDefDebt_t^{G-S} = \frac{\sum_{j=1}^{N_t^G} SpendDefDebt_{j,t}^G}{\max_t \sum_{j=1}^{N_t^G} SpendDefDebt_{j,t}^G} - \frac{\sum_{j=1}^{N_t^S} SpendDefDebt_{j,t}^S}{\max_t \sum_{j=1}^{N_t^S} SpendDefDebt_{j,t}^S}$$

[18] I scale by the maximum absolute difference rather than the sum of demand and supply articles (or its maximum) because doing so can assign a larger weight to months with only one article with a demand or supply narrative.

$$ComEne_t^{G-S} = \frac{\sum_{j=1}^{N_t^G} ComEne_{j,t}^G}{\max_t \sum_{j=1}^{N_t^G} ComEne_{j,t}^G} - \frac{\sum_{j=1}^{N_t^S} ComEne_{j,t}^S}{\max_t \sum_{j=1}^{N_t^S} ComEne_{j,t}^S}$$

$$Labor_t^{G-S} = \frac{\sum_{j=1}^{N_t^G} Labor_{j,t}^G}{\max_t \sum_{j=1}^{N_t^G} Labor_{j,t}^G} - \frac{\sum_{j=1}^{N_t^S} Labor_{j,t}^S}{\max_t \sum_{j=1}^{N_t^S} Labor_{j,t}^S}$$

$$SupplyChain_t^{G-S} = \frac{\sum_{j=1}^{N_t^G} SupplyChain_{j,t}^G}{\max_t \sum_{j=1}^{N_t^G} SupplyChain_{j,t}^G} - \frac{\sum_{j=1}^{N_t^S} SupplyChain_{j,t}^S}{\max_t \sum_{j=1}^{N_t^S} SupplyChain_{j,t}^S}$$

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.

3.4 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 initially developed by Apel and Grimaldi (2012) and extended by Apel, Blix Grimaldi, and Hull (2022) to discern whether narratives mention inflation as increasing or decreasing.

The dictionary presented in Apel et al. (2022) 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 from the difference between the counts of hawkish and dovish 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.

In contrast to the approach adopted by Apel et al. (2022), I use a dependency parser[19] to precisely identify when a modifier is used in conjunction with an inflation expression. A dependency parser analyzes the grammatical structure of a sentence.[20] Therefore, it can verify whether adjective modifiers refer to inflation. For instance, in the sentence *"Supply-chain bottlenecks might lead to high inflation and unemployment"*, a narrative is identified and marked as hawkish because the hawkish modifier "high" refers to inflation. In contrast, in the sentence *"Tighter monetary policy might lead to high inflation and low unemployment"*, a narrative is identified and marked as dovish because the dovish modifier "low" refers to inflation. However, the dictionary method by Apel et al. (2022) would mark it as neither hawkish nor dovish because the hawkish and dovish modifiers "high" and "low" cancel each other out. In addition, 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 increasing because of supply-chain bottlenecks"*, a narrative is identified and marked as hawkish because the hawkish modifier "increasing" has an inflation keyword as its subject. In contrast, 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 mark it as neither hawkish nor dovish because the hawkish and dovish modifiers "higher" and "decelerating" cancel each other out.

Via this approach, I count all the hawkish and dovish narratives mentioned for each article and take the difference in their counts. I classify an article as hawkish/dovish if this difference is positive/negative. Instead, if an article contains neither hawkish nor dovish narratives, or their count difference is zero, it is classified as neither hawkish nor dovish. For an article j published in month t by newspaper n, this step yields two dummies, $Hawkish_{j,t}^{n}$ and $Dovish_{j,t}^{n}$, taking value one if the article is hawkish and dovish, respectively. Naturally, $Hawkish_{j,t}^{n}$ and $Dovish_{j,t}^{n}$ cannot both take value one. Next, I compute monthly newspaper-specific hawkish and dovish narrative

[19]I use the Python implementation of the spaCy dependency parser.

[20]Appendix B provides a detailed description of dependency parsing.

indicators as:

$$Hawkish_t^n = \sum_{j=1}^{N_t^n} Hawkish_{j,t}^n$$

$$Dovish_t^n = \sum_{j=1}^{N_t^n} Dovish_{j,t}^n$$

Finally, I compute the net hawkish indicator at the level of newspaper n as:[21]

$$NetHawkish_t^n = \frac{Hawkish_t^n - Dovish_t^n}{\max_t |Hawkish_t^n - Dovish_t^n|}$$

Positive values represent months when there are more articles with hawkish narratives than dovish narratives and, hence, more narratives that describe inflation as increasing rather than decreasing. The opposite holds for negative values. Finally, I compute hawkish-dovish narrative disagreement between general and specialized newspapers as:

$$NetHawkish_t^{G-S} = NetHawkish_t^G - NetHawkish_t^S$$

3.5 Observed vs. Expected Inflation Narratives

An important question is whether inflation narratives discuss past/present inflation episodes rather than future/potential ones. To discriminate between these two cases, I follow Baele et al. (2023) in their formulation of three non-exclusive conditions under which a text-based causal relation is about future/expected events. Applied to my inflation narratives, they describe when an inflation narrative can be related to future/potential inflation episodes. I identify these inflation narratives as expected and all the others 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.

[21] I scale by the maximum absolute difference rather than by the sum of hawkish and dovish articles (or its maximum) because doing so can assign a larger weight to months with only one article with a hawkish or dovish narrative. $NetHawkish_t^n$ can take any value between -1 and 1.

Second, inflation narratives extracted from causal relations mentioning modal verbs[22] 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. In the case of causal verbs and resultative constructions, I check if a modal verb is used between the verb and its subject. In the case of causal links, I check if a modal verb is used between the subject and the verb of the inflation narrative or its effect.

Finally, inflation narratives extracted from causal relations that mention the verb *"to expect"* or any of its synonyms[23] are identified as expected. I identify expected inflation narratives from the verb *"to expect"* and its synonyms as I do with modal verbs.

Via this approach, I count all the observed and expected narratives mentioned for each article and take the difference in their counts. I classify an article as observed/expected if this difference is positive/negative. Instead, if an article contains neither observed nor expected narratives or their count difference is zero, it is classified as neither observed nor expected. For an article j published in month t by newspaper n, this step yields two dummies, $Observed^n_{j,t}$ and $Expected^n_{j,t}$, taking value one if the article is classified as observed and expected, respectively. Naturally, $Observed^n_{j,t}$ and $Expected^n_{j,t}$ cannot both take the value one. Next, I compute monthly newspaper-specific observed and expected narrative indicators as:

$$Observed^n_t = \sum_{j=1}^{N^n_t} Observed^n_{j,t}$$

$$Expected^n_t = \sum_{j=1}^{N^n_t} Expected^n_{j,t}$$

[22] The list of the modal verbs I use comes from here.

[23] Synonyms of the verb "to expect" are from the thesaurus by Merriam-Webster.

Finally, I compute the net hawkish indicator at the level of newspaper n as:[24]

$$NetObserved_t^n = \frac{Observed_t^n - Expected_t^n}{\max_t |Observed_t^n - Expected_t^n|}$$

Positive values represent months when there are more articles with observed narratives than expected narratives, and vice versa for negative values. Finally, I compute observed-expected narrative disagreement between general and specialized newspapers as:

$$NetObserved_t^{G-S} = NetObserved_t^G - NetObserved_t^S$$

4 Results

This section discusses the main empirical findings. Section 4.1 describes the results of the CE algorithm, section 4.2 the results from the classification of my inflation narratives, and section 4.3 the model used to test the hypotheses from section 2 and the test results. Finally, section 4.4 examines how the relationship between the absolute expectation gap and narrative disagreement differs based on the individual demand and supply narratives underlying narrative disagreement or the direction of narrative disagreement, whether the results differ when individual general newspapers are used, when introducing expert disagreement, whether narrative disagreement is also positively related to households' forecast errors, and how general and specialized newspapers' publication of demand and supply narratives aligns with realized inflation and unemployment.

4.1 Inflation Narratives

Applying the CE method outlined in section 3.2 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 about 77% of publication days of inflation articles (8,880 out of 11,555). Of

[24] I scale by the maximum absolute difference rather than by the sum of observed and expected articles (or its maximum) because doing can assign a larger weight to months with only one article with an observed or expected narrative. $NetObserved_t^n$ can take any value between -1 and 1.

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 method 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, the performance of their CE method might differ when extracting inflation narratives. This is why I repeat their validation exercise within this study. For each causal keyword, I inspect forty inflation sentences: twenty where an inflation narrative is extracted through that causal keyword and twenty where that causal keyword is found, but an inflation narrative cannot be extracted (e.g., an inflation expression can be the cause or can be used neither as the cause nor as the effect). For instance, I inspect twenty sentences with inflation narratives found via the causal keyword *"because"* and twenty sentences where *"because"* is mentioned, but no inflation narrative is found. Therefore, I exclude all causal keywords appearing in fewer than twenty causal inflation sentences. Finally, I manually annotate the causal relationships identified by each causal keyword and evaluate them, as 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 et al. (2022) show that studies using comparable CE methods generally achieve F-scores ranging between 54 and 71%. I use these manual annotations to select the causal keywords used to identify the inflation narratives I study in the rest of the analysis. Based on a minimum F-score of 54%, I retain all resultative constructions, four out of nine causal link keywords, and seventeen out of twenty-two causal verb keywords. As a consequence of this, the number of inflation narratives is drastically reduced to 5,204. These narratives appear in 4,896 causal inflation articles, 1,582 published by general newspapers, and 3,314 by specialized newspapers. Causal inflation articles from this set are published on about 27% of publication days of inflation articles (3,086 of 11,555).

Table 3 reproduces table 1 for the final set of causal inflation articles. Interestingly, specialized newspapers publish causal inflation articles almost twice as frequently and as much as general newspapers. However, causal inflation articles are similarly long and complex across general and specialized newspapers and are slightly longer than non-causal inflation articles. The Jaccard index indicates that general and specialized newspapers publish causal inflation articles on the same day only 13% of the time.

Finally, as done for inflation articles, I construct press coverage intensity measures for causal inflation articles and show them in figure 4. Causal press coverage about inflation is at its highest for general newspapers in June 2022 and for specialized newspapers in September 2019. The causal inflation press coverage of specialized newspapers is more volatile than that of general newspapers and reaches beyond half its maximum multiple times. Causal inflation press coverage is high in the same period of high inflation press coverage highlighted in section 1. Additional periods of heightened causal inflation press coverage by specialized newspapers are October 2005, June 2009, the months between the last quarter of 2010 and the first quarter of 2011, November and December 2016, the months between the last quarter of 2017 and the first quarter of 2018, and March and July 2022. Most of these periods coincide with spikes in the expectation gap, particularly in June 2009, the first quarter of 2011, and the second quarter of 2022.

4.2 Narrative Features and Disagreement

I now describe the results of classifying my inflation narratives along all the three narrative dimensions described in sections 3.3 to 3.5. Table 4 provides some summary statistics on $NetDemand$, $NetHawkish$, and $NetObserved$, as well as their disagreement measures. In addition, it shows 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 than general newspapers, which suggests narrative disagreement between them. However, a two-sided t-test reveals the $NetDemand^{G-S}$ is statistically different from zero only with a significance level of 10%. Therefore, there is only weak evidence of systematic disagreement about demand and supply narratives between general and specialized newspapers; hence, the third hypothesis cannot be accepted.

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 is at its highest in April-May 2006, March 2018, and February 2022. The top panel of figure 5 shows that, in April and May 2006, general newspapers publish relatively more supply narratives than specialized newspapers. Delving deeper into the types of narratives behind these two spikes, figure 8 shows specialized newspapers published relatively more commodities/energy narratives than specialized newspapers 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 than general newspapers 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. This is also showcased in the central panel of figure 5 and aligns with existing evidence that the media tends to report more on rising inflation (Lamla and Lein, 2014). Moreover, the median of $NetHawkish^{S}$ is positive while the median of $NetHawkish^{G}$ equals zero. This means that specialized newspapers publish relatively more hawkish narratives than general newspapers, suggesting another form of narrative disagreement between them. However, a two-sided t-test reveals the null of $NetHawkish^{G-S}$ 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 of realized inflation episodes than on 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 publish relatively more observed narratives than general newspapers, suggesting another form of narrative disagreement. However, a two-sided t-test reveals the null of $NetObserved^{G-S}$ cannot be rejected.

Overall, this section shows that general and specialized newspapers predominantly publish narratives that attribute realized inflationary episodes to supply factors. While general and specialized newspapers usually agree along their three aggregate narrative dimensions, there is evidence of disagreement about the importance of different types of narratives, namely monetary policy, government spending/deficit/debt, and labor narratives.

4.3 Narratives and Expectations

4.3.1 Inflation Press Coverage and Expectations

The first hypothesis from section 2 states that the absolute expectation gap becomes narrower when inflation press coverage by general or specialized newspapers becomes more intense. To test this hypothesis, I estimate the following models:

$$\begin{aligned} GAP_t = \alpha_1 + \alpha_2 * News^G_{t-1} + \alpha_3 * News^S_{t-1} \\ + \alpha_4 * News^P_t + \alpha_5 * \pi_{t-1} + \alpha_6 * \sigma^2_{\pi,t-1} + \epsilon_t \end{aligned} \tag{1}$$

$$\begin{aligned} GAP_{i,t} = \alpha_1 + \alpha_2 * News^G_{t-1} + \alpha_3 * News^S_{t-1} \\ + \alpha_4 * News^P_{i,t} + \alpha_5 * \pi_{t-1} + \alpha_6 * \sigma^2_{\pi,t-1} + x_{i,t}\alpha_7 + \epsilon_{i,t} \end{aligned} \tag{2}$$

Where $GAP_t = |\pi^{MSC}_{t,t+12} - \pi^{SPF}_{t,t+12}|$ and $GAP_{i,t} = |\pi^{MSC}_{i,t,t+12} - \pi^{SPF}_{t,t+12}|$ respectively represent the aggregate and individual measures of the absolute expectation gap. $\pi^{MSC}_{t,t+12}$ and $\pi^{MSC}_{i,t,t+12}$ respectively represent the MSC mean and individual inflation forecasts, while $\pi^{SPF}_{t,t+12}$ is the SPF mean inflation forecast. $News^G$ and $News^S$, respectively, represent the monthly volumes of inflation articles published by general and specialized newspapers scaled by their maxima in any month. They appear with a lag to ensure I use articles that households and experts could have accessed 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, $\sigma^2_{\pi,t-1}$ 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 because of their publication lag. x_i is a vector of socioeconomic characteristics for MSC households (namely gender, age, income, education, marital status, and location in the United States).[25] 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 the aggregate absolute expectation gap is unrelated to inflation press coverage. On the other hand, the individual absolute expectation gap widens when inflation press coverage by either general or specialized newspapers gets more intense, contrary to the theoretical prediction by Carroll (2003). Moreover, when both news variables are included, the results point to a positive correlation with general newspapers and a negative one with specialized newspapers. The adjusted R-squared changes from one press coverage measure to two indicate that the largest

[25] Household 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.

explanatory power resides with general newspapers' press coverage. This result suggests that the absolute expectation gap widens with press coverage.

The distinction between causal and non-causal inflation articles from section 4.1 suggests their relationship with the absolute expectation gap might also differ. If households pay attention to 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 2. $CausalNews^G$ and $CausalNews^S$ respectively represent the monthly volumes of causal inflation articles published by general and specialized newspapers. These measures are also scaled by their maxima in any month. The results in table 6 indicate that causal inflation press coverage intensity positively correlates with the absolute expectation gap. In particular, the aggregate absolute expectation gap becomes wider when specialized newspapers publish more causal inflation articles. In contrast, it does not change with the volume of causal inflation articles published by general newspapers. In addition, the volume of causal inflation articles from both newspapers positively relates to the individual absolute expectation gap. This result aligns with the evidence by Mazumder (2021) that the relation between the expectations and press coverage of the Fed changes across newspapers.

Overall, the evidence aligns with Pfajfar and Santoro (2013) in concluding that the absolute expectation gap does not narrow when inflation press coverage becomes more intense. Therefore, the evidence does not confirm the first hypothesis.

4.3.2 Inflation Press Coverage and Households' Inflation News Perceptions

The second hypothesis from section 2 states that households are more likely to report having heard of news about inflation when inflation press coverage by general or specialized newspapers becomes more intense. To test this hypothesis, I estimate the following models:

$$\begin{aligned} News^P_t = \alpha_1 + \alpha_2 * News^G_{t-1} + \alpha_3 * News^S_{t-1} \\ + \alpha_4 * \pi_{t-1} + \alpha_5 * \sigma^2_{\pi,t-1} + \epsilon_t \end{aligned} \tag{3}$$

$$\begin{aligned} News^P_{i,t} = \alpha_1 + \alpha_2 * News^G_{t-1} + \alpha_3 * News^S_{t-1} \\ + \alpha_4 * \pi_{t-1} + \alpha_5 * \sigma^2_{\pi,t-1} + x_{i,t}\alpha_6 + \epsilon_{i,t} \end{aligned} \tag{4}$$

All variables are defined as before. To test the second hypothesis, I test whether α_2 and α_3 are positive in equations 3 and 4. For this purpose, I also estimate modified versions of equations 3 and 4 where the news-based measures appear individually. 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 7.

When the news measures appear separately, the results show that households are more likely to report having heard of inflation news when inflation press coverage by general or specialized newspapers intensifies both at the aggregate and individual levels. However, when the news measures appear together, this positive relationship survives only for the general newspapers' news measure. This result suggests that households pay more attention to inflation news published by general newspapers and again supports the evidence by Mazumder (2021).

The distinction between causal and non-causal inflation articles from section 4.1 suggests their relationship with households' perceptions of news about inflation might also differ. If households pay attention to 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 3 and 4. The results in table 8 confirm those from table 7.

Overall, the evidence confirms the second hypothesis only for general newspapers.

4.3.3 Narratives Disagreement and Differences in Inflation Expectations

The fourth hypothesis from section 2 states that the absolute expectation gap becomes wider with increasing disagreement about demand and supply narratives between newspapers. To test this hypothesis both at the aggregate and individual levels, I estimate the following two models based on Carroll (2003) and Pfajfar and Santoro (2013):

$$GAP_t = \alpha_1 + \alpha_2 * |NetDemand^{G-S}_{t-1}| + \alpha_3 * News^P_t + \alpha_4 * \pi_{t-1} + \alpha_5 * \sigma^2_{\pi,t-1} + \epsilon_t \tag{5}$$

$$GAP_{i,t} = \alpha_1 + \alpha_2 * |NetDemand_{t-1}^{G-S}| + \alpha_3 * News_{i,t}^{P} + \alpha_4 * \pi_{t-1} + \alpha_5 * \sigma^2_{\pi,t-1} + x_{i,t}\alpha_6 + \epsilon_{i,t} \tag{6}$$

$|NetDemand_{t-1}^{G-S}|$ measures disagreement about demand and supply narratives between newspapers, and all other variables are defined as before. To test the fourth hypothesis, I test whether α_2 is positive in equations 5 and 6. 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 and then together. For equation 5, standard errors are computed with the Huber–White sandwich estimator. For equation 6, standard errors are clustered at the individual level. The results are shown in table 9.

The results show that the aggregate and individual absolute expectation gaps widen when newspapers' demand-supply narrative disagreement increases. This is the only form of narrative disagreement related to the expectation gap at the aggregate level, as the coefficients for $|NetHawkish_{t-1}^{G-S}|$ and $|NetObserved_{t-1}^{G-S}|$ are not statistically different from zero. However, the slope coefficient of $|NetDemand_{t-1}^{G-S}|$ is statistically significant only at the 10% level when the dependent variable is the aggregate absolute expectation gap, while it is statistically significant at the 1% level when the dependent variable is the individual absolute expectation gap. On the other hand, both measures of disagreement about hawkish/dovish narratives and observed/expected narratives are positively related to the individual absolute expectation gap. This result suggests households and experts disagree relatively more on expected inflation when general and specialized newspapers disagree along more than one narrative dimension.

Overall, the results confirm the fourth hypothesis that the absolute expectation gap widens with increasing narrative disagreement between newspapers.

4.3.4 Narratives and Expectations across Individual Characteristics

The fifth hypothesis from section 2 states that households' characteristics moderate the correlation of the absolute expectation gap with disagreement about demand and supply narratives between newspapers. To test this hypothesis, I estimate a modified version of equation 6 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 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.[26] Table 10 shows the results.

The main results are that the positive correlation between narrative disagreement and the individual absolute expectation gap increases with age and decreases with education. In particular, it is smaller for households with a college degree. This is a novel and intuitive result, as college-educated households are more similar to experts than non-college-educated households. On the one hand, the evidence of an increase in the correlation with age lines up with what Ehrmann et al.

[26] Married, 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.

(2018) show for inflation press coverage intensity. On the other hand, the moderating role of education is a novel and intuitive result, as college-educated households are more similar to experts than non-college-educated households. In particular, college-educated households are more likely to read inflation articles published by specialized newspapers than non-college-educated households (Pew Research Center, 2012).

Minor results are the absence of a clear pattern in the coefficients of the interaction terms with the income dummies. In fact, the positive correlation between narrative disagreement and the individual absolute expectation gap decreases for individuals moving from the middle-income quintile to the second-income quintile and the fifth one. In addition, the correlation does not change with sex. The evidence of no change in the correlation with sex lines up with what Ehrmann et al. (2018) show for inflation press coverage intensity.

All in all, the evidence partially confirms the fifth hypothesis, specifically for three out of four demographic characteristics.

4.3.5 Newspaper Demand and Supply Narratives and Individual Demand and Supply Expectations

The sixth hypothesis from section 2 states that households and experts expect inflation and unemployment to move in opposite directions by a larger degree when general and specialized newspapers publish relatively more demand narratives. The seventh hypothesis adds that the expectations of households are more strongly correlated with the narratives of general newspapers than of specialized newspapers, while the opposite holds for the expectations of experts. To test these hypotheses, I estimate the following two models:

$$\begin{aligned}\Delta\pi^{MSC}_{i,t,t+12} * \Delta u^{MSC}_{i,t,t+12} = \alpha_1 + \alpha_2 * NetDemand^{G}_{t-1} + \alpha_3 * NetDemand^{S}_{t-1} \\ + \alpha_4 * \pi_{t-1} + \alpha_5 * \sigma^2_{\pi,t-1} + \alpha_6 * News^{P}_{i,t} + x_{i,t}\alpha_7 + \epsilon^{MSC}_{i,t}\end{aligned} \tag{7}$$

$$
\begin{aligned}
\Delta\pi_{i,t,t+4}^{SPF} * \Delta u_{i,t,t+4}^{SPF} = \alpha_1 + \alpha_2 * NetDemand_{t-1}^{G} + \alpha_3 * NetDemand_{t-1}^{S} \\
+ \alpha_4 * \pi_{t-1} + \alpha_5 * \sigma_{\pi,t-1}^2 + \epsilon_{i,t}^{SPF}
\end{aligned}
\tag{8}
$$

$\Delta\pi_{i,t,t+12}^{MSC}$ 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 these 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 ($\widetilde{\pi}_t$). $\Delta u_{i,t,t+12}^{MSC}$ 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.

$NetDemand^G$ ($NetDemand^S$) is positive when general (specialized) newspapers publish relatively more demand articles. Demand articles predominantly contain demand narratives, which describe a negative relationship between changes in inflation and unemployment. In contrast, $NetDemand^G$ ($NetDemand^S$) is negative when general (specialized) newspapers publish relatively more supply articles. Supply articles predominantly contain supply narratives, which describe a positive relationship between changes in inflation and unemployment. Therefore, to test the sixth hypothesis, I test whether α_2 and α_3 are negative in equations 7 and 8. In addition, to test the seventh hypothesis, I test whether the magnitude of α_2 is lower than that of α_3 in equation 8, and vice versa in 7. For this purpose, I also estimate modified versions of equations 7 and 8 where the narrative measures appear one by one. As the MSC is conducted monthly and the SPF is run quarterly, equations 7 and 8 are estimated at the monthly and quarterly frequency. Therefore, $NetDemand^G$ and $NetDemand^S$ are computed using quarterly numbers in equation 8. Standard errors are clustered at the individual level as respondents in the MSC and SPF can be reinterviewed. The results are shown in table 11.

The first three columns show that households expect inflation and unemployment to move in the same direction by a larger degree when both newspapers publish relatively more supply articles. Since supply articles describe a positive relationship between inflation and unemployment, the evidence suggests that the narratives of newspapers align correctly with the expectations of households. A one-sided t-test shows the null hypothesis that the coefficient of general newspapers is not higher than that of specialized newspapers cannot be rejected (t-test = -0.195). Therefore, the narratives of both newspapers align equally with the expectations of households. On the other hand, the fifth column shows experts expect inflation and unemployment to move in the same direction by a larger degree when specialized newspapers publish relatively more supply articles. Therefore, the evidence suggests that the narratives of specialized newspapers are correctly aligned with the expectations of experts. In contrast, the fourth column shows experts' joint expectations of inflation and unemployment changes are uncorrelated with the publication of demand and supply narratives by specialized newspapers. However, they are correctly correlated in the fifth column when both general and specialized newspapers' narrative measures are used. Therefore, the evidence suggests that the narratives of general newspapers are not correctly aligned with the expectations of experts.

All in all, the evidence confirms the sixth hypothesis fully for households and only partially for experts, as only the narratives of specialized newspapers are correctly aligned with experts' expectations. In addition, the evidence confirms the seventh hypothesis only for experts. This is because the narratives of general and specialized newspapers are equally aligned with households' expectations.

4.4 Robustness checks

4.4.1 Narratives and Expectations across Narrative Types

My hypotheses so far focus on disagreement about demand and supply narratives between general and specialized newspapers. However, newspapers might also disagree about the importance of different types of demand and supply narratives. For instance, general newspapers might publish mostly narratives about energy prices (a supply narrative) in a given period. In contrast, specialized newspapers might publish mainly narratives about loose monetary policy (a demand narrative). Importantly, narrative disagreement can also occur when general newspapers publish relatively as many demand and supply narratives as specialized newspapers while mentioning different types of demand and supply narratives. For instance, general newspapers might publish mostly narratives about energy prices, while specialized newspapers publish mainly narratives about supply chain disruptions (also a supply narrative). Therefore, the absolute expectation gap might also widen when general and specialized newspapers pay different attention to the same type of narrative. As there are multiple demand and supply narrative categories, the relationship between narrative disagreement and the absolute expectation gap might differ across the kinds of narratives on which newspapers disagree. For instance, the relationship with disagreement over commodities/energy narratives might be stronger than with disagreement over monetary policy narratives.

To verify whether the relationship between narrative disagreement and the absolute expectation gap changes based on the individual types of narratives newspapers disagree about, I estimate the following models:

$$\begin{aligned} GAP_t = \alpha_1 + \alpha_2 * |MonPol_{t-1}^{G-S}| + \alpha_3 * |SpendDefDebt_{t-1}^{G-S}| \\ + \alpha_4 * |ComEne_{t-1}t^{G-S}| + \alpha_5 * |Labor_{t-1}^{G-S}| \\ + \alpha_6 * News_t^P + \alpha_7 * \pi_{t-1} + \alpha_8 * \sigma_{\pi,t-1}^2 + \epsilon_t \end{aligned} \tag{9}$$

$$\begin{aligned} GAP_{i,t} = \alpha_1 + \alpha_2 * |MonPol_{t-1}^{G-S}| + \alpha_3 * |SpendDefDebt_{t-1}^{G-S}| \\ + \alpha_4 * |ComEne_{t-1}t^{G-S}| + \alpha_5 * |Labor_{t-1}^{G-S}| \\ + \alpha_6 * News_{i,t}^P + \alpha_7 * \pi_{t-1} + \alpha_8 * \sigma_{\pi,t-1}^2 + x_{i,t}\alpha_9 + \epsilon_{i,t} \end{aligned} \tag{10}$$

$MonPol^S$, $SpendDefDebt^S$, $ComEne^S$, and $Labor^S$ represent narrative disagreement about monetary policy, government spending/deficit/debt, commodities/energy, and labor, respectively.[27] All other variables are defined as before. I also estimate modified versions of equations 9 and 10

[27] I 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.

in which the measures of disagreement about narrative types appear individually. 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 A4 in Appendix D shows the results.

The results show that the aggregate absolute expectation gap widens when newspapers' narrative disagreement around monetary policy narratives increases. In addition, most narrative disagreement measures are positively related to the individual absolute expectation gap. In particular, the null hypothesis that the coefficient of $|MonPol^{G-S}|$ is not higher than the coefficient of $|ComEn^{G-S}|$ can be rejected at the 1% significance level (t-test = 4.352). In contrast, the null hypothesis that the coefficient of $|ComEne^{G-S}|$ is not higher than that of $|SpenDefDebt^{G-S}|$ cannot be rejected (t-test = 0.363). The results for monetary policy disagreement are particularly interesting in line with the recent survey evidence by Stantcheva (2024) that higher-income and college-educated individuals hold monetary policy narratives more often than lower-income respondents. Surprisingly, the aggregate and individual absolute expectation gaps become narrower when disagreement about labor narratives increases. More work is needed to understand why this is the case.

All in all, the evidence shows that the relationship between narrative disagreement and the absolute expectation gap changes based on the individual types of narratives newspapers disagree about.

4.4.2 Asymmetric Narratives Disagreement and Differences in Inflation Expectations

The sixth hypothesis from section 2 is tested in section 4.3.3 using the absolute value of $NetDemand^{G-S}$. The absolute value is used because the positive relationship postulated in the sixth hypothesis concerns the magnitude of narrative disagreement, not its sign. An important question is whether the sign also matters. That is, whether the results change when general or specialized newspapers publish relatively more demand narratives than the other. To answer this question, I estimate the following modified versions of the models specified in equations 5 and 6:

$$\begin{aligned} GAP_t = \alpha_1 &+ \alpha_2 * |NetDemand^{G-S}_{t-1}| * 1_{NetDemand^{G-S}_{t-1}>0} \\ &+ \alpha_3 * |NetDemand^{G-S}_{t-1}| * 1_{NetDemand^{G-S}_{t-1}<0} \\ &+ \alpha_4 * News^P_t + \alpha_5 * \pi_{t-1} + \alpha_6 * \sigma^2_{\pi,t-1} + \epsilon_t \end{aligned} \tag{11}$$

$$\begin{aligned} GAP_{i,t} = \alpha_1 &+ \alpha_2 * |NetDemand^{G-S}_{t-1}| * 1_{NetDemand^{G-S}_{t-1}>0} \\ &+ \alpha_3 * |NetDemand^{G-S}_{t-1}| * 1_{NetDemand^{G-S}_{t-1}<0} \\ &+ \alpha_4 * News^P_{i,t} + \alpha_5 * \pi_{t-1} + \alpha_6 * \sigma^2_{\pi,t-1} + x_{i,t}\alpha_7 + \epsilon_{i,t} \end{aligned} \tag{12}$$

$1_{NetDemand^{G-S}_{t-1}>0}$ is a dummy variable that takes value one only when $NetDemand^{G-S}_{t-1}$ is positive, i.e., when general newspapers publish relatively more demand narratives that specialized newspapers. Instead,$1_{NetDemand^{G-S}_{t-1}<0}$ is a dummy variable that takes value one only when $NetDemand^{G-S}_{t-1}$ is negative, i.e., when general newspapers publish relatively fewer demand narratives that specialized newspapers. All other variables are defined as before. For equation 11, standard errors are computed with the Huber–White sandwich estimator. For equation 12, standard errors are clustered at the individual level.

The results in table A5 in Appendix E indicate that the absolute expectation gap is positively related to narrative disagreement between newspapers, no matter the sign. In fact, none of the estimates of both α_2 and α_3 is negative. However, there are differences in the magnitudes of the coefficients across signs. The aggregate absolute expectation gap widens only when general newspapers publish relatively more demand 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, differences in inflation expectations are positively related to narrative disagreement but differently based on its sign.

4.4.3 Between-General-Newspaper Heterogeneity

The analysis so far measures general newspapers' narratives using the narratives published by the New York Times, USA Today, and Washington Post together. An important question is whether the aggregation of these three newspapers hides any between-general-newspaper heterogeneity that matters for the positive relationship between the absolute expectation gap and narrative disagreement. For instance, this might be the case if the readership demographics of the three general newspapers differ (Pew Research Center, 2012). In addition, disagreement between general and specialized newspapers might capture partisan bias. In fact, the New York Times and Washington Post are typically considered liberal, whereas the Wall Street Journal is considered conservative (Gallup/Knight, 2020). In contrast, USA Today is considered moderate. Therefore, if partisan bias explains narrative disagreement, then only narrative disagreement between liberal and conservative newspapers should be positively correlated to the absolute expectation gap.

To answer this question, I construct three general-newspaper-specific measures of demand and supply narratives based on the methods described in sections 3.2 and 3.3. I call these measures $NetDemand_t^{NYT}$, $NetDemand_t^{USAT}$, and $NetDemand_t^{WaPo}$ and use them to derive three distinct measures of narrative disagreement as follows:

$$NetDemand_t^{NYT-S} = NetDemand_t^{NYT} - NetDemand_t^{S} \tag{13}$$

$$NetDemand_t^{USAT-S} = NetDemand_t^{USAT} - NetDemand_t^{S} \tag{14}$$

$$NetDemand_t^{WaPo-S} = NetDemand_t^{WaPo} - NetDemand_t^{S} \tag{15}$$

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 $NetDemand_t^{WaPo-S}$ respectively measure demand-supply narrative disagreement between the Washington Post and specialized newspapers.

Then, I estimate modified versions of the models specified at equations 5 and 6 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 $NetDemand^{WaPo-S}$). I show the results in table A6 in Appendix F and 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.

Finally, I estimate modified versions of the models specified at equations 7 and 8 in which the general newspapers' narrative measure ($NetDemand^{G}$) is replaced by each of the three general newspaper-specific narrative measures ($NetDemand^{NYT}$, $NetDemand^{USAT}$, and $NetDemand^{WaPo}$). I show the results in table A7 in Appendix F. The main result is that the results obtained with the general newspapers' narrative measure broadly extend to those obtained with each of the three general newspaper-specific narrative measures. In particular, narratives never incorrectly (correctly) align with households' (experts') demand and supply expectations. Overall, this finding is reassuring as it also suggests that the aggregate general measure does hide differences between its components.

Overall, the results obtained by treating the three general newspapers independently align with those from their aggregation in one general newspaper.

4.4.4 Expert Disagreement

A branch of the literature on inflation expectations looks at the second moment in the cross-sectional distribution of expectations and documents substantial dispersion of inflation expectations, especially among households with respect to experts (Weber et al., 2022). The evidence from section 4.3.4 shows the relevance of narrative disagreement for expectation disagreement among households, particularly for older vs. younger households and for college-educated vs. non-college-educated households. Narrative disagreement might also be related to disagreement among experts if general and specialized newspapers publish the narratives of different sets of experts. To test this, I regress the interquartile range of the SPF point forecasts on the absolute value of lagged narrative disagreement, controlling for the most recent level and volatility of inflation.[28] The results in table A8 in Appendix G indicate that narrative disagreement is positively related to expert disagreement. This result suggests that expert disagreement might subsume the predictive of narrative disagreement toward the absolute expectation gap. To verify this, I estimate modified versions of equations 5 and 6 in which I interact narrative disagreement with the interquartile range of the SPF point forecasts.[29] The results are shown in table A9 in Appendix G. When the dependent variable is the aggregate absolute expectation gap, the main slope coefficient of narrative disagreement is positive and statistically significant at the 10% level, while the coefficient of its interaction with expert disagreement is not statistically significant. On the other hand, narrative disagreement becomes completely unrelated to the individual absolute expectation gap.

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

4.4.5 Incentives to Gather Information about Inflation

Cavallo, Cruces, and Perez-Truglia (2017) show that households living in high-inflation environments are more informed about inflation than households living in low-inflation environments. Since inflation ranges between -2 and 10% in the sample I study, the evidence by Cavallo et al. (2017) begs the question of whether the relationship between narrative disagreement and the expectation gap changes based on the level of inflation. In particular, households might pay more attention to general newspapers when inflation is high and, hence, the relationship between narrative disagreement and the absolute expectation gap might become stronger with respect to when inflation is low. To verify this, I estimate modified versions of equations 5 and 6 in which I interact narrative disagreement with the most recently observed level of inflation. The results are shown in table A10 in Appendix H. When the dependent variable is the aggregate absolute expectation gap,

[28] I use quarterly observations.

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

the slope coefficient of the interaction term is not statistically significant, while the slope coefficient of the main term of narrative disagreement remains positive and statistically significant at the 10% level. Therefore, the relationship between narrative disagreement and the absolute expectation gap does not change based on the recent level of inflation. On the other hand, when the dependent variable is the individual absolute expectation gap, the slope coefficient of the interaction term is positive and statistically significant at the 5% level. In contrast, the slope coefficient of the main term of narrative disagreement is not statistically significant. Therefore, the relationship between narrative disagreement and the individual expectation gap is positive when inflation is positive, as it is generally the case in my sample. In addition, it becomes stronger as inflation increases.

Overall, the results show that the relationship between narrative disagreement and the individual absolute expectation becomes stronger as inflation increases.

4.4.6 Forecast Errors

The analysis so far benchmarks households' inflation expectations against those of experts and documents a positive relationship between their differences and narrative disagreement between newspapers. An important question is whether households' forecast errors are also positively correlated with narrative disagreement between newspapers. If so, households' forecasts would be both closer to those of experts and more accurate when narrative disagreement between newspapers declines. To verify this, I estimate the following models:

$$FE_t = \alpha_1 + \alpha_2 * |NetDemand_{t-1}^{G-S}| + \alpha_3 * News_t^P + \alpha_4 * \pi_{t-1} + \alpha_5 * \sigma^2_{\pi,t-1} + \epsilon_t \tag{16}$$

$$\begin{aligned} FE_{i,t} = \alpha_1 + \alpha_2 * |NetDemand_{t-1}^{G-S}| \\ + \alpha_3 * News_{i,t}^P + \alpha_4 * \pi_{t-1} + \alpha_5 * \sigma^2_{\pi,t-1} + x_{i,t}\alpha_6 + \epsilon_{i,t} \end{aligned} \tag{17}$$

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 16, standard errors are computed with the Huber–White sandwich estimator. For equation 17, standard errors are clustered at the individual level.

The results in table A11 in Appendix I indicate that household-level forecast errors widen with rising narrative disagreement between newspapers. Therefore, individual households produce better inflation forecasts when narrative disagreement between newspapers declines. However, this positive relationship loses statistical significance once I control for the most recent level and volatility of inflation.

Overall, households' forecast errors are positively related to narrative disagreement between newspapers, albeit the evidence is not robust to macroeconomic controls.

4.4.7 Narratives and Macroeconomic Dynamics

An important question is whether the content of demand and supply narratives is captured by observable macroeconomic variables, namely inflation and unemployment. Assessing this relationship is particularly important due to the evidence from section 4.3.5 that the narratives of general newspapers align differently with households' and experts' demand and supply views. To this end, I estimate the following models:

$$Demand_t^j = \alpha_1 + \alpha_2 * \pi_t + \alpha_3 * u_t + \alpha_4 * \pi_t * u_t + \alpha_5 * Demand_{t-1}^j + \epsilon_t \tag{18}$$

$$Supply_t^j = \alpha_1 + \alpha_2 * \pi_t + \alpha_3 * u_t + \alpha_4 * \pi_t * u_t + \alpha_5 * Supply_{t-1}^j + \epsilon_t \tag{19}$$

$$NetDemand_t^j = \alpha_1 + \alpha_2 * \pi_t + \alpha_3 * u_t + \alpha_4 * \pi_t * u_t + \alpha_5 * NetDemand_{t-1}^j + \epsilon_t \tag{20}$$

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 3.3, and u_t is the seasonally adjusted civilian unemployment rate.[30]

The content of demand articles predominantly describes a negative relationship between changes

[30]Barnichon 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.

in inflation and unemployment, while the opposite holds for supply articles. Therefore, α_4 should be non-positive in equation 18 and non-negative in equation 19. In addition, since $NetDemand^j$ is the scaled difference between $Demand^j$ and $Supply^j$, α_4 should be non-positive in equation 20. Table A12 in Appendix J shows the results.

The second and fifth columns indicate that the publication volume of demand narratives by both general and specialized newspapers increases relatively less if inflation and unemployment jointly increase. This result aligns with the type of inflation-unemployment relationship described by demand articles. However, the third column indicates a negative coefficient for the interaction term coefficient when the dependent variable is general newspapers' publication of supply articles. This result goes against the type of inflation-unemployment relationship described by supply articles. As the interaction term coefficient in the third column is marginally smaller than in the second column, the fourth column displays an incorrectly positive sign for α_5 when the dependent variable is general newspapers' NetDemand measure, though the statistical significance is limited to the 10% level. In contrast, the sixth column shows that the interaction term coefficient is not statistically significant when the dependent variable is specialized newspapers' publication of supply articles. Consequently, the last column shows a correctly negative interaction term coefficient when the dependent variable is specialized newspapers NetDemand measure. Overall, the results suggest that general newspapers' narratives are not only in contrast with experts' views but also with macroeconomic dynamics.

On a minor note, the main terms in the second and third columns show that general newspapers publish more demand and supply narratives when inflation or unemployment increases. In addition, the fifth column shows that a similar result holds for the demand narratives published by specialized newspapers. In contrast, the sixth column indicates that the publication of supply narratives by specialized newspapers is only weakly positively correlated with inflation.

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

5 Conclusion

In this paper, I study the connection between disagreement about demand and supply narratives between newspapers and households' absolute gap in inflation expectations with experts. I measure disagreement between newspapers by applying a causality extraction (CE) algorithm and dictionary-based method to inflation articles published by general and specialized newspapers. I collect more than 180,000 articles published by three major general newspapers (NYT, Washington Post, and USA Today) and one specialized (WSJ) between 1991 and 2022 that mention "inflation," "cpi," "consumer price," "ppi," or "producer price." CE is designed to extract causal relations and, hence, can be used to construct measures of inflation narratives that attribute inflation to its triggers. After applying CE to my inflation articles, I classify the inflation narratives extracted into demand and supply narratives using dictionaries of demand and supply factors previously used in the literature.

Using households' and experts' expectations from the MSC and the SPF, this study shows that general and specialized newspapers' demand and supply narratives predict households' and experts' demand and supply views, respectively. In particular, general newspapers' narratives align correctly with households' views and incorrectly with experts' views, whereas specialized newspapers' narratives align correctly with both. This finding suggests that the narratives of general and specialized newspapers capture the beliefs of different audiences. Interestingly, general newspapers' narratives also incorrectly align with the joint dynamics of realized inflation and unemployment, while specialized newspapers' narratives do not. Therefore, this finding further suggests that general newspapers might communicate incorrect narratives to households. Moreover, the evidence further indicates that using a single source of inflation articles is insufficient to study how the inflation expectations of distinct groups of individuals covary with inflation news.

Next, I use general and specialized newspapers' demand and supply narratives to measure narrative disagreement. General and specialized newspapers tend to agree in their demand and supply narratives. However, when their narrative disagreement increases, the absolute expectation gap widens, regardless of the direction of narrative disagreement. Across different households, the absolute gap grows relatively more for households with low education levels. When looking at which types of demand and supply narratives newspapers disagree on, this study highlights the role of narrative disagreement about the importance of monetary policy narratives. Importantly, the results are unchanged when accounting for the direction of narrative disagreement, the level of inflation, the type of general newspaper used to measure narrative disagreement, and disagreement among experts. In addition, the predictive power of demand-supply narrative disagreement is captured by narrative disagreement about whether inflation is increasing or decreasing and whether articles talk about realized or future inflation episodes. Finally, narrative disagreement between newspapers is also positively related to households' forecast errors with respect to realized inflation, albeit the evidence is not robust to macroeconomic controls.

These findings have interesting implications for policymakers and the media, suggesting that more reporting about inflation brings households' inflation expectations closer to those of experts only when there is low narrative disagreement in the media landscape. In particular, these findings suggest that if central banks want to lower the dispersion of inflation forecasts among different groups of individuals, they need to communicate their inflation narratives across a broad range of channels, particularly those that cater to non-college-educate individuals.

References

G. A. Akerlof and D. J. Snower. Bread and bullets. *Journal of Economic Behavior & Organization*, 126:58–71, 2016.

B. Altenberg. Causal linking in spoken and written english. *Studia linguistica*, 38(1):20–69, 1984.

P. Andre, C. Pizzinelli, C. Roth, and J. Wohlfart. Subjective models of the macroeconomy: Evidence from experts and representative samples. *The Review of Economic Studies*, 89(6):2958–2991, 2022.

P. Andre, I. Haaland, C. Roth, and J. Wohlfart. Narratives about the macroeconomy. 2023.

M. Apel and M. Grimaldi. The information content of central bank minutes. *Riksbank Research Paper Series*, (92), 2012.

M. Apel, M. Blix Grimaldi, and I. Hull. 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, 2022.

L. Baele, F. De Jong, and G. Trebbi. What triggers flights to safety? *Available at SSRN 4503287*, 2023.

S. R. Baker, N. Bloom, S. J. Davis, and K. Kost. Policy news and stock market volatility. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2019-53), 2021.

R. Barnichon and A. H. Shapiro. How much has the cooling economy reduced inflation? *FRBSF Economic Letter*, 2024(30):1–5, 2024.

W. Bruine de Bruin, W. Vanderklaauw, J. S. Downs, B. Fischhoff, G. Topa, and O. Armantier. Expectations of inflation: The role of demographic variables, expectation formation, and financial literacy. *Journal of Consumer Affairs*, 44(2):381–402, 2010.

M. F. Bryan and G. Venkatu. The demographics of inflation opinion surveys. *Economic Commentary*, (10/15/2001), 2001.

L. Bybee, B. T. Kelly, A. Manela, and D. Xiu. Business news and business cycles. *Journal of Finance, Forthcoming*, 2023.

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

A. Cavallo, G. Cruces, and R. Perez-Truglia. Inflation expectations, learning, and supermarket prices: Evidence from survey experiments. *American Economic Journal: Macroeconomics*, 9 (3):1–35, 2017.

C. Charles and C. Kendall. Causal narratives. *Available at SSRN 4669371*, 2023.

O. Coibion, Y. Gorodnichenko, and M. Weber. Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy*, 130(6):000–000, 2022.

C. Dougal, J. Engelberg, D. Garcia, and C. A. Parsons. Journalists and the stock market. *The Review of Financial Studies*, 25(3):639–679, 2012.

L. Dräger, M. J. Lamla, and D. Pfajfar. Are survey expectations theory-consistent? the role of central bank communication and news. *European Economic Review*, 85:84–111, 2016.

F. D'Acunto, D. Hoang, M. Paloviita, and M. Weber. Cognitive abilities and inflation expectations. In *AEA Papers and Proceedings*, volume 109, pages 562–566. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, 2019.

F. D'Acunto, U. Malmendier, J. Ospina, and M. Weber. Exposure to grocery prices and inflation expectations. *Journal of Political Economy*, 129(5):1615–1639, 2021a.

F. D'Acunto, U. Malmendier, and M. Weber. Gender roles produce divergent economic expectations. *Proceedings of the National Academy of Sciences*, 118(21):e2008534118, 2021b.

M. Ehrmann, D. Pfajfar, and E. Santoro. Consumers' attitudes and their inflation expectations. *47th issue (February 2017) of the International Journal of Central Banking*, 2018.

K. Eliaz and R. Spiegler. A model of competing narratives. *American Economic Review*, 110(12): 3786–3816, 2020.

N. Ersahin, M. Giannetti, and R. Huang. Supply chain risk: Changes in supplier composition and vertical integration. *Journal of International Economics*, 147:103854, 2024.

Gallup/Knight. American views 2020: Trust, media and democracy. *Knight Foundation*, 2020.

D. Garcia, X. Hu, and M. Rohrer. The colour of finance words. *Journal of Financial Economics*, 147(3):525–549, 2023.

R. Girju. Automatic detection of causal relations for question answering. In *Proceedings of the ACL 2003 workshop on Multilingual summarization and question answering*, pages 76–83, 2003.

B. Hayo, K. Henseler, M. S. Rapp, and J. Zahner. Complexity of ecb communication and financial market trading. *Journal of International Money and Finance*, 128:102709, 2022.

C. S. Khoo, J. Kornfilt, R. N. Oddy, and S. H. Myaeng. Automatic extraction of cause-effect information from newspaper text without knowledge-based inferencing. *Literary and Linguistic Computing*, 13(4):177–186, 1998.

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

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

A. Manela and A. Moreira. News implied volatility and disaster concerns. *Journal of Financial Economics*, 123(1):137–162, 2017.

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

S. Mazumder. The reaction of inflation forecasts to news about the fed. *Economic Modelling*, 94: 256–264, 2021.

Pew Research Center. In changing news landscape, even television is vulnerable, 2012. URL `https://www.pewresearch.org/politics/2012/09/27/section-4-demographics-and-political-views-of-news-audiences/#attitudes-about-the-news`.

Pew Research Center. News platform fact sheet, 2023. URL `https://www.pewresearch.org/journalism/fact-sheet/news-platform-fact-sheet/`.

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

R. Prasad, N. Dinesh, A. Lee, E. Miltsakaki, L. Robaldo, A. K. Joshi, and B. L. Webber. The penn discourse treebank 2.0. In *LREC*. Citeseer, 2008.

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

L. Smales and N. Apergis. Understanding the impact of monetary policy announcements: The importance of language and surprises. *Journal of Banking & Finance*, 80:33–50, 2017.

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

P. C. Tetlock. Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3):1139–1168, 2007.

P. C. Tetlock, M. Saar-Tsechansky, and S. Macskassy. More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance*, 63(3):1437–1467, 2008.

M. Weber, F. D'Acunto, Y. Gorodnichenko, and O. Coibion. The subjective inflation expectations of households and firms: Measurement, determinants, and implications. *Journal of Economic Perspectives*, 36(3):157–84, August 2022.

M. Weber, Y. Gorodnichenko, and O. Coibion. The expected, perceived, and realized inflation of us households before and during the covid19 pandemic. *IMF Economic Review*, 71(1):326–368, 2023.

J. Yang, S. C. Han, and J. Poon. A survey on extraction of causal relations from natural language text. *Knowledge and Information Systems*, pages 1–26, 2022.

6 Tables

Table 1: Inflation articles - Summary statistics

	Source	Mean	SD	Q1	Median	Q3
# 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	9.0	2.0	7.8	9.0	10.2
	General	9.5	2.0	8.3	9.5	10.7
	Specialized	8.7	1.8	7.5	8.6	9.8

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 are NYT, USAT, and WaPo, whereas WSJ is the 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) + {*"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"*}
Demand	Spending/Deficit/Debt	"Spending/Deficit/Debt" dictionary from Baker et al. (2021) + {*"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"*}
Supply	Commodities/Energy	"Commodity Markets" dictionary from Baker et al. (2021) + {*"commodity", "crop", "crude", "diesel", "electric", "electricity", "energy", "fuel", "gasoline", "grain", "lumber", "opec", "petroleum", "soybean"*}
Supply	Labor	"Labor Markets" and "Labor Disputes" dictionaries from Baker et al. (2021) + {*"collective bargaining agreement", "job creation", "job market", "jobless", "pay", "pay raise", "paycheck", "union", "work force", "workforce", "worker"*}
Supply	Supply chain	Top 100 supply-chain risk bigrams from Ersahin et al. (2024)

Table 3: Causal inflation articles - Summary statistics

	Source	Mean	SD	Q1	Median	Q3
# 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	9.2	1.7	8.1	9.1	10.2
	General	9.6	1.8	8.5	9.6	10.7
	Specialized	9.0	1.6	7.9	8.9	10.0

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 are NYT, USAT, and WaPo, whereas WSJ is the 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
$NetDemand^{G}$	384.0	-0.073***	0.227	-1.000	-0.125	0.000	0.000	0.875
$NetDemand^{S}$	384.0	-0.049***	0.166	-1.000	-0.143	-0.048	0.000	0.714
$NetDemand^{G-S}$	384.0	-0.023*	0.242	-0.952	-0.125	0.000	0.109	0.780
$NetHawkish^{G}$	384.0	0.064***	0.163	-0.500	0.000	0.000	0.100	1.000
$NetHawkish^{S}$	384.0	0.081***	0.174	-0.350	0.000	0.050	0.150	1.000
$NetHawkish^{G-S}$	384.0	-0.017	0.202	-0.750	-0.100	0.000	0.100	0.850
$NetObserved^{G}$	384.0	0.045***	0.165	-0.500	-0.056	0.000	0.111	1.000
$NetObserved^{S}$	384.0	0.048***	0.225	-0.632	-0.105	0.053	0.158	1.000
$NetObserved^{G-S}$	384.0	-0.003	0.256	-0.868	-0.158	0.000	0.115	1.091
$ConsSpendSent^{G-S}$	384.0	0.005	0.144	-1.000	0.000	0.000	0.000	1.000
$MonPol^{G-S}$	384.0	-0.031***	0.161	-0.846	-0.077	0.000	0.000	0.701
$SpendDefDebt^{G-S}$	384.0	0.029***	0.162	-0.364	0.000	0.000	0.000	1.000
$ComEne^{G-S}$	384.0	-0.004	0.143	-0.636	-0.052	0.000	0.043	0.810
$Labor^{G-S}$	384.0	-0.030***	0.138	-0.545	-0.082	0.000	0.000	0.727
$SupplyChain^{G-S}$	384.0	0.002	0.106	-1.000	0.000	0.000	0.000	1.000

Note: *This table shows summary statistics 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 3.3. NetHawkish is measured as described in section 3.4. NetObserved is measured as described in section 3.5. General newspapers are NYT, USAT, and WaPo, whereas WSJ is the 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 expectations

	GAP_t			$GAP_{i,t}$		
$News^G_{t-1}$	0.452		0.981	1.244***		1.693***
	(0.530)		(0.725)	(0.068)		(0.090)
$News^S_{t-1}$		-0.187	-0.624		0.343***	-0.544***
		(0.306)	(0.448)		(0.053)	(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 are NYT, USAT, and WaPo, whereas WSJ is the 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 expectations

	GAP_t			$GAP_{i,t}$		
$CausalNews^G_{t-1}$	0.412		0.136	1.133***		1.086***
	(0.453)		(0.482)	(0.067)		(0.073)
$CausalNews^S_{t-1}$		0.531**	0.502**		0.367***	0.085*
		(0.206)	(0.238)		(0.044)	(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 are NYT, USAT, and WaPo, whereas WSJ is the 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: Inflation press coverage and news perceptions

	$News_t^P$			$News_{i,t}^P$		
$News_{t-1}^G$	0.316***		0.340***	0.294***		0.317***
	(0.025)		(0.034)	(0.010)		(0.011)
$News_{t-1}^S$		0.157***	-0.030		0.140***	-0.028***
		(0.024)	(0.028)		(0.006)	(0.006)
Demographics control	-	-	-	Yes	Yes	Yes
Past inflation control	Yes	Yes	Yes	Yes	Yes	Yes
Past inflation volatility control	Yes	Yes	Yes	Yes	Yes	Yes
N	382	382	382	161641	161641	161641

Note: *This table shows the results obtained from estimating the models specified in equations 3 and 4, as well as modified versions of them where the news measures appear one by one. General newspapers are NYT, USAT, and WaPo, whereas WSJ is the specialized newspaper. Demographic controls include gender, age, income, education, marital status, and location in the United States. 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: Causal inflation press coverage and news perceptions

	$News_t^P$			$News_{i,t}^P$		
$CausalNews_{t-1}^G$	0.297***		0.294***	0.271***		0.276***
	(0.026)		(0.029)	(0.009)		(0.010)
$News_{t-1}^S$		0.088***	0.005		0.062***	-0.008*
		(0.021)	(0.020)		(0.004)	(0.005)
Demographics control	-	-	-	Yes	Yes	Yes
Past inflation control	Yes	Yes	Yes	Yes	Yes	Yes
Past inflation volatility control	Yes	Yes	Yes	Yes	Yes	Yes
N	382	382	382	161641	161641	161641

Note: *This table shows the results obtained from estimating modified versions of the models specified in equations 3 and 4, where measures of causal inflation press coverage intensity replace measures of inflation press coverage intensity. General newspapers are NYT, USAT, and WaPo, whereas WSJ is the specialized newspaper. Demographic controls include gender, age, income, education, marital status, and location in the United States. For the modified version of equation 3, standard errors are computed with the Huber–White sandwich estimator. Standard errors are clustered at the individual level for the modified version of equation 4. The sample includes all months between 1991 and 2022.*

Table 9: Inflation narratives and expectations

(a) Aggregate

	GAP_t			
$\lvert NetDemand_{t-1}^{G-S}\rvert$	0.376*	0.371*	0.376*	0.373*
	(0.200)	(0.201)	(0.206)	(0.206)
$\lvert NetHawkish_{t-1}^{G-S}\rvert$		0.024		0.027
		(0.216)		(0.227)
$\lvert NetObserved_{t-1}^{G-S}\rvert$			-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

	$GAP_{i,t}$			
$\lvert NetDemand_{t-1}^{G-S}\rvert$	0.133***	0.097**	0.090**	0.071*
	(0.037)	(0.038)	(0.038)	(0.038)
$\lvert NetHawkish_{t-1}^{G-S}\rvert$		0.179***		0.124***
		(0.043)		(0.045)
$\lvert NetObserved_{t-1}^{G-S}\rvert$			0.202***	0.174***
			(0.038)	(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 5 and 6.* 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 3.3.* $NetHawkish^{G-S}$ *is the hawkish-dovish narrative disagreement measure and is constructed as described in section 3.4.* $NetObserved^{G-S}$ *is the observed-expected narrative disagreement measure and is constructed as described in section 3.4. For equation 5, standard errors are computed with the Huber–White sandwich estimator. For equation 6, standard errors are clustered at the individual level. The results are shown in table 9. The sample includes all months between 1991 and 2022.*

Table 10: Inflation narratives and expectations across individual characteristics

	$GAP_{i,t}$			
$\lvert NetDemand_{t-1}^{G-S}\rvert$	0.144***	0.078*	0.314***	0.320***
	(0.047)	(0.043)	(0.083)	(0.083)
$\lvert NetDemand_{t-1}^{G-S}\rvert * FEMALE_{i,t}$	-0.022			
	(0.073)			
$\lvert NetDemand_{t-1}^{G-S}\rvert * AGE_{i,t}$		0.005**		
		(0.002)		
$\lvert NetDemand_{t-1}^{G-S}\rvert * EDUC1_{i,t}$			-0.087	
			(0.333)	
$\lvert NetDemand_{t-1}^{G-S}\rvert * EDUC2_{i,t}$			0.128	
			(0.269)	
$\lvert NetDemand_{t-1}^{G-S}\rvert * EDUC4_{i,t}$			-0.133	
			(0.111)	
$\lvert NetDemand_{t-1}^{G-S}\rvert * EDUC5_{i,t}$			-0.360***	
			(0.106)	
$\lvert NetDemand_{t-1}^{G-S}\rvert * EDUC6_{i,t}$			-0.267**	
			(0.109)	
$\lvert NetDemand_{t-1}^{G-S}\rvert * INC1_{i,t}$				-0.128
				(0.141)
$\lvert NetDemand_{t-1}^{G-S}\rvert * INC2_{i,t}$				-0.343***
				(0.122)
$\lvert NetDemand_{t-1}^{G-S}\rvert * INC4_{i,t}$				-0.183*
				(0.110)
$\lvert NetDemand_{t-1}^{G-S}\rvert * INC5_{i,t}$				-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 6. The model differs because I sequentially interact the variable* $|NetDemand_{t-1}^{G-S}|$ *with a number of the consumer characteristics represented in* x_i. $FEMALE_{i,t}$ *is a dummy taking value 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 of the respondents in the MSC are reinterviewed. The sample includes all months between 1991 and 2022.*

Table 11: Newspaper and individual narratives

	$\Delta\pi^{MSC}_{i,t,t+12} * \Delta u^{MSC}_{i,t,t+12}$			$\Delta\pi^{SPF}_{i,t,t+4} * \Delta u^{SPF}_{i,t,t+4}$		
$NetDemand^{G}_{t-1}$	-0.149***		-0.131***	0.282***		0.405***
	(0.033)		(0.033)	(0.040)		(0.049)
$NetDemand^{S}_{t-1}$		-0.157***	-0.119***		-0.226**	-0.345***
		(0.044)	(0.044)		(0.036)	(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 7 and 8, as well as modified versions of them where the narrative measures appear one by one.* $\pi^{MSC}_{i,t,t+12}$ *measures the MSC's household expected change in inflation, whereas* $\Delta u^{MSC}_{i,t,t+12}$ *measures the MSC's household expected change in unemployment.* $\Delta\pi^{SPF}_{i,t,t+4}$ *and* $\Delta u^{SPF}_{i,t,t+4}$ *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 3.3. General newspapers are NYT, USAT, and WaPo, whereas WSJ is the 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 of the respondents in the MSC and SPF are reinterviewed. The sample includes all months and quarters between 1991 and 2022.*

7 Figures

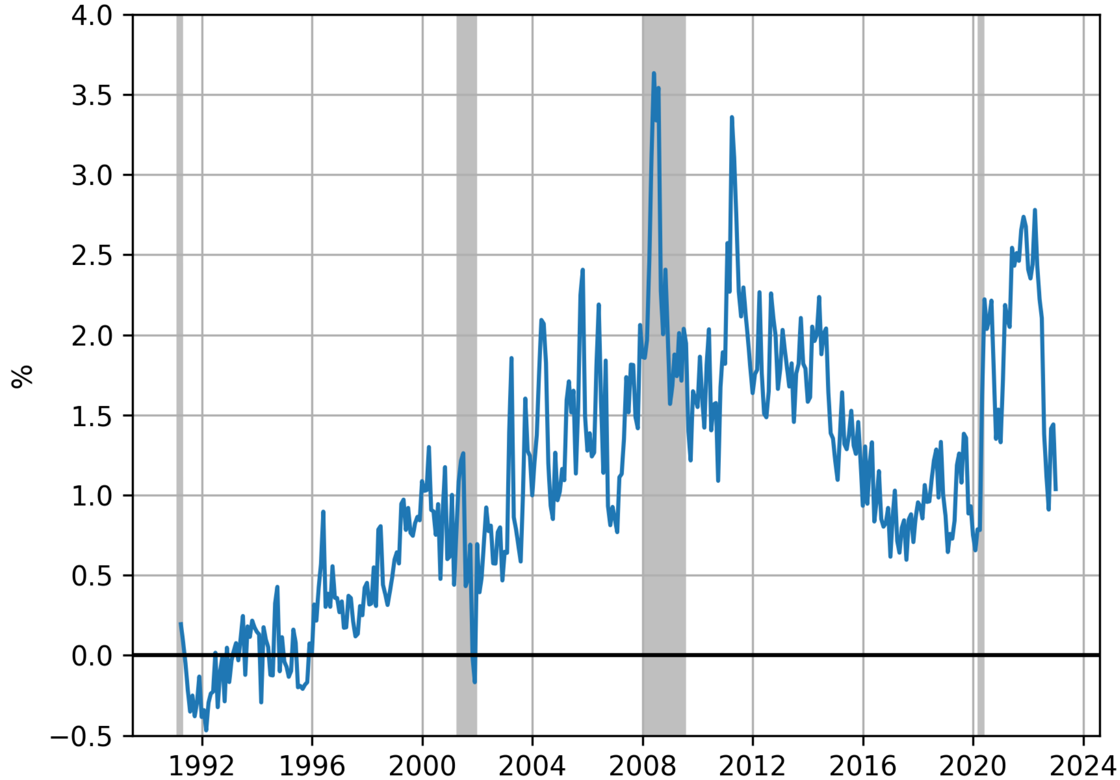

Figure 1: Absolute expectation gap

Note: *This figure shows the difference between the MSC and SPF mean forecasts. Shaded areas represent NBER recession periods. The sample includes all months between January 1991 and December 2022.*

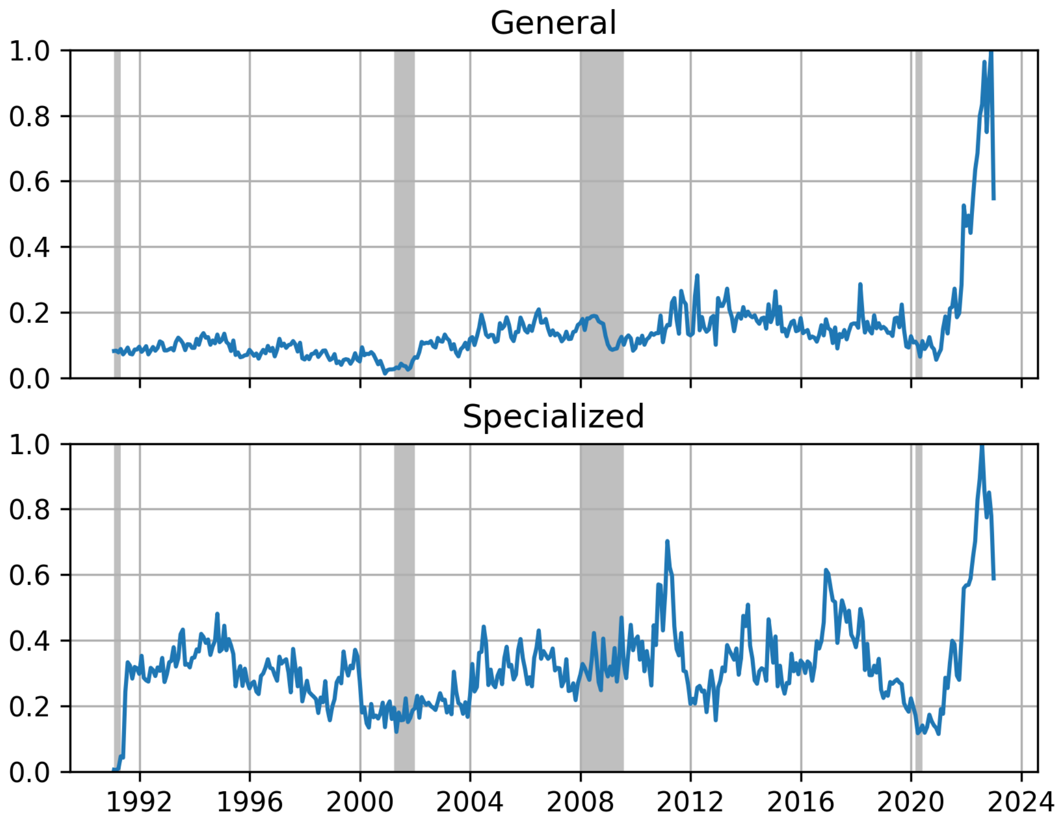

Figure 2: Inflation press coverage intensity

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 are NYT, USAT, and WaPo, whereas WSJ is the specialized newspaper. Shaded areas represent NBER recession periods. The sample includes all months between January 1991 and December 2022.*

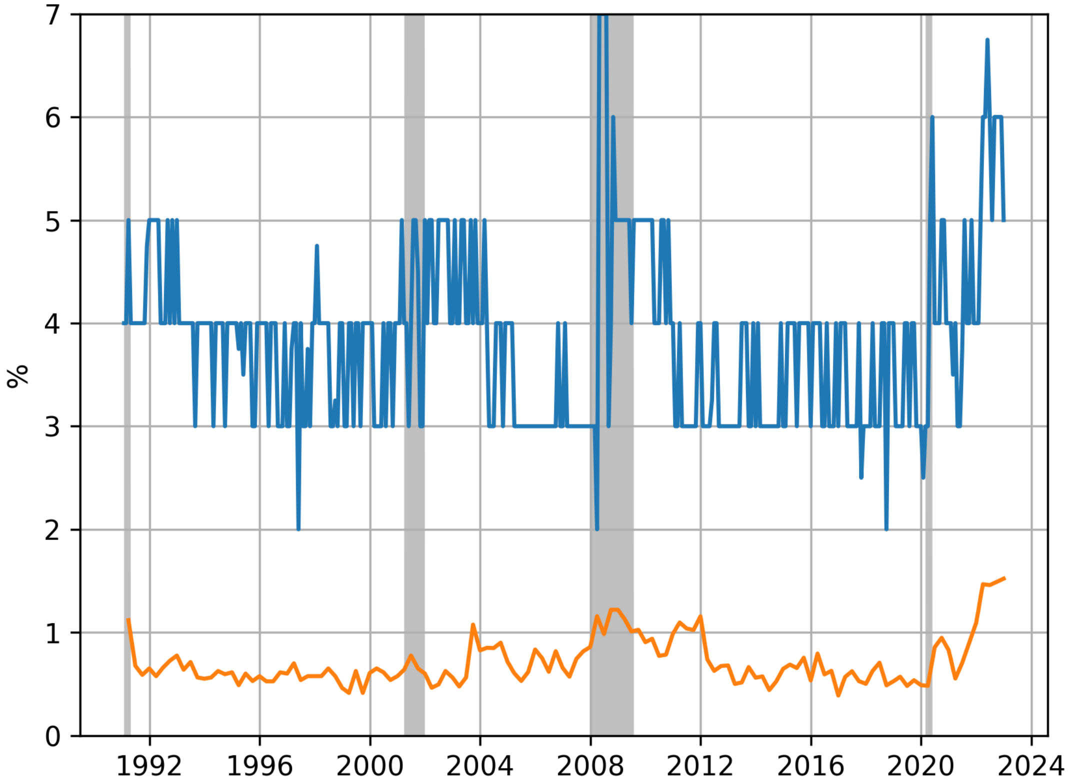

Figure 3: Inflation expectations disagreement

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

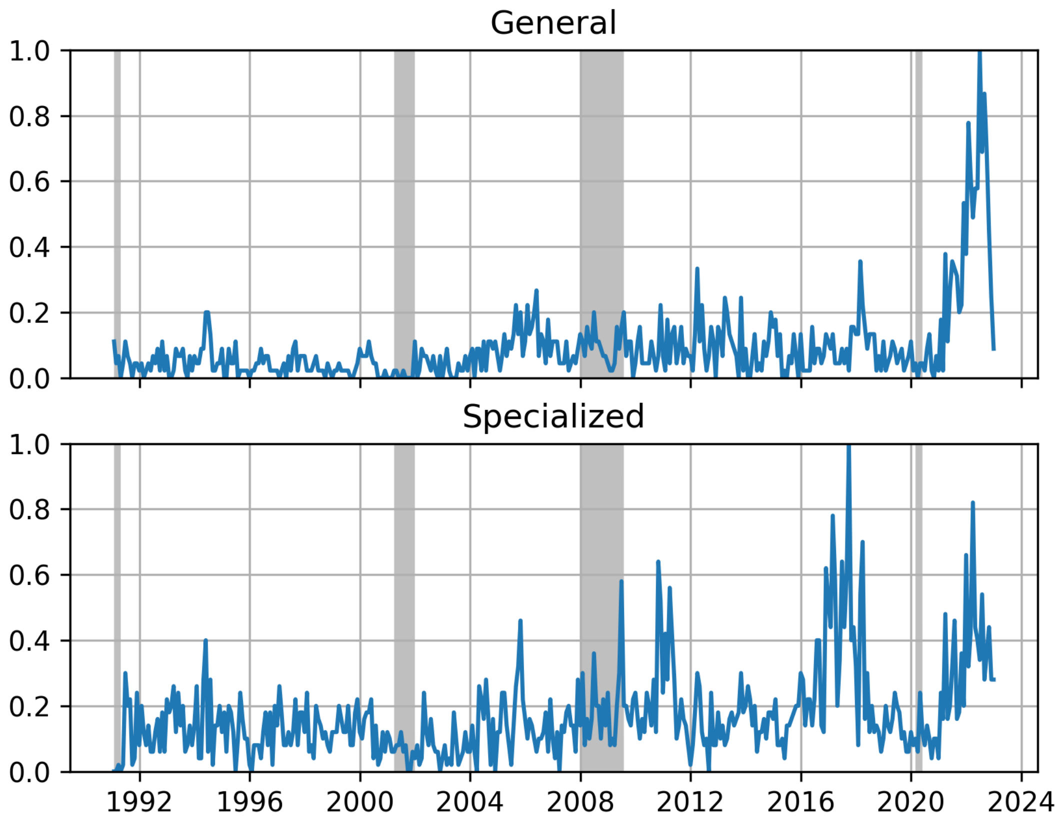

Figure 4: Causal inflation press coverage intensity

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 are NYT, USAT, and WaPo, whereas WSJ is the specialized newspaper. Shaded areas represent NBER recession periods. The sample includes all months between January 1991 and December 2022.*

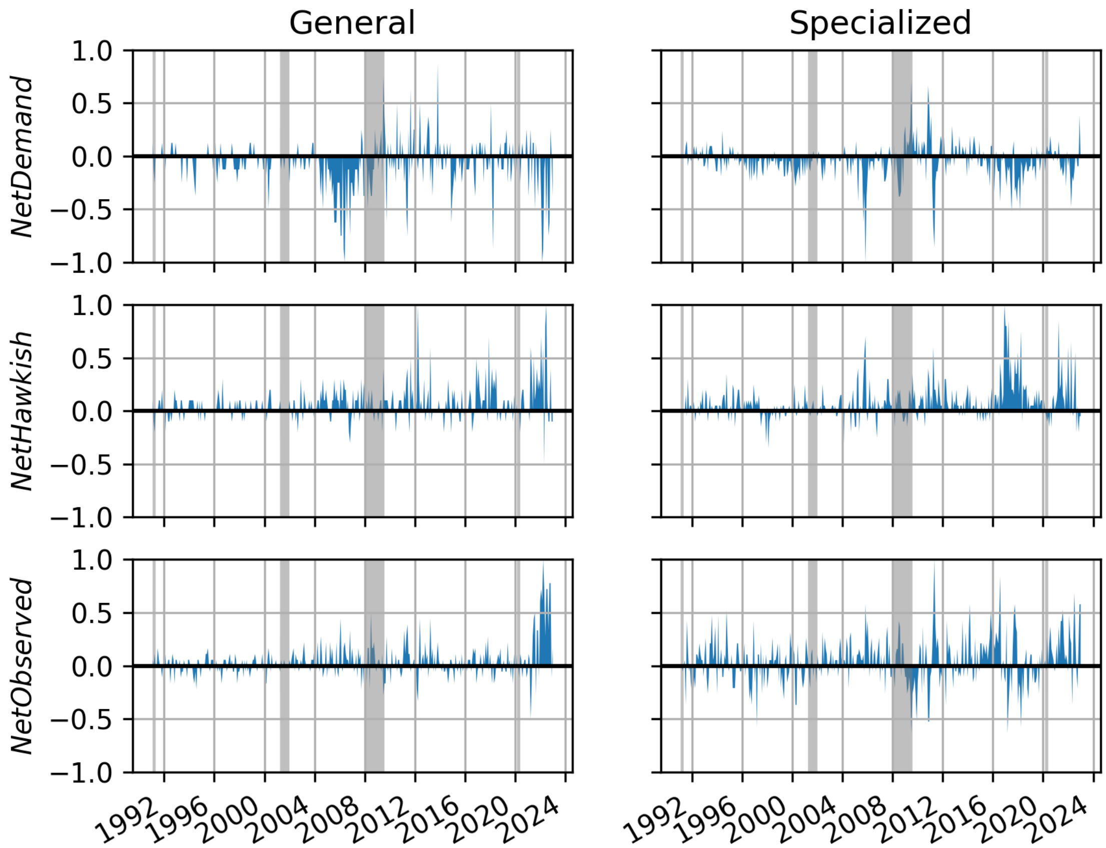

Figure 5: Demand-Supply, Hawkish-Dovish, and Observed-Expected narratives

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 3.3. NetHawkish is measured as described in section 3.4. NetObserved is measured as described in section 3.5. General newspapers are NYT, USAT, and WaPo, whereas WSJ is the specialized newspaper. Shaded areas represent NBER recession periods. The sample includes all months between January 1991 and December 2022.*

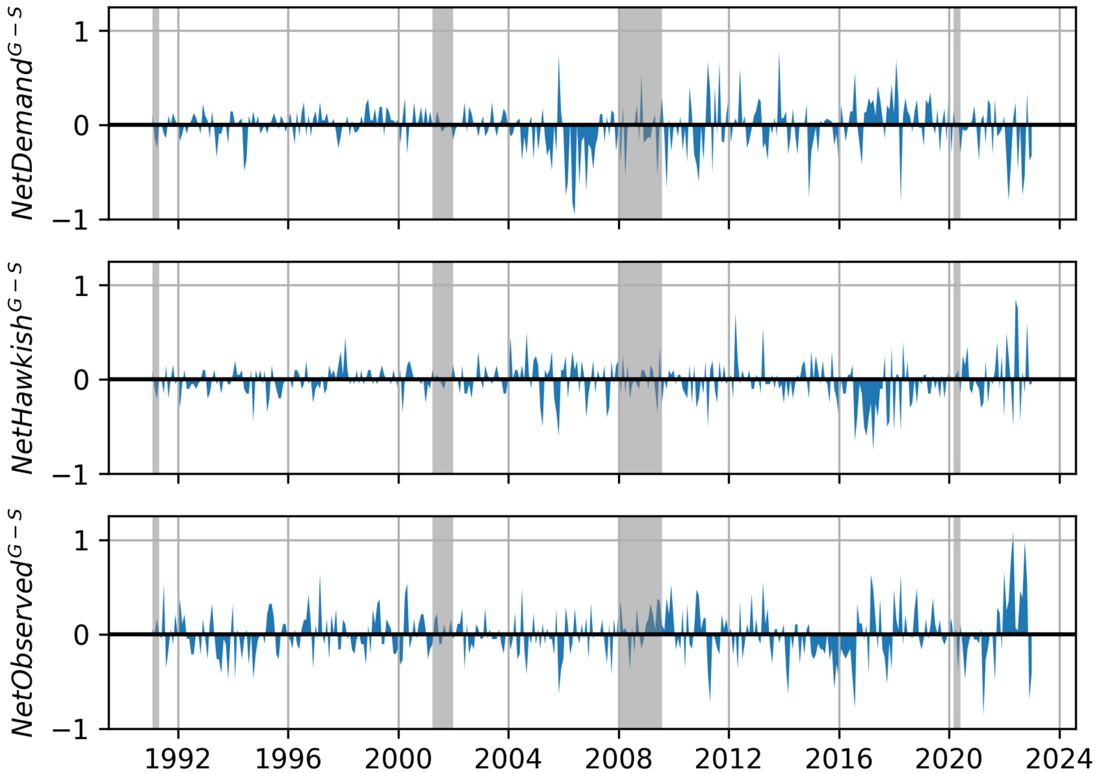

Figure 6: Disagreement about demand-supply, hawkish-dovish, and observed-expected narratives

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 3.3. NetHawkish is measured as described in section 3.4. NetObserved is measured as described in section 3.5. General newspapers are NYT, USAT, and WaPo, whereas WSJ is the specialized newspaper. Shaded areas represent NBER recession periods. The sample includes all months between January 1991 and December 2022.*

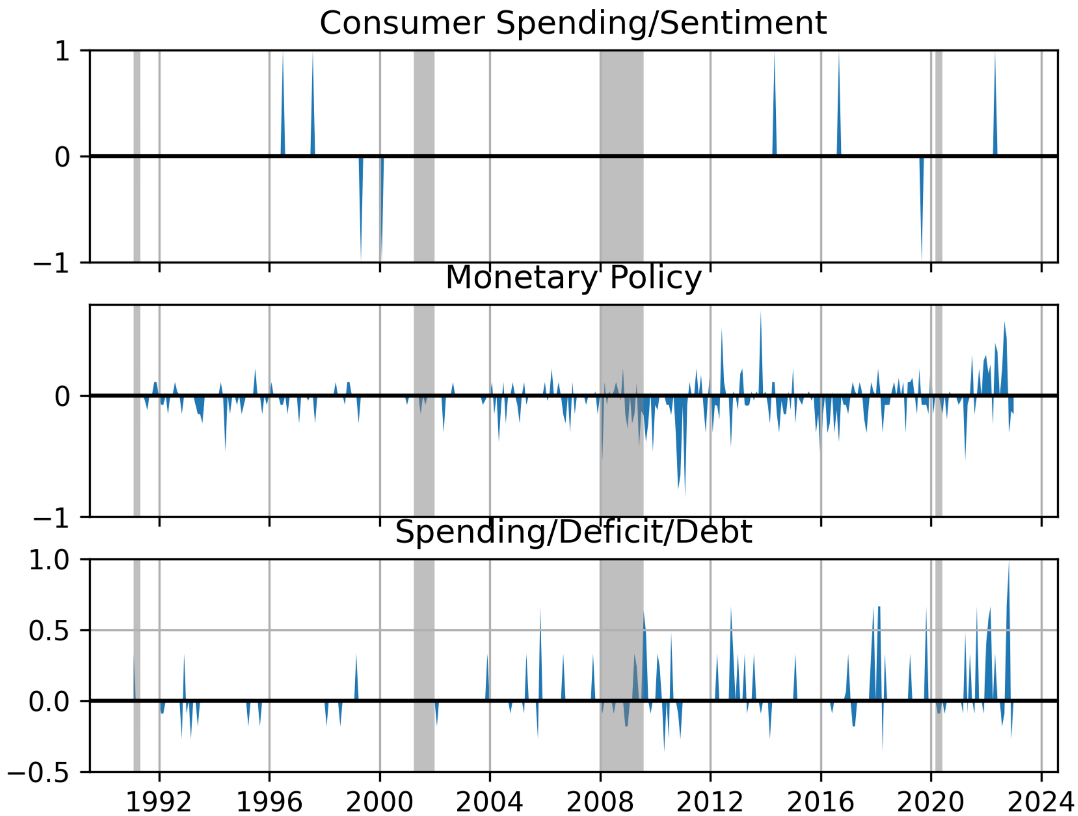

Figure 7: Disagreement about demand narratives

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 3.3. General newspapers are NYT, USAT, and WaPo, whereas WSJ is the specialized newspaper. Shaded areas represent NBER recession periods. The sample includes all months between January 1991 and December 2022.*

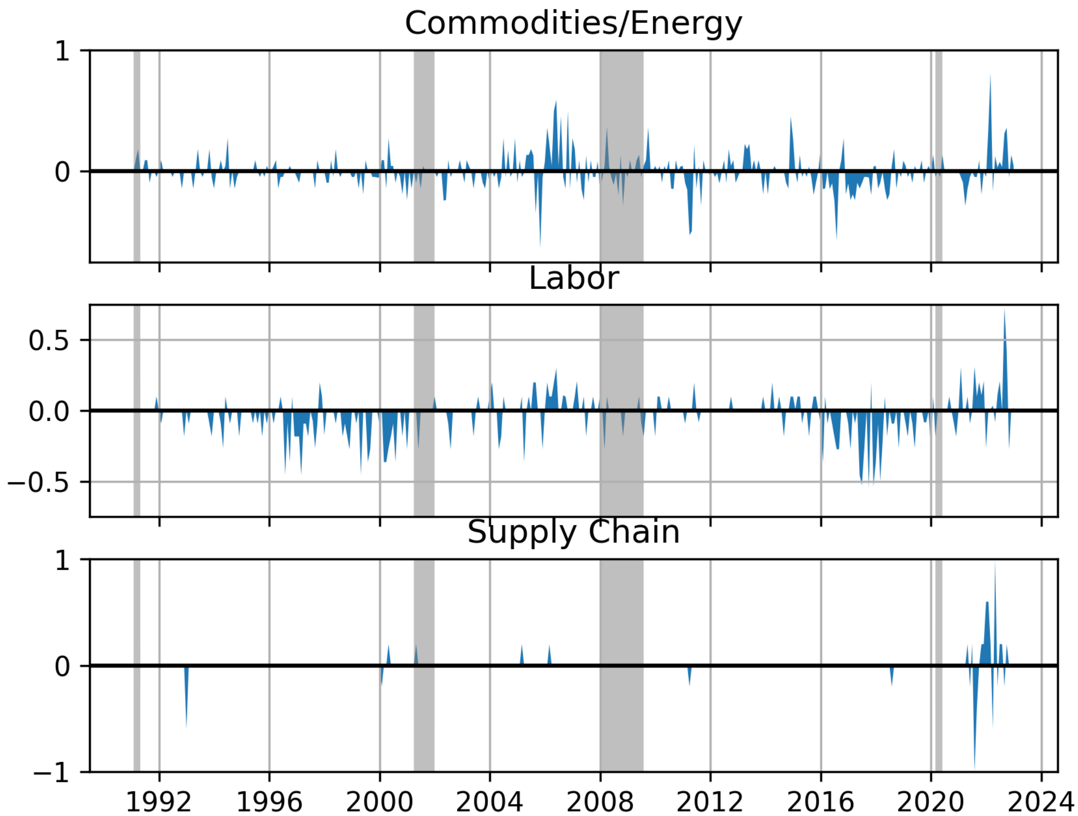

Figure 8: Disagreement about supply narratives

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 3.3. General newspapers are NYT, USAT, and WaPo, whereas WSJ is the 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
a consequence of	Before Cause
as a result of	Before Cause
because	Before Cause
because of	Before Cause
due to	Before Cause
for	Before Cause
for the sake of	Before Cause
on account of	Before Cause
on grounds of	Before Cause
on the grounds of	Before Cause
owing to	Before Cause
since	Before Causes
the conclusion of	Before Cause
the consequence of	Before Cause
the result of	Before Cause
by reason of	Before Effect
so that	Before Effect
the reason for	Before Effect
the reason why	Before Effect
why	Before Effect

Causal Verb	When
result	C-V-E (Active)
give birth to	C-V-E (Active)
activate	C-V-E (Active) or E-V-C (Passive)
actuate	C-V-E (Active) or E-V-C (Passive)
arouse	C-V-E (Active) or E-V-C (Passive)
begin	C-V-E (Active) or E-V-C (Passive)
bring	C-V-E (Active) or E-V-C (Passive)
call	C-V-E (Active) or E-V-C (Passive)
cause	C-V-E (Active) or E-V-C (Passive)
commence	C-V-E (Active) or E-V-C (Passive)
conduce	C-V-E (Active) or E-V-C (Passive)
contribute	C-V-E (Active) or E-V-C (Passive)
create	C-V-E (Active) or E-V-C (Passive)
develop	C-V-E (Active) or E-V-C (Passive)
educe	C-V-E (Active) or E-V-C (Passive)
effect	C-V-E (Active) or E-V-C (Passive)
effectuate	C-V-E (Active) or E-V-C (Passive)
elicit	C-V-E (Active) or E-V-C (Passive)
entail	C-V-E (Active) or E-V-C (Passive)
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 respectively report the true positive rate, true negative rate, positive predicted value, accuracy, and F1 score, computed as shown 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 called 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 starts from the smallest independent sentence and expands it by sequentially adding words based on their importance for 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.

Concerning the objective of the dependency parser, a dependency parser 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 OntoNotes[31], which is a large annotated corpus of various genres of text (news, conversational telephone speech, weblogs, Usenet newsgroups, broadcast, talk shows).

[31] `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 table 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 all inflation sentences where neither the manual annotator nor the CE algorithm can find an inflation narrative.
- False positives 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 all inflation sentences with a manually annotated inflation narrative but no inflation drive 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:

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FN} + \mathrm{FP}} \tag{21}$$

$$\mathrm{F1} = \frac{2 * \mathrm{TP}}{2 * \mathrm{TP} + \mathrm{FN} + \mathrm{FP}} \tag{22}$$

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

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{23}$$

$$\mathrm{TNR} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \tag{24}$$

$$\mathrm{PPV} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{25}$$

D Narratives and Expectations across Narrative Types

Table A4: Inflation narratives and expectations across narrative types

(a) Aggregate

	GAP_t				
$\lvert MonPol_{t-1}^{G-S}\rvert$	0.589**				0.598**
	(0.241)				(0.233)
$\lvert SpendDefDebt_{t-1}t^{G-S}\rvert$		0.210			0.140
		(0.212)			(0.220)
$\lvert ComEne_{t-1}^{G-S}\rvert$			0.312		0.347
			(0.342)		(0.337)
$\lvert Labor_{t-1}^{G-S}\rvert$				-0.597***	-0.652***
				(0.227)	(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

	$GAP_{i,t}$				
$\lvert MonPol_{t-1}^{G-S}\rvert$	0.591***				0.573***
	(0.048)				(0.048)
$\lvert SpendDefDebt_{t-1}t^{G-S}\rvert$		0.276***			0.209***
		(0.042)			(0.042)
$\lvert ComEne_{t-1}^{G-S}\rvert$			0.252***		0.237***
			(0.060)		(0.060)
$\lvert Labor_{t-1}^{G-S}\rvert$				-0.510***	-0.582***
				(0.049)	(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 9 and 10, as well as modified versions of them where the narrative factor disagreement measures appear one by one.* $\lvert MonPol^{G-S}\rvert$ *measures disagreement between general and specialized newspapers on Monetary Policy as an inflation narrative,* $\lvert SpendDefDebt^{G-S}\rvert$ *measures disagreement between general and specialized newspapers on Spending/Deficit/Debt as an inflation narrative,* $\lvert ComEne^{G-S}\rvert$ *measures disagreement between general and specialized newspapers on Commodities/Energy as an inflation narrative, and* $\lvert Labor^{G-S}\rvert$ *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 3.3. In the bottom panel, standard errors are clustered at the individual level as some of the respondents in the MSC are reinterviewed. The sample includes all months between 1991 and 2022.*

E Asymmetric narrative disagreement

Table A5: Inflation narratives and expectations

	GAP_t	$GAP_{i,t}$
$\lvert NetDemand_{t-1}^{G-S}\rvert * 1_{NetDemand_{t-1}^{G-S}>0}$	0.688**	0.057
	(0.292)	(0.052)
$\lvert NetDemand_{t-1}^{G-S}\rvert * 1_{NetDemand_{t-1}^{G-S}<0}$	0.225	0.168***
	(0.212)	(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: Inflation narratives and expectations

	GAP_t				$GAP_{i,t}$			
$\lvert NetDemand_{t-1}^{NYT-S}\rvert$	0.297			-0.103	0.236***			0.041
	(0.265)			(0.275)	(0.042)			(0.048)
$\lvert NetDemand_{t-1}^{USAT-S}\rvert$		0.580***		0.403***		0.395***		0.312***
		(0.185)		(0.204)		(0.038)		(0.041)
$\lvert NetDemand_{t-1}^{WaPo-S}\rvert$			0.778***	0.693***			0.364***	0.231***
			(0.229)	(0.238)			(0.041)	(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

Table A7: Newspaper and individual narratives

	$\Delta\pi_{i,t,t+12}^{MSC} * \Delta u_{i,t,t+12}^{MSC}$			$\Delta\pi_{i,t,t+4}^{SPF} * \Delta u_{i,t,t+4}^{SPF}$		
$NetDemand_{t-1}^{NYT}$	-0.064			0.047		
	(0.043)			(0.036)		
$NetDemand_{t-1}^{USAT}$		-0.084**			0.278***	
		(0.039)			(0.049)	
$NetDemand_{t-1}^{WaPo}$			-0.215***			0.364***
			(0.042)			(0.077)
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	0.49	0.49	0.50	7.73	8.21	8.11
N	161044	161044	161044	4614	4614	4614

G Expert disagreement

Table A8: Narrative and expert disagreement

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

Table A9: Inflation narratives and expectations

	GAP_t	$GAP_{i,t}$
$\lvert NetDemand^{G-S}_{t-1}\rvert$	1.204*	-0.073
	(0.660)	(0.117)
$\lvert NetDemand^{G-S}_{t-1}\rvert * var(\pi^{SPF}_{i,t,t+12})$	-1.152	0.080
	(0.918)	(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

H Incentives to gather information about inflation

Table A10: Inflation narratives and expectations

	GAP_t	$GAP_{i,t}$
$\lvert NetDemand_{t-1}^{G-S}\rvert$	0.634*	0.002
	(0.336)	(0.064)
$\lvert NetDemand_{t-1}^{G-S}\rvert * \pi_{t-1}$	-0.090	0.045**
	(0.102)	(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

I Forecast errors

Table A11: Inflation narratives and forecast errors

	FE_t		$FE_{i,t}$	
$\lvert NetDemand_{t-1}^{G-S}\rvert$	-0.037	-0.113	0.087**	-0.060
	(0.396)	(0.385)	(0.038)	(0.039)
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	14.26	16.28	2.45	3.33
N	380	380	160666	160666

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***	0.914***	-0.063***	0.933***	0.452*	0.015
	(0.104)	(0.143)	(0.019)	(0.180)	(0.231)	(0.013)
u_t	0.295***	0.185***	0.005	0.609***	-0.080	0.031***
	(0.047)	(0.061)	(0.009)	(0.089)	(0.109)	(0.007)
$\pi_t * u_t$	-0.114***	-0.118***	0.006*	-0.181***	-0.043	-0.005**
	(0.018)	(0.024)	(0.003)	(0.032)	(0.041)	(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