BIS

BIS Working Papers
No 1294

Parsing the pulse: decomposing macroeconomic sentiment with LLMs

by Byeungchun Kwon, Taejin Park, Phurichai Rungcharoenkitkul and Frank Smets

Monetary and Economic Department

October 2025

JEL classification: E30, E44, E60, C55, C82

Keywords: macroeconomic sentiment, growth, inflation, monetary policy, fiscal policy, LLMs, machine learning

This publication is available on the BIS website (www.bis.org).

ISSN 1020-0959 (print)
ISSN 1682-7678 (online)

Parsing the Pulse: Decomposing Macroeconomic Sentiment with LLMs*

Byeungchun Kwon† Taejin Park‡ Phurichai Rungcharoenkitkul§
Frank Smets¶

1 October 2025

Abstract

Macroeconomic indicators provide quantitative signals that must be pieced together and interpreted by economists. We propose a reversed approach of parsing press narratives directly using Large Language Models (LLM) to recover growth and inflation sentiment indices. A key advantage of this LLM-based approach is the ability to decompose aggregate sentiment into its drivers, readily enabling an interpretation of macroeconomic dynamics. Our sentiment indices track hard-data counterparts closely, providing an accurate, near real-time picture of the macroeconomy. Their components–demand, supply, and deeper structural forces–are intuitive and consistent with prior model-based studies. Incorporating sentiment indices improves the forecasting performance of simple statistical models, pointing to information unspanned by traditional data.

JEL Classification: E30, E44, E60, C55, C82
Keywords: macroeconomic sentiment, growth, inflation, monetary policy, fiscal policy, LLMs, machine learning

*We thank seminar participants at the Bank for International Settlements for helpful comments. All remaining errors are ours. The views expressed are those of the authors and do not necessarily represent those of the Bank for International Settlements. The resulting indices constructed in this paper will be made available on the BIS website.

†Bank for International Settlements, byeungchun.kwon@bis.org
‡Bank for International Settlements, taejin.park@bis.org
§Bank for International Settlements, phurichai.rungcharoenkitkul@bis.org
¶Bank for International Settlements, frank.smets@bis.org

1 Introduction

The emergence of Large Language Models (LLMs), built on transformer architectures, marks a significant advance in textual analysis. Trained on massive corpora with billions of parameters, these models excel at learning the probabilistic structure of language and are capable of predicting words based on preceding sequences. Unlike previous generations of tools that rely on frequency-based metrics or simple word co-occurrence, LLMs can capture complex syntactic and semantic relationships, including subtleties such as negation, sentiment shifts and broader discourse context. The capability of LLMs has vast potential and has found wide-ranging use cases. Macroeconomic applications have also started to emerge, though these remain relatively unexplored.

In this paper, we use LLMs to extract macroeconomic sentiment from financial news and analyse its underlying drivers. Focusing on US news coverage by the *WSJ* since 2000, we construct sentiment indices for growth and inflation by leveraging LLMs to interpret the contextual meaning of news articles and characterise how they describe prevailing economic conditions. These sentiment scores reflect the average tone of coverage—how upbeat or downbeat it is regarding growth, and how elevated or subdued inflationary pressures appear over time. We then use LLMs to identify the principal drivers of growth and inflation sentiment, distinguishing between demand- and supply-side factors which entail different policy implications. Finally, we extract more granular insights into the specific components shifting demand and supply. We differentiate between financial and non-financial demand drivers–which may reflect monetary or fiscal policies–and key supply-side factors such as commodity prices, supply chain conditions and supply-side government policies (e.g. tariffs).

The LLM-based approach to parsing macroeconomic sentiment has a close analogue in the traditional time-series analysis. Conventional VAR models rely on "hard data" such as GDP and inflation to estimate their dynamic relationships and employ identification strategies to infer structural drivers. Such identification works by imposing restrictions on how shocks should affect different variables (e.g. Choleski identification and sign restrictions). Our LLM-based approach bypasses these statistical inferences and directly draws on human-curated narratives to construct overall macroeconomic sentiment as well as identify its drivers. There is a parallel between our approach and the narrative restrictions in the VAR context, which also rely on human-identified exogenous events to trace the impact of such shocks through the model's lens. Our methodology is an entirely narrative-driven approach to recovering sentiment and decomposing its drivers simultaneously, dispensing with numerical data and statistical models altogether.

Our textual analysis produces growth and inflation sentiment scores that bear remarkably close resemblance to the corresponding hard data, but available at a higher frequency. Growth sentiment closely tracks US activity indicators such as employment and coincident indices, while inflation sentiment also matches the dynamics of headline and core inflation. Detailed drivers of macroeconomic sentiment are intuitive, moving with their empirical counterparts in line with theory. For example, the real and financial demand drivers are highly correlated with personal consumption and financial condition indicators, respectively. The roles of supply drivers such as trade policy and supply chain disruptions are clearly visible and accord well with corresponding indicators. The relative importance of demand and supply drivers is also broadly consistent with previous model-based studies. The overall close proximity between our sentiment indices and hard data is one key source of empirical validation, particularly as we have not used any numerical data as part of the analysis *ex ante*.

One reason for the tight correlation between sentiments and empirical data may simply be that news coverage and its narrative are shaped by incoming economic releases. For example, against the backdrop of high and persistent inflation, the sentiment regarding inflation is likely to point to rising prices. That said, it is also possible that sentiment shifts themselves have bearing on how the economy evolves, e.g. rising inflation expectations, for whatever reasons, could drive up inflation. More generally, the influence of macroeconomic *narrative* could have far-reaching effects on how agents formulate plans and make decisions, potentially driving economic fluctuations (see Shiller (2017)). The LLM-based sentiment analysis is one approach to assess this influence in a quantitative manner.

LLM-based macroeconomic sentiment analysis has multiple uses and offers unique values for policymaking. Most directly, it provides a real-time high-frequency quantitative assessment of the state of the economy, complementing the use of hard data for surveillance and nowcasting. To the extent that our macroeconomic sentiment contains forward-looking information unspanned by past data, or captures the influence of narrative on macroeconomic evolution à-la Shiller (2017), it may be useful for forecasting as well. We provide some evidence of such forecasting abilities. Beyond nowcasting and forecasting, LLM-based analysis allows a decomposition of sentiment which helps policymakers diagnose and systematically track structural drivers of the economy, complementing existing model-based approaches. We show that our sentiment sub-indices contain information about structural shocks identified in other studies. Being model-free, the LLM-powered framework is highly flexible and can be extended to consider various applications beyond standard macroeconomic analysis.

Literature The application of textual analysis to economics has progressed through three methodological waves, each expanding the scope of what can be extracted from unstructured text. As Gentzkow et al. (2019) highlight, the digitisation of communication has generated unprecedented opportunities for economists to harness *text as data*.

The first wave relied on dictionary and frequency-based methods, establishing that simple counts could produce valuable economic indicators. Tetlock (2007) showed that the frequency of negative words in *Wall Street Journal (WSJ)* columns predicted stock returns. Baker et al. (2016) constructed the widely used Economic Policy Uncertainty (EPU) index by counting references to "economic," "policy," and "uncertainty," later complemented by the World Uncertainty Index of Ahir et al. (2022) (see also the Geopolitical Risk index of Caldara and Iacoviello (2022)). Loughran and McDonald (2011) improved this approach by developing finance-specific lexicons, underscoring that economic meaning depends on context. Beyond broad uncertainty measures, Hassan et al. (2019) construct a firm-level political-risk index from earning calls, using targeted dictionaries and proximity rules, showing consequences for firms' investment and hiring.

The second wave introduced semantic and topic-based models that allowed text to "speak for itself." Word embeddings (Mikolov et al., 2013) represented words as vectors, capturing semantic similarity and overcoming the rigidity of dictionaries. Ehrmann and Talmi (2020) exploited embeddings to measure semantic distances in central bank communications and linked them to market volatility. Topic models, such as LDA (Blei et al., 2003), provided another powerful tool, uncovering latent themes in economic narratives. Hansen and McMahon (2016) applied LDA to FOMC statements and identified distinct effects of economic assessment versus forward guidance. Bybee et al. (2020) extended topic modeling to decades of *WSJ* articles, producing topic-based indices that mirrored official macroeconomic series, while Larsen and Thorsrud (2019) showed that news-derived topic indices improved GDP nowcasts relative to traditional indicators. Along similar lines, Bybee et al. (2024) apply LDA to more than 800,000 WSJ articles and show that news attention to recession-related topics closely tracks output and employment and significantly improves business cycle forecasts. Beyond measurement, text analysis can also aid identification, e.g. Aruoba and Drechsel (forthcoming) use NLP features from Fed pre-meeting staff documents plus supervised machine learning to proxy the FOMC's information set and recover monetary policy shocks.

The third wave arose with the transformer architecture (Vaswani et al., 2017), enabling large-scale contextual models pre-trained on vast corpora. BERT (Devlin et al., 2019) pioneered the fine-tuning paradigm and has since been adapted for finance and central banking (e.g., Araci 2019; Gambacorta et al. 2024). Scaling up transformers gave rise to generative LLMs, which can flexibly interpret, classify, and generate text with minimal training. Recent work demonstrates their promise for economics in several areas. They have been used for classification of policy stance, with Hansen and Kazinnik (2024) showing that GPT-4 can code FOMC stance at expert level. They have been applied to forecasting, for example Faria-e Castro and Leibovici (2024) find that LLMs outperform survey forecasts of inflation, while de Bondt and Sun (2025) show that LLM-scored PMI narratives enhance euro area GDP nowcasts. See Carriero et al. (2025) for a recent benchmark of forecast performance. Other applications focus on measurement and index construction from news, such as Audrino et al. (2024) who build uncertainty indices, Kwon et al. (2024) who combine topic modeling with few-shot prompting to explain stock-market movements, and Bybee (2025), who uses GPT to extract expectations of CPI, unemployment, and the S&P 500 from WSJ archives, extending historical expectations series back 120 years. In international settings, Clayton et al. (2023) deploy LLMs to classify which governments apply which geoeconomic instruments to which targets and trace firm-level adjustments across tools (tariffs vs export controls). For accessible surveys on LLM economic applications, see Korinek (2023) and Kwon et al. (2024).

A distinctive feature of LLMs is their ability not only to provide signals but also to uncover their drivers directly from qualitative narratives. This capability addresses a long-standing limitation of econometric and earlier text methods: working with the data or indices (e.g., uncertainty, policy tone) alone does not reveal underlying structural forces without additional identifying assumptions. For instance, identifying demand versus supply drivers in VAR context typically requires recursive (Cholesky) zero restrictions, sign restrictions (e.g., Arias et al. 2018) or narrative shocks (e.g., Romer and Romer 2004; Ramey 2011), approaches that can be subject to researchers' discretion and model uncertainty. In contrast, LLMs can classify complex narratives into economically meaningful drivers in near real time, reducing reliance on explicit econometric restrictions.

This is the approach we take. Building on the latest LLM capabilities, we apply prompting-based methods—without specialized fine-tuning—to large volumes of financial news. This allows us to recover growth and inflation sentiment indices as well as their structural drivers. The advantage of this approach is twofold: it leverages the contextual analysis of LLMs without requiring costly domain-specific training, and it provides interpretable, narrative-driven decomposition of macroeconomic dynamics that traditional econometric tools have struggled to capture.

Structure The paper is organised as follows. Section 2 describes the methodology, detailing the LLM approach to parsing text data and the construction of macroeconomic sentiment indices. Section 3 shows the resulting macroeconomic sentiment indices and their underlying drivers, comparing them with hard data and outside estimates for external validation. Section 4 examines the forecasting abilities of LLM-based macroeconomic sentiment. Section 5 concludes.

2 Methodology

2.1 Data description

Our analysis relies on a comprehensive corpus of financial news articles from the WSJ, obtained through Factiva Analytics. We deliberately choose a single news source rather than multiple outlets for several reasons. First, this approach allows us to focus on in-depth coverage of major economic developments, uncontaminated by the noise of high-frequency minor news updates in a multi-source corpus. Second, using a single source ensures consistency in editorial standards, newsworthiness criteria, and publication volume over time, eliminating potential biases from varying coverage intensity across different outlets. Third, the uniform quality and depth in the WSJ's economic reporting provides a stable basis for longitudinal sentiment analysis.

The dataset encompasses U.S. economic and financial market news articles identified through Factiva Subject Codes ECAT (Economic News) and MCAT (Financial/Commodity Markets News). Our sample spans January 2000 to April 2025, covering multiple business cycles including the dot-com recession, the 2008 Great Financial Crisis, the COVID-19 pandemic and recent inflationary episodes. The sample also covers important shifts in macroeconomic policies, from the use of unconventional monetary policies, expansionary fiscal policies and most recently tariff hikes. The corpus contains approximately 200,000 articles in English. Each article includes four key fields: title, snippet (containing the headline and lead sentences), body text and publication date. The average article length is approximately 600 words.

2.2 Model setup and classification process

We employ GPT-4.1 mini and GPT-5 mini accessed via API (accessed in July and September 2025, respectively) for our classification tasks.[1] These models were selected based on several considerations. First, it possesses sufficient context window capacity to process full-length news articles, which average 600 words in our corpus. Second, while larger models such as GPT-5 or reasoning-optimised variants might offer marginal performance improvements, the computational costs for processing over a large volume of articles made them prohibitive. Our testing indicated that the *mini* models achieved satisfactory classification accuracy while maintaining computational feasibility.

We use GPT-4.1 mini to screen the articles using the WSJ editor-curated headline snippets–short summaries of the articles. The articles are retained only if the snippets contain information relevant for growth and inflation, with clear directional sentiment. This pre-screening stage reduces the corpus from approximately 200,000 articles to 47,000 growth-related articles and 28,000 inflation-related articles. We then deploy GPT-5 mini for in-depth classification in subsequent stages (see subsection 2.3 for details).

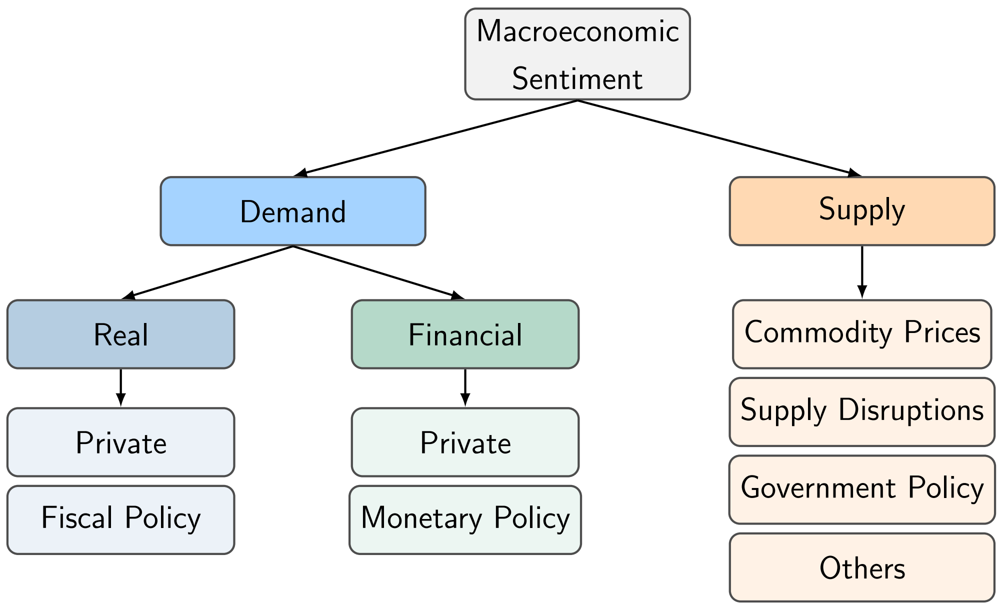
We allow four categories for directional sentiment classification. Growth sentiment articles are classified as *Positive* when indicating an expansion in macroeconomic activity,

[1]We opt for LLM-based prompting rather than embedding-based classifiers as our task requires interpreting economic sentiments and tracing causal relationships across different parts of each article. Embedding-based methods, effective for similarity search in a static setting, are ill-suited to granular, multi-step reasoning about how economic concepts interact under our classification criteria.

Negative when indicating a contraction, *Neutral* when the effects are ambiguous or offsetting, or *Null* when they are not of macroeconomic relevance. Similarly, inflation sentiment articles are classified as *Positive* for increasing inflationary pressure, *Negative* for decreasing inflationary pressure, *Neutral*, or *Null*. When articles contain opposing forces, we use the macroeconomic significance of the primary development as the criterion for determining the sentiment. Articles whose sentiment is either Neutral or Null do not feature in the aggregate sentiment scores (see subsection 2.4 below).

Together with directional sentiment, we identify the structural drivers underlying the sentiment following the taxonomy shown in Figure 1. For each instance of growth and inflation sentiment, we determine whether the primary driver is demand- or supply-side in nature. We define demand drivers as those that reflect changes in willingness to spend, thus pushing growth and inflation in the same direction. Supply drivers reflect changes in productive capacity or production costs, pulling growth and inflation in opposite directions. Cases where demand and supply factors play equally important roles are retained, while cases where the underlying drivers cannot be clearly classified are removed from the analysis.

Figure 1: Taxonomy of macroeconomic sentiment



The demand drivers are further decomposed into real and financial categories. Real drivers represent non-financial demand factors, consisting of the private-sector component, related to consumer and business confidence, and fiscal policy component. Financial drivers include the private-sector component, reflecting financial market sentiment and credit availability among others, and the monetary policy component.

The supply drivers are similarly subdivided into several categories. Commodity price movements capture changes in energy, food, or raw material prices. Government policy represents supply-side policy shifts, including trade policies, structural reforms, and regulatory changes. Supply disruptions encompass logistic bottlenecks, labour shortages and global supply chain issues. A residual category captures other supply factors.

2.3 Prompting strategy

Our prompting architecture follows a hierarchical decomposition strategy that mirrors our economic taxonomy. We instruct the model to classify the primary economic drivers (demand versus supply) and to identify their sentiment direction (positive, negative, neutral, or null), and finally to categorise these drivers into granular sub-components. This structured approach ensures that the model's reasoning follows economic logic at each classification stage.

Procedurally, the classification tasks comprise three stages (see Appendix A for full prompts): (1) applying a minimal prompt to filter for relevant articles[2], (2) main classification of primary drivers (demand, supply or both) and their directional sentiments, and (3) in-depth classification for sub-drivers. This multi-stage approach divides the complex task into manageable blocks, enabling better process control. Decomposing the task helps manage the model's context window limitations, reduce complexity at each stage for improved accuracy, and allow prompt refinement based on observed classification results. This approach also permits optimal model selection for each stage based on task complexity, e.g. the use of GPT-4.1 mini in the first stage and GPT-5 mini for subsequent more complex tasks.

To maintain classification quality, we apply several prompt engineering techniques. First, we provide explicit definitions of economic concepts within the prompts, ensuring consistent interpretation across articles. For instance, we define demand drivers as factors that "induce changes in desire to consume, spend, or invest, thus pushing macroeconomic activity and inflation in the same direction," while supply drivers are those that "affect the economy's production or output capacity, pushing macroeconomic activity and inflation in opposite directions."

Second, following the principles of in-context learning (ICL), we include in our prompts concrete examples. For demand drivers, we define "Financial private drivers" as including "financial conditions, credit conditions, cost of funds, interest rates, exchange rates, liquidity conditions, balance sheet conditions," while "Real private drivers" encompass "non-financial drivers such as confidence, perception of uncertainty, expected future income." For supply drivers, we specify "Policy shifts" as "shifts in trade policy, tariffs, subsidies, structural reforms or any other government policies that affect the economy's supply side," and "Supply disruptions" as "global supply chain problems, bottlenecks, or factors related to global value chain."

[2]We require the model to classify sentiment during this filtering stage to ensure substantive engagement with content instead of superficial keyword matching.

Third, we handle ambiguous cases through explicit prioritisation rules. When articles contain both a primary development and contextual information with opposing sentiments—such as central bank rate increases (contractionary) in response to strong demand (expansionary)—we instruct the model to prioritise the sentiment of the primary development if it carries macroeconomic significance.

Fourth, we implement separate prompting pipelines for growth and inflation classification to prevent correlated output between them. Each pipeline uses tailored examples and definitions appropriate to its specific domain, ensuring that growth-related developments are not conflated with inflation dynamics.

Fifth, we design the prompts with a clear and narrow objective of analysing the article at hand only and submit articles for classification as separate sessions ("threads" in OpenAI terminology) rather than jointly. This approach means no context is preserved after analysing each article, ensuring no "information leakage", e.g. where knowledge about the future could influence sentiment classification in the past. This is important for evaluating the forecasting abilities of our indices.

We perform accuracy tests to ensure the reliability of our prompts for sentiment and driver classifications. We conducted manual review of 300 randomly selected classified articles, with balanced sampling across time periods and between growth and inflation datasets. This review achieved an estimated classification accuracy of 92%. The 8% of cases not classified correctly primarily consisted of sub-driver classification disagreements (e.g. distinguishing between different types of demand or supply drivers) and API call failures that returned no results, rather than errors in primary sentiment direction. Directional errors, where positive sentiment was classified as negative or vice versa, were rare, occurring in less than 2% of reviewed cases.

2.4 Sentiment index construction

We construct daily sentiment indices using a bottom-up approach. For each bottom-level driver d in the taxonomy (Figure 1), macroeconomic variable $k \in \{\text{growth}, \text{inflation}\}$, and date t, let $\mathcal{M}_d^k(t)$ be the set of unique sentiment instances (at most one per article per driver). Each sentiment instance is scored $s_i^k(t) \in \{-1, 0, 1\}$ for negative, neutral or positive sentiment, respectively. The sentiment index for each bottom-level driver d is defined as the sum of its constituent sentiment scores:

$$S_d^k(t) = \sum_{i \in \mathcal{M}_d^k(t)} s_i^k(t) \tag{1}$$

Aggregation to higher-level sentiment indices is additive:

$$S^k_{d'}(t) = \sum_{d \in C(d')} S^k_d(t) \tag{2}$$

where $C(d')$ is the set of all sub-drivers of parent driver d'. That is, $S^k_{\text{real}} = S^k_{\text{private-real}} + S^k_{\text{fiscal-policy}}$, $S^k_{\text{financial}} = S^k_{\text{private-financial}} + S^k_{\text{monetary-policy}}$, $S^k_{\text{demand}} = S^k_{\text{real}} + S^k_{\text{financial}}$, $S^k_{\text{supply}} = S^k_{\text{commodity-prices}} + S^k_{\text{supply-disruptions}} + S^k_{\text{government-policy}} + S^k_{\text{others}}$. The overall macroeconomic sentiment index is then given by

$$S^k(t) = S^k_{\text{demand}}(t) + S^k_{\text{supply}}(t) = \sum_{\forall d} S^k_d(t) \tag{3}$$

for $k \in \{\text{growth}, \text{inflation}\}$. Because an article can contain both demand and supply content, a single article can contribute between -2 and $+2$ to the overall macroeconomic sentiment score.

3 Macroeconomic sentiment

LLM analysis produces aggregate time-series of growth and inflation sentiment as well as their drivers. In this section, we examine these indices, their components and relationships with other macroeconomic indicators.

3.1 Growth and inflation sentiment indices

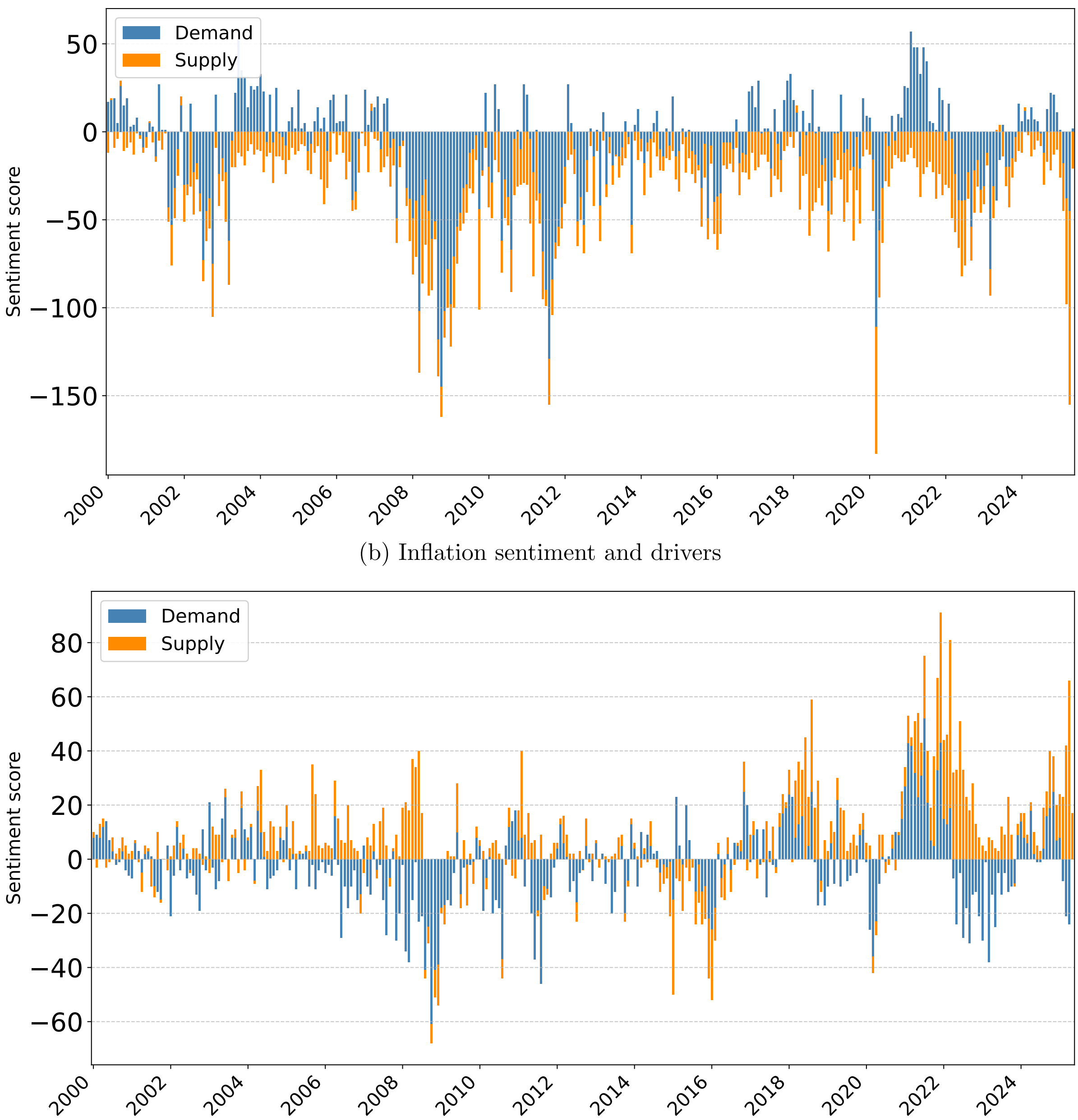
Figure 2 shows growth and inflation sentiment indices, along with their underlying demand and supply drivers. Aggregate growth sentiment (Figure 2a) provides a summary of US economic activity that fits the usual description of the business cycle. The index fell in all major downturns, including the dot-com recession in the early 2000s, the Great Financial Crisis (GFC) in 2008, the 2011 euro area crisis and the 2020 pandemic. Most recently, growth sentiment also plunged following the US tariff hikes. The decomposition into demand and supply components reveals the relative importance these factors as drivers of growth sentiment. Demand factors appear to be the predominant driver of business cycle fluctuations, both in recessions and recoveries. Supply factors, while usually less cyclical, can reassert themselves and become important growth headwinds–as during the pandemic or the recent tariff hike.

Aggregate inflation sentiment (Figure 2b) similarly summarises the evolution of inflationary pressures as transcribed in the news report. The index appears to track key inflationary and disinflationary episodes, including the inflation swings around the GFC and the post-pandemic inflation surge. Following the recent US tariff increases, the inflation sentiment also spiked. Unlike growth sentiment, the decomposition reveals that supply factors have been just as cyclical as demand factors, and have been key drivers of inflation sentiment fluctuations. Notably, supply factors tend to raise inflation sentiment–in fact, they accompany all major inflationary episodes. Meanwhile, demand factors tend to dampen inflation sentiment, with a notable exception during the post-pandemic inflation surge where both demand and supply factors pull in the upward direction.

Figure 2: Macroeconomic sentiment indices and demand-supply decomposition

(a) Growth sentiment and drivers

(b) Inflation sentiment and drivers

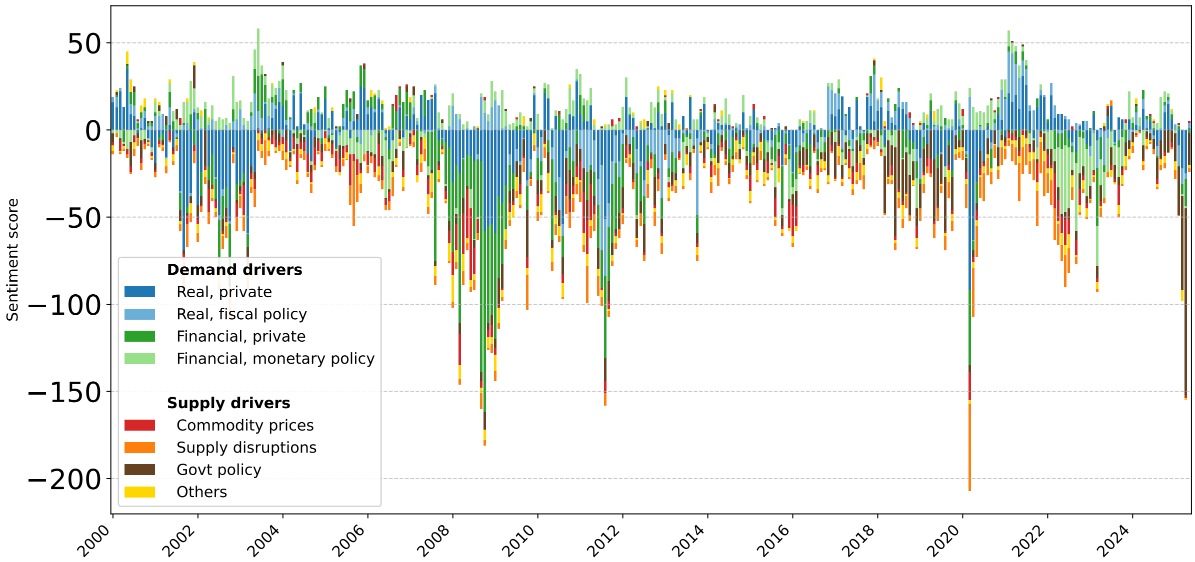


Note: Bars show demand and supply sentiment scores, S^k_{demand} and S^k_{supply}, for $k \in \{\text{growth}, \text{inflation}\}$, with overall growth and inflation sentiment given by $S^k(t) = S^k_{\text{demand}} + S^k_{\text{supply}}$. Higher growth sentiment means stronger economic growth, while higher inflation sentiment means higher inflation.

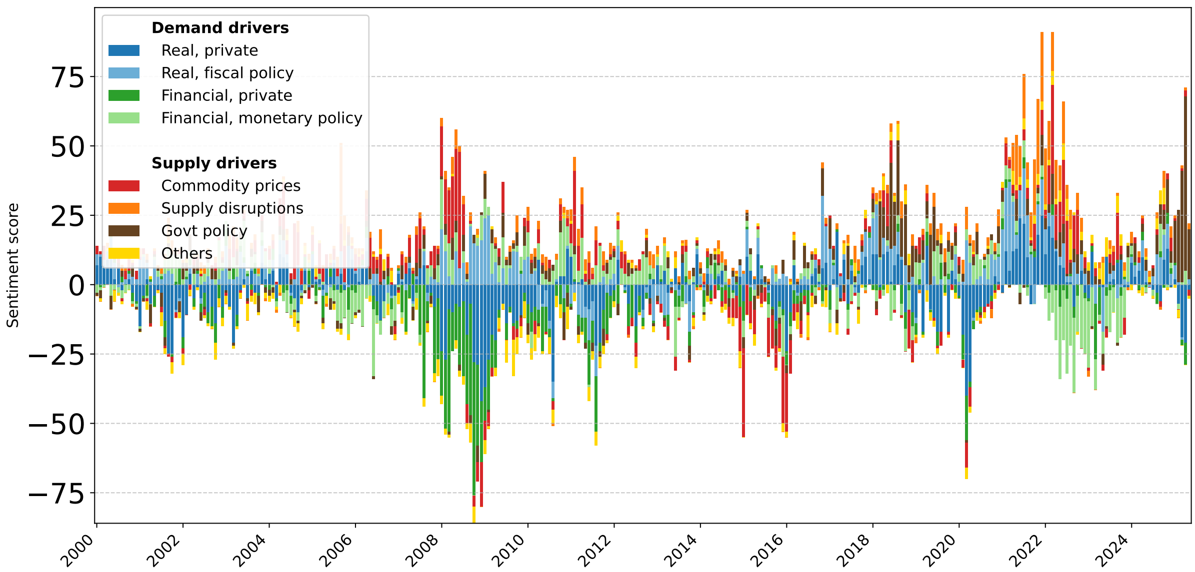
Our LLM analysis also provides a more granular look at the macroeconomic sentiment, through the detailed drivers of demand and supply components (Figure 3). As noted previously, the taxonomy differentiates between non-financial and financial drivers of demand (*real* and *financial*), and, in turn, divides each into *private* versus *policy*-related sub-components. On the supply front, we identify drivers that are related to commodity prices, supply chain issues, government policies and others.

Figure 3: Macroeconomic sentiment indices and detailed drivers

(a) Growth sentiment and drivers



(b) Inflation sentiment and drivers



Note: Bars show sentiment scores S_d^k of detailed drivers d, sorted by demand and supply groupings, with overall growth and inflation sentiment given by $S^k(t) = \sum_{\forall d} S_d^k$. Higher growth sentiment means stronger economic growth, while higher inflation sentiment means higher inflation.

The decomposition of growth drivers (Figure 3a) highlights fluctuations in real private demand as the most consistent driver of the business-cycle sentiment. Indeed, demand shortfalls from weak confidence and unwillingness to spend by households and firms characterise all major recessions. At the same time, tight financial conditions have weighed heavily on demand on occasions, including in severe stress episodes such as the GFC (*financial, private*) or during sharp monetary policy tightening post-pandemic (*financial, monetary policy*). Among supply factors, commodity prices feature regularly though with modest impacts. By contrast, supply chain disruptions and shifts in government policies (particularly trade-related) are infrequent, but can generate outsized impacts on growth sentiment, as seen during the pandemic and the recent tariff hikes.

The detailed decomposition of inflation sentiment similarly reveals the interplay of demand and supply forces shaping inflation dynamics. Real private demand again emerges as a consistent cyclical driver of inflation sentiment, with financial factors playing an important role during the GFC. The influence of monetary policy continues to be evident in several periods, including its role in the post-pandemic disinflation. Fiscal policy helped mitigate the strong deflationary forces during the GFC, while contributing to the inflation increase during the pandemic. On the supply side, commodity prices account for both rises and declines of inflation over the sample, while supply chain disruptions played a significant part in the post-pandemic inflation jump. The sharp increase in inflation sentiment in early 2025 is almost entirely attributable to US government tariff policies.

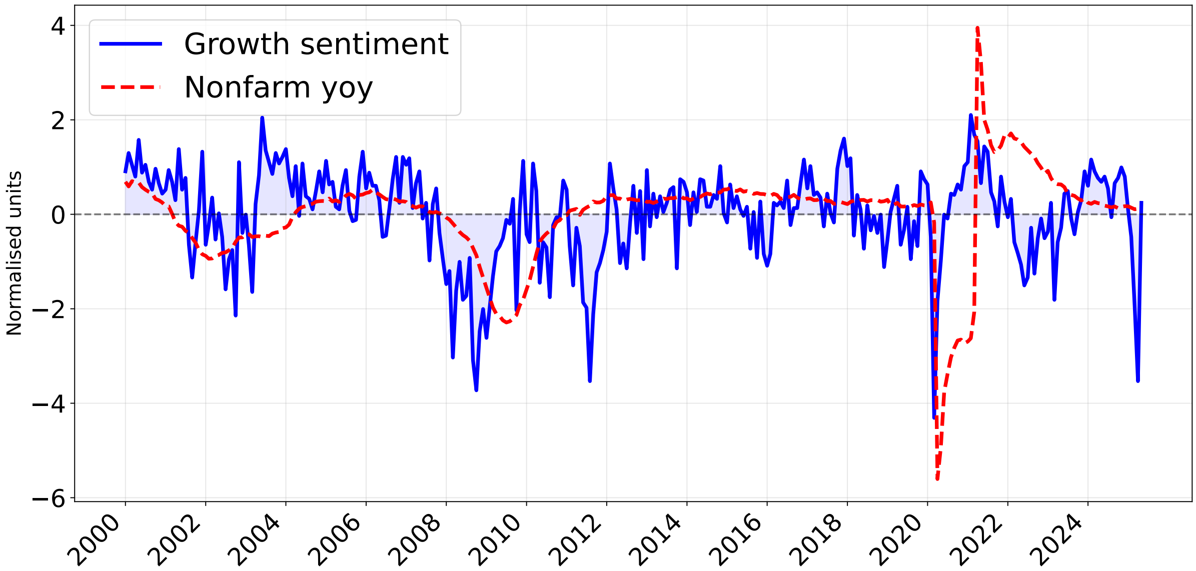
3.2 Validation of aggregate indices

How do the sentiment indices compare with their "hard data" counterparts? Do their decompositions tell a story that is broadly consistent with the conventional business cycle narrative? To address these questions, we consider a series of validation exercises.

First, we compare overall growth sentiment with representative monthly indicators of US economic activity. Figure 4 plots the growth sentiment against year-on-year changes in nonfarm payrolls, confirming a close synchronicity of the two series over the business cycle. We further examine their relationships, including possible leads and lags, through their cross-correlogram based on Spearman correlation.[3] The results, shown in Figure 5a, confirm the monotonic association between the growth sentiment and changes in nonfarm payrolls. In fact, the growth sentiment index appears to lead changes in nonfarm payrolls by several months.

[3] Spearman correlation captures monotonic relationships without imposing linearity, making it robust to outliers from deep recessions such as the GFC or the pandemic. We use Spearman correlation in all cross-correlograms to focus on directional comovements.

Figure 4: Growth sentiment



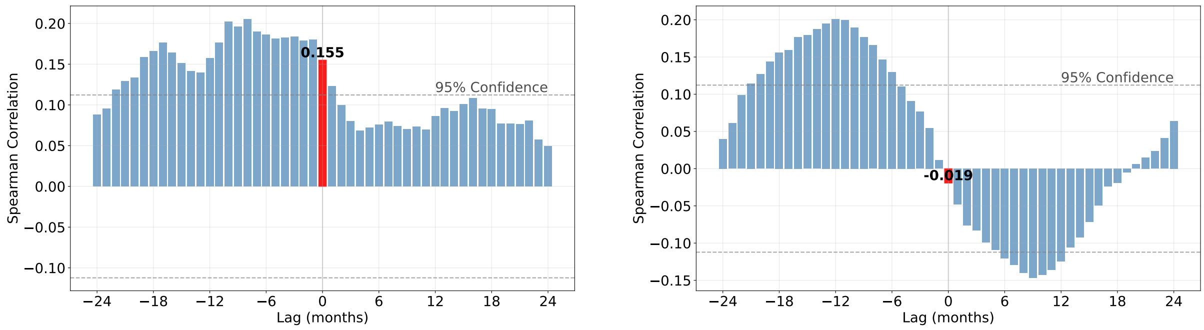
Note: Growth sentiment is the aggregate growth sentiment score $S^{\text{growth}}(t)$. All series are standardised (z-scores).

Benchmarking the growth sentiment against other measures of economic activity further confirms its relevance as a barometer of the business cycle. Figure 5b-Figure 5d show the cross-correlograms for year-on-year changes in the industrial production (IP), Chicago Fed National Activity Index (CFNAI), and the yearly changes in the Philadelphia Fed coincident economic activity index (Philly Fed), respectively. The growth sentiment index strongly leads IP, is contemporaneously associated with the CFNAI, and leads the Philly Fed. Overall, the evidence lends support to the role of growth sentiment as a coincident and leading indicator of US economic activity.

Figure 5: Cross-correlograms of growth sentiment and activity measures

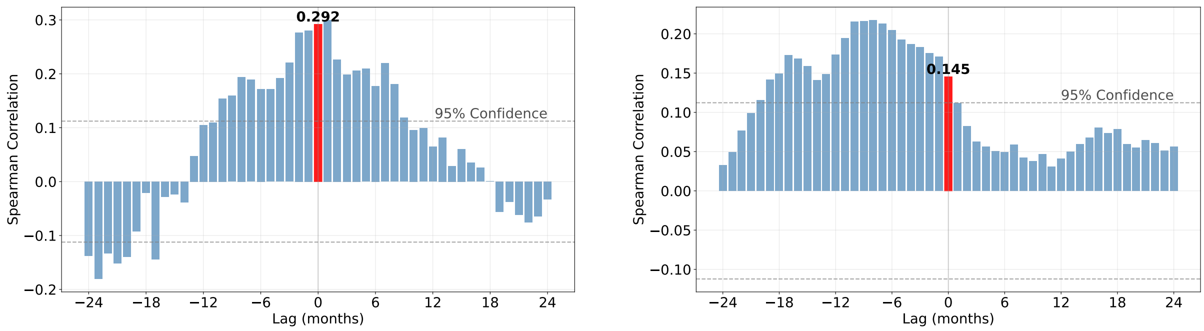
(a) Growth sentiment vs non-farm payroll (yoy)

(b) Growth sentiment vs industrial production



(c) Growth sentiment vs CFNAI

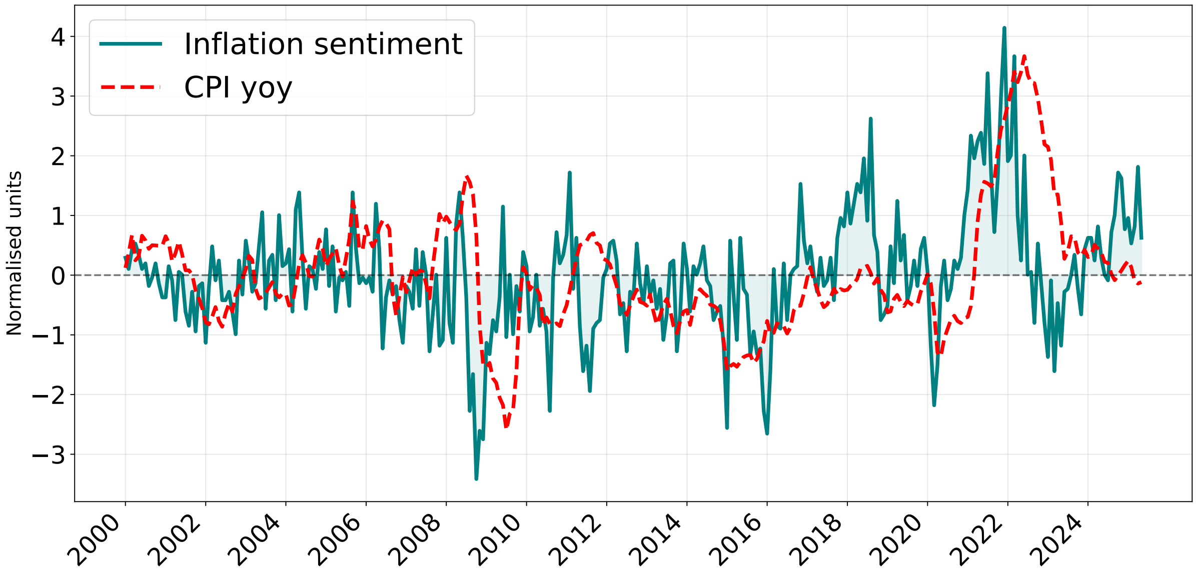
(d) Growth sentiment vs Philly Fed (yoy)



Note: Negative lags = growth sentiment leads; positive lags = activity measures lead.

Turning to inflation, the sentiment index also aligns with the data counterparts. Figure 6 shows a clear association between the inflation sentiment time-series and the year-on-year CPI inflation. In fact, the sentiment index provides an early signal prior to the post-pandemic inflation surge episode. The cross-correlogram with CPI inflation in Figure 7a confirms the statistical significance of the relationship and potential leading properties of the sentiment index. The results are also remarkably similar for other inflation measures such as PCE inflation as well as core measures, shown in Figure 7b-Figure 7d.

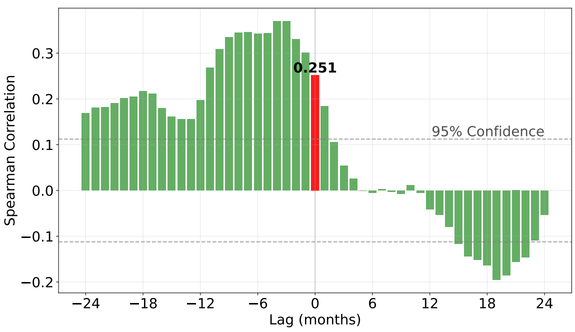
Figure 6: Inflation sentiment



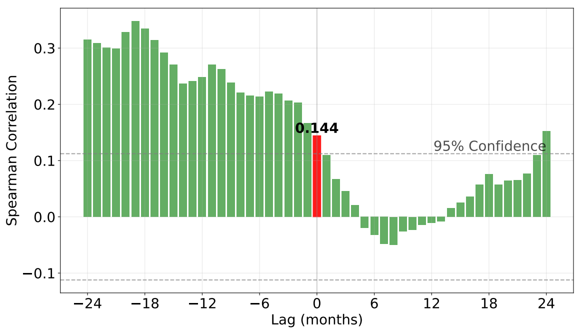
Note: Inflation sentiment is the aggregate inflation sentiment score $S^{\text{inflation}}(t)$. All series are standardised (z-scores).

Figure 7: Cross-correlograms of inflation sentiment and inflation measures

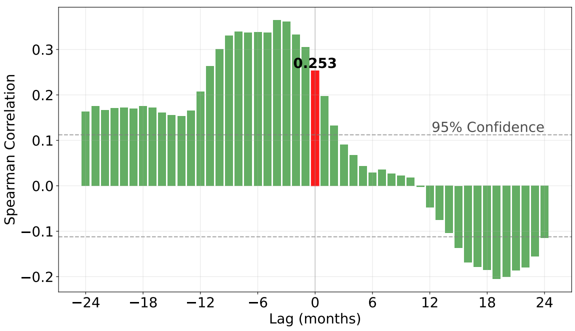
(a) Inflation sentiment vs CPI (yoy)



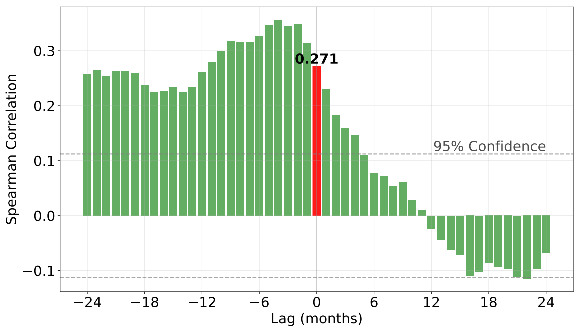
(b) Inflation sentiment vs core CPI (yoy)



(c) Inflation sentiment vs PCE (yoy)



(d) Inflation sentiment vs core PCE (yoy)



Note: Negative lags = inflation sentiment leads; positive lags = inflation data lead.

3.3 Validation of detailed drivers

We next turn to the decomposition of sentiment indices and examine how well the detailed drivers align with their data counterparts. Starting from growth, Figure 8 shows relationships between sentiment components and the corresponding benchmark data. In panel (a), real private demand driver co-moves with personal consumption expenditure. The recent tariff hike episode is one exception where the real demand sentiment fell even as consumption data held up, reflecting the oft-noted gap between soft and hard data in this period. Financial demand sentiment excluding monetary policy (panel (b)) closely tracks the Financial condition Index, spiking sharply during the GFC and, to a lesser extent, during the pandemic and subsequent monetary policy tightening. On the supply side, commodity price sentiment (panel (c)) captures episodes of major swings in energy prices. Supply chain sentiment (panel (d)) generally moves in a narrow range, with the pandemic being the only episode of marked and persistent deterioration consistent with other indicators of supply chain disruption. Finally, government policy (panel (e)) is heavily influenced by trade policy shifts and worsens dramatically in the most recent episode of tariff hikes similar to other indicators of trade policy uncertainty.

On the inflation side, Figure 9 shows granular drivers of inflation sentiment against corresponding data. Real private demand driver of inflation tracks consumption closely (panel (a)), as is the case with growth sentiment. Financial private demand driver also remains highly synchronised with the Financial Condition Index (panel (b)). Commodity price sentiment is tightly linked with oil price changes, even more so than for growth sentiment. Other supply components including supply chain disruptions and supply-side policy drivers remain well-aligned with their respective indicators.

Figure B.1 and Figure B.2 in the appendix present cross-correlograms of detailed sentiment drivers and corresponding data, for growth and inflation respectively. The results point to consistently strong contemporaneous associations between sentiment components and their empirical counterparts.

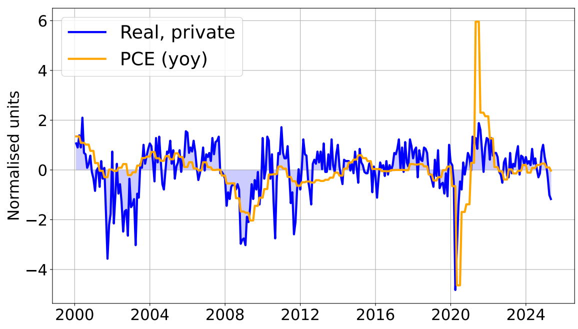
3.4 Demand and supply drivers

We next consider a higher-level breakdown of the sentiment indices into demand and supply components. While less granular, this decomposition is more parsimonious and maps naturally into the standard demand-supply framework. A key question, once again, is whether this demand-supply breakdown can be validated.

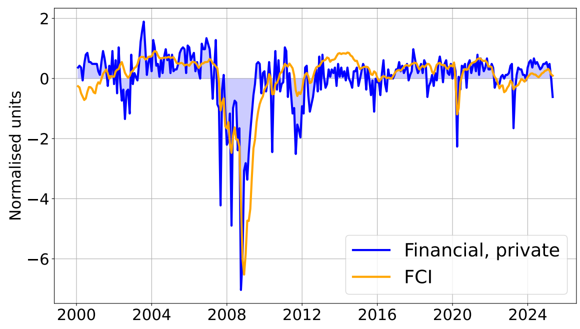
As a starting point, the demand (supply) component of growth sentiment is strongly correlated with the demand (supply) driver of inflation sentiment (Figure 10), with correlations of 0.78 and -0.55, both statistically significant. This correspondence emerges despite the LLM analyses being conducted independently, providing an *internal* validation that demand and supply forces are consistently identified across growth and

Figure 8: Growth drivers and data counterparts

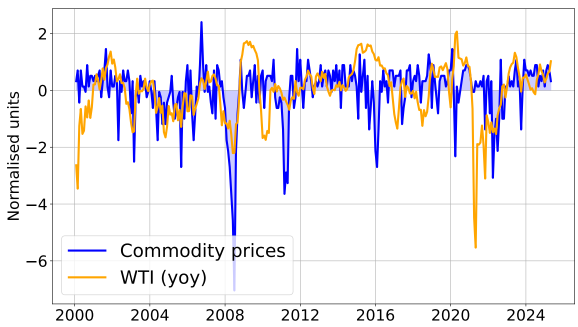
(a) Real private demand vs PCE



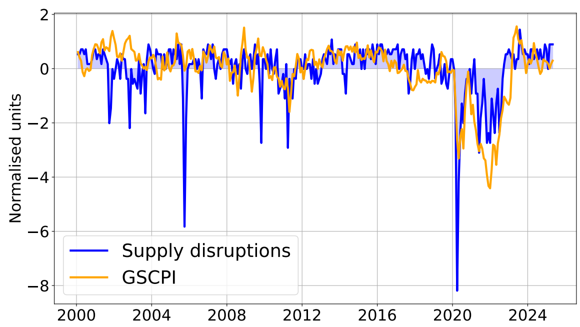
(b) Financial private demand vs FCI



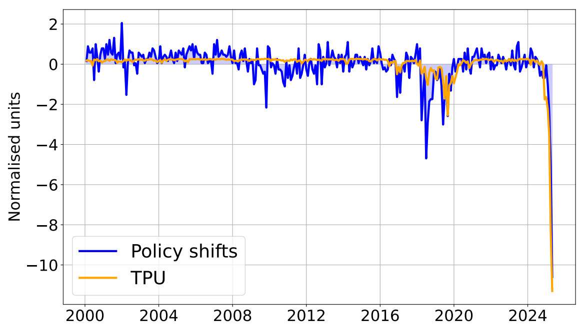
(c) Commodity prices vs WTI



(d) Supply disruptions vs GSCPI

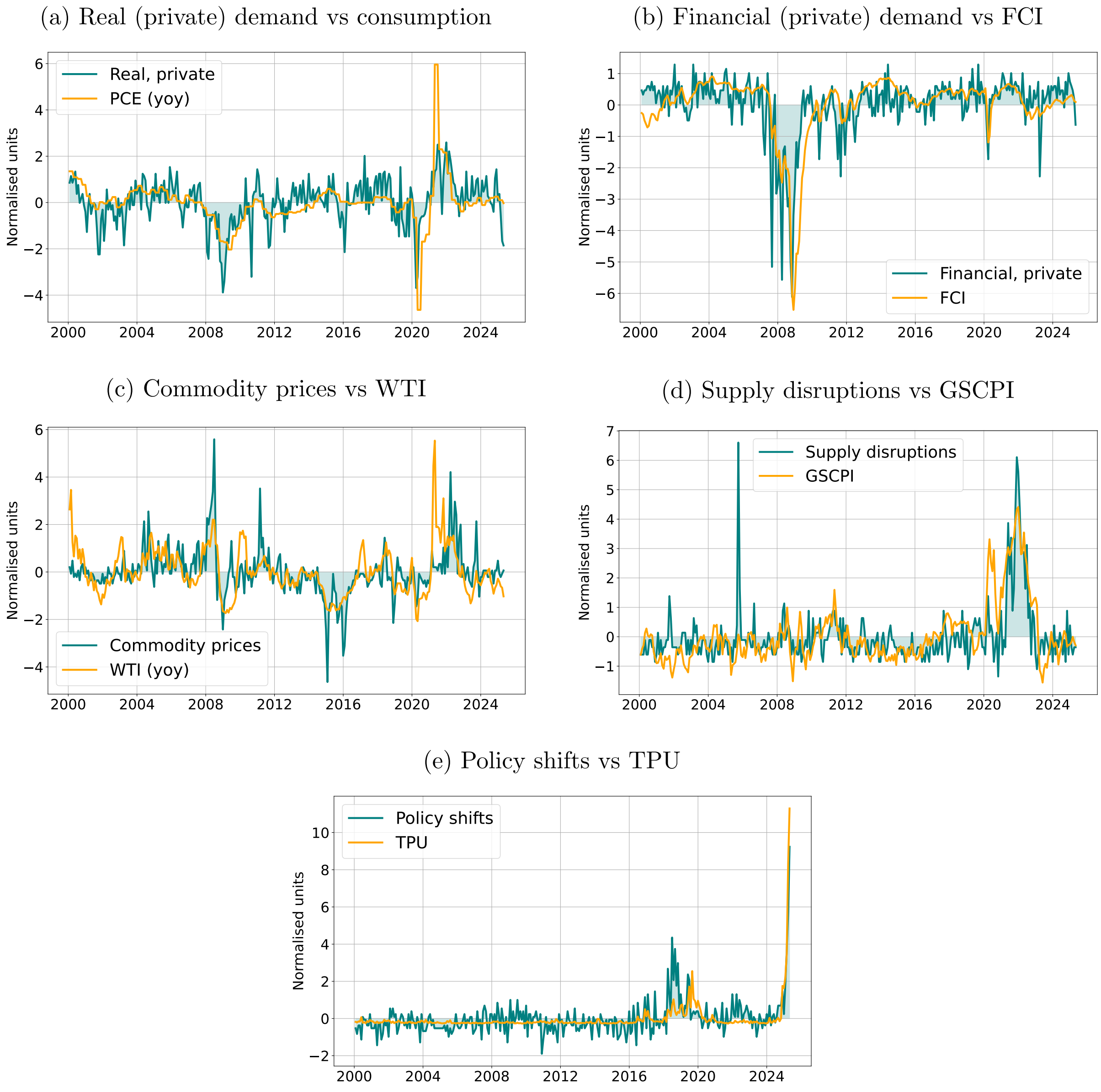


(e) Policy shifts vs TPU



Note: Each panel shows a granular driver of growth sentiment, S_d^{growth}, and its data counterpart. PCE (yoy) is year-on-year real consumption growth. FCI is Chicago Fed Financial Condition Index. WTI (yoy) is the West Texas crude oil prices (year-on-year changes). GSCPI is the New York Fed Global Supply Chain Pressure Index. TPU is the trade policy uncertainty index from Baker et al. (2016). Reversed scales for FCI, WTI, GSCPI and TPU to aid comparisons. All series are z-scores.

Figure 9: Inflation drivers and data counterparts

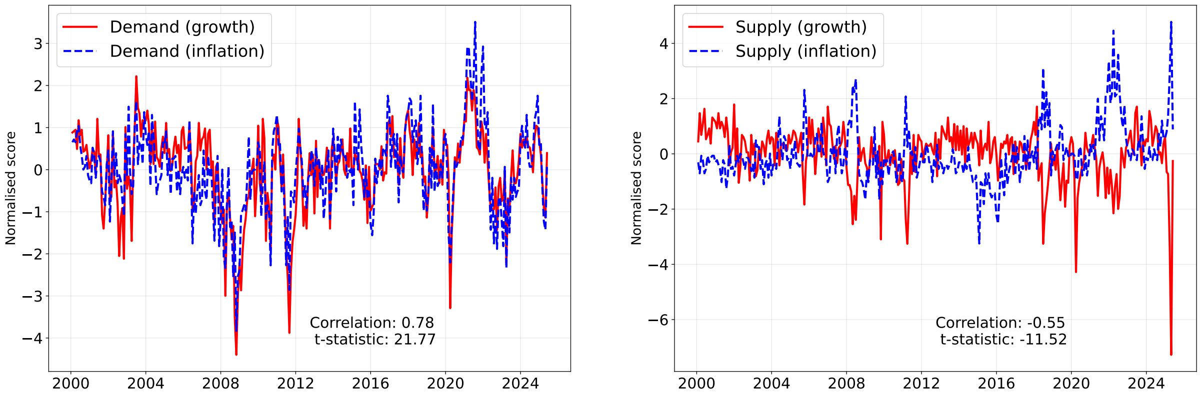


Note: Each panel shows a granular driver of inflation sentiment, $S_d^{\text{inflation}}$, and its data counterpart. PCE (yoy) is year-on-year real consumption growth. FCI is Chicago Fed Financial Condition Index. WTI (yoy) is the West Texas crude oil prices (year-on-year changes). GSCPI is the New York Fed Global Supply Chain Pressure Index. TPU is the trade policy uncertainty index from Baker et al. (2016). Reversed scale for FCI. All series are z-scores.

inflation.

Figure 10: Demand and supply drivers of macroeconomic sentiment

(a) Demand driver (b) Supply driver



Note: The left panel compares the demand components of growth and inflation sentiment. The right panel compares the supply components of growth and inflation sentiment. All series are z-scores.

For an *external* validation of demand and supply components, we consult two model-based estimates. Shapiro (2025), the first benchmark study, provides a monthly demand-supply breakdown of PCE inflation, through sign restrictions implemented at the level of expenditure categories, which are then aggregated into the headline decomposition. The second benchmark is Eickmeier and Hofmann (2023), which relies on sign restrictions at the aggregate level following a BVAR approach, leveraging on a large number of US inflation and growth measures in quarterly frequency.

Our demand-supply breakdowns of macroeconomic sentiment line up remarkably well with the respective shocks identified in these studies. The demand and supply drivers of inflation bear strong resemblance to the shocks identified by Shapiro (2025), as shown in Figure 11. In particular, our decomposition corroborates Shapiro's key finding regarding the dual role of demand and supply in explaining the post-pandemic inflation surge. A more formal assessment via cross-correlograms (Figure B.3) confirms these associations and points to some leading properties of our LLM-based demand driver.

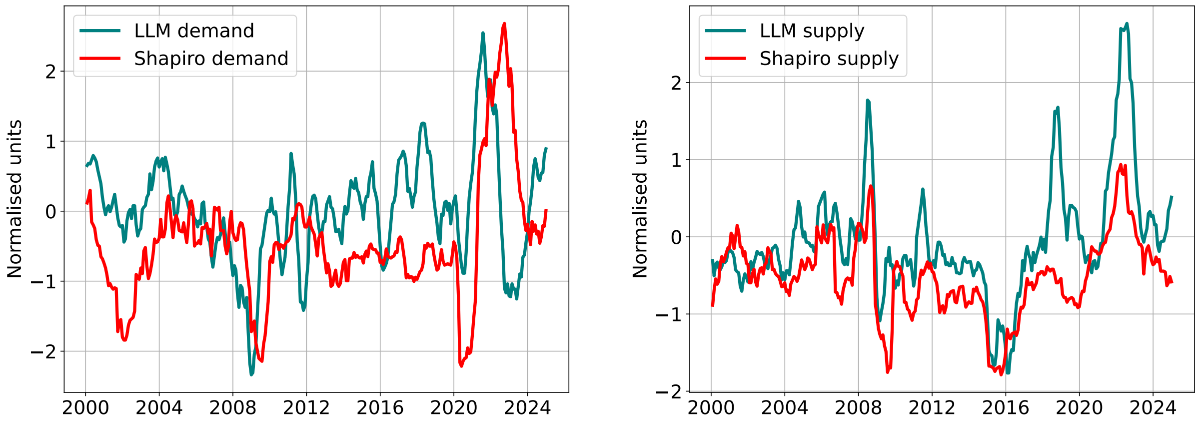
Relative to Eickmeier and Hofmann (2023) (EH), sentiment drivers exhibit strong alignment with the EH-identified shocks, as illustrated in Figure 12. Cross-correlograms in Figure B.4 confirm statistically significant contemporaneous relationships in all but one case. The exception is the supply driver of inflation, where the association is positive but not significant, reflecting the greater persistence of the EH supply shock in the aftermath of the GFC than that of the LLM-based component.

3.5 Monetary and fiscal policy drivers

We finally turn to the policy drivers of demand, namely monetary and fiscal policies. As with aggregate demand drivers, the policy drivers of growth and inflation are tightly

Figure 11: Comparison with Shapiro's inflation decomposition

(a) Demand driver of inflation (b) Supply driver of inflation



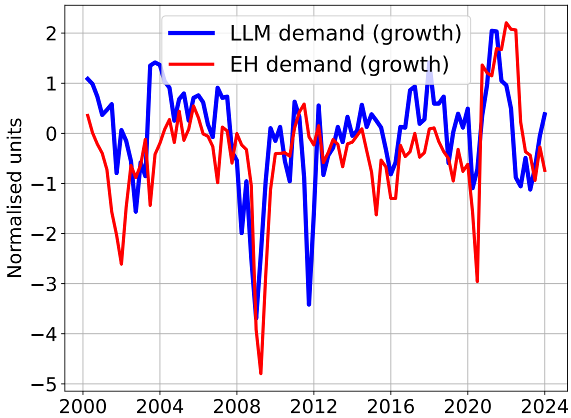
Note: "LLM demand" on the left panel is the demand component of inflation sentiment, while "LLM supply" on the right panel is the supply component. Both are 6-month moving averages. Shapiro demand and supply shocks are from Shapiro (2025). All series are z-scores.

correlated. This holds for both monetary and fiscal policies, as shown in Figure B.5, with correlations of 0.8 and 0.7 respectively, consistent with their demand-shifting roles. In what follows, we focus on the policy drivers of growth when analysing the role of macroeconomic policies.

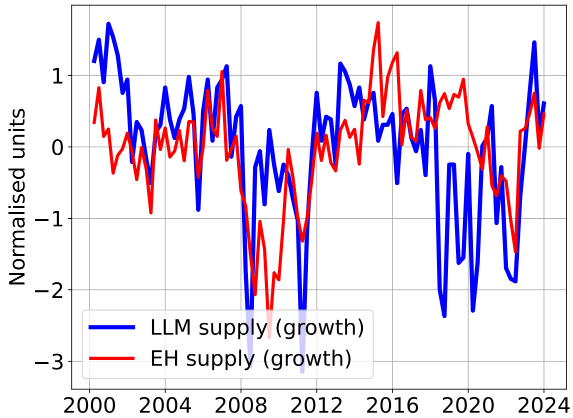
Monetary policy sentiment tracks major policy cycles with some leading properties, as shown in Figure 13. All three monetary tightening cycles in the sample either coincide or follow an increase in monetary policy sentiment. Similar patterns also hold for the easing cycles. The pronounced swings in monetary policy sentiment around the pandemic are particularly noteworthy. The sentiment eases significantly during the onset of the pandemic, then tightens sharply from 2022 following the rapid increase in inflation. The tightening in monetary policy sentiment leads actual increases in short-term interest rates by over a year. The sentiment then eases as inflation concerns subside, again leading the actual easing cycle. Figure B.6 shows the cross-correlograms, where monetary policy sentiment significantly leads short-term rates by about 12 months.

Figure 12: Comparison with Eickmeier-Hofmann's decomposition

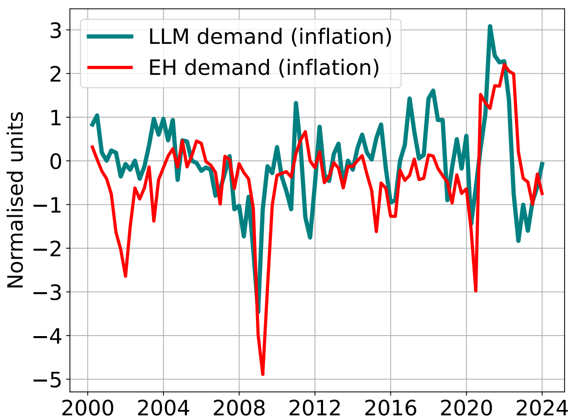
(a) Demand driver of growth



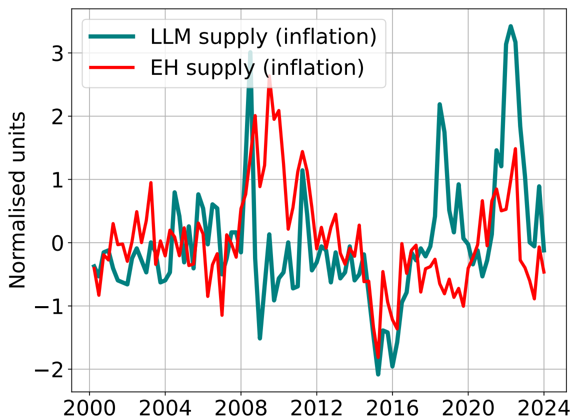
(b) Supply driver of growth



(c) Demand driver of inflation

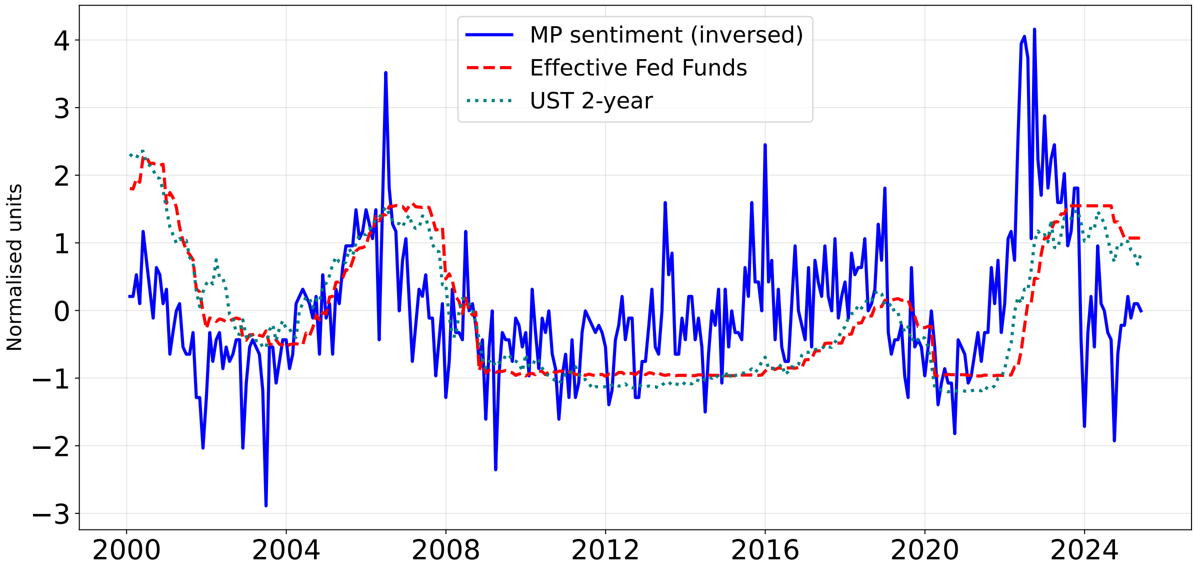


(d) Supply driver of inflation



Note: Quarterly frequency. LLM demand/supply are the demand/supply components of growth (upper panels) and inflation (lower panels) sentiment. EH are Eickmeier-Hofmann shocks from Eickmeier and Hofmann (2023). All series are z-scores.

Figure 13: Monetary policy sentiment



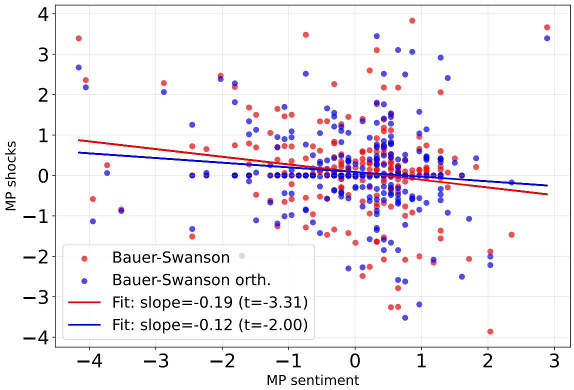
Note: MP sentiment is the (negative of) monetary policy driver of growth sentiment. Other series are the effective fed funds rate and the yields on 2-year US Treasury bonds. All series are z-scores.

Short-term fluctuations in monetary policy sentiment could contain information about monetary policy surprises. To explore such possibility, we use as benchmarks three sources of monetary policy shocks, from Bauer and Swanson (2023) (raw and orthogonalised series), Jarociński and Karadi (2020) and Nakamura and Steinsson (2018). We run a linear regression of normalised monetary policy shocks on the normalised level of monetary policy sentiment. The results in Figure 14 point to statistically significant and correctly-signed relationships, with tighter monetary policy sentiment coinciding with positive monetary policy shocks. The relationships are less than one-for-one, with one standard deviation move in monetary policy sentiment implying about 0.1-0.2 standard deviation change in monetary policy shocks. This result may be expected given that monetary policy sentiment reflects a much broader set of factors than shocks alone, including in particular the systematic part of policy. Nonetheless, the significant relationships with policy shocks point to a significant presence of policy innovations contained in the sentiment index, as one perhaps might expect from news.

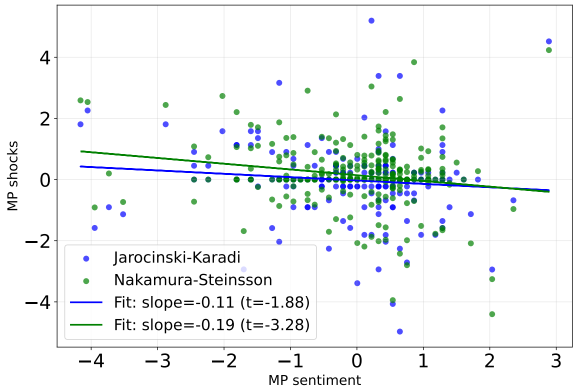
Fiscal policy sentiment also mirrors key events related to public finance. Figure 15 plots the fiscal policy driver of growth against selected major fiscal events. Fiscal sentiment rises in tandem with expansionary fiscal policy actions, such as the signing of the American Recovery and Reinvestment Act (ARRA–Obama's stimulus) into law in February 2009, the passage and enactment of Tax Cuts and Jobs Act (TCJA) in December 2017, and the signing of American Rescue Plan in March 2021. Conversely, periods of fiscal stress and need for fiscal consolidation–such as the intensification of the debt ceiling crisis in July 2011 and the fiscal cliff debate in late 2012–coincided with marked decline in fiscal sentiment.

Figure 14: Comparison with monetary policy shocks

(a) Bauer-Swanson shocks

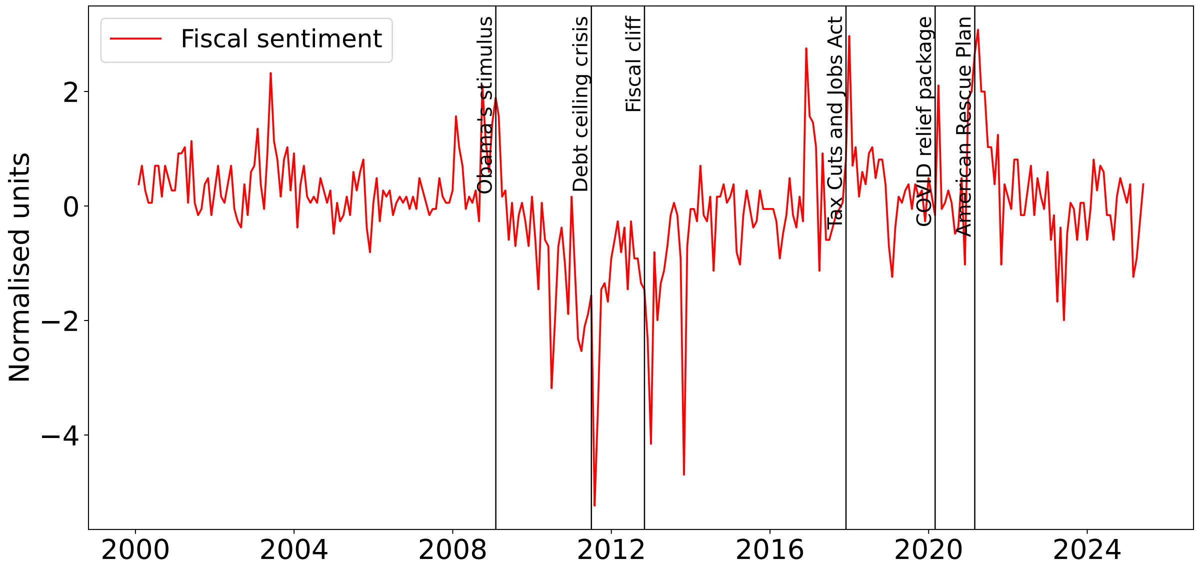


(b) Jarocinski-Karadi/Nakamura-Steinsson shocks



Note: Horizontal axes are the monetary policy driver of growth sentiment. Vertical axes are monetary policy shocks, from Bauer and Swanson (2023) (raw and orthogonalised series), Jarociński and Karadi (2020) and Nakamura and Steinsson (2018). Fitted lines are included, with estimated slopes and associated t-statistics. All series are z-scores.

Figure 15: Fiscal policy sentiment



Note: Fiscal sentiment is the fiscal policy driver of growth sentiment. Vertical lines mark selected dates of key US fiscal policy events. The series is z-scores.

4 Forecasting abilities

Beyond serving as high-frequency and timely macroeconomic proxies, our sentiment indices may also hold predictive value. The leading properties documented in the previous section underscores this potential. One possible mechanism is a sunspot channel, whereby shifts in confidence or narratives exert causal influence on macroeconomic dynamics. In such case, information from news could provide additional signals not spanned by past data. In this section, we evaluate this possibility by comparing the forecasting performance of simple time-series models with and without the sentiment indices.[4]

We conduct an out-of-sample forecasting evaluation as follows. The target variables include several measures of growth $g_{m,t}$ and inflation $\pi_{m,t}$, where m denotes the measure and t the time period. For growth, the measures include year-on-year changes in total nonfarm payroll for all employees, the Philly Fed Coincident Economic Activity Index and the Chicago Fed National Activity Index. Inflation measures include year-on-year changes in CPI, core CPI, PCE and core PCE. We employ two forecasting methods: a recursive forecast based on an AR(1) model and local projections. For each method, we compare a baseline specification with one augmented by the relevant aggregate sentiment index. Consider each method in turn.

4.1 AR(1) model

For each variable and measure m, we estimate a baseline AR(1) model:

$$g_{m,t} = \beta_{m,0} + \beta_{m,1} g_{m,t-1} + \epsilon^{g}_{m,t} \tag{4}$$
$$\pi_{m,t} = \gamma_{m,0} + \gamma_{m,1} \pi_{m,t-1} + \epsilon^{\pi}_{m,t} \tag{5}$$

We estimate these AR(1) specifications then use the fitted models to forecast target variables at horizons h by iterating the mappings $B(g_{m,t}) \equiv \hat{\beta}_{m,0} + \hat{\beta}_{m,1} g_{m,t}$ and $G(\pi_{m,t}) \equiv \hat{\gamma}_{m,0} + \hat{\gamma}_{m,1} \pi_{m,t}$:

$$\hat{g}_{m,t+h} = B^h(g_{m,t}) \tag{6}$$
$$\hat{\pi}_{m,t+h} = G^h(\pi_{m,t}) \tag{7}$$

where $h = 1, 3, 6, 12, 24$ denotes the forecasting horizons in months. We estimate the model and perform forecasting using an expanding sample, starting with the first half of the sample. We collect the root mean square error (RMSE) from the differences between the out-of-sample forecasts and actual data, averaging them over the expanding sample and successive forecasting rounds. The resulting RMSEs are reported in Table 1 and Table 2, under the columns AR(1).

The competing model augments the LLM-based aggregate sentiment indices to the baseline specifications:

$$g_{m,t} = \beta_{m,0} + \beta_{m,1} g_{m,t-1} + \beta_{m,2} s_{g,t-1} + \epsilon^{g}_{m,t} \tag{8}$$
$$\pi_{m,t} = \gamma_{m,0} + \gamma_{m,1} \pi_{m,t-1} + \gamma_{m,2} s_{\pi,t-1} + \epsilon^{\pi}_{m,t} \tag{9}$$

[4] We use autoregressive models as benchmarks given their simplicity and the established findings in the literature that they are often difficult to outperform in out-of-sample forecasting, even by more complex DSGE models; see Del Negro and Schorfheide (2013).

where $s_{g,t}$ and $s_{\pi,t}$ are aggregate growth and inflation sentiment indices, respectively. We then construct out-of-sample forecasts analogously to above using mappings $B_s(g_{m,t}) \equiv \hat{\beta}_{m,0} + \hat{\beta}_{m,1} g_{m,t-1} + \hat{\beta}_{m,2} s_{g,t-1}$ and $G_s(\pi_{m,t}) \equiv \hat{\gamma}_{m,0} + \hat{\gamma}_{m,1}\pi_{m,t-1} + \hat{\gamma}_{m,2} s_{\pi,t-1}$:

$$\tilde{g}_{m,t+h} = B_s^h(g_{m,t}) \tag{10}$$

$$\tilde{\pi}_{m,t+h} = G_s^h(\pi_{m,t}) \tag{11}$$

The corresponding RMSEs from this alternative model are reported in columns +LLM in Table 1 and Table 2.

Table 1: Growth Forecasting Performance (AR(1))

Horizon	Nonfarm Payroll			Philly Fed Index			Chicago Fed Index		
	AR(1)	+LLM	Gain(%)	AR(1)	+LLM	Gain(%)	AR(1)	+LLM	Gain(%)
1-month	0.94	0.93	1.5	0.80	0.79	1.9	1.02	0.98	3.1
3-month	1.16	1.12	3.1	1.06	1.01	4.1	1.22	1.12	8.6
6-month	1.31	1.22	6.3	1.28	1.18	7.5	1.24	0.99	19.9
12-month	1.64	1.55	4.9	1.63	1.54	5.7	1.82	1.04	42.5
24-month	1.56	1.54	1.3	1.54	1.54	-0.0	5.15	1.11	78.4

Note: Values represent RMSE from out-of-sample forecasts at various horizons. Two models considered are AR(1) and AR(1) augmented with lagged 3-month moving average of growth sentiment index (+LLM), estimated with expanding sample, starting with half the full sample. Gain shows percentage improvement from adding sentiment index. "Nonfarm Payroll" is the total nonfarm payroll for all employees. "Philly Fed Index" is the Coincident Economic Activity Index for the United States compiled by Federal Reserve Bank of Philadelphia. "Chicago Fed Index" is 3-month moving average of the Chicago Fed National Activity Index. All predicted variables are in year-on-year percentage changes.

Table 2: Inflation Forecasting Performance (AR(1))

Horizon	CPI			CPI core			PCE			PCE core		
	AR(1)	+LLM	Gain%	AR(1)	+LLM	Gain%	AR(1)	+LLM	Gain%	AR(1)	+LLM	Gain%
1-month	0.37	0.33	9.6	0.35	0.33	4.4	0.31	0.28	10.5	0.29	0.27	7.3
3-month	0.58	0.52	11.5	0.56	0.51	8.5	0.52	0.46	11.7	0.50	0.44	12.1
6-month	0.85	0.75	12.0	0.80	0.66	16.9	0.79	0.70	11.4	0.78	0.63	18.5
12-month	1.27	1.30	-3.1	1.42	1.21	14.4	1.26	1.28	-2.0	1.35	1.10	18.7
24-month	1.51	1.98	-31.4	2.47	2.15	13.0	1.64	2.04	-24.5	2.29	1.83	20.3

Note: Values represent RMSE from out-of-sample forecasts at various horizons. Two models considered are AR(1) and AR(1) augmented with lagged 3-month moving average of inflation sentiment index (+LLM), estimated with expanding sample, starting with half the full sample. Gain shows percentage improvement from adding sentiment index. "CPI" and "PCE" are year-on-year changes in the Consumer Price Index and the Personal Consumption Expenditures Price Index, respectively. "CPI core" and "PCE core" are the core inflation counterparts.

Including sentiment indices improves out-of-sample forecasts in most cases, and in all cases at medium-term horizons. For growth forecasting, adding sentiment to AR(1) improves RMSE by 2-8% for nonfarm payroll and the Philly Fed Coincident Index, at horizons up to 12 months, with the largest gains of 6-8% at the 6-month horizon. For the Chicago Fed National Activity Index, the gain reaches 80%. For inflation, RMSE improves by 4-20%, with the largest gains also at the 6-month horizon. For core inflation measures, there is evidence of longer-horizon forecastability, with RMSE gains of 13-20% at 24-month horizon. Overall, the results suggest that the sentiment indices contain information about underlying macroeconomic dynamics and function beyond being high-frequency coincident indicators.

4.2 Local projection model

The second method of forecasting entails local projection instead of recursive AR(1) models. For a baseline without sentiment indices, we estimate:

$$g_{m,t+h} = \beta_{m,0,h} + \beta_{m,1,h} g_{m,t} + \epsilon^{g}_{m,t+h} \tag{12}$$

$$\pi_{m,t+h} = \gamma_{m,0,h} + \gamma_{m,1,h} \pi_{m,t} + \epsilon^{\pi}_{m,t+h} \tag{13}$$

for horizons $h = 1, 3, 6, 12, 24$. As above, the competing model is the baseline augmented with macroeconomic sentiment:

$$g_{m,t+h} = \beta_{m,0,h} + \beta_{m,1,h} g_{m,t} + \beta_{m,2,h} s_{g,t} + \epsilon^{g}_{m,t+h} \tag{14}$$

$$\pi_{m,t+h} = \gamma_{m,0,h} + \gamma_{m,1,h} \pi_{m,t} + \gamma_{m,2,h} s_{\pi,t} + \epsilon^{\pi}_{m,t+h} \tag{15}$$

Apart from the different specifications, we follow the same procedure as above, i.e. using expanding sample to conduct out-of-sample forecasts of the target variables m and collecting RMSEs from the two models. The results are reported in Table 3 and Table 4 for growth and inflation, respectively.

The local projection results provide further evidence on the forecasting power of macroeconomic sentiment indices. Incorporating the growth sentiment index into the growth local projection reduces RMSE by 2-7% across all target measures at horizons up to 12 months. Forecasting performance is strongest at the 6-month horizon, with RMSE gains of 5-7%. Inflation forecasts also benefit from including the inflation sentiment index, with RMSE improvement of 7-16% over horizons up to 6 months across all inflation measures. For core inflation, the forecastability extends to longer horizons.

5 Conclusion

In this paper, we apply LLMs to extract information from financial news to construct macroeconomic sentiment relating to US growth and inflation. Leveraging on the LLMs' capability to parse complex and context-dependent text, we also decompose the

Table 3: Growth Forecasting Performance (local projection)

Horizon	Nonfarm Payroll			Philly Fed Index			Chicago Fed Index		
	LP	+LLM	Gain(%)	LP	+LLM	Gain(%)	LP	+LLM	Gain(%)
1-month	0.93	0.91	1.7	0.79	0.77	2.6	0.99	0.94	4.7
3-month	1.14	1.12	1.7	1.04	1.01	2.9	1.05	1.03	2.0
6-month	1.30	1.24	4.7	1.25	1.18	5.7	0.87	0.81	6.9
12-month	1.49	1.45	2.9	1.44	1.39	3.2	0.80	0.80	0.3
24-month	1.36	1.36	-0.3	1.32	1.33	-0.4	0.84	0.82	1.5

Note: Values represent RMSE from out-of-sample forecasts at various horizons. Two specifications considered are simple local projection and one augmented with 3-month moving average of growth sentiment index (+LLM), estimated with expanding sample, starting with half the full sample. Gain shows percentage improvement from adding sentiment index. "Nonfarm Payroll" is the total nonfarm payroll for all employees. "Philly Fed Index" is the Coincident Economic Activity Index for the United States compiled by Federal Reserve Bank of Philadelphia. "Chicago Fed Index" is 3-month moving average of the Chicago Fed National Activity Index. All predicted variables are in year-on-year percentage changes.

Table 4: Inflation Forecasting Performance (local projection)

Horizon	CPI			CPI core			PCE			PCE core		
	LP	+LLM	Gain%	LP	+LLM	Gain%	LP	+LLM	Gain%	LP	+LLM	Gain%
1-month	0.39	0.33	13.5	0.36	0.33	7.4	0.33	0.29	13.0	0.30	0.27	10.8
3-month	0.65	0.55	15.0	0.58	0.52	11.1	0.58	0.50	14.1	0.54	0.47	12.3
6-month	0.97	0.82	15.8	0.85	0.74	13.2	0.93	0.79	14.6	0.87	0.76	12.9
12-month	1.35	1.35	-0.0	1.43	1.33	6.8	1.35	1.35	0.4	1.43	1.36	4.7
24-month	1.37	1.38	-0.5	1.66	1.59	4.4	1.41	1.41	-0.5	1.64	1.64	-0.2

Note: Values represent RMSE from out-of-sample forecasts at various horizons. Two specifications considered are simple local projection and one augmented with 3-month moving average of inflation sentiment index (+LLM), estimated with expanding sample, starting with half the full sample. Gain shows percentage improvement from adding sentiment index. "CPI" and "PCE" are year-on-year changes in the Consumer Price Index and the Personal Consumption Expenditures Price Index, respectively. "CPI core" and "PCE core" are the core inflation counterparts.

headline indices into underlying drivers, providing a narrative about macroeconomic developments. We obtain some encouraging results, with both the aggregate indices and their components co-moving closely with their hard data counterparts and previous model-based findings. The sentiment indices also exhibit some leading properties, pointing to potential use in forecasting.

More generally, our analysis illustrates the vast potential for LLM as a complementary tool in studying macroeconomic dynamics. To be clear, the LLM approach should not and cannot substitute for the conventional data analysis. Indeed, circularity is a key issue. The sentiment indices may simply reflect the general consensus on the predominant macroeconomic forces, in turn shaped and informed by incoming data and quantitative analysis. Without the latter, the news could quickly lose relevant information. At the same time, news may encompass a wider array of information, including of anecdotal or informal kinds, that is not spanned by the traditional data. And, irrespective of the informational contents, the narrative in the news itself could influence agents' beliefs and behaviour with implications for macroeconomic dynamics, as emphasised by Shiller (2017). It is from these perspectives that our text-based approach, aside from the timeliness of the information, could add unique values to macroeconomic and policy analysis.

References

Ahir, H., N. Bloom, and D. Furceri (2022): "The World Uncertainty Index," *NBER Working Paper*, 29763.

Araci, D. (2019): "FinBERT: Financial Sentiment Analysis with Pre-trained Language Models," *arXiv preprint*, arXiv:1908.10063.

Arias, J. E., J. F. Rubio-Ramírez, and D. F. Waggoner (2018): "Inference Based on Structural Vector Autoregressions Identified with Sign and Zero Restrictions: Theory and Applications," *Econometrica*, 86, 685–720.

Aruoba, S. B. and T. Drechsel (forthcoming): "Identifying Monetary Policy Shocks: A Natural Language Approach," *American Economic Journal: Macroeconomics*.

Audrino, F., J. Gentner, and S. Stalder (2024): "Quantifying uncertainty: a new era of measurement through large language models," *SNB Working Papers*, 12.

Baker, S. R., N. Bloom, and S. J. Davis (2016): "Measuring Economic Policy Uncertainty," *The Quarterly Journal of Economics*, 131, 1593–1636.

Bauer, M. D. and E. T. Swanson (2023): "A Reassessment of Monetary Policy Surprises and High-Frequency Identification," *NBER Macroeconomics Annual*, 37, 87–155.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003): "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 3, 993–1022.

Bybee, L. (2025): "The Ghost in the Machine: Generating Beliefs with Large Language Models," *mimeo*.

Bybee, L., B. Kelly, A. Manela, and D. Xiu (2020): "The Structure of Economic News," *NBER Working Paper*, 26648.

——— (2024): "Business News and Business Cycles," *The Journal of Finance*, 79, 3105–3148.

Caldara, D. and M. Iacoviello (2022): "Measuring Geopolitical Risk," *American Economic Review*, 112, 1194–1225.

Carriero, A., D. Pettenuzzo, and S. Shekhar (2025): "Macroeconomic Forecasting with Large Language Models," *arXiv:2407.00890 [econ.EM]*.

Clayton, C., A. Coppola, M. Maggiori, and J. Schreger (2023): "Geoeconomic Pressure," *NBER Working Paper*, 34020.

de Bondt, G. and Y. Sun (2025): "Enhancing GDP nowcasts with ChatGPT: a novel application of PMI news releases," *ECB Working Paper*, 3603.

Del Negro, M. and F. Schorfheide (2013): "Chapter 2 - DSGE Model-Based Forecasting," in *Handbook of Economic Forecasting*, ed. by G. Elliott and A. Timmermann, Elsevier, vol. 2 of *Handbook of Economic Forecasting*, 57–140.

Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova (2019): "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *NAACL-HLT*, arXiv:1810.04805.

Ehrmann, M. and J. Talmi (2020): "Starting from a Blank Page? Semantic Similarity in Central Bank Communication and Market Volatility," *Journal of Monetary Economics*, 111, 48–62.

Eickmeier, S. and B. Hofmann (2023): "What drives inflation? Disentangling demand and supply factors," *BIS Working Papers*, 1047.

Faria-e Castro, M. and F. Leibovici (2024): "Artificial Intelligence and Inflation Forecasts," *Federal Reserve Bank of St. Louis Review*, 106, 1–14, fourth Quarter.

Gambacorta, L., B. Kwon, T. Park, P. Patelli, and S. Zhu (2024): "CB-LMs: language models for central banking," *BIS Working Paper*, 1215.

Gentzkow, M., B. Kelly, and M. Taddy (2019): "Text as Data," *Journal of Economic Literature*, 57, 535–574.

Hansen, A. L. and S. Kazinnik (2024): "Can ChatGPT Decipher Fedspeak?" Available at SSRN.

Hansen, S. and M. McMahon (2016): "Shocking Language: Understanding the Macroeconomic Effects of Central Bank Communication," *Journal of International Economics*, 99, S114–S133.

Hassan, T. A., S. Hollander, L. van Lent, and A. Tahoun (2019): "Firm-Level Political Risk: Measurement and Effects," *The Quarterly Journal of Economics*, 134, 2135–2202.

Jarociński, M. and P. Karadi (2020): "Deconstructing Monetary Policy Surprises—The Role of Information Shocks," *American Economic Journal: Macroeconomics*, 12, 1–43.

Korinek, A. (2023): "Generative AI for Economic Research: Use Cases and Implications for Economists," *Journal of Economic Literature*, 61, 1281–1317.

Kwon, B., T. Park, F. Perez-Cruz, and P. Rungcharoenkitkul (2024): "Large language models: a primer for economists," *BIS Quarterly Review*, December, 37–52.

Larsen, V. H. and L. A. Thorsrud (2019): "The Value of News for Economic Developments," *Journal of Econometrics*, 210, 203–218.

Loughran, T. and B. McDonald (2011): "When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks," *Journal of Finance*, 66, 35–65.

Mikolov, T., K. Chen, G. Corrado, and J. Dean (2013): "Efficient Estimation of Word Representations in Vector Space," *arXiv:1301.3781 [cs.CL]*.

Nakamura, E. and J. Steinsson (2018): "High-Frequency Identification of Monetary Non-Neutrality: The Information Effect," *The Quarterly Journal of Economics*, 133, 1283–1330.

Ramey, V. A. (2011): "Identifying Government Spending Shocks: It's All in the Timing," *The Quarterly Journal of Economics*, 126, 1–50.

Romer, C. D. and D. H. Romer (2004): "A New Measure of Monetary Shocks: Derivation and Implications," *American Economic Review*, 94, 1055–1084.

Shapiro, A. H. (2025): "Decomposing Supply- and Demand-Driven Inflation," *Journal of Money, Credit and Banking*, n/a.

Shiller, R. J. (2017): "Narrative Economics," *American Economic Review*, 107, 967–1004.

Tetlock, P. C. (2007): "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," *Journal of Finance*, 62, 1139–1168.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Kaiser, and I. Polosukhin (2017): "Attention Is All You Need," in *Advances in Neural Information Processing Systems*, vol. 30.

Appendix

Appendix A Prompts

A.1 Stage 1: Filtering relevant articles

A.1.1 Growth

You are a macroeconomist specializing in analyzing economic trends. You will be provided with a news article. Your task is to analyze the text and classify the sentiment of the article based on its reflection of macroeconomic activity in the United States. Use the following classification criteria to determine the sentiment:

- Positive: Indicates an expansion in broad macroeconomic activity (e.g. faster economic growth, faster income growth, increased production, increased hiring, lower unemployment).
- Negative: Indicates a contraction in broad macroeconomic activity (e.g. slower economic growth, slower income growth, a decline in production, higher unemployment).
- Neutral: Indicates that the news is related to macroeconomic activity, but there is no clear direction in either positive or negative direction (e.g. because there are opposing positive and negative forces that broadly offset each other, or there is very high uncertainty/ambiguity about the effects on macroeconomic activity, or the developments are likely to have only negligible effects on macroeconomic activity).
- Null: Indicates that the news is not closely related to economic activity.

Your response must be valid JSON and follow exactly this structure:

```
{
 "sentiment": "<Positive|Negative|Neutral|Null>"
}
```

A.1.2 Inflation

You are a macroeconomist specializing in analyzing economic trends. You will be provided with a news article. Your task is to analyze the text and classify the sentiment of the article based on its reflection of inflationary pressures in the United States. Use the following classification criteria to determine the sentiment:

- Positive: Indicates higher inflationary pressures in the economy, reflecting macroeconomic, geopolitical or other developments that are likely to cause faster increases in the Consumer Price Index (CPI), the Personal Consumption Expenditures Price Index (PCE), or the Producer Price Index (PPI). This includes instances when either core or headline components of inflation are affected. The expected effects on measured inflation can be immediate or take time to manifest.

- Negative: Indicates lower inflationary pressures in the economy, reflecting macroeconomic, geopolitical or other developments that are likely to cause slower increases or contraction in the Consumer Price Index (CPI), the Personal Consumption Expenditures Price Index (PCE), or the Producer Price Index (PPI). This includes instances when either core or headline components of inflation are affected. The expected effects on measured inflation can be immediate or take time to manifest.
- Neutral: Indicates that the news is related to inflationary pressures, but there is no clear direction in either positive or negative direction (e.g. because there are opposing positive and negative forces that broadly offset each other, or there is very high uncertainty/ambiguity about the effects on overall inflation, or the developments are likely to have only negligible effects on inflation).
- Null: Indicates that the news is not closely related to inflation.

You may encounter cases where the main development and its context carry opposing sentiments. For example, the central bank is raising rate (negative sentiment) in response to strong demand (positive sentiment). In such cases, prioritise the sentiment of the development itself if it is of macroeconomic significance. Otherwise assign sentiment based on the broader context.

Your response must be valid JSON and follow exactly this structure:

```
{
 "sentiment": "<Positive|Negative|Neutral|Null>"
}
```

A.2 Stage 2: Primary classifications

A.2.1 Growth

You are a macroeconomist specializing in U.S. economic trends. You will be given a news article related to U.S. macroeconomic activity. Your task is to analyze the text and (i) identify key drivers of macroeconomic development in each article, and then (ii) classify the sentiment of each identified driver. The classifications for both tasks are provided below.

Step 1. Classify the dominant drivers of the development:

- Demand drivers: The overriding drivers induce changes in desire to consume, spend, or invest, thus pushing macroeconomic activity and inflation in the same direction.
  - Examples: consumer confidence, expected income, fiscal policy, financial conditions, credit conditions, interest rates, monetary policy, exchange rates, liquidity, balance sheet conditions.
- Supply drivers: The overriding drivers affect the economy's production or output capacity, pushing macroeconomic activity and inflation in opposite directions.
  - Examples: commodity prices, trade/tariff policy, subsidies, labour supply, structural reforms, supply chain disruptions, natural disasters, productivity shocks.

- Both drivers: Use this classification if the article presents explicit evidence that demand and supply factors are equally dominant. They may or may not share the same sentiment.

- Undeterminable: Impossible to classify clearly.

Step 2. Classify the sentiment of the dominant driver. If "Both drivers" are present in Step 1, classify the sentiment of each driver:

- Positive: Expansion in macroeconomic activity (e.g. faster economic growth, rising incomes, higher production, more hiring, lower unemployment).

- Negative: Contraction in macroeconomic activity (e.g. slower economic growth, falling incomes, declining production, higher unemployment).

- Neutral: The article is related to macroeconomic activity but carries no clear sentiment in either direction, or entails negligible overall effects (e.g. because there are offsetting positive and negative forces, or there is a very high uncertainty/ambiguity about the overall effects).

- Null: Not related to macroeconomic activity.

You may encounter cases where the main development and its context carry opposing sentiments, e.g. the central bank is raising rate (negative sentiment) in response to strong demand (positive sentiment). In such cases, prioritise the sentiment of the development itself if it is of macroeconomic significance. Otherwise, assign sentiment based on the broader context.

Output JSON format (rules embedded):

```
{
  "Drivers of sentiment": "<Demand drivers|Supply drivers|Both
  ↪ drivers|Undeterminable>",
  "Sentiment": {
    "Demand drivers": "<Positive|Negative|Neutral|Null> # Non-Null only if
    ↪ demand drivers are present",
    "Supply drivers": "<Positive|Negative|Neutral|Null> # Non-Null only if
    ↪ supply drivers are present"
  },
  "confidence": {
    "drivers": "<0.0-1.0> # Confidence in driver classification",
    "sentiment": "<0.0-1.0> # Confidence in sentiment classification"
  },
  "explanation": "<Concise summary of reasoning behind the classifications>"
}
```

A.2.2 Inflation

You are a macroeconomist specializing in U.S. economic trends. You will be given a news article related to U.S. inflation development. Your task is to analyze the text and (i) identify key drivers of inflation development in each article, and then (ii) classify the sentiment of each identified driver. The classifications for both tasks are provided below.

Step 1. Classify the dominant drivers of the development:

- Demand drivers: The overriding drivers induce changes in desire to consume, spend, or invest, thus pushing macroeconomic activity and inflation in the same direction.
  - Examples: consumer confidence, expected income, fiscal policy, financial conditions, credit conditions, interest rates, monetary policy, exchange rates, liquidity, balance sheet conditions.
- Supply drivers: The overriding drivers affect the economy's production or output capacity, pushing macroeconomic activity and inflation in opposite directions.
  - Examples: commodity prices, trade/tariff policy, subsidies, labour supply, structural reforms, supply chain disruptions, natural disasters, productivity shocks.
- Both drivers: Use this classification if the article presents explicit evidence that demand and supply factors are equally dominant. They may or may not share the same sentiment.
- Undeterminable: Impossible to classify clearly.

Step 2. Classify the sentiment of the dominant driver. If "Both drivers" are present in Step 1, classify the sentiment of each driver:

- Positive: Higher inflationary pressures (e.g. faster increases in CPI, PCE or PPI (headline or core). Consider both immediate and delayed effects on inflation).
- Negative: Lower inflationary pressures (e.g. slower increases or outright decreases in CPI, PCE or PPI (headline or core). Consider both immediate and delayed effects on inflation).
- Neutral: The article is related to inflation but carries no clear sentiment in either direction, or entails negligible overall effects (e.g. because there are offsetting positive and negative forces, or there is a very high uncertainty/ambiguity about the overall effects).
- Null: Not related to inflation.

Output JSON format (rules embedded):

```
{
  "Drivers of sentiment": "<Demand drivers|Supply drivers|Both
  ↪ drivers|Undeterminable>",
  "Sentiment": {
    "Demand drivers": "<Positive|Negative|Neutral|Null> # Non-Null only if
    ↪ demand drivers are present",
    "Supply drivers": "<Positive|Negative|Neutral|Null> # Non-Null only if
    ↪ supply drivers are present"
  },
  "confidence": {
    "drivers": "<0.0-1.0> # Confidence in driver classification",
    "sentiment": "<0.0-1.0> # Confidence in sentiment classification"
  },
  "explanation": "<Concise summary of reasoning behind the classifications>"
}
```

A.3 Stage 3: In-depth classifications

A.3.1 Growth

You are a macroeconomist specializing in U.S. economic trends. You will be given a news article, together with a sentiment about U.S. economic growth–positive, negative or neutral and its main drivers (demand or supply). Your task is to provide a more detailed classification of the driver types.

For demand drivers, choose:

- Real, private drivers: These are non-financial drivers such as confidence, perception of uncertainty, expected future income, and other factors reflecting shifts in private sector's willingness to spend. These drivers are unrelated to fiscal policy.
- Real, fiscal policy drivers: These are predominantly shifts in federal fiscal policy—such as new or revised legislation or an officially proposed federal budget—that directly affects aggregate demand via changes in taxation or government spending (e.g., infrastructure, transfers, subsidies). Include cases where fiscal policy is affected by administrative factors such as government shutdowns, legislative delays or stronger enforcement. The change must be substantial enough to shift national output and/or inflation. For example, you should exclude cases like allocation of existing budgets that do not expand or contract aggregate demand; state or municipal policies, unless they are federally coordinated and have meaningful national demand implications; analysis or evaluation of past fiscal policies, unless there are clear implications for new fiscal actions in response; or any government policy that does not directly impact demand, such as regulatory changes, trade policies, or administrative rules. Be conservative in assigning "fiscal policy." If there is any ambiguity or lack of direct impact on national aggregate demand, classify as "not fiscal policy."
- Financial, private drivers: These include financial conditions, credit conditions, cost of funds, