A Century of Inflation Narratives*

Mourad Heddaya† Chenhao Tan† Rob Voigt‡ Qingcheng Zeng§
Alexander K. Zentefis¶

September 19, 2025

Abstract

We study the evolution of U.S. inflation narratives in American newspapers since 1923. An inflation narrative is an explanation of the causes and/or effects of inflation. Using natural language processing to analyze 4.2 million sentences, we document significant shifts in narrative prevalence across economic eras. We find sharp regional differences as well, with Northern papers emphasizing fiscal causes while Southern papers focusing on supply factors and interest rate effects. Narrative framing also predicts heterogeneity in both short- and long-term consumer inflation expectations across income and education groups, with lower-income households showing greater sensitivity to narratives about the social/political consequences and cost-of-living effects of inflation. These narrative effects in some cases exceed the predictive power of realized inflation itself, suggesting exposure to different media framing may contribute to persistent gaps in inflation expectations across households.

JEL classification: C88, D83, D85, E31, L82
Keywords: inflation, newspapers, natural language processing, beliefs

*We thank Nick Barberis, John Cochrane, David Figlio, Will Goetzmann, Gary Gorton, Aniko Öry, Jung Sakong, and for their helpful comments. Part of this research was conducted while Zentefis was visiting the Hoover Institution, Stanford University.

†Department of Computer Science and Data Science, University of Chicago, 5730 S Ellis Ave, Chicago, IL 60637. Emails: mourad@uchicago.edu, chenhao@uchicago.edu.

‡Department of Linguistics, UC Davis, One Shields Avenue, Davis, CA 95616. Email: robvoigt@ucdavis.edu.

§Department of Linguistics, Northwestern University, 2000 Sheridan Rd, Evanston, IL 60208. Email: qingchengzeng2027@u.northwestern.edu.

¶Leavey School of Business, Santa Clara University; 500 El Camino Real, Santa Clara, CA 95053. Email: azentefis@scu.edu. Corresponding author.

1 Introduction

Social science researchers increasingly recognize that narratives influence beliefs, behaviors, and policy decisions (Shiller 2017, 2019). Drawing from the literature spanning from Labov and Waletzky (1967a)'s foundational work on narratives as temporal accounts to contemporary economic formalizations by Akerlof and Snower (2016) and Eliaz and Spiegler (2020), narratives can be understood as causal explanations that connect events and outcomes through interpretable sequences. These frameworks enable the public to process complex economic information and form expectations about future conditions (Akerlof and Snower 2016; Barron and Fries 2023).

Among economic phenomena, inflation stands as particularly narrative-dependent (Andre, Haaland, Roth and Wohlfart 2023). An inflation narrative can be defined as an explanation of the causes and/or effects of inflation. This focus on causality is crucial—narratives are not merely descriptions of price changes but explanations of *why* prices change and what consequences follow.

But two key dimensions of inflation narratives remain underexplored. First, how have inflation narratives evolved across time and space throughout American economic history? Understanding this evolution is crucial for interpreting how economic ideas develop and spread through public discourse during different economic regimes. Second, do specific narrative frames predict heterogeneity in inflation expectations across demographic groups? This relationship is particularly important for monetary policy effectiveness, as Coibion, Gorodnichenko and Weber (2022) demonstrate that communication strategies significantly affect inflation expectations and spending behavior. This mechanism complements D'Acunto, Malmendier, Ospina and Weber (2021)'s finding that personal shopping experiences influence inflation beliefs, indicating both media narratives and lived experiences jointly shape the inflation expectation formation process that affects economic decisions.

In this paper, we conduct a large-scale, century-spanning analysis of U.S. inflation narratives by applying natural language processing methods to American newspapers published from 1923 to 2025. To systematically identify and classify inflation narratives, we build on our methodological framework developed in Heddaya, Zeng, Tan, Voigt and Zentefis (2024). We first segment newspaper articles into sentences and filter for those containing "inflation." We then implement a two-step classification process: (1) detecting whether a sentence contains an inflation narrative, and (2) classifying identified narratives into one or more of 19 categories—8 causes (e.g., monetary, fiscal,

supply-side factors) and 11 effects (e.g., reduced purchasing power, interest rate changes, social impacts). This classification is performed using a fine-tuned, open-source large language model that achieves strong performance. The model does better than alternative approaches including keyword-based methods, which fail to capture the causal structure inherent in narratives. This comprehensive approach allows us to track how inflation is explained across different time periods, regions, and demographic contexts with fine granularity.

Our analysis reveals two main findings. First, inflation narratives have evolved substantially over the past century, with fiscal explanations dominating pre-1980 discussion, while monetary and supply-side factors gained prominence thereafter. We also document significant regional differences, with Northern papers emphasizing fiscal causes whereas Southern papers focus more on supply factors and interest rate effects. These patterns reveal how economic discourse adapts to changing institutional frameworks and regional economic structures.

Second, we analyze how narrative frames predict both short-run and long-run inflation expectations across demographic groups. We regress standardized Michigan Consumer Survey inflation expectations on lagged realized inflation, inflation news volume, and shares of different narratives, with separate specifications for various demographic groups. We find narrative frames strongly predict expectation heterogeneity, with effects in some cases exceeding the predictive power of realized inflation itself. For example, a one standard deviation increase in the share of narratives in newspapers explaining the social/political consequences of inflation predicts a 0.13 standard deviation increase for the lowest income quartile's one-year inflation expectations but minimal effect on the highest income quartile (coefficient 0.03), suggesting exposure to different media frames contributes to persistent expectation gaps. Similarly, narratives emphasizing cost-of-living effects show dramatically stronger associations with expectation dispersion than realized inflation, with social narratives predicting dispersion effects 1.8 times larger than actual inflation for short-term expectations.

These findings have important implications for monetary policy, media framing, and our understanding of inflation expectation formation. The regional differences in inflation narratives suggest that policymakers may want to consider geographic variation when crafting communication strategies, with regional Federal Reserve banks potentially monitoring local media narratives as leading indicators of expectation formation in their districts. Our findings on expectation heterogeneity suggest that the framing of inflation in news media may contribute to persistent

gaps in inflation expectations across demographic groups—gaps that may influence consumption, saving, and investment decisions. Notably, narrative effects in some cases exceed the predictive power of realized inflation itself, and lower-income households show heightened sensitivity to media narratives, suggesting that central banks control only part of the message that reaches households, with media interpretation and retransmission being equally consequential for expectation formation. The stark differences in narrative sensitivity across income and education groups point toward the need for demographic-targeted communication strategies that account for how different populations process economic information. By documenting the evolution and impact of inflation narratives over a century of American economic history, the paper provides new insights into how economic ideas are communicated to the public in the news and demonstrates the significant impact of media framing on economic beliefs.

2 Related Literature

The study of economic narratives has emerged as an important area of research examining how stories and causal explanations shape economic beliefs and behaviors. Several strands of this literature have developed distinct approaches to defining, measuring, and analyzing economic narratives.

Foundational Definitions of Narratives. Early work by Labov and Waletzky (1967b) defined narratives as temporal accounts of events, providing a structural framework for analyzing personal experience stories. More recent economic research has refined these definitions with special attention to causality. Akerlof and Snower (2016) characterize narratives as "sequences of causally linked events and their underlying sources," while Morag and Loewenstein (2023) similarly emphasize the temporal and causal dimensions: "stories [that] place selected events on a timeline and establish causal links between them." Barron and Fries (2023) highlight the interpretive aspect, defining narratives as subjective interpretations, explanations, or models of a data set that provide causal explanations for collections of events.

Theoretical Economic Models of Narratives. A significant theoretical literature has emerged modeling how narratives function in economic contexts. Eliaz and Spiegler (2020) represent narratives as directed acyclic graphs (DAGs) that "manipulate correlations between

different sets of variables," drawing on Bayesian networks to study equilibrium narrative-policy pairs. Benabou, Falk and Tirole (2018) model narratives as signals that can alter agent beliefs about tradeoffs between private benefits and social costs. Shiller's influential work (Shiller 2017, 2019, 2020) emphasizes the contagious dimension, defining narratives as "stories that offer interpretations of economic events...[that] go viral just as diseases do," proposing that narrative contagion can drive economic fluctuations.

Empirical Measurement and Analysis of Narratives. Empirical work has sought to measure and analyze narratives in various contexts. Jalil and Rua (2016) track inflation-related language in newspapers during the Great Depression to study expectation formation. Flynn and Sastry (2024) analyze sentiment in firm 10-Ks to extract narratives and incorporate them into a macroeconomic model. Andre et al. (2023) use open-ended surveys to elicit causal accounts of inflation from households and experts, using DAGs to represent these narratives. They find that household narratives significantly shape inflation expectations, with news media serving as an important transmission channel. Lange, Reccius, Schmidt, Müller, Roos and Jentsch (2022) extended the computational linguistic method of Ash, Gauthier and Widmer (2021) to extract narratives based on Roos and Reccius (2021)'s definition. Bybee, Kelly and Su (2023) extract systematic risk factors from news text using topic modeling, showing narrative-based factors outperform traditional models in asset pricing. More recently, Trebbi (2025) studies how demand-supply narrative disagreement between general and specialized newspapers can explain households' absolute gap in inflation expectations with experts.

Expectation Formation and Household Beliefs. Recent work has specifically examined how inflation news and narratives shape inflation expectations. Coibion et al. (2022) conduct a randomized controlled trial showing that different forms of monetary policy communication have varying effects on inflation expectations and subsequent household spending. Their findings suggest that how central banks communicate to the broader public matters significantly. Lamla and Lein (2014) further demonstrate that both the volume and tone of media coverage significantly influence consumers' inflation expectations, with neutrally toned news improving forecast accuracy while negatively framed reporting can impair it. Building on this evidence, Ryngaert, Binder and Frank (2025) find that increased media coverage raises expectations with effects peaking

within a few days and fading after approximately 10 days. Chahrour, Shapiro and Wilson (2025) show that media's systematic selection of reporting topics creates asymmetric effects on household inflation expectations, with high-inflation news having larger impacts than low-inflation news. Complementing these studies, D'Acunto et al. (2021) demonstrate that consumers rely heavily on price changes in their personal grocery bundles when forming expectations about aggregate inflation, with frequency of purchase and positive price changes having outsized influence. As comprehensive surveys by Weber, d'Acunto, Gorodnichenko and Coibion (2022) and Kose, Matsuoka, Panizza and Vorisek (2019) document, subjective inflation expectations are shaped by a complex interplay of information sources, cognitive limitations, and institutional factors that vary significantly across demographic groups and economic contexts.

This Paper's Contribution. Our research extends this literature by examining the evolution of U.S. inflation narratives in American newspapers since 1923, providing a much longer historical perspective than previous studies. Unlike work focused on contemporary periods or specific episodes, our century-long analysis allows us to identify persistent patterns in narrative formation and transmission across diverse inflation regimes. While Andre et al. (2023) examine narratives in survey responses and Jalil and Rua (2016) track keyword frequency, we employ advanced natural language processing techniques to systematically classify causal explanations in news media at scale. Furthermore, our research distinctively focuses on two dimensions: (1) the geographic distribution of inflation narratives through time, and (2) whether specific narrative frames predict heterogeneity in inflation expectations across demographic groups. This comprehensive approach bridges theoretical models of narrative formation with empirical measurement of their real-world impacts on inflation expectations.

Outline. This paper proceeds as follows. In Section 3, we describe our data sources. Section 4 details our methodology for extracting inflation narratives from newspapers, drawing from our earlier work (Heddaya et al. 2024). Section 5 presents our main findings, examining inflation narratives across time and space (Section 5.1) and their relationship with inflation expectations across demographic groups (Section 5.2). Section 6 concludes.

3 Data

3.1 Newspapers

We draw on several data sources. For newspaper articles, we use ProQuest's newspaper database, which provides access to U.S. newspaper articles. Our sample ranges from January 1923 to January 2025, covering all 50 states plus DC and Guam. It includes 807 unique newspapers, from major national publications such as *The New York Times*, *Wall Street Journal*, and *Washington Post* (46 in total), to hundreds of local and regional newspapers (761 in total). We distinguish national newspapers/newswires from local publications for some of the analyses. The list of national publications is in Table 1, with roughly 47% of the total number of inflation sentences originating from national sources. Over the full sample period, our dataset averages 81.6 newspapers per month, with 8.6 national newspapers and 73.0 local newspapers per month on average. The coverage varies substantially over time, ranging from a minimum of 13 newspapers (4 national, 9 local) in early months to a maximum of 324 newspapers (24 national, 300 local) during peak coverage periods.

3.2 Economic Variables

For inflation expectations, we use the University of Michigan Survey of Consumers, which surveys household expectations about future price changes. For realized inflation, we use Consumer Price Index (CPI) data from the U.S. Bureau of Labor Statistics, retrieved through the Federal Reserve Economic Data (FRED) system. This includes both national inflation measures and regional CPI data for the Northeast, Midwest, South, and West regions, allowing us to examine potential geographic patterns in both inflation narratives and actual price changes.

4 Extracting Inflation Narratives from Newspapers

To extract inflation narratives from newspapers, we apply the classification approach we developed in our previous work (Heddaya et al. 2024). This methodology enables us to methodically identify and categorize sentence-level explanations of the causes and effects of inflation across our corpus.

4.1 Narrative Framework

We define an inflation narrative as a sentence-level explanation of the cause(s) and/or effect(s) of inflation. This granular approach allows us to identify narratives within individual sentences that capture implicit or explicit cause-effect relationships around inflation. These sentence-level narratives reveal how newspapers conceptualize the economic mechanisms behind inflation and its consequences.

It is important to note that our classification framework identifies the presence of causal relationships between factors and inflation, but does not take a stance on directional effects (i.e., whether a particular cause increases or decreases inflation, or whether inflation increases or decreases a particular effect). While we did collect directional information during annotation, we exclude it from our analysis for simplicity. Our primary research objective is to track which explanatory mechanisms dominate public discourse across regions and time, not to determine whether each factor is described as increasing or decreasing prices. Analyzing the directional effects of these causal relationships would be an interesting avenue for future work.

Our approach uses a comprehensive ontology of causes and effects specific to inflation. We developed the ontology, detailed in the next section, consisting of 8 causes of inflation and 11 effects that could follow from inflation. We curated this ontology based on a combination of domain knowledge, online searches, and LLM interactions. When using a LLM (Open AI ChatGPT 3.5, Google Gemini, Anthropic Claude), the prompt was: "What are the causes (effects) of inflation? Describe the economic mechanisms and give examples." If we wanted to expand on a cause (effect), the prompt was: "Explain the economic mechanisms and examples of X as a cause (effect) of inflation." The ontology of the causes and effects with a brief description of each follows.

4.2 Inflation Narrative Categories

Our classification framework organizes inflation narratives into a comprehensive ontology of causes and effects. Each category represents a distinct explanatory framework through which inflation is understood and communicated in media discourse.

4.2.1 Causes of Inflation

1. **Demand-side Factors (demand):** Pull-side or demand-pull inflation narratives attribute rising prices to excessive aggregate demand in the economy, often due to increased consumer spending, business investment, or government expenditure that outpaces productive capacity.
2. **Supply-side Factors (supply):** Push-side or cost-push inflation narratives focus on constraints in production capacity, supply chain disruptions, or resource shortages that drive up costs independent of demand conditions.
3. **Built-in Wage Inflation (wage):** Also known as wage-price spiral narratives, these explanations center on the self-reinforcing cycle where workers demand higher wages to compensate for rising prices, which in turn leads businesses to raise prices further.
4. **Monetary Factors (monetary):** These narratives highlight central bank policies as primary drivers of inflation, especially through excessive money supply growth, low interest rates, or quantitative easing measures.
5. **Fiscal Factors (fiscal):** Government spending and taxation policies are emphasized in these narratives, which typically link budget deficits, stimulus programs, or expansionary fiscal policy to rising inflation.
6. **Expectations (expect):** These narratives focus on how anticipation of future inflation can itself cause inflation, as economic actors make decisions based on their inflation expectations rather than current conditions.
7. **International Trade & Exchange Rates (international):** Cross-border factors including currency exchange fluctuations, global commodity prices, international capital flows, or trade imbalances feature in these explanations.
8. **Other Causes (other-cause):** Additional causal explanations that don't fit the above categories.

4.2.2 Effects of Inflation

1. **Reduced Purchasing Power (purchase):** Narratives emphasizing how inflation erodes the buying power of money, reducing what consumers can purchase with the same nominal amount.
2. **Cost of Living Increases (cost):** These narratives focus on how inflation raises everyday

expenses, highlighting impacts on individuals with fixed incomes, pensioners, and lower-wage earners.

3. **Uncertainty Increases (uncertain):** Explanations centering on how inflation, especially when volatile or unpredictable, creates economic uncertainty that affects planning, investment, and consumption decisions.
4. **Interest Rates Raised (rates):** Narratives that highlight how central banks respond to inflation by raising interest rates, with subsequent effects on borrowing costs, investment, and economic growth.
5. **Income or Wealth Redistribution (redistribution):** These narratives focus on how inflation redistributes wealth between different economic groups, such as borrowers versus lenders, or wage earners versus asset holders.
6. **Impact on Savings (savings):** Explanations emphasizing how inflation affects various forms of savings and financial investments, underscoring the erosion of value in fixed-return assets.
7. **Impact on Global Trade (trade):** Narratives discussing how domestic inflation affects international trade competitiveness, export performance, or trade balances.
8. **Cost-Push on Businesses (cost-push):** These narratives focus on how businesses respond to inflationary pressures, including impacts on profit margins, investment decisions, employment, and pricing strategies.
9. **Social and Political Impact (social):** Explanations that highlight broader societal consequences of inflation, including effects on social stability, political developments, or public trust in institutions.
10. **Government Policy & Public Finances Impact (govt):** Narratives focusing on how inflation affects government programs, public debt servicing, or fiscal policy options.
11. **Other Effects (other-effect):** Additional consequences of inflation that don't fit into the above categories.

4.3 Example Narratives

Fig. 4 shows examples of how inflation narratives are identified and classified in our framework. It displays two sample sentences about inflation. The first sentence, "A year ago the administration assumed inflation would run at 7.5 percent in 1979," is labeled as having "No narrative" because

it merely states a fact without attributing causes or effects to inflation.

The second example, "If these fiscal policies materialize, they are likely to boost economic growth and spur inflation, potentially forcing the Fed to hike rates more quickly to keep up," contains two distinct narratives:

1. Narrative 1: "Fiscal Factors" (cause) → "Inflation"
2. Narrative 2: "Inflation" → "Interest Rates Raised" (effect)

This illustrates how the framework identifies and categorizes inflation narratives by analyzing the causal relationships expressed in sentences, showing both what causes inflation and what effects inflation has.

4.4 Data Preprocessing

Following our earlier approach, we segment all articles from the ProQuest corpus into sentences and filter for those containing the keyword "inflation." This initial filtering yields a corpus of approximately 4.2 million sentences spanning a century of economic reporting. Fig. 4 illustrates a time series of the volume of inflation sentences, revealing distinctive patterns in media attention to inflation across major economic episodes throughout the past century. For sentence segmentation, we use BlingFire, which we found to be effective for processing news text with its varied formatting and typographical conventions.

4.5 Classification Approach

To identify and categorize inflation narratives across our large corpus of news articles, we developed an automated classification system using machine learning techniques. This approach allows us to process millions of sentences while maintaining high accuracy in identifying causal narratives about inflation.

4.5.1 Two-Stage Classification Framework

Our classification system operates in two sequential stages, designed to first filter relevant content and then categorize the narratives present. In the first stage, the model determines whether a sentence contains any causal narrative about inflation. This binary classification step distinguishes between sentences that merely mention inflation factually (e.g., "Inflation rose to 3.2%

in March") versus those that express causal relationships (e.g., "Rising energy costs are driving inflation higher"). This filtering step is crucial because only about half of inflation-mentioning sentences actually contain causal narratives, making this preliminary screening essential for focused analysis.

For sentences identified as containing narratives, the model then classifies them into one or more of our 19 narrative categories consisting of 8 causes and 11 effects. Since sentences can express multiple causal relationships at once, this constitutes a multi-label classification problem. For example, a sentence might express both that fiscal policy causes inflation and that inflation leads to higher interest rates.

4.5.2 Model Architecture and Training

We employed a fine-tuned open-source large language model (Llama 3.1 8B) for this classification task. Large language models have demonstrated superior performance on complex text understanding tasks compared to traditional machine learning approaches, especially for nuanced tasks like identifying implicit causal relationships in natural language.

Our training data consisted of approximately 2,100 annotated sentences from two distinct corpora: contemporary news articles spanning 2010-2023 from the News on the Web corpus (Davies 2016) and historical articles covering 1960-1980 from ProQuest. This dual-corpus approach ensures our model can handle both modern and historical language patterns. The training sets comprised approximately 1,000 examples from each corpus, providing sufficient data for effective model fine-tuning while maintaining computational feasibility.

Three expert annotators independently labeled each sentence following detailed annotation guidelines that define our 19 narrative categories. The annotation process achieved strong inter-annotator agreement with Krippendorff's alpha values of 0.80 for narrative detection and 0.66 for multi-class classification. The higher agreement on binary detection compared to multi-class classification reflects the inherent complexity of distinguishing between semantically similar narrative categories, though both metrics indicate reliable annotation quality for training purposes.

Rather than using the model out-of-the-box, we fine-tuned it specifically for our inflation narrative task. Fine-tuning involves adjusting the model's internal parameters based on our training data, adapting it from general language understanding to our specific classification objectives. We employed standard machine learning practices including a learning rate of 1e-4,

AdamW optimizer, 600 training steps, and an effective batch size of 16. The model was trained to focus on predicting the correct narrative classifications while ignoring formatting elements, concentrating learning on the substantive classification task.

4.5.3 Testing Protocol and Ground Truth Establishment

Our testing approach required careful consideration of how to handle disagreements between human annotators when establishing ground truth labels. For the test sets, all three annotators labeled the same subset of data: 201 sentences from contemporary news and 488 sentences from historical news (reduced from an initial 500 after filtering sentences longer than 150 words where sentence segmentation failed).

To resolve annotator disagreements and establish reliable ground truth labels, we employed majority rule across the three annotators. In practice, 97% of test set instances achieved agreement between at least two annotators, allowing us to retain nearly the entire test set for evaluation. The remaining 14 test instances with no majority agreement were excluded from performance calculations. This approach balances the need for reliable ground truth with the pragmatic reality that even expert annotators disagree on ambiguous cases, particularly when dealing with implicit causal relationships or subtle linguistic cues.

The use of majority rule also provides insight into the inherent difficulty of the classification task. Most annotator disagreements stemmed from differing judgments about whether a narrative existed in the text at all, rather than disagreements about which specific narrative category to apply. This pattern informed our two-stage classification approach and highlighted the importance of the initial narrative detection step.

4.5.4 Performance and Validation

Our final model achieved strong performance metrics that demonstrate its reliability for large-scale analysis. For narrative detection, we obtained F1 scores of 0.87 on contemporary data and 0.78 on historical data. For narrative classification, the model achieved F1 scores of 0.71 on contemporary data and 0.62 on historical data. These F1 scores, which range from 0 to 1 with 1 representing perfect accuracy, indicate strong performance for complex text classification tasks, with scores above 0.7 considered robust for such applications. The lower performance on historical data reflects the challenges of processing older language patterns and different journalistic styles.

A critical validation involved testing the model's ability to generalize across different time periods. When we trained on one era and tested on another, performance declined by only 3-4 percentage points, confirming the model's robustness to linguistic changes over our century-long analysis period. This cross-temporal validation was essential given our goal of analyzing narrative evolution from 1923 to 2025. Additionally, our fine-tuned model outperformed larger general-purpose models like GPT-4o when applied to our specific task, demonstrating the value of domain-specific training while also being more computationally efficient for processing millions of sentences.

4.5.5 Error Analysis and Model Limitations

To understand model limitations, we conducted detailed error analysis comparing model predictions to human annotations. Model errors often mirrored human annotator disagreements, especially for implicit causal relationships and social or political impacts of inflation. This suggests the model captures genuine ambiguities in language rather than failing in arbitrary ways. The model struggled most with sentences requiring external context to interpret causal relationships, performing well on explicit cause-effect statements. This limitation is important for interpreting results, as subtle or highly contextual narratives may be underrepresented in our analysis.

For our full corpus analysis, we applied our best-performing model—trained on the combined historical ProQuest (1960-1980) and contemporary News on the Web (2010-2023) datasets—to process the entire ProQuest corpus spanning 1923 to 2025. We conducted additional validation by reviewing random samples of classified sentences across different time periods, ensuring consistent quality throughout our historical analysis. This classification approach enables us to identify and analyze inflation narratives across over 4 million sentences spanning more than a century.

4.6 Narrative vs Keyword Classification

To evaluate our narrative classification against simpler methods, we compare it with keyword-based classification. Table 3 presents keywords chosen to reflect the most prominent words associated with each narrative category. Because the vectors of labeled sentences are sparse under both classifications, we compute the Jaccard similarity between the two instead of the Pearson correlation. The magnitude signifies the percentage of sentences labeled with either classification

is labeled with both. Particularly, a value of 0 similarity indicates that the two classifications have no sentences in common, and a value of 1 implies perfect overlap.

The similarity between these approaches varies across categories (Fig. 2). Fig. 2A displays the sentence-level similarity between each narrative category and only its associated keywords, providing a direct measure of how well simple keyword matching identifies each narrative type. Interest rate narratives show the highest similarity with keywords (0.45), while international trade narratives show the lowest (0.06). This range demonstrates how certain inflation discussions use consistent terminology, while others employ diverse language that keywords cannot capture.

Figure 2B presents the full similarity matrix between all narrative categories and all keyword sets, revealing patterns of cross-similarity. The matrix shows stronger similarities along the diagonal as expected, but also reveals high off-diagonal values. For example, monetary narratives share similarity with fiscal keywords, and cost-push narratives are similar with cost keywords. The "No narrative" category shares similarity with multiple keyword classifications, revealing a key limitation: keyword approaches often classify sentences that mention inflation-related terms without expressing causal relationships.

These patterns justify our narrative classification approach over keyword methods. While keyword classification offers simplicity, it lacks the capacity to identify causal structures in sentences and distinguish between mere mentions and actual explanations. The narrative classification approach captures these nuances, providing deeper insight into how inflation is explained across sources and time periods.

4.7 Alternative Narrative Discovery Approach

While manually curated taxonomies offer the advantage of domain expertise, they face inherent limitations that may compromise their comprehensiveness and objectivity. Expert-driven classification systems risk introducing researcher bias, potentially overlooking narrative categories that emerge organically from the data or emphasizing frameworks that align with existing theoretical expectations rather than empirical patterns. Additionally, manual taxonomies may inadvertently reflect the particular disciplinary backgrounds or conceptual frameworks of their creators, potentially missing causal relationships that are prominent in actual discourse but less salient in academic literature.

Data-driven approaches to narrative discovery offer a complementary perspective that can

validate expert taxonomies while identifying gaps or alternative organizational structures present in the corpus itself. Unlike traditional topic modeling methods such as Latent Dirichlet Allocation (LDA) (Blei, Ng and Jordan 2003) or BERTopic (Grootendorst 2022), which cannot incorporate explicit instructions about causal relationships and often produce generic thematic clusters, TopicGPT provides a framework specifically designed to extract interpretable, structured topics from unstructured text (Pham, Hoyle, Sun, Resnik and Iyyer 2024). This approach leverages large language models' capacity to understand causal relationships and generate human-interpretable category descriptions, making it particularly well-suited for discovering the nuanced cause-effect structures that characterize inflation narratives.

To validate our human-constructed narrative taxonomy, we implemented an unsupervised causal narrative discovery approach based on the TopicGPT framework. This computational method can automatically identify causal narratives directly from text corpora without relying on predefined categories, allowing us to assess whether our ontology captures the full range of inflation causal relationships presented in newspapers.

Our discovery approach operates in two stages: taxonomy generation and refinement. In the generation phase, we randomly sampled 10,000 sentences from our corpus and employed asynchronous calls to a large language model (GPT 4o, November 20, 2024 update) to generate individual causal ontologies for each sentence. This comprehensive approach, processing each sentence individually rather than relying on a smaller representative subset, produced 10,000 distinct causal ontologies that thoroughly capture the range of narrative structures in our dataset. In the refinement phase, we used batch-level processing to consolidate these individual ontologies. We iteratively fed 200 first-step narratives into the LLM and obtained refined versions, continuing this process until the total number of narratives was reduced to fewer than 200. A final refinement round ultimately identified 11 distinct causal narratives (Table 4) from the original set of 10,000.[1]

The unsupervised discovery results (Fig. 3) reveals two key findings: (i) no novel causal mechanisms emerged beyond our original taxonomy, and (ii) discovered categories generally consolidated our granular classifications into broader clusters.

Fig. 3A displays semantic similarity scores between our 19 manual categories and TopicGPT's 11 discovered categories, measuring conceptual alignment at the category definition level. Semantic similarity was determined using an LLM (GPT 4o, Nov. 20, 2024 update), with the prompt

[1]The prompts we used at each stage are presented in Appendix A.

presented in Appendix A.4. The highest scores occur between wage and "Labor Market Pressures" (8), demand and "Demand-Supply Imbalance" (9), and monetary and "Monetary and Fiscal Drivers" (9). All TopicGPT categories map to existing frameworks in our original ontology.

Fig. 3B presents Jaccard similarity coefficients measuring sentence-level overlap between the two classification schemes. Peak coefficients include wage/"Labor Market Pressures" (0.56), supply/"Supply-Side Constraints" (0.51), and rates/"Economic Slowdown Risks" (0.49). The figure reveals systematic consolidation: "Exchange Rate and Inflation Expectations" draws from expectations (0.27) and international (0.22) categories, while "Rising Living Costs and Income Inequality" combines cost (0.40) and redistribution (0.23) classifications. TopicGPT consistently aggregates multiple granular categories into broader thematic clusters rather than identifying distinct causal mechanisms.

These results imply that our manual classification generally captures the full range of inflation explanations found in our corpus. While the automated approach covers the same conceptual ground, it groups related causes together rather than maintaining important economic distinctions. This validation supports using our more detailed classification system, which preserves economically relevant differences that automated methods tend to blur together.

5 Main Results

In this section, we present our main analysis of inflation narratives. We first examine the temporal evolution of different narrative types across the century in Section 5.1, documenting how the prevalence of various cause-effect explanations changed during different economic periods. Next, in the same section, we explore the spatial variation in these narratives, analyzing regional differences in how inflation is explained in the media. In Section 5.2, we investigate the relationship between narrative prevalence and survey-based measures of inflation expectations, testing whether shifts in media discourse predict changes in household expectations.

5.1 Inflation Narratives Through Time and Space

In this section, we present our analysis of how inflation narratives have evolved over time and varied across U.S. geographic regions. We first examine the prevalence of different narrative types across our century-long sample, documenting which causal explanations and effect discussions

dominated public discourse during different economic periods. We then explore geographic variation in narrative emphasis, analyzing how inflation is explained differently in Northern versus Southern newspapers, urban versus rural publications, and across newspapers from areas with different political orientations. These patterns reveal not only the changing economic understanding of inflation through time but also persistent regional differences in how inflation's causes and consequences are framed for public consumption.

5.1.1 Narrative Prevalence

Our analysis of inflation narratives from 1923 to 2025 reveals several important patterns in how inflation is discussed in U.S. newspapers. About 54.5% of inflation-related sentences in our corpus do not express any specific narrative about inflation's causes or effects, suggesting that the majority of media coverage typically reports on inflation without exploring its underlying mechanisms or consequences.

Table 5 reveals interesting differences in inflation coverage between national and local newspapers. National publications tend to emphasize certain causal narratives more strongly, primarily fiscal (9.3% vs. 7.8%), monetary (5.2% vs. 3.9%), and international factors (2.2% vs. 1.3%). Overall, national newspapers dedicate substantially more coverage to causal narratives (26.6% vs. 21.4% in local papers). However, local newspapers give relatively greater attention to certain direct effects of inflation, such as impacts on savings (6.4% vs. 5.6%), cost of living (3.9% vs. 2.8%), and purchasing power (1.7% vs. 1.1%), although the total coverage of effect narratives is similar between local (25.4%) and national papers (24.6%). These patterns suggest that while national outlets focus more on macroeconomic policy discussions and the underlying causes of inflation, local papers emphasize practical implications that directly affect their communities. Note that the slight overlap in percentages (exceeding 100% when combined with the "None" category) indicates that some sentences express multiple narratives simultaneously.

The time series plots in Fig. 5 reveal substantial time variation in narrative prominence across all newspapers in our sample. Panel A presents the top three most prominent cause narratives. The fiscal causes of inflation narrative dominated discussions in the pre-1980 period, with pronounced peaks during the Great Depression, World War II, and the inflationary episode of the 1970s, when it reached nearly 25% of inflation sentences. Since the 1980s, however, fiscal causal explanations have declined in prominence, with monetary and supply-side factors becoming relatively more

important in the inflation discourse.

The evolving pattern of effect narratives in Panel B shows a significant shift in how inflation consequences are framed over time across the entire newspaper sample. While savings effects consistently appeared in discussions throughout the century, interest rate effects became substantially more prominent after 1970. This coincides with the Federal Reserve's more aggressive monetary policy stance under Volcker and the subsequent era of inflation targeting. Before 1970, interest rate effects were nearly absent from inflation discussions, highlighting a fundamental change in how the media expresses inflation's implications. During the post-COVID inflation episode of 2021-2022, we observe a pronounced spike in cost-of-living effect narratives, reflecting heightened public concern about inflation's direct impact on household expenses during this period of rapidly rising prices for essential goods and services.

5.1.2 Regional Differences

To investigate how inflation narratives differ across regions, we analyze two key dimensions: geographic location (North vs. South) and level of urbanization (Urban vs. Rural). Our classification methodology assigns each newspaper to these categories based on its headquarters's geographic position relative to the Mason-Dixon line, the population density of its location, and the historical voting patterns of its area, respectively. To ensure we capture genuine regional differences rather than national editorial trends, we exclude national newspapers (listed in Table 1) and focus exclusively on local publications across the sample period.

North vs. South. The geographic analysis reveals differences in how Northern and Southern newspapers frame inflation narratives. Figure 6 illustrates these differences across both cause and effect narratives.

Northern newspapers show a higher propensity to attribute inflation to fiscal policy, with approximately a 0.92 percentage point higher share of fiscal narratives compared to Southern newspapers. International and inflation expectations also feature more prominently in Northern publications, though the differences are small.

In contrast, Southern newspapers emphasize supply factors as causes of inflation, with approximately 0.26 percentage points higher share compared to Northern publications. Southern publications also show slightly higher emphasis on monetary, demand, and wage-based explanations

of inflation, though these differences are mildly smaller in magnitude.

For effect narratives, the most notable difference is in the discussion of interest rates, which Southern newspapers emphasize substantially more (1.5 percentage points higher) than Northern counterparts. Southern publications also discuss costs of living, uncertainty, and purchasing effects of inflation more frequently. Northern newspapers, conversely, focus more on government program (0.48 percentage points higher) savings effects (0.33 percentage points higher).

These patterns suggest differences in regional economic concerns, with Northern regions more focused on government fiscal policy as both a cause and solution to inflation, while Southern regions place greater emphasis on supply constraints and the impact of inflation on interest rates.

Urban vs. Rural. Urban-rural differences in inflation narratives reveal distinct patterns that may reflect the economic structures of these areas. Figure 7 demonstrates these contrasts.

Urban newspapers emphasize fiscal factors when discussing causes of inflation, with a 3.4 percentage point higher share compared to rural publications. International factors, monetary policy, trade effects, and wage pressures also receive more attention in urban newspapers. Rural newspapers do not show significantly higher emphasis on any causal narrative categories.

In effect narratives, a sharp contrast appears in the discussion of interest rates and costs of living, which rural newspapers emphasize more heavily (3.3 and 2.0 percentage points higher, respectively). Urban newspapers, however, focus more on savings effects (2.4 percentage points higher) and modestly more on social and effects on trade.

These differences likely reflect the economic makeup of urban versus rural areas, with urban areas more connected to global markets and government policy, while rural areas may experience more direct impacts of inflation on borrowing costs and consumer prices.

5.2 Narratives and Household Expectations

Understanding how consumers form inflation expectations is a central question in macroeconomics with significant implications for monetary policy effectiveness, household financial decisions, and overall economic stability. Inflation expectations influence wage negotiations, investment choices, and consumption patterns, making them a critical transmission mechanism for economic policy. While previous work emphasize the role of past inflation (Gaspar, Smets and Vestin 2010), consumption bundle price changes (D'Acunto et al. 2021), and official cen-

tral bank communications (Coibion et al. 2022) on inflation expectation formation, the news media—and specifically the narrative frames through which inflation is discussed—may materially shape public perceptions and expectations (Carroll 2003; Lamla and Lein 2014). This section empirically investigates how different inflation narratives predict consumer inflation expectations. By examining how media coverage translates economic data into cause and effect stories that resonate with different audiences, we gain insight into the persistent gaps in inflation expectations across demographic groups.

5.2.1 Empirical Design

Our analysis employs a uniform framework to investigate the relationship between inflation expectations, realized inflation, news coverage, and narrative frames. Specifications vary by demographic group or expectation measure while maintaining a consistent empirical approach. The baseline model is:

$$\mathrm{E}_t[\pi_{t+h}] = \alpha + \sum_{j=1}^{p} \beta_j \mathrm{E}_{t-j}[\pi_{t-j+h}] + \gamma \pi_{t-1} + \delta \log(N_{t-1}) + \eta S_{k,t-1} + \varepsilon_t \tag{1}$$

Where $\mathrm{E}_t[\pi_{t+h}]$ represents inflation expectations at time t for horizon h (either 1-year or 5-10 years), π_{t-1} is lagged realized inflation, $\log(N_{t-1})$ is the log number of sentences containing the word "inflation" (a proxy for inflation news volume), and $S_{k,t-1}$ represents the share of inflation sentences containing a specific inflation narrative k. The number of autoregressive lags p is selected to minimize AIC for each specification.

We study three specifications for each horizon: (1) a model with only realized inflation and optimal autoregressive lags, (2) adding the log number of inflation sentences, and (3) results from separate regressions that include *one* narrative share at a time. In the regressions, we exclude the shares of "other-cause" and "other-effect" narratives, as well as the share of sentences without identifiable narratives ("none"), to allow for a clearer interpretation of the specific narrative associations with inflation expectations. Columns 1-3 in each table of results presents the findings for 1-year expectations, while columns 4-6 present the findings for 5-10 year (long-run) expectations.

All regressions use monthly data with standardized variables across specifications. T-statistics reported in brackets are calculated using Newey and West (1994) standard errors with lag

truncation parameters automatically selected according to the optimal bandwidth.

5.2.2 National Inflation Expectations

Table 6 presents our baseline results for national average inflation expectations from the Michigan Survey of Consumers. The dependent variables are standardized weighted average inflation expectations across all survey respondents for both short-term (1-year) and long-term (5-10 year) horizons.

For 1-year expectations, realized inflation demonstrates significant predictive power with a coefficient of 0.116 (t-statistic 2.92) in the baseline specification (Column 1). This suggests that a one standard deviation increase in realized inflation is associated with a 0.116 standard deviation increase in short-term inflation expectations. The relationship persists when controlling for inflation news volume in Column 2, with the coefficient remaining stable at 0.119 (t-stat 3.00).

In contrast, long-run expectations (5-10 years) exhibit a substantially weaker relationship with realized inflation. The coefficient drops to 0.028 in Column 4 (t-stat 1.77) and further to 0.022 in Column 5 (t-stat 0.99). This pattern suggests that long-run expectations are more firmly anchored and less responsive to current inflation conditions.

The volume of inflation-related news, measured by the log number of sentences containing the word "inflation," shows minimal predictive power for national expectations after controlling for realized inflation. The coefficient is exactly 0 (t-stat 0.01) for 1-year expectations and 0.017 (t-stat 0.69) for 5-10 year horizons, both statistically insignificant, suggesting that the mere quantity of inflation coverage has limited impact on average expectations.

Specific narrative frames show some predictive power beyond realized inflation and news volume, though they fall shy of statistical significance at the 5% conventional level in all cases. For 1-year expectations, the "social" narrative emerges as the most influential predictor (coefficient 0.034, t-stat 1.82), indicating that news coverage emphasizing inflation's social/political consequences are associated with higher short-term expectations. While this coefficient is approximately 29% of the magnitude of the lagged realized inflation effect (0.034 vs. 0.116), it represents a substantial additional impact that persists even after controlling for actual inflation levels. The "wage" narrative also shows a strong positive association (coefficient 0.031, t-stat 1.67) with 1-year expectations, approximately 27% of the magnitude of the realized inflation effect, suggesting that discussions of wage-price dynamics are associated with a higher inflation outlook among

households. The "trade" narrative demonstrates similar influence (coefficient 0.027, t-stat 1.73), representing about 23% of the realized inflation effect.

For long-run expectations, the pattern shifts. The most strongest narrative predictor becomes "redistribution" (coefficient 0.037, t-stat 1.73), suggesting that media emphasis on inflation's redistributive effects is associated with higher long-term expectations. This narrative effect is roughly 1.32 times the magnitude of the already weak realized inflation effect (0.037 vs. 0.028) on long-run expectations. Interestingly, the "expect" narrative shows a negative coefficient (-0.025, t-stat -1.87) for long-term expectations, contrasting with its positive effect on short-term expectations. This finding suggests that the way media frames inflation narratives may have opposing effects on short- versus long-term expectations.

5.2.3 Dispersion in Inflation Expectations

Table 7 examines how various factors affect the heterogeneity in inflation expectations across households. The dependent variable is the standardized cross-sectional standard deviation of individual inflation expectations, representing the disagreement or dispersion among consumers about future inflation rates.

For 1-year expectations, realized inflation (Column 1) shows a moderate relationship with expectation dispersion, with a coefficients of 0.064 (t-stat 2.06). This indicates that periods of higher inflation are associated with somewhat greater disagreement about future inflation. For long-run (5-10 year) expectations in column 4, the relationship is weaker but still precisely estimated, with a coefficient of 0.038 (t-stat 2.16).

The volume of inflation news shows minimal association with expectation dispersion, with small and imprecisely estimated coefficients (0.016 for 1-year and 0.008 for 5-10 year horizons), suggesting that mere coverage volume does not substantially affect disagreement.

A salient result from the analysis is that specific narrative frames demonstrate dramatically stronger associations with expectation dispersion than realized inflation itself. Social narratives about inflation—those emphasizing societal impacts and political consequences—strongly predict increased dispersion in both short and long-run expectations. For 1-year expectations, the social narrative coefficient (0.115) is approximately 1.8 times larger than the realized inflation coefficient (0.064), with a dramatically higher t-statistic (5.36 vs. 2.06). Similarly, for long-run expectations, the social narrative coefficient (0.084) is about 2.2 times the magnitude of the realized inflation

coefficient (0.038), with a t-statistic nearly twice as large (3.86 vs. 2.16).

Cost narratives, stressing the pernicious effect of inflation on the cost of living, also show substantial predictive power for expectation dispersion, with coefficients of 0.069 (t-stat 2.85) for 1-year and 0.064 (t-stat 2.67) for 5-10 year horizons. These effects are approximately 1.1 and 1.7 times larger than the realized inflation effects for the respective horizons. Cost-push narratives specifically (emphasizing how rising input costs such as materials, labor, and energy force businesses to raise prices) similarly show significant positive associations with dispersion, especially for 1-year expectations (coefficient 0.041, t-stat 2.42).

These patterns suggest that when media coverage emphasizes social impacts and political consequences of inflation, or focuses on cost pressures, households form more divergent expectations—potentially reflecting heterogeneous interpretations of these narratives based on personal circumstances. The fiscal narrative (coefficient 0.052, t-stat 2.13 for 1-year expectations) and supply narrative (coefficient 0.030, t-stat 2.66 for 1-year expectations) also significantly predict increased disagreement, though with smaller magnitudes than the social and cost narratives.

Interestingly, the savings narrative is associated with reduced dispersion in long-run expectations (coefficient -0.036, t-stat -2.01), suggesting that news emphasizing savings behaviors may create more consensus about distant future inflation.

The economic significance of these narrative effects on expectation dispersion substantially exceeds that of realized inflation itself, highlighting how media framing may be more consequential than underlying inflation conditions in explaining why households disagree about future inflation.

5.2.4 Income-Based Heterogeneity in Expectations

Inflation expectation associations with media narratives vary considerably by income level, with the strongest differences between households in the lowest and highest income quartiles. Tables 8 through 11 present these patterns in detail. For one-year ahead expectations, lower-income households (1st quartile) exhibit significantly stronger responsiveness to inflation narratives compared to higher-income households. The lowest income quartile shows strong reactions to social narratives (coefficient 0.130, t-stat 3.32), expectation narratives (coefficient 0.111, t-stat 4.10), cost narratives (coefficient 0.080, t-stat 2.08), and uncertainty narratives (coefficient 0.074, t-stat 2.39). This heightened sensitivity may reflect greater information frictions among lower-income households and their reliance on general news media when forming expectations. By contrast,

the highest income quartile (4th quartile) displays much more muted responses, with their strongest reactions being to trade narratives (coefficient 0.032, t-stat 1.97) and social narratives (coefficient 0.028, t-stat 1.32), suggesting higher-income households may have access to more diverse information sources that make individual news narratives less influential in updating their expectations.

Lower-income households also show greater sensitivity to the volume of inflation news. The coefficient on the lagged log number of inflation sentences is 0.057 (t-stat 1.47) for the 1st quartile in the baseline specification, compared to -0.020 (t-stat -1.05) for the 4th quartile. For the lowest-income households, this coefficient reaches statistical significance in some specifications, with t-statistics approaching 2.0 in some cases, while the highest-income quartile typically observe negative coefficients, ranging from -0.027 to -0.012, though none are precisely estimated (t-stats -1.48 to -0.57). This suggests that lower-income households may increase their inflation expectations simply based on greater media coverage of inflation, regardless of content, while higher-income households may actually hold steady or mildly reduce their expectations in response to increased coverage—perhaps interpreting greater discussion as a sign that policymakers are addressing the issue or because they have stronger prior beliefs about inflation that are less easily shifted by news volume.

For 5-10 year ahead expectations, the patterns change slightly. The lowest income quartile shows strong responses to cost narratives (coefficient 0.101, t-stat 2.17) and social narratives (coefficient 0.092, t-stat 1.80), but exhibit negative responses to demand narratives (coefficient -0.080, t-stat -2.56) and international narratives (coefficient -0.048, t-stat -1.82). The highest income quartile shows minimal responsiveness to most narratives for long-term expectations, with only fiscal narratives (coefficient 0.037, t-stat 1.23) and redistribution narratives (coefficient 0.042, t-statistic 1.30) showing modest positive effects, though both are imprecisely estimated.

The sensitivity to realized inflation also varies systematically across income groups. For short-term expectations, the coefficient on lagged realized inflation is 0.215 (t-stat 3.99) for the 1st quartile versus 0.243 (t-stat 6.06) for the 4th quartile. However, for long-term expectations, lower-income households demonstrate stronger anchoring to realized inflation (coefficient 0.078, t-stat 2.75) compared to higher-income households (coefficient 0.027, t-stat 0.78). This pattern suggests higher-income households may more effectively distinguish between transitory and permanent inflation shocks.

The gap in inflation expectations between the bottom and top income quartiles also varies with the prevalence of different inflation narratives, as shown in Table 12. The lowest income quartile expects higher inflation rates (5.03% for 1-year and 4.54% for 5-10 years) on average compared to the highest income quartile (3.45% for 1-year and 3.40% for 5-10 years). Cost narratives are associated with larger disparities, with a coefficient of 0.128 (t-stat 2.30) for short-term expectations and 0.126 (t-stat 2.28) for long-term expectations. Social narratives similarly widen the gap, with coefficients of 0.114 (t-stat 1.87) for one-year and 0.117 (t-stat 2.03) for 5-10 year expectations. International narratives significantly reduce the short-term expectation gap (coefficient -0.108, t-stat -3.07) while demand narratives show their strongest convergence effects for long-run expectations (coefficient -0.107, t-stat -3.00). This suggests these narrative types may provide common information that helps align expectations across income groups, but at different time horizons. The volume of inflation news itself widens the expectations gap, though the impact is imprecisely estimated, with coefficients of 0.097 (t-stat 1.94) for short-term and 0.041 (t-stat 1.17) for long-term expectations in the baseline specification (columns 2 and 5, respecitvely).

5.2.5 Education-Based Heterogeneity in Expectations

Our analysis also reveals meaningful heterogeneity in how narratives predict inflation expectations across different education groups, notably college and non-college graduates. Tables 13 through 15 present the results. For one-year ahead expectations, non-college graduates exhibit significantly stronger sensitivity to inflation narratives compared to college graduates. Non-college graduates show reactions to expectations narratives (coefficient 0.052, t-stat 2.14), supply narratives (coefficient 0.027, t-stat 1.36), and wage narratives (coefficient 0.022, t-stat 0.95). This heightened responsiveness may reflect non-college graduates' greater reliance on general news media when forming expectations. By contrast, college graduates display more muted responses, with their strongest reactions being to uncertainty narratives (coefficient 0.028, t-stat 2.00), expectations narratives (coefficient 0.026, t-stat 1.64), and demand narratives (coefficient 0.019, t-stat 1.61), suggesting they may have access to more diverse information sources or rely less on the media when forming inflation expectations.

Non-college graduates also show greater sensitivity to the volume of inflation news. The coefficient on the lagged log number of inflation sentences is 0.028 (t-stat 1.04) for non-college graduates in the baseline specification (column 2), compared to -0.014 (t-stat -0.90) for college

graduates, though neither are precisely estimated. For non-college graduates, this coefficient ranges from 0.020 to 0.035 across specifications that include separate narrative shares, while college graduates show negative coefficients ranging from -0.025 to -0.006. This suggests that non-college graduates may increase their inflation expectations based on greater media coverage of inflation, regardless of content, while college graduates may actually reduce their expectations, or at least keep them steady, in response to increased coverage.

For 5-10 year ahead expectations, non-college graduates show strong, but imprecisely estimated, responses to fiscal narratives (coefficient 0.054, t-stat 1.65) and cost narratives (coefficient 0.073, t-stat 1.01), but exhibit negative responses to supply narratives (coefficient -0.043, t-stat -1.50). College graduates show minimal responsiveness to most narratives for long-term expectations, with only uncertainty narratives (coefficient 0.054, t-stat 2.06) and demand narratives (coefficient 0.036, t-statistic 1.29) showing modest positive effects.

The sensitivity to realized inflation also varies across education groups. For short-term expectations, the coefficient on lagged realized inflation is 0.100 (t-stat 1.80) for non-college graduates versus 0.128 (t-stat 2.49) for college graduates. However, for long-term expectations, non-college graduates demonstrate stronger anchoring to realized inflation (coefficient 0.080, t-stat 2.41) compared to college graduates (coefficient 0.044, t-stat 1.12). This difference suggests college graduates may more effectively distinguish between transitory and permanent inflation shocks, similar to the income results from the previous section.

The prevalence of different narratives leads to significant widening of expectation gaps between college and non-college graduates, as shown in Table 15. On average, non-college graduates report consistently higher inflation expectations than college graduates, with a gap of 0.76 percentage points for one-year ahead expectations (4.69% vs. 3.93%) and 0.62 percentage points for 5-10 year expectations (4.14% vs. 3.52%). Social narratives especially exacerbate these disparities, with a one standard deviation increase in social narratives associated with a 0.169 standard deviation increase in the expectation gap between non-college and college graduates for short-term horizons (t-stat 3.43) and 0.123 for long-term horizons (t-stat 2.98). Cost narratives similarly widen the gap, with coefficients of 0.130 (t-stat 2.62) for one-year expectations. Higher supply narrative prevalence predicts a reduction in the gap for long-run expectations (coefficient -0.067, t-stat -1.80), potentially indicating that supply-side discussions reduce variation across education groups. The volume of inflation news itself widens the expectations gap, with a coefficient of 0.076 (t-stat

1.99) for short-term expectations in the baseline specification (column 2) and 0.052 (t-stat 1.23) for long-run expectations (column 5).

5.2.6 Narratives and Realized Inflation

Importantly, Table 16 demonstrates that neither the volume of inflation news nor specific narrative frames significantly predict actual realized inflation. This suggests that while narratives have limited predictive power for actual inflation, they substantially influence expectations across demographic groups.

This finding highlights a potential disconnect between media narratives and economic fundamentals. Narratives appear to shape expectations through channels other than those conveying superior information about future inflation. The heterogeneous effects across demographic groups further suggest that media framing may contribute to differences in expectations, with certain narratives resonating differently with different segments of the population.

6 Conclusion

This paper presents a large scale analysis of U.S. inflation narratives across a century of newspaper coverage, providing a detailed look at how inflation's causes and effects are explained to the American public in the media.

We find significant shifts in narrative prevalence across economic eras, with fiscal explanations dominating pre-1980 discourse and monetary narratives gaining prominence thereafter. We also document substantial regional differences, with Northern papers emphasizing fiscal causes while Southern papers focus more on supply factors and interest rate effects.

Importantly, narrative frames strongly predict heterogeneity in inflation expectations across demographic groups, with effects that in some cases exceed the predictive power of realized inflation itself. Lower-income households show greater sensitivity to narratives about the social/political and cost-of-living effects of inflation, while higher-income households remain relatively insulated from media narrative influence. When media coverage emphasizes social impacts and political consequences of inflation, households form more divergent expectations, with social narratives predicting dispersion effects 1.8 times larger than realized inflation for short-term expectations. These findings suggest media narrative exposure contributes to persistent expectation gaps

across demographic groups—gaps that have implications for consumption, saving, and investment decisions.

These results have important implications for monetary policy effectiveness and the theoretical understanding of expectation formation. Central banks dedicate substantial resources to communication strategies, but our research demonstrates they control only part of the message that reaches households. Media interpretation and narrative framing may be equally consequential for expectation formation. This suggests that effective monetary policy communication may need to account for how messages are interpreted and retransmitted through media channels to different demographic groups. Regional Federal Reserve banks could benefit from monitoring local media narratives as leading indicators of expectation formation, given the persistent geographic differences we document. The heightened sensitivity of lower-income and non-college educated households to media narratives suggests that demographic-specific communication strategies may improve policy transmission.

Our findings also point toward extensions for macroeconomic modeling. The documented relationship between narrative frames and expectation formation across demographic groups suggests that media-mediated information processing could be incorporated into heterogeneous agent models to better understand expectation dynamics, monetary policy transmission, and the persistence of inflation expectation disagreement. Future research could develop structural models that explicitly account for how narratives serve as information intermediaries, transforming economic data into household beliefs through demographic-specific channels.

By documenting the evolution and impact of inflation narratives over a century of American economic history, this paper provides new insights into how economic ideas are communicated to the public and demonstrates the significant role of media framing in shaping economic beliefs and expectations.

References

Akerlof, George A and Dennis J Snower, "Bread and bullets," *Journal of Economic Behavior & Organization*, 2016, *126*, 58–71.

Andre, Peter, Ingar Haaland, Christopher Roth, and Johannes Wohlfart, "Narratives about the Macroeconomy," 2023. Working paper.

Ash, Elliott, Germain Gauthier, and Philine Widmer, "Relatio: Text semantics capture political and economic narratives," *arXiv preprint arXiv:2108.01720*, 2021.

Barron, Kai and Tilman Fries, "Narrative persuasion," 2023. Working paper.

Benabou, Roland, Armin Falk, and Jean Tirole, "Narratives, Imperatives, and Moral Reasoning," 2018.

Blei, David M, Andrew Y Ng, and Michael I Jordan, "Latent dirichlet allocation," *Journal of machine Learning research*, 2003, *3* (Jan), 993–1022.

Bybee, Leland, Bryan Kelly, and Yinan Su, "Narrative asset pricing: Interpretable systematic risk factors from news text," *The Review of Financial Studies*, 2023, *36* (12), 4759–4787.

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

Chahrour, Ryan, Adam Hale Shapiro, and Daniel Wilson, "News selection and household inflation expectations," 2025. Working paper.

Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber, "Monetary policy communications and their effects on household inflation expectations," *Journal of Political Economy*, 2022, *130* (6), 1537–1584.

Davies, Mark, "Corpus of News on the Web (NOW)," https://www.english-corpora.org/now/ 2016. Available online at https://www.english-corpora.org/now/.

D'Acunto, Francesco, Ulrike Malmendier, Juan Ospina, and Michael Weber, "Exposure to grocery prices and inflation expectations," *Journal of Political Economy*, 2021, *129* (5), 1615–1639.

Eliaz, Kfir and Ran Spiegler, "A model of competing narratives," *American Economic Review*, 2020, *110* (12), 3786–3816.

Flynn, Joel P and Karthik Sastry, "The Macroeconomics of Narratives," Working Paper 32602, National Bureau of Economic Research June 2024.

Gaspar, Vitor, Frank Smets, and David Vestin, "Inflation expectations, adaptive learning and optimal monetary policy," in "Handbook of monetary economics," Vol. 3, Elsevier, 2010, pp. 1055–1095.

Grootendorst, Maarten, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," 2022.

Heddaya, Mourad, Qingcheng Zeng, Chenhao Tan, Rob Voigt, and Alexander Zentefis, "Causal Micro-Narratives," in "Proceedings of the 6th Workshop on Narrative Understanding" Association for Computational Linguistics Miami, Florida, USA 2024, pp. 67–84.

Jalil, Andrew J. and Gisela Rua, "Inflation expectations and recovery in spring 1933," *Explorations in Economic History*, 2016, *62*, 26–50.

Kose, M Ayhan, Hideaki Matsuoka, Ugo Panizza, and Dana Vorisek, "Inflation expectations: review and evidence," 2019.

Labov, William and Joshua Waletzky, "Narrative analysis: Oral versions of personal experience," *Journal of Narrative and Life History*, 1967.

_ and _, "Narrative analysis: Oral versions of personal experience.," *Journal of Narrative and Life History*, 1967.

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

Lange, Kai-Robin, Matthias Reccius, Tobias Schmidt, Henrik Müller, Michael WM Roos, and Carsten Jentsch, *Towards extracting collective economic narratives from texts* number 963, Ruhr Economic Papers, 2022.

Morag, Dor and George Loewenstein, "Narratives and valuations," 2023.

Newey, Whitney K and Kenneth D West, "Automatic lag selection in covariance matrix estimation," *The Review of Economic Studies*, 1994, *61* (4), 631–653.

Pham, Chau Minh, Alexander Hoyle, Simeng Sun, Philip Resnik, and Mohit Iyyer, "TopicGPT: A prompt-based topic modeling framework," 2024.

Roos, Michael W.M. and Matthias Reccius, *Narratives in Economics*, RWI, Sep 2021.

Ryngaert, Jane, Carola Binder, and Pascal Frank, "The Causal Effect of News on Inflation Expectations," 2025. Working paper.

Shiller, R.J., *Narrative Economics: How Stories Go Viral and Drive Major Economic Events*, Princeton University Press, 2020.

Shiller, Robert J, "Narrative economics," *American Economic Review*, 2017, *107* (4), 967–1004.

_ , "Popular Economic Narratives Advancing the Longest U.S. Economic Expansion 2009-2019," Working Paper 26857, National Bureau of Economic Research March 2019.

Trebbi, Giovanni, "Inflation Narratives and Expectations," June 2025. Working paper.

Weber, Michael, Francesco d'Acunto, Yuriy Gorodnichenko, and Olivier Coibion, "The subjective inflation expectations of households and firms: Measurement, determinants, and implications," *Journal of Economic Perspectives*, 2022, *36* (3), 157–184.

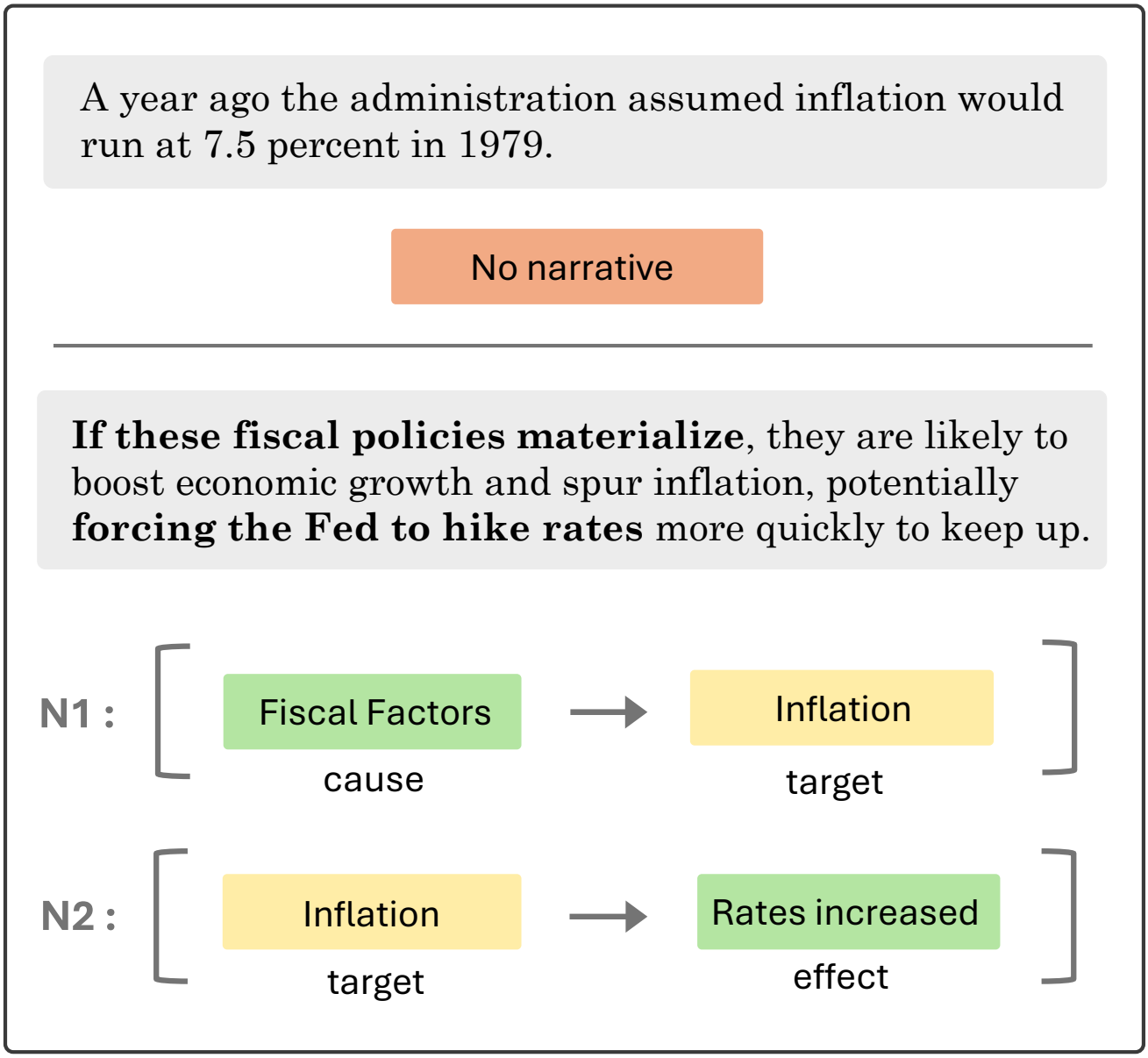
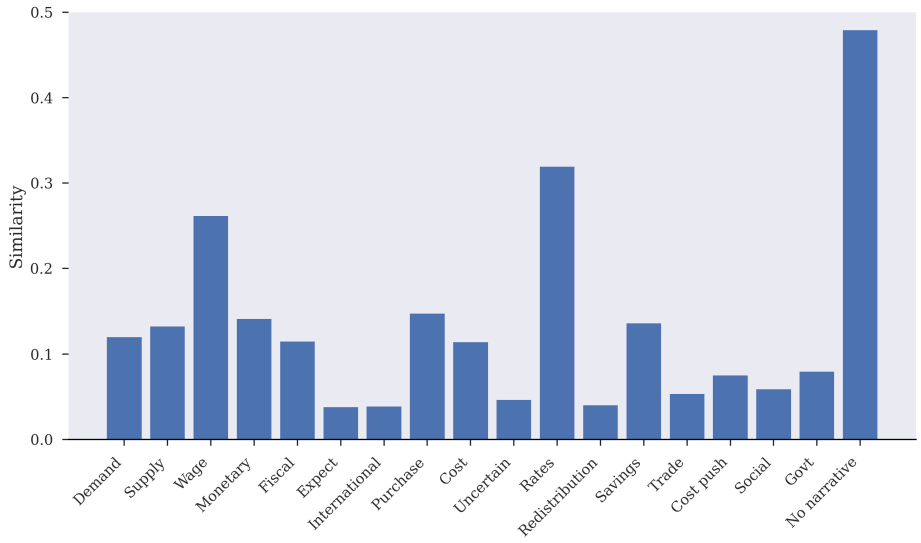
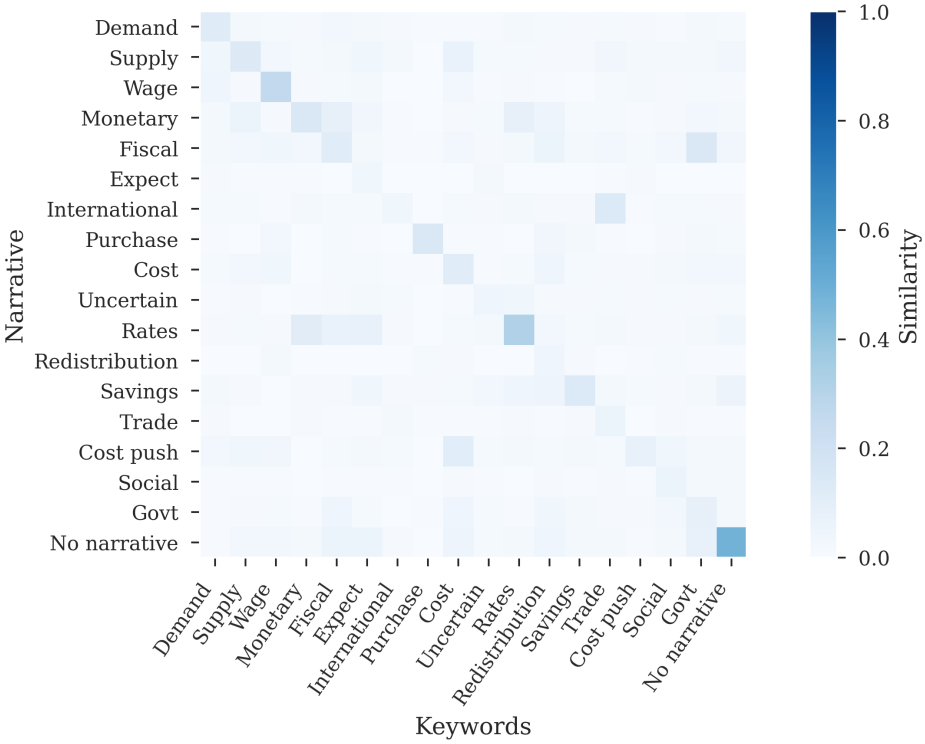


FIGURE 1
EXAMPLE INFLATION NARRATIVES

In the first sentence, no narratives are identified; in the second, two narratives (N1 and N2) are identified, one representing a cause of inflation (fiscal) and the other representing an effect of it (rates).



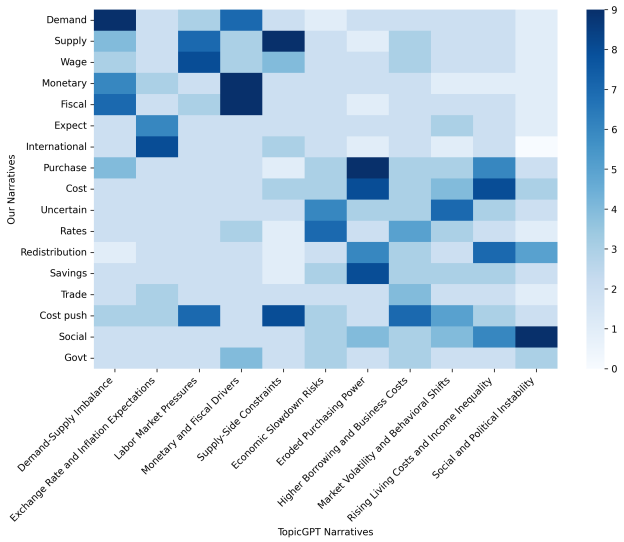
(A) Single-Narrative Similarity



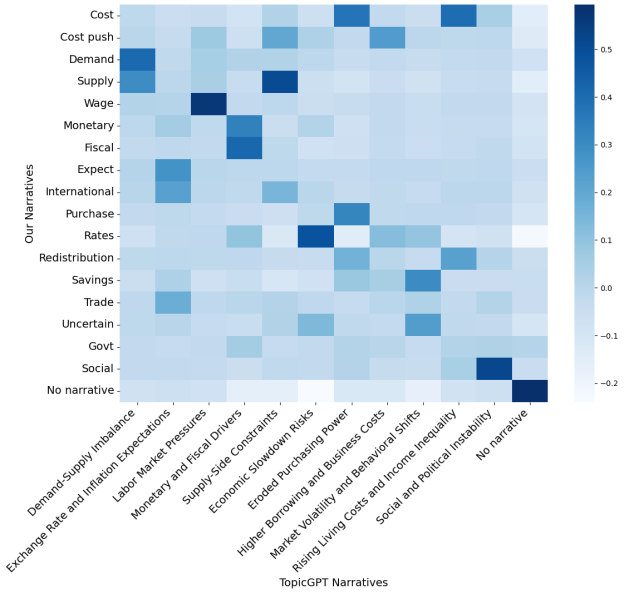
(B) Full Similarity Matrix

FIGURE 2
SIMILARITY BETWEEN KEYWORD AND NARRATIVE CLASSIFICATION

The figure presents Jaccard similarities between a keyword-based classification and our narrative classification. Fig. 2A displays the sentence-level similarity between each narrative category and its associated keywords. Fig. 2B presents the similarity matrix between all narrative categories and keyword classifications, where darker shades of blue indicate that the two classifications have more sentences in common. The keywords associated with each narrative category are in Table 3.



(A) Semantic Similarity



(B) Sentence Similarity

FIGURE 3
SIMILARITY BETWEEN TOPICGPT AUTOMATED NARRATIVES AND OUR MANUAL TAXONOMY

The figure compares our manually constructed narrative taxonomy with TopicGPT's automated narrative discovery approach. Fig. 3A presents semantic similarity at the category level between the two taxonomies, where darker blue shades indicate more positive correlations. Semantic similarity was determined using an LLM (GPT 4o, Nov. 20, 2024 update), with the prompt presented in Appendix A.4. Fig. 3B displays Jaccard similarity at the sentence level, showing how sentences classified under our 19 original narrative categories (rows) overlap with TopicGPT's 11 discovered categories (columns). Darker blue cells indicate higher sentence overlap counts.

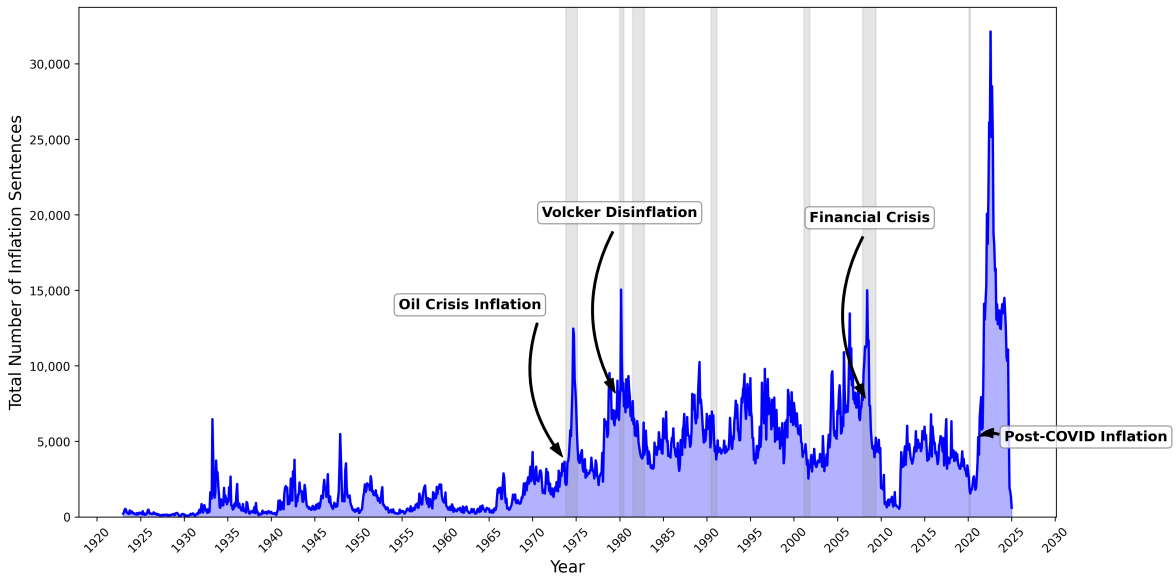
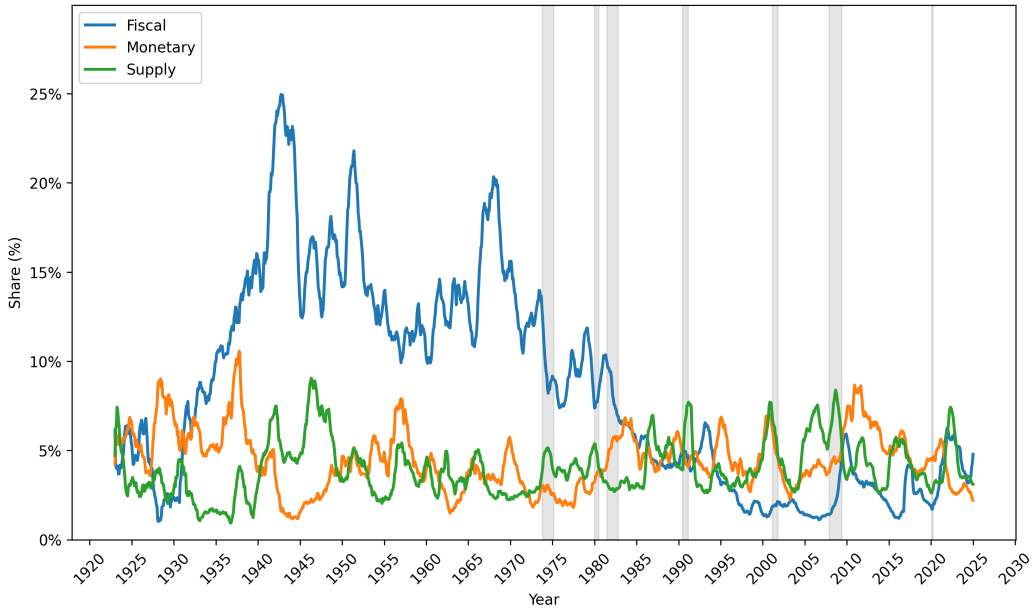
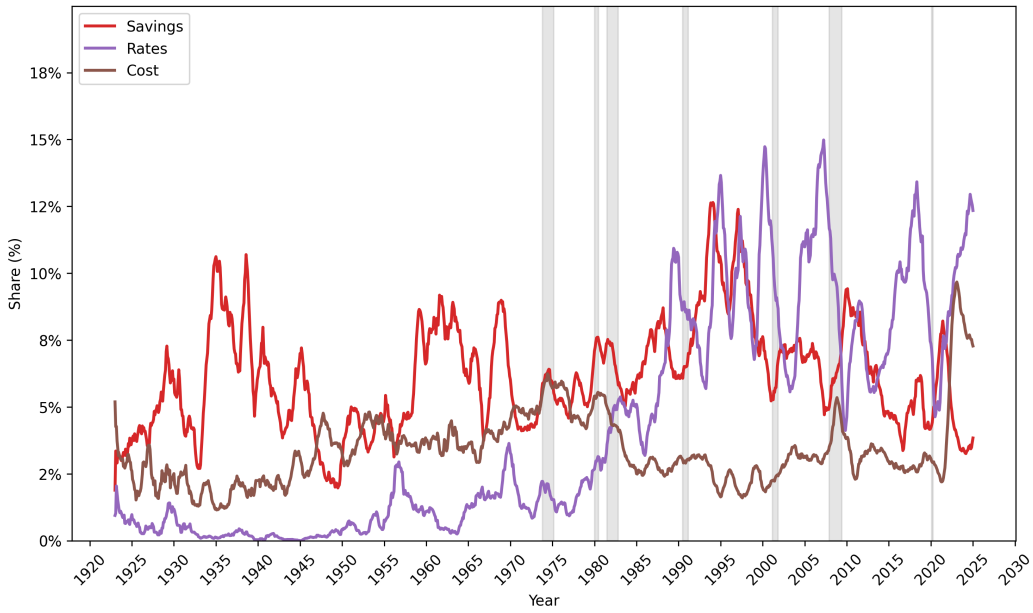


FIGURE 4
INFLATION NEWS OVER TIME

This figure shows the total number of sentences containing the word "inflation" in the ProQuest news corpus from 1923 to 2025. Gray shaded regions indicate NBER recession periods. Major inflation episodes are annotated, including the 1970s Oil Crisis inflation, the Volcker Disinflation of the early 1980s, the 2008 Financial Crisis, and the Post-COVID inflation surge beginning in 2021.



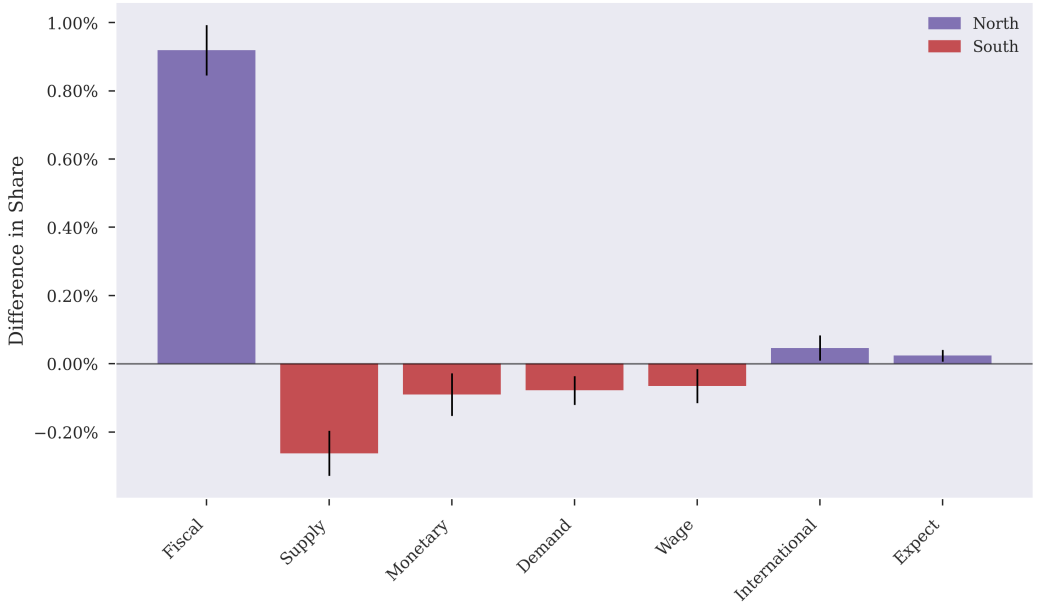
(A) Top 3 Cause Narratives



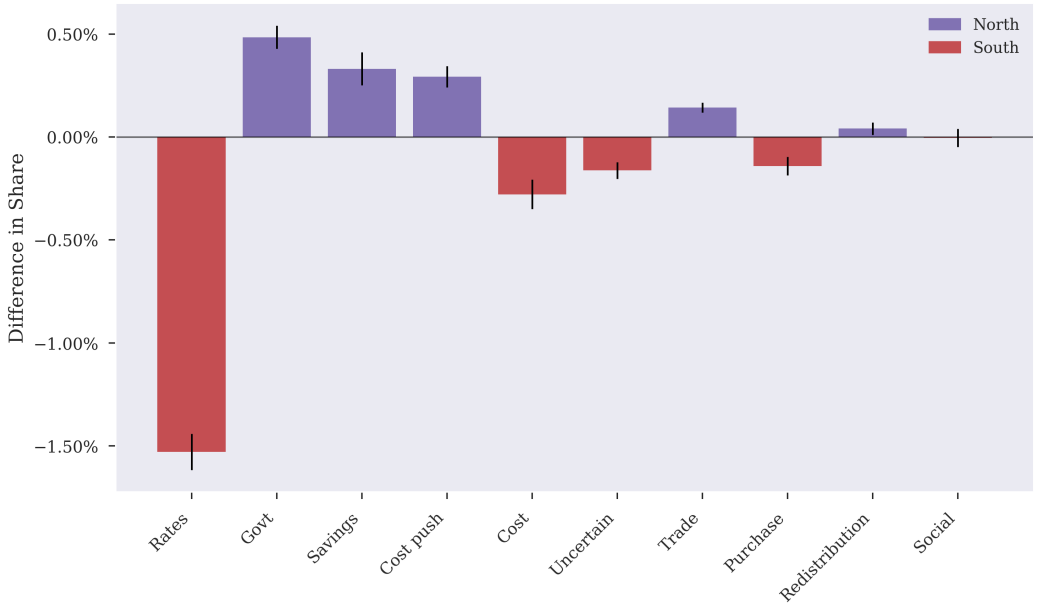
(B) Top 3 Effect Narratives

FIGURE 5
TOP NATIONAL INFLATION NARRATIVES

The figure shows the 1-year rolling average monthly shares of the top three inflation cause narratives (Panel A) and effect narratives (Panel B). Narrative measures represent the rolling 12-month average of monthly shares of all inflation sentences from ProQuest newspaper articles that we classify as expressing that particular narrative about inflation. Gray shaded regions indicate NBER recession periods. The sample period is January 1923 to January 2025.



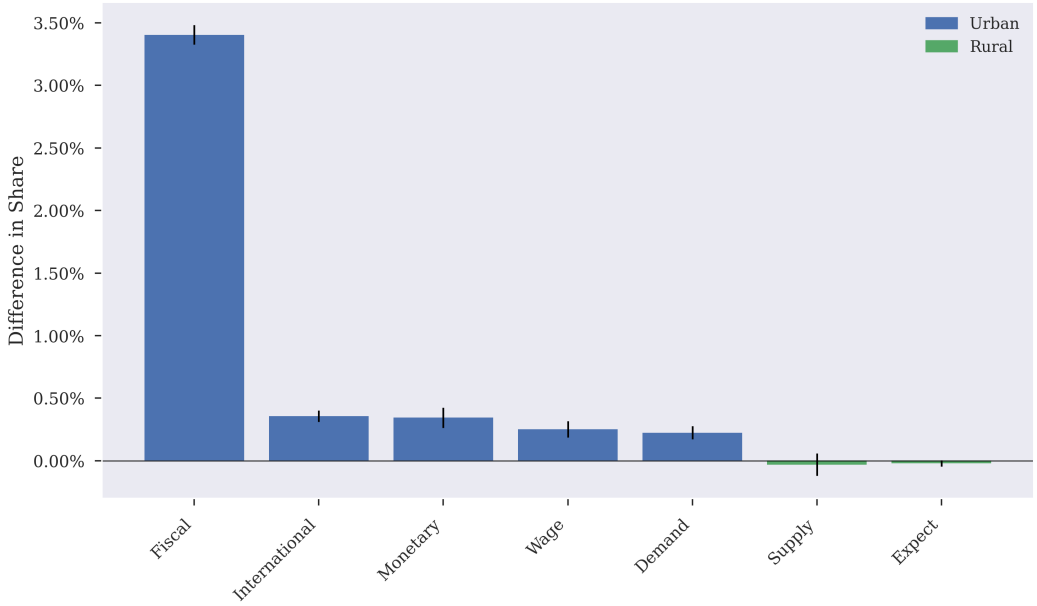
(A) Cause Narratives



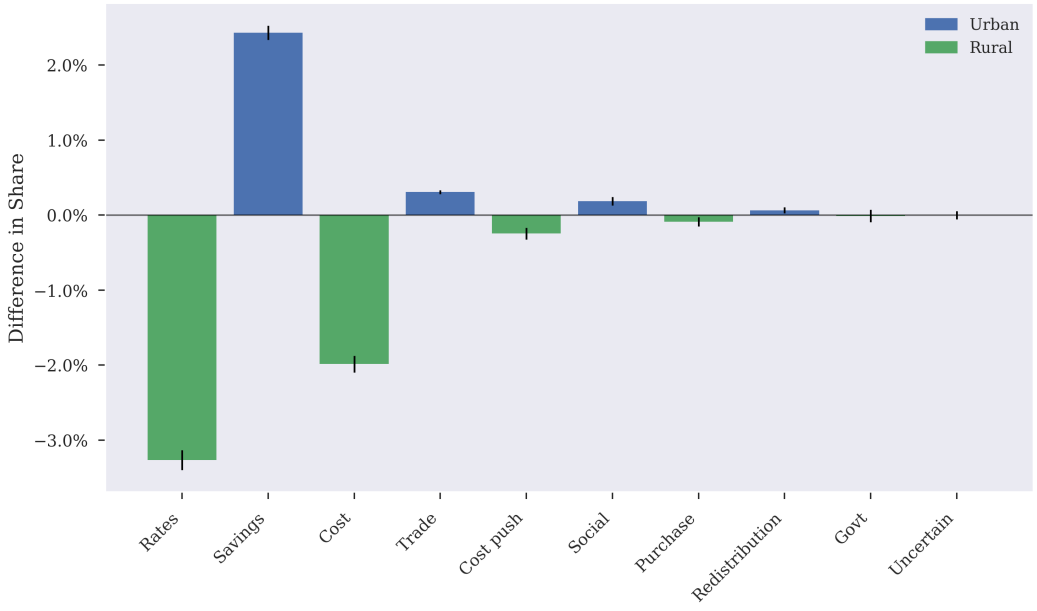
(B) Effect Narratives

FIGURE 6
NORTH VS SOUTH INFLATION NARRATIVES

This figure shows the differences in inflation narrative shares between newspapers headquartered in northern and southern cities. The top panel shows differences in causal narratives of inflation, while the bottom panel shows differences in effect narratives. Error bars represent 95% confidence intervals. Purple bars indicate narratives more prevalent in northern newspapers, while red bars indicate narratives more prevalent in southern newspapers. Cities are classified as North or South based on their geographical location relative to the Mason-Dixon line. The analysis uses only local newspaper articles (excluding national newspapers listed in Table 1) from the ProQuest corpus from January 1923 to January 2025.



(A) Cause Narratives



(B) Effect Narratives

FIGURE 7
URBAN VS RURAL INFLATION NARRATIVES

This figure shows the differences in inflation narrative shares between newspapers headquartered in urban and rural areas. The top panel shows differences in causal narratives of inflation, while the bottom panel shows differences in effect narratives. Error bars represent 95% confidence intervals. Blue bars indicate narratives more prevalent in urban newspapers, while green bars indicate narratives more prevalent in rural newspapers. Urban and rural classifications are based on population density thresholds, with urban areas defined as those with population density above the median. The analysis uses only local newspaper articles (excluding national newspapers listed in Table 1) from the ProQuest corpus from January 1923 to January 2025.

Table 1
National Newspapers

Newspaper	Inflation Sentences
New York Times	642,633
Wall Street Journal	543,075
Washington Post	253,378
Targeted News Service	169,065
News Wire Services	115,031
Tribune/McClatchy Services	83,698
Christian Science Monitor	47,022
USA TODAY	32,541
University Wire	25,879
Newsday	22,282
Politico	19,303
Spanish Language Services	12,203
Voice of America News / FIND	2,461
New York Post	852
ProPublica	168
Religion News Service	1
National Total Count	1,969,592
National Total Share	46.92%

The table lists all newspapers in the sample that are classified as national publications, along with the count of inflation-related sentences from each source. Publications from the same family have been consolidated (e.g., online and print versions of the same newspaper, as well as different time periods of the same publication). The sample period covers January 1923 to January 2025. Sentences were selected if they contained the word "inflation."

Table 2
Example Sentences by Narrative Category

Narrative	Example Sentence
Causes	
Demand	Recessions historically have been caused when demand for goods and services has exceeded supply, fueling inflation and higher interest rates.
Supply	An expected sharp jump in gasoline prices accounted for almost half of the month's increase and produced the worst monthly inflation report in 14 months.
Wage	Bruce Steinberg of Merrill Lynch & Co. said inflation will continue to be tame as long as labor costs stay low.
Monetary	The Federal Reserve Board, analysts said, has contributed to the rising interest rates by its policy of restricting the supply of money in an effort to control inflation.
Fiscal	Ronald Reagan wrung out inflation by quickly putting the economy through the worst contraction since the Depression.
Expect	Perhaps even more important, consumer and business expectations for future inflation are starting to rise quite a bit.
International	One factor behind muted inflation could be the stronger U.S. dollar, which makes imports cheaper.
Effects	
Purchase	But because of inflation, there is no way of telling in advance how much purchasing power those dollars will have – for example, whether they will cover a college tuition payment or not.
Cost	One reason our credit card bills have been rising is that inflation, too, is on the rise, sometimes exceeding the growth of our paychecks.
Uncertainty	If those dynamics should lead this "transitory" inflation to continue too long, it could affect the planning of households and businesses and unanchor their inflation expectations.
Rates	An alarmed Federal Reserve then began to boost short-term interest rates in the spring of 1988 to cool off the economy enough to slow wage inflation.
Redistribution	He referred to inflation as a tax on the consumer dollar that places the greatest financial burden on persons in the lowest income group.
Savings	High yield funds' performance and the lower yields available from investment-grade bonds in this period of low inflation combined to attract many investors.
Trade	The result, Roach added, will be continued low inflation, growing U.S. competitiveness in overseas markets and an eventual improvement in the American standard of living.
Cost-Push	Companies across the tech sector have been trying to control costs to adapt to high inflation, economic-growth concerns and a slowdown in advertising spending.
Social	The White House is monitoring inflation closely as rising prices for food, rent and gas shake consumer confidence and dent President Biden's approval ratings ahead of midterm elections in November.
Govt	Cities generally rely on local property taxes for their basic revenue, and inflation has sent property taxes through the roof in some areas where citizens will approve no further tax hikes.

The table presents example sentences corresponding to each inflation narrative category described in Section 4.2.

TABLE 3
KEYWORDS ASSOCIATED WITH NARRATIVE CATEGORIES

Narrative	Keywords
Causes	
Demand	demand
Supply	supply, energy
Wage	wage inflation, wage spiral, wage
Monetary	monetary policy, central bank, monetary
Fiscal	policy, spending, fiscal
Expect	expect, expectations, predict
International	exchange rates, international trade, global
Effects	
Purchase	purchasing power, money value
Cost	cost, expense
Uncertainty	uncertainty, future
Rates	interest rates, rate hike
Redistribution	income, wealth, redistribution, debt, credit
Savings	savings, investment, retirement, assets
Trade	trade, export, import
Cost-Push	businesses, profits
Social	impact, political, social
Govt	government, spending, budget

The table presents the keywords (n-grams) associated with each narrative category that we use to classify sentences in a keyword-based classification. The sentence-level correlations between our narrative classification and the keyword-based classification are presented in Fig. 2.

Table 4
Inflation Narrative Discovery Categories

Discovered Narrative	Description
Causes	
Demand-Supply Imbalance	When demand for goods and services exceeds supply due to production bottlenecks, economic recovery, or other constraints, prices rise, fueling inflation.
Exchange Rate and Inflation Expectations	Depreciation of a currency raises import costs and domestic prices, while anticipated future price increases drive businesses to raise prices and consumers to spend more, reinforcing inflation.
Labor Market Pressures	Wage growth and labor shortages increase production costs, driving higher consumer prices and contributing to inflation.
Monetary and Fiscal Drivers	Excessive fiscal spending, expansionary monetary policies, or significant increases in the money supply boost aggregate demand relative to supply, contributing to inflationary pressures.
Supply-Side Constraints	Rising production costs, such as energy, raw materials, or labor costs, along with supply chain disruptions, geopolitical conflicts, or other external shocks, limit supply, increase business expenses, and elevate consumer prices, driving inflation.
Effects	
Economic Slowdown Risks	Inflation and efforts to combat it, such as interest rate hikes, dampen consumer and business spending, increase debt burdens, slow economic growth, and heighten the risks of recession or stagflation.
Eroded Purchasing Power	Inflation decreases the real value of incomes, wages, savings, and investments, limiting affordability and disproportionately affecting lower-income households.
Higher Borrowing and Business Costs	Inflation raises interest rates and increases input, production, and operational expenses, reducing borrowing, spending, profitability, and economic competitiveness.
Market Volatility and Behavioral Shifts	Inflation triggers uncertainty in financial markets, alters consumer spending patterns, and drives businesses to adapt through workforce adjustments or cost-saving measures.
Rising Living Costs and Income Inequality	Inflation increases the cost of essential goods and services while exacerbating financial strain for lower-income groups and widening income disparities.
Social and Political Instability	Inflation worsens inequality, redistributes economic burdens, erodes trust in leadership, and contributes to political unrest and societal strain.

This table presents the 11 distinct causal narratives identified through unsupervised computational discovery using the TopicGPT framework, described in Section 4.7.

Table 5
Inflation Narrative Prevalence

Narrative	All Newspapers	National Newspapers	Local Newspapers
None	54.5%	52.3%	56.3%
Cause Narratives			
Fiscal	8.4%	9.3%	7.8%
Monetary	4.4%	5.2%	3.9%
Supply	4.0%	4.4%	3.7%
Wage	2.5%	2.8%	2.4%
International	1.7%	2.2%	1.3%
Demand	1.6%	1.8%	1.5%
Other Cause	0.5%	0.5%	0.5%
Expect	0.4%	0.4%	0.3%
Effect Narratives			
Savings	6.1%	5.6%	6.4%
Rates	4.2%	4.5%	3.8%
Cost	3.4%	2.8%	3.9%
Govt	2.9%	3.0%	2.9%
Cost-Push	2.4%	2.6%	2.3%
Social	1.8%	1.8%	1.8%
Purchase	1.5%	1.1%	1.7%
Trade	1.0%	1.4%	0.7%
Uncertain	1.0%	1.0%	0.9%
Redistribution	0.7%	0.7%	0.8%
Other Effect	0.2%	0.1%	0.2%

This table presents the average share of each inflation narrative across newspaper sentences from January 1923 - January 2025. The narrative prevalence is calculated as the proportion of inflation-related sentences expressing each narrative type. The "All Newspapers" column includes all newspapers in the ProQuest data, "National Newspapers" includes only national newspapers listed in Table 1, and "Local Newspapers" includes only local/regional newspapers (i.e., all sources except those listed in Table 1). Narratives are ranked from highest to lowest share based on the overall average over the sample period.

Table 6
National Average Inflation Expectations and Narrative Shares

Dep. Var.: Avg. Inflation Expectations	Horizon: 1 year			Horizon: 5-10 years		
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline						
L. YoY Realized Inflation	0.116	0.119	0.114 to 0.124	0.028	0.022	0.007 to 0.027
	[2.92]	[3.00]	[2.83 to 3.15]	[1.77]	[0.99]	[0.23 to 1.22]
L. Log(No. of Infl. Sentences)		0.000	-0.008 to 0.007		0.017	0.012 to 0.022
		[0.01]	[-0.51 to 0.42]		[0.69]	[0.52 to 0.84]
Cause Narratives						
Wage			0.031			-0.016
			[1.67]			[-1.09]
Expect			0.025			-0.025
			[1.66]			[-1.87]
Demand			0.014			-0.012
			[1.31]			[-0.67]
Fiscal			0.028			0.029
			[1.23]			[1.38]
Supply			0.014			0.007
			[0.90]			[0.37]
International			0.009			-0.025
			[0.70]			[-1.20]
Monetary			0.007			-0.004
			[0.53]			[-0.24]
Effect Narratives						
Social			0.034			0.032
			[1.82]			[0.81]
Trade			0.027			0.010
			[1.73]			[0.52]
Uncertain			0.017			0.001
			[1.16]			[0.04]
Savings			-0.013			-0.029
			[-1.13]			[-1.19]
Govt			0.023			0.017
			[1.03]			[1.07]
Purchase			-0.015			-0.006
			[-1.02]			[-0.34]
Cost-Push			-0.013			0.014
			[-0.62]			[0.54]
Redistribution			0.006			0.037
			[0.40]			[1.73]
Rates			-0.004			-0.003
			[-0.27]			[-0.18]
Cost			0.003			0.020
			[0.14]			[0.56]
Observations	558	557	557	416	415	415
Adj. R^2	0.891	0.891	0.891 to 0.892	0.867	0.865	0.865 to 0.866

The dependent variables are standardized weighted average one-year ahead and 5-10 year ahead inflation expectations from the Michigan Survey of Consumers. For each model, we include the optimal number of autoregressive lags that minimizes AIC. Columns (1) and (4) regress expectations on lagged year-over-year realized inflation and the optimal number of autoregressive lags. Columns (2) and (5) add the lagged log number of inflation sentences to the model (i.e., newspaper sentences containing the word "inflation"). Columns (3) and (6) report results from separate regressions that each include the optimal number of autoregressive lags, lagged realized inflation, the lagged log number of inflation sentences, and one narrative at a time. For lagged realized inflation and log number of inflation sentences, columns (3) and (6) show the range of coefficients and t-statistics [in brackets] across the narrative regressions. For narratives, they show the coefficient and t-statistic from the regression including that narrative. For columns (3) and (6), the adjusted R-squared values shown represent the range across all narrative regressions. Realized Inflation is the standardized year-over-year percent change in the Consumer Price Index for All Urban Consumers (All Items, U.S. City Average) from the Bureau of Labor Statistics. Narrative measures are lagged standardized monthly shares of inflation sentences from ProQuest newspaper articles that we classify as expressing that particular narrative about inflation. Sample period is September 1978 to February 2025 for one-year expectations and July 1990 to February 2025 for 5-10 year expectations. T-statistics shown in brackets are calculated using Newey and West (1994) standard errors, optimally selected using the rule-of-thumb bandwidth estimator, computed to be 5 lags. All specifications include a constant.

Table 7
Standard Deviation of Inflation Expectations and Narrative Shares

Dep. Var.: Std. Dev. Inflation Expectations	Horizon: 1 year			Horizon: 5-10 years		
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline						
L. YoY Realized Inflation	0.064	0.068	0.053 to 0.071	0.038	0.038	0.001 to 0.044
	[2.06]	[2.37]	[1.99 to 2.55]	[2.16]	[1.85]	[0.05 to 2.22]
L. Log(No. of Infl. Sentences)		0.016	0.000 to 0.026		0.008	-0.007 to 0.017
		[1.00]	[0.00 to 1.54]		[0.57]	[-0.62 to 1.15]
Cause Narratives						
Supply			0.030			0.018
			[2.66]			[1.47]
Fiscal			0.052			0.018
			[2.13]			[1.23]
Monetary			-0.021			0.005
			[-1.37]			[0.33]
Demand			-0.018			-0.013
			[-1.32]			[-1.00]
Wage			-0.019			-0.009
			[-1.27]			[-0.93]
Expect			0.013			-0.000
			[1.10]			[-0.02]
International			0.008			-0.006
			[0.67]			[-0.42]
Effect Narratives						
Social			0.115			0.084
			[5.36]			[3.86]
Cost			0.069			0.064
			[2.85]			[2.67]
Cost-Push			0.041			0.024
			[2.42]			[1.33]
Savings			-0.015			-0.036
			[-0.95]			[-2.01]
Govt			-0.015			-0.016
			[-0.84]			[-1.00]
Uncertain			0.008			-0.028
			[0.61]			[-1.49]
Redistribution			-0.007			0.005
			[-0.40]			[0.29]
Trade			0.006			-0.012
			[0.40]			[-0.93]
Rates			-0.006			-0.004
			[-0.34]			[-0.27]
Purchase			-0.001			0.008
			[-0.11]			[0.51]
Observations	552	551	551	416	415	405
Adj. R^2	0.901	0.901	0.901 to 0.907	0.921	0.917	0.917 to 0.921

The dependent variables are standardized standard deviations of one-year ahead and 5-10 year ahead inflation expectations from the Michigan Survey of Consumers. For each model, we include the optimal number of autoregressive lags that minimizes AIC. Columns (1) and (4) regress the standard deviation on lagged year-over-year realized inflation and the optimal number of autoregressive lags. Columns (2) and (5) add the lagged log number of inflation sentences (i.e., newspaper sentences containing the word "inflation"). Columns (3) and (6) report results from separate regressions that each include the optimal number of autoregressive lags, lagged realized inflation, the lagged log number of inflation sentences, and one narrative at a time. For lagged realized inflation and log number of inflation sentences, columns (3) and (6) show the range of coefficients and t-statistics [in brackets] across the narrative regressions. For narratives, they show the coefficient and t-statistic from the regression including that narrative. For columns (3) and (6), the adjusted R-squared values shown represent the range across all narrative regressions. Realized Inflation is the standardized year-over-year percent change in the Consumer Price Index for All Urban Consumers (All Items, U.S. City Average) from the Bureau of Labor Statistics. Narrative measures are lagged standardized monthly shares of inflation sentences from ProQuest newspaper articles that we classify as expressing that particular narrative about inflation. Sample period is March 1979 to February 2025 for one-year expectations and July 1990 to February 2025 for 5-10 year expectations. T-statistics shown in brackets are calculated using Newey and West (1994) standard errors, optimally selected using the rule-of-thumb bandwidth estimator, computed to be 5 lags. All specifications include a constant.

TABLE 8
1ST INCOME QUARTILE AVERAGE INFLATION EXPECTATIONS AND NARRATIVE SHARES

Dep. Var.: 1st (Lowest) Quartile Avg. Inflation Expectations	Horizon: 1 year			Horizon: 5-10 years		
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline						
L. YoY Realized Inflation	0.215	0.210	0.197 to 0.221	0.078	0.070	0.031 to 0.078
	[3.99]	[3.95]	[3.77 to 4.08]	[2.75]	[2.28]	[0.71 to 2.61]
L. Log(No. of Infl. Sentences)		0.057	0.028 to 0.070		0.026	0.009 to 0.038
		[1.47]	[0.70 to 1.93]		[0.95]	[0.31 to 1.30]
Cause Narratives						
Expect			0.111			-0.000
			[4.10]			[-0.02]
Supply			0.040			-0.005
			[1.45]			[-0.19]
International			-0.027			-0.048
			[-0.99]			[-1.82]
Demand			0.027			-0.080
			[0.90]			[-2.56]
Monetary			0.025			-0.034
			[0.79]			[-1.12]
Wage			-0.010			-0.037
			[-0.39]			[-1.40]
Fiscal			0.006			0.024
			[0.17]			[0.79]
Effect Narratives						
Social			0.130			0.092
			[3.32]			[1.80]
Uncertain			0.074			-0.005
			[2.39]			[-0.13]
Cost			0.080			0.101
			[2.08]			[2.17]
Purchase			-0.038			-0.009
			[-1.28]			[-0.33]
Rates			0.036			-0.002
			[0.97]			[-0.07]
Savings			0.022			-0.050
			[0.82]			[-1.19]
Cost-Push			0.025			0.063
			[0.67]			[1.33]
Trade			0.018			-0.000
			[0.56]			[-0.01]
Govt			0.013			0.044
			[0.41]			[1.62]
Redistribution			-0.001			0.038
			[-0.03]			[1.20]
Observations	522	521	521	395	394	394
Adj. R^2	0.584	0.586	0.585 to 0.596	0.683	0.676	0.676 to 0.681

The dependent variables are standardized weighted average one-year ahead and 5-10 year ahead inflation expectations from respondents in the 1st (Lowest) Quartile income group in the Michigan Survey of Consumers. For each model, we include the optimal number of autoregressive lags that minimizes AIC. Columns (1) and (4) regress expectations on lagged year-over-year realized inflation and the optimal number of autoregressive lags. Columns (2) and (5) add the lagged log number of inflation sentences to the model (i.e., newspaper sentences containing the word "inflation"). Columns (3) and (6) report results from separate regressions that each include the optimal number of autoregressive lags, lagged realized inflation, the lagged log number of inflation sentences, and one narrative at a time. For lagged realized inflation and log number of inflation sentences, columns (3) and (6) show the range of coefficients and t-statistics [in brackets] across the narrative regressions. For narratives, they show the coefficient and t-statistic from the regression including that narrative. For columns (3) and (6), the adjusted R-squared values shown represent the range across all narrative regressions. Realized Inflation is the standardized year-over-year percent change in the Consumer Price Index for All Urban Consumers (All Items, U.S. City Average) from the Bureau of Labor Statistics. Narrative measures are lagged standardized monthly shares of inflation sentences from ProQuest newspaper articles that we classify as expressing that particular narrative about inflation. Sample period is September 1981 to January 2025 for one-year expectations and April 1992 to January 2025 for 5-10 year expectations. T-statistics shown in brackets are calculated using Newey and West (1994) standard errors, optimally selected using the rule-of-thumb bandwidth estimator, computed to be 5 lags. All specifications include a constant.

TABLE 9
2ND INCOME QUARTILE AVERAGE INFLATION EXPECTATIONS AND NARRATIVE SHARES

Dep. Var.: 2nd Quartile Avg. Inflation Expectations	Horizon: 1 year			Horizon: 5-10 years		
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline						
L. YoY Realized Inflation	0.283	0.282	0.271 to 0.292	0.055	0.064	0.029 to 0.067
	[6.53]	[6.58]	[6.06 to 6.74]	[1.40]	[1.29]	[0.48 to 1.39]
L. Log(No. of Infl. Sentences)		0.014	-0.007 to 0.029		-0.007	-0.038 to 0.001
		[0.52]	[-0.28 to 1.03]		[-0.18]	[-1.08 to 0.03]
Cause Narratives						
Expect			0.045			-0.018
			[1.81]			[-0.57]
Monetary			0.043			-0.038
			[1.80]			[-1.04]
Demand			0.021			-0.014
			[0.83]			[-0.45]
Fiscal			0.023			0.056
			[0.65]			[1.31]
International			-0.008			-0.068
			[-0.35]			[-2.33]
Wage			0.007			-0.005
			[0.32]			[-0.20]
Supply			0.004			-0.004
			[0.16]			[-0.12]
Effect Narratives						
Cost			0.058			0.109
			[2.08]			[1.82]
Uncertain			0.050			-0.000
			[2.02]			[-0.00]
Social			0.058			0.051
			[1.85]			[0.94]
Trade			0.041			-0.018
			[1.53]			[-0.60]
Cost-Push			0.035			0.055
			[1.06]			[1.33]
Redistribution			0.009			0.045
			[0.37]			[1.16]
Purchase			-0.006			0.014
			[-0.22]			[0.35]
Rates			-0.005			0.037
			[-0.20]			[1.00]
Savings			-0.003			-0.089
			[-0.13]			[-1.78]
Govt			0.000			0.013
			[0.00]			[0.43]
Observations	543	542	542	396	395	395
Adj. R^2	0.700	0.701	0.700 to 0.702	0.616	0.609	0.608 to 0.615

The dependent variables are standardized weighted average one-year ahead and 5-10 year ahead inflation expectations from respondents in the 2nd Quartile income group in the Michigan Survey of Consumers. For each model, we include the optimal number of autoregressive lags that minimizes AIC. Columns (1) and (4) regress expectations on lagged year-over-year realized inflation and the optimal number of autoregressive lags. Columns (2) and (5) add the lagged log number of inflation sentences to the model (i.e., newspaper sentences containing the word "inflation"). Columns (3) and (6) report results from separate regressions that each include the optimal number of autoregressive lags, lagged realized inflation, the lagged log number of inflation sentences, and one narrative at a time. For lagged realized inflation and log number of inflation sentences, columns (3) and (6) show the range of coefficients and t-statistics [in brackets] across the narrative regressions. For narratives, they show the coefficient and t-statistic from the regression including that narrative. For columns (3) and (6), the adjusted R-squared values shown represent the range across all narrative regressions. Realized Inflation is the standardized year-over-year percent change in the Consumer Price Index for All Urban Consumers (All Items, U.S. City Average) from the Bureau of Labor Statistics. Narrative measures are lagged standardized monthly shares of inflation sentences from ProQuest newspaper articles that we classify as expressing that particular narrative about inflation. Sample period is December 1979 to January 2025 for one-year expectations and March 1992 to January 2025 for 5-10 year expectations. T-statistics shown in brackets are calculated using Newey and West (1994) standard errors, optimally selected using the rule-of-thumb bandwidth estimator, computed to be 5 lags. All specifications include a constant.

TABLE 10
3RD INCOME QUARTILE AVERAGE INFLATION EXPECTATIONS AND NARRATIVE SHARES

Dep. Var.: 3rd Quartile Avg. Inflation Expectations	Horizon: 1 year			Horizon: 5-10 years		
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline						
L. YoY Realized Inflation	0.427	0.425	0.393 to 0.440	0.033	0.041	0.017 to 0.049
	[7.40]	[7.25]	[6.79 to 7.67]	[1.07]	[1.21]	[0.39 to 1.38]
L. Log(No. of Infl. Sentences)		0.002	-0.012 to 0.020		-0.016	-0.039 to -0.008
		[0.06]	[-0.45 to 0.82]		[-0.55]	[-1.21 to -0.29]
Cause Narratives						
Monetary			0.056			-0.010
			[2.97]			[-0.38]
Fiscal			0.061			0.074
			[2.19]			[2.48]
Expect			0.047			-0.027
			[2.02]			[-1.23]
International			0.019			-0.023
			[0.93]			[-0.91]
Supply			0.017			0.009
			[0.66]			[0.29]
Demand			-0.005			0.015
			[-0.26]			[0.59]
Wage			-0.001			-0.011
			[-0.07]			[-0.52]
Effect Narratives						
Govt			0.054			0.020
			[2.37]			[0.66]
Social			0.064			0.049
			[2.30]			[1.07]
Trade			0.051			-0.009
			[2.13]			[-0.32]
Uncertain			0.030			0.066
			[1.45]			[2.40]
Savings			-0.026			0.016
			[-1.44]			[0.44]
Purchase			-0.032			0.000
			[-1.34]			[0.00]
Rates			-0.029			0.032
			[-1.33]			[1.09]
Cost			0.028			0.004
			[1.28]			[0.07]
Cost-Push			0.009			-0.013
			[0.29]			[-0.30]
Redistribution			0.003			0.008
			[0.12]			[0.22]
Observations	544	543	543	415	414	414
Adj. R^2	0.767	0.768	0.768 to 0.771	0.707	0.702	0.701 to 0.706

The dependent variables are standardized weighted average one-year ahead and 5-10 year ahead inflation expectations from respondents in the 3rd Quartile income group in the Michigan Survey of Consumers. For each model, we include the optimal number of autoregressive lags that minimizes AIC. Columns (1) and (4) regress expectations on lagged year-over-year realized inflation and the optimal number of autoregressive lags. Columns (2) and (5) add the lagged log number of inflation sentences to the model (i.e., newspaper sentences containing the word "inflation"). Columns (3) and (6) report results from separate regressions that each include the optimal number of autoregressive lags, lagged realized inflation, the lagged log number of inflation sentences, and one narrative at a time. For lagged realized inflation and log number of inflation sentences, columns (3) and (6) show the range of coefficients and t-statistics [in brackets] across the narrative regressions. For narratives, they show the coefficient and t-statistic from the regression including that narrative. For columns (3) and (6), the adjusted R-squared values shown represent the range across all narrative regressions. Realized Inflation is the standardized year-over-year percent change in the Consumer Price Index for All Urban Consumers (All Items, U.S. City Average) from the Bureau of Labor Statistics. Narrative measures are lagged standardized monthly shares of inflation sentences from ProQuest newspaper articles that we classify as expressing that particular narrative about inflation. Sample period is November 1979 to January 2025 for one-year expectations and August 1990 to January 2025 for 5-10 year expectations. T-statistics shown in brackets are calculated using Newey and West (1994) standard errors, optimally selected using the rule-of-thumb bandwidth estimator, computed to be 5 lags. All specifications include a constant.

Table 11
4th Income Quartile Average Inflation Expectations and Narrative Shares

Dep. Var.: 4th (Highest) Quartile Avg. Inflation Expectations	Horizon: 1 year			Horizon: 5-10 years		
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline						
L. YoY Realized Inflation	0.243	0.251	0.238 to 0.261	0.027	0.022	0.003 to 0.030
	[6.06]	[6.01]	[5.34 to 6.20]	[0.78]	[0.55]	[0.07 to 0.79]
L. Log(No. of Infl. Sentences)		-0.020	-0.027 to -0.012		0.013	0.005 to 0.031
		[-1.05]	[-1.48 to -0.57]		[0.33]	[0.14 to 0.80]
Cause Narratives						
Supply			0.023			-0.002
			[1.32]			[-0.06]
Monetary			0.020			-0.022
			[1.29]			[-0.72]
Expect			0.023			-0.051
			[1.27]			[-1.77]
Wage			0.017			-0.021
			[1.14]			[-1.02]
International			0.015			-0.047
			[0.98]			[-1.85]
Fiscal			0.022			0.037
			[0.86]			[1.23]
Demand			0.011			0.009
			[0.77]			[0.34]
Effect Narratives						
Trade			0.032			-0.001
			[1.97]			[-0.03]
Social			0.028			0.009
			[1.32]			[0.16]
Uncertain			0.018			-0.015
			[1.25]			[-0.52]
Rates			-0.015			-0.044
			[-1.06]			[-1.35]
Purchase			-0.015			-0.005
			[-0.70]			[-0.20]
Savings			-0.009			0.010
			[-0.68]			[0.30]
Redistribution			0.005			0.042
			[0.24]			[1.30]
Govt			0.003			0.014
			[0.17]			[0.58]
Cost			-0.002			-0.010
			[-0.11]			[-0.20]
Cost-Push			0.002			0.032
			[0.06]			[0.83]
Observations	544	543	543	413	412	412
Adj. R^2	0.840	0.840	0.840 to 0.841	0.683	0.683	0.682 to 0.684

The dependent variables are standardized weighted average one-year ahead and 5-10 year ahead inflation expectations from respondents in the 4th (Highest) Quartile income group in the Michigan Survey of Consumers. For each model, we include the optimal number of autoregressive lags that minimizes AIC. Columns (1) and (4) regress expectations on lagged year-over-year realized inflation and the optimal number of autoregressive lags. Columns (2) and (5) add the lagged log number of inflation sentences to the model (i.e., newspaper sentences containing the word "inflation"). Columns (3) and (6) report results from separate regressions that each include the optimal number of autoregressive lags, lagged realized inflation, the lagged log number of inflation sentences, and one narrative at a time. For lagged realized inflation and log number of inflation sentences, columns (3) and (6) show the range of coefficients and t-statistics [in brackets] across the narrative regressions. For narratives, they show the coefficient and t-statistic from the regression including that narrative. For columns (3) and (6), the adjusted R-squared values shown represent the range across all narrative regressions. Realized Inflation is the standardized year-over-year percent change in the Consumer Price Index for All Urban Consumers (All Items, U.S. City Average) from the Bureau of Labor Statistics. Narrative measures are lagged standardized monthly shares of inflation sentences from ProQuest newspaper articles that we classify as expressing that particular narrative about inflation. Sample period is November 1979 to January 2025 for one-year expectations and October 1990 to January 2025 for 5-10 year expectations. T-statistics shown in brackets are calculated using Newey and West (1994) standard errors, optimally selected using the rule-of-thumb bandwidth estimator, computed to be 5 lags. All specifications include a constant.

TABLE 12
LOW - HIGH INCOME QUARTILE INFLATION EXPECTATIONS GAP AND NARRATIVE SHARES

Dep. Var.: Lowest - Highest Income Quartile Inflation Expectations Gap	Horizon: 1 year			Horizon: 5-10 years		
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline						
L. YoY Realized Inflation	0.188	0.149	0.087 to 0.164	0.115	0.096	0.040 to 0.106
	[4.29]	[3.65]	[1.94 to 3.79]	[2.88]	[2.53]	[0.86 to 2.98]
L. Log(No. of Infl. Sentences)		0.097	0.073 to 0.112		0.041	0.021 to 0.056
		[1.94]	[1.53 to 2.14]		[1.17]	[0.56 to 1.49]
Cause Narratives						
International			-0.108			-0.038
			[-3.07]			[-1.09]
Wage			-0.048			-0.055
			[-1.77]			[-1.66]
Supply			-0.053			-0.018
			[-1.51]			[-0.56]
Expect			0.039			0.012
			[1.14]			[0.47]
Fiscal			-0.031			-0.003
			[-0.84]			[-0.08]
Demand			-0.017			-0.107
			[-0.49]			[-3.00]
Monetary			0.020			-0.033
			[0.42]			[-0.95]
Effect Narratives						
Cost			0.128			0.126
			[2.30]			[2.28]
Social			0.114			0.117
			[1.87]			[2.03]
Trade			-0.048			0.011
			[-1.34]			[0.33]
Govt			0.047			0.065
			[1.21]			[1.91]
Cost-Push			0.061			0.050
			[1.19]			[0.90]
Purchase			0.036			0.000
			[1.06]			[0.00]
Uncertain			-0.029			0.004
			[-0.79]			[0.10]
Rates			0.025			0.015
			[0.64]			[0.40]
Savings			-0.017			-0.042
			[-0.52]			[-0.81]
Redistribution			-0.003			0.039
			[-0.07]			[1.02]
Observations	524	523	523	395	394	394
Adj. R^2	0.279	0.278	0.277 to 0.287	0.528	0.522	0.521 to 0.531

The dependent variables are standardized differences between weighted average inflation expectations of the lowest income quartile and the highest income quartile from the Michigan Survey of Consumers. For each model, we include the optimal number of autoregressive lags that minimizes AIC. Columns (1) and (4) regress the expectation gap on its lagged realized inflation and the optimal number of autoregressive lags. Columns (2) and (5) add the lagged log number of inflation sentences to the model (i.e., newspaper sentences containing the word "inflation"). Columns (3) and (6) report results from separate regressions that each include the optimal number of autoregressive lags, lagged realized inflation, the lagged log number of inflation sentences, and one narrative at a time. For lagged realized inflation and log number of inflation sentences, columns (3) and (6) show the range of coefficients and t-statistics [in brackets] across the narrative regressions. For narratives, they show the coefficient and t-statistic from the regression including that narrative. For columns (3) and (6), the adjusted R-squared values shown represent the range across all narrative regressions. Realized Inflation is the standardized year-over-year percent change in the Consumer Price Index for All Urban Consumers (All Items, U.S. City Average) from the Bureau of Labor Statistics. Narrative measures are lagged standardized monthly shares of inflation sentences from ProQuest newspaper articles that we classify as expressing that particular narrative about inflation. Sample period is July 1981 to January 2025 for one-year expectations and April 1992 to January 2025 for 5-10 year expectations. T-statistics shown in brackets are calculated using Newey and West (1994) standard errors, optimally selected using the rule-of-thumb bandwidth estimator, computed to be 5 lags. All specifications include a constant.

Table 13
College Educated Average Inflation Expectations and Narrative Shares

Dep. Var.: College Graduates Avg. Inflation Expectations	Horizon: 1 year			Horizon: 5-10 years		
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline						
L. YoY Realized Inflation	0.128	0.136	0.122 to 0.142	0.044	0.047	0.037 to 0.051
	[2.49]	[2.58]	[2.22 to 2.67]	[1.12]	[1.19]	[0.94 to 1.31]
L. Log(No. of Infl. Sentences)		-0.015	-0.027 to -0.007		-0.006	-0.025 to 0.000
		[-0.96]	[-1.67 to -0.39]		[-0.19]	[-0.74 to 0.01]
Cause Narratives						
Expect			0.026			-0.026
			[1.64]			[-0.93]
Demand			0.019			0.036
			[1.61]			[1.27]
Wage			0.028			-0.021
			[1.52]			[-0.85]
International			0.013			-0.043
			[1.07]			[-1.70]
Fiscal			0.021			0.040
			[0.94]			[1.00]
Supply			0.012			0.021
			[0.69]			[0.75]
Monetary			0.004			0.003
			[0.24]			[0.12]
Effect Narratives						
Uncertain			0.028			0.055
			[2.00]			[2.10]
Trade			0.019			-0.007
			[1.24]			[-0.21]
Rates			-0.017			0.003
			[-1.05]			[0.10]
Cost-Push			-0.018			0.012
			[-0.86]			[0.41]
Redistribution			0.013			-0.010
			[0.70]			[-0.36]
Purchase			-0.011			-0.009
			[-0.69]			[-0.39]
Social			0.013			-0.006
			[0.68]			[-0.17]
Savings			-0.007			0.029
			[-0.54]			[0.84]
Govt			0.003			0.013
			[0.13]			[0.44]
Cost			-0.002			-0.009
			[-0.12]			[-0.23]
Observations	546	546	546	399	399	399
Adj. R^2	0.873	0.873	0.872 to 0.873	0.668	0.667	0.666 to 0.668

The dependent variables are standardized weighted average one-year ahead and 5-10 year ahead inflation expectations from respondents in the College Graduates group in the Michigan Survey of Consumers. For each model, we include the optimal number of autoregressive lags that minimizes AIC. Columns (1) and (4) regress expectations on lagged year-over-year realized inflation and the optimal number of autoregressive lags. Columns (2) and (5) add the lagged log number of inflation sentences to the model (i.e., newspaper sentences containing the word "inflation"). Columns (3) and (6) report results from separate regressions that each include the optimal number of autoregressive lags, lagged realized inflation, the lagged log number of inflation sentences, and one narrative at a time. For lagged realized inflation and log number of inflation sentences, columns (3) and (6) show the range of coefficients and t-statistics [in brackets] across the narrative regressions. For narratives, they show the coefficient and t-statistic from the regression including that narrative. For columns (3) and (6), the adjusted R-squared values shown represent the range across all narrative regressions. Realized Inflation is the standardized year-over-year percent change in the Consumer Price Index for All Urban Consumers (All Items, U.S. City Average) from the Bureau of Labor Statistics. Narrative measures are lagged standardized monthly shares of inflation sentences from ProQuest newspaper articles that we classify as expressing that particular narrative about inflation. Sample period is January 1978 to June 2024 for one-year expectations and April 1990 to June 2024 for 5-10 year expectations. T-statistics shown in brackets are calculated using Newey and West (1994) standard errors with 5 lags.

Table 14
Non-College Educated Average Inflation Expectations and Narrative Shares

Dep. Var.: Non-College Graduates Avg. Inflation Expectations	Horizon: 1 year			Horizon: 5-10 years		
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline						
L. YoY Realized Inflation	0.100	0.108	0.100 to 0.128	0.080	0.075	0.041 to 0.089
	[1.80]	[1.92]	[1.70 to 2.37]	[2.41]	[2.12]	[0.77 to 2.54]
L. Log(No. of Infl. Sentences)		0.028	0.018 to 0.033		0.046	0.026 to 0.054
		[1.04]	[0.60 to 1.25]		[1.49]	[1.00 to 1.73]
Cause Narratives						
Expect			0.052			0.012
			[2.14]			[0.54]
Supply			0.027			-0.040
			[1.36]			[-1.41]
Wage			0.022			-0.022
			[0.95]			[-0.88]
Fiscal			0.022			0.054
			[0.78]			[1.66]
Demand			0.004			-0.034
			[0.23]			[-0.90]
Monetary			0.003			-0.005
			[0.14]			[-0.16]
International			0.000			-0.040
			[0.01]			[-1.63]
Effect Narratives						
Cost			0.038			0.075
			[1.22]			[1.03]
Trade			0.022			-0.004
			[1.17]			[-0.10]
Govt			0.029			0.030
			[1.14]			[1.09]
Social			0.040			0.039
			[1.05]			[0.53]
Purchase			-0.019			-0.028
			[-0.99]			[-1.17]
Uncertain			0.022			-0.017
			[0.98]			[-0.47]
Savings			-0.013			-0.036
			[-0.87]			[-0.77]
Rates			0.011			0.024
			[0.47]			[0.78]
Cost-Push			0.006			0.072
			[0.19]			[1.29]
Redistribution			-0.001			0.044
			[-0.05]			[1.01]
Observations	554	553	553	407	406	406
Adj. R^2	0.803	0.804	0.803 to 0.806	0.618	0.590	0.589 to 0.593

The dependent variables are standardized weighted average one-year ahead and 5-10 year ahead inflation expectations from respondents in the Non-College Graduates group in the Michigan Survey of Consumers. For each model, we include the optimal number of autoregressive lags that minimizes AIC. Columns (1) and (4) regress expectations on lagged year-over-year realized inflation and the optimal number of autoregressive lags. Columns (2) and (5) add the lagged log number of inflation sentences to the model (i.e., newspaper sentences containing the word "inflation"). Columns (3) and (6) report results from separate regressions that each include the optimal number of autoregressive lags, lagged realized inflation, the lagged log number of inflation sentences, and one narrative at a time. For lagged realized inflation and log number of inflation sentences, columns (3) and (6) show the range of coefficients and t-statistics [in brackets] across the narrative regressions. For narratives, they show the coefficient and t-statistic from the regression including that narrative. For columns (3) and (6), the adjusted R-squared values shown represent the range across all narrative regressions. Realized Inflation is the standardized year-over-year percent change in the Consumer Price Index for All Urban Consumers (All Items, U.S. City Average) from the Bureau of Labor Statistics. Narrative measures are lagged standardized monthly shares of inflation sentences from ProQuest newspaper articles that we classify as expressing that particular narrative about inflation. Sample period is January 1978 to February 2025 for one-year expectations and April 1990 to February 2025 for 5-10 year expectations. T-statistics shown in brackets are calculated using Newey and West (1994) standard errors with 5 lags.

TABLE 15
NON-COLLEGE/COLLEGE EDUCATED INFLATION EXPECTATIONS GAP AND NARRATIVE SHARES

Dep. Var.: Non-College - College Graduates Inflation Expectations Gap	Horizon: 1 year			Horizon: 5-10 years		
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline						
L. YoY Realized Inflation	0.005	-0.032	-0.155 to -0.025	0.119	0.098	0.033 to 0.123
	[0.11]	[-0.71]	[-2.98 to -0.56]	[2.50]	[2.19]	[0.65 to 2.61]
L. Log(No. of Infl. Sentences)		0.076	0.061 to 0.090		0.052	0.031 to 0.067
		[1.99]	[1.68 to 2.41]		[1.23]	[0.81 to 1.49]
Cause Narratives						
Expect			0.058			0.022
			[1.87]			[0.72]
Wage			-0.045			-0.023
			[-1.56]			[-0.64]
International			-0.045			-0.021
			[-1.51]			[-0.55]
Fiscal			0.035			0.069
			[0.91]			[1.59]
Demand			-0.032			-0.053
			[-0.84]			[-1.45]
Monetary			0.018			0.014
			[0.61]			[0.52]
Supply			0.019			-0.067
			[0.53]			[-1.80]
Effect Narratives						
Social			0.169			0.123
			[3.43]			[2.98]
Cost			0.130			0.087
			[2.62]			[1.65]
Cost-Push			0.086			0.051
			[1.89]			[0.85]
Savings			-0.032			-0.056
			[-1.22]			[-1.49]
Govt			0.028			0.017
			[0.74]			[0.44]
Purchase			0.010			-0.052
			[0.31]			[-1.42]
Rates			0.004			-0.001
			[0.11]			[-0.02]
Uncertain			0.003			-0.043
			[0.08]			[-1.02]
Trade			-0.002			0.034
			[-0.06]			[0.84]
Redistribution			-0.002			0.033
			[-0.05]			[1.07]
Observations	546	546	546	399	399	399
Adj. R^2	0.418	0.421	0.420 to 0.435	0.463	0.464	0.462 to 0.472

The dependent variables are standardized differences between weighted average inflation expectations of non-college graduates and college graduates from the Michigan Survey of Consumers. For each model, we include the optimal number of autoregressive lags that minimizes AIC. Columns (1) and (4) regress the expectation gap on lagged year-over-year realized inflation and the optimal number of autoregressive lags. Columns (2) and (5) add the lagged log number of inflation sentences to the model (i.e., newspaper sentences containing the word "inflation"). Columns (3) and (6) report results from separate regressions that each include the optimal number of autoregressive lags, lagged realized inflation, the lagged log number of inflation sentences, and one narrative at a time. For lagged realized inflation and log number of inflation sentences, columns (3) and (6) show the range of coefficients and t-statistics [in brackets] across the narrative regressions. For narratives, they show the coefficient and t-statistic from the regression including that narrative. For columns (3) and (6), the adjusted R-squared values shown represent the range across all narrative regressions. Realized Inflation is the standardized year-over-year percent change in the Consumer Price Index for All Urban Consumers (All Items, U.S. City Average) from the Bureau of Labor Statistics. Narrative measures are lagged standardized monthly shares of inflation sentences from ProQuest newspaper articles that we classify as expressing that particular narrative about inflation. Sample period is January 1978 to June 2024 for one-year expectations and April 1990 to June 2024 for 5-10 year expectations. T-statistics shown in brackets are calculated using Newey and West (1994) standard errors with 5 lags.

Table 16
Realized Inflation and Narrative Shares

Dep. Var.: Year-over-Year Inflation		
	(1)	(2)
Baseline		
L. Log(No. of Infl. Sentences)	0.002	0.001-0.013
	[0.41]	[0.16-1.62]
Cause Narratives		
Wage		0.008
		[1.85]
Fiscal		0.014
		[1.58]
Demand		0.006
		[1.14]
Supply		0.005
		[0.87]
Monetary		-0.003
		[-0.75]
Expect		0.002
		[0.44]
International		-0.000
		[-0.02]
Effect Narratives		
Savings		0.007
		[1.50]
Rates		-0.009
		[-1.26]
Redistribution		0.006
		[1.06]
Social		0.004
		[0.84]
Govt		0.003
		[0.58]
Purchase		0.003
		[0.51]
Cost		0.003
		[0.47]
Uncertain		0.002
		[0.45]
Trade		0.002
		[0.38]
Cost-Push		-0.000
		[-0.01]
Observations	922	922
Adj. R^2	0.981	0.980-0.981

The dependent variable is standardized year-over-year percent change in the Consumer Price Index for All Urban Consumers (All Items, U.S. City Average) from the Bureau of Labor Statistics. For each model, we include the optimal number of autoregressive lags that minimizes AIC. Column (1) regresses realized inflation on the lagged log number of inflation sentences with the optimal number of autoregressive lags. Column (2) reports results from separate regressions that each include the optimal number of autoregressive lags, the lagged log number of inflation sentences, and one narrative at a time. For the log number of inflation sentences, column (2) shows the range of coefficients and t-statistics [in brackets] across the narrative regressions. For narratives, it shows the coefficient and t-statistic from the regression including that narrative. The adjusted R-squared values in column (2) represent the range across all narrative regressions. Narrative measures are lagged standardized monthly shares of inflation sentences from ProQuest newspaper articles that we classify as expressing that particular narrative about inflation. Sample period is April 1948 to January 2025. T-statistics shown in brackets are calculated using Newey and West (1994) standard errors, optimally selected using the rule-of-thumb bandwidth estimator, computed to be 6 lags. All specifications include a constant.

For Online Publication: Appendix

A TopicGPT Narrative Discovery Prompts

This sections contains the complete prompt templates used for narrative discovery using the TopicGPT framework, described in Section 4.7.

A.1 Generation Prompt

The generation prompt identifies causal narratives about inflation from news articles, extracting both causes and effects of inflation.

A.1.1 Task Description and Seed Narrative

Generation Prompt - Task & Seed Narrative

```
generation_prompt = """
You will receive a document related to inflation. Your task is to identify and summarize causal narratives expressed in the document. Specifically, determine:
1. Causes of inflation     Broad, generalizable factors that contribute to inflation.
2. Effects of inflation     The main consequences or outcomes of inflation.

Your response should not focus on document-specific details but should generalize findings into broad, widely applicable causes and effects.

[Seed Narrative]
[1] [Cause] Monetary Policy: Excessive money supply or low interest rates can increase inflation by boosting demand beyond supply capacity.

In this seed narrative, the "[1]" represents the hierarchy level of the narrative, which is the first level of the narrative hierarchy in this case. Then, "[Cause]" represents the type of the narrative, which is "Cause" in this case. Finally, "Monetary Policy: Excessive money supply or low interest rates can increase inflation by boosting demand beyond supply capacity." is the cause label and the explanation of the cause.
```

A.1.2 Examples

Generation Prompt - Examples

```
[Examples]
Example 1:
Document:
"The surge in energy prices and disruptions in global supply chains have ↩
    ↪ significantly contributed to rising inflation rates."

Your response:
[1] [Cause] Energy Price Surge: Higher energy costs increase production ↩
    ↪ expenses, leading to inflation.
[1] [Cause] Supply Chain Disruptions: Limited supply of goods raises prices ↩
    ↪ across industries.

Example 2:
Document:
"Inflation has led to increased borrowing costs and a decline in consumer ↩
    ↪ purchasing power."

Your response:
[1] [Effect] Higher Borrowing Costs: Central banks raise interest rates to ↩
    ↪ curb inflation, making loans more expensive.
[1] [Effect] Declining Purchasing Power: Rising prices reduce the real ↩
    ↪ value of consumers' income.
```

A.1.3 Instructions and Response Format

Generation Prompt - Instructions & Template

```
[Instructions]
Step 1: Analyze the Document
- Identify whether the document discusses causes of inflation or effects of ↩
    ↪ inflation (or both).
- Ensure the identified causes and effects are generalizable and not ↩
    ↪ specific to a single case.

Step 2: Summarize the Findings in the Following Format:
Please ONLY return the relevant or modified cause and effect relationships. ↩
    ↪ Your response should be in the following format:
[1] [Cause] Cause Label: Brief explanation of how this factor contributes ↩
    ↪ to inflation.
[1] [Effect] Effect Label: Brief explanation of how inflation leads to this ↩
    ↪ outcome.
If the document contains no causal relationships, return None.

[Document]
{Document}

Your response:
""".strip()
```

A.2 Refinement Prompt

The refinement prompt merges similar or duplicate narratives within the same hierarchical level to reduce redundancy and consolidate conceptually similar ideas.

Refinement Prompt Template

```
refinement_prompt = """
You will receive a list of narratives that belong to the same level of a narrative hierarchy. Your task is to analyze these narratives and merge those that are paraphrases, near duplicates, or that express the same underlying concept. However, ensure that the merged narrative is generalized and conceptually unified rather than merely combining details.

[Examples]
Example 1:
Narrative List:
[1] [Cause] Monetary Policy: Excessive money supply increases inflation.
[1] [Cause] Central Bank Policy: Loose monetary policy, such as a low interest rate environment, contributes to rising inflation.
Your response:
[1] [Cause] Monetary Policy: Overly expansionary monetary and central bank policies can drive inflation by significantly increasing aggregate demand.

Example 2:
Narrative List:
[1] [Effect] Higher Borrowing Costs: Inflation leads to higher interest rates, making loans more expensive.
[1] [Effect] Reduced Purchasing Power: Inflation erodes the real value of income, limiting consumer spending.
Your response:
None

[Instructions]
- Each line represents a narrative, identified by a level indicator and a narrative type (either [Cause] or [Effect]).
- Merge narratives only when they refer to the same underlying idea.
- When merging, generate a generalized narrative that captures the shared concept.
- If no appropriate merging is identified, simply return None.
- DO NOT include any additional commentary or notes beyond the required output. ONLY return the merged narrative separated by newline.

[Narrative List]
{Narratives}

Your response:
""".strip()
```

A.3 Assignment Prompt

The assignment prompt classifies documents by selecting the most relevant narrative(s) from the final TopicGPT taxonomy.

A.3.1 Task Description and Example Document

Assignment Prompt - Task & Example Document

```
assignment_prompt = """
You will receive a document related to inflation. Additionally, you will ↩
    ↪ receive a list of causal narratives. Your task is to assign the ↩
    ↪ document the most relevant causal narrative(s) from the list.

[Examples]
Example 1:
Document:
"The surge in energy prices and disruptions in global supply chains have ↩
    ↪ significantly contributed to rising inflation rates."
```

A.3.2 Cause Narratives

Assignment Prompt - Cause Narratives

```
Narrative List:
[1] [Cause] Demand-Supply Imbalance: When demand for goods and services ↩
    ↪ exceeds supply due to production bottlenecks, economic recovery, or ↩
    ↪ other constraints, prices rise, fueling inflation.
[1] [Cause] Exchange Rate and Inflation Expectations: Depreciation of a ↩
    ↪ currency raises import costs and domestic prices, while anticipated ↩
    ↪ future price increases drive businesses to raise prices and consumers ↩
    ↪ to spend more, reinforcing inflation.
[1] [Cause] Labor Market Pressures: Wage growth and labor shortages ↩
    ↪ increase production costs, driving higher consumer prices and ↩
    ↪ contributing to inflation.
[1] [Cause] Monetary and Fiscal Drivers: Excessive fiscal spending, ↩
    ↪ expansionary monetary policies, or significant increases in the money ↩
    ↪ supply boost aggregate demand relative to supply, contributing to ↩
    ↪ inflationary pressures.
[1] [Cause] Supply-Side Constraints: Rising production costs, such as ↩
    ↪ energy, raw materials, or labor costs, along with supply chain ↩
    ↪ disruptions, geopolitical conflicts, or other external shocks, limit ↩
    ↪ supply, increase business expenses, and elevate consumer prices, ↩
    ↪ driving inflation.
```

A.3.3 Effect Narratives and Example Response

Assignment Prompt - Effect Narratives & Example

```
[1] [Effect] Economic Slowdown Risks: Inflation and efforts to combat it, ↩
    ↪ such as interest rate hikes, dampen consumer and business spending, ↩
    ↪ increase debt burdens, slow economic growth, and heighten the risks ↩
    ↪ of recession or stagflation.
[1] [Effect] Eroded Purchasing Power: Inflation decreases the real value of ↩
    ↪ incomes, wages, savings, and investments, limiting affordability and ↩
    ↪ disproportionately affecting lower-income households.
[1] [Effect] Higher Borrowing and Business Costs: Inflation raises interest ↩
    ↪ rates and increases input, production, and operational expenses, ↩
    ↪ reducing borrowing, spending, profitability, and economic ↩
    ↪ competitiveness.
[1] [Effect] Market Volatility and Behavioral Shifts: Inflation triggers ↩
    ↪ uncertainty in financial markets, alters consumer spending patterns, ↩
    ↪ and drives businesses to adapt through workforce adjustments or ↩
    ↪ cost-saving measures.
[1] [Effect] Rising Living Costs and Income Inequality: Inflation increases ↩
    ↪ the cost of essential goods and services while exacerbating financial ↩
    ↪ strain for lower-income groups and widening income disparities.
[1] [Effect] Social and Political Instability: Inflation worsens ↩
    ↪ inequality, redistributes economic burdens, erodes trust in ↩
    ↪ leadership, and contributes to political unrest and societal strain.

Your response:
[1] [Cause] Demand-Supply Imbalance: When demand for goods and services ↩
    ↪ exceeds supply due to production bottlenecks, economic recovery, or ↩
    ↪ other constraints, prices rise, fueling inflation.
```

A.3.4 Instructions and Response Template

Assignment Prompt - Final Instructions & Template

```
[Instructions]
- If the document contains no causal relationships, return None.
- It is possible that the document contains multiple causal relationships. ↩
    ↪ In this case, return all the causal relationships that are relevant ↩
    ↪ to the document, separated by \n.
- DO NOT include any additional commentary or notes beyond the required ↩
    ↪ output. ONLY return the causal relationships.

[Document]
{Document}

[Narrative List]
{Narratives}

Your response:
""".strip()
```

A.4 Evaluation Prompt

The evaluation prompt measures the semantic similarity between the TopicGPT discovered narratives and our manual taxonomy based on a 0-10 scale.

Evaluation Prompt Template

```
eval_prompt = """
Task: Check two summarized causal relations and determine how similar they ↩
    ↪ are.

Instructions:
1. You will be given two causal relations about inflation. For example:
     Input 1: [Cause] Inflation Dynamics: Inflation arises from a ↩
    ↪ combination of factors, including increased demand, supply chain ↩
    ↪ disruptions, rising production costs, and monetary or fiscal policy ↩
    ↪ decisions, collectively driving upward pressure on prices.
     Input 2: [Cause] Demand-side factors: Pull-side or demand-pull ↩
    ↪ inflation. Occurs when (aggregate) demand for goods and services ↩
    ↪ exceeds (aggregate) supply. Leads to rise in or upward pressure on ↩
    ↪ prices due to the scarcity (or imbalance) of goods or services ↩
    ↪ relative to demand.
2. Your job is to determine how similar these two causal relations are.
3. Output the score ranging from 0 to 10. 0 means the two causal relations ↩
    ↪ are completely different, 10 means the two causal relations are ↩
    ↪ exactly the same.
4. You must ONLY output the score without any other text.

Input 1: {input_1}
Input 2: {input_2}

Score:
""".strip()
```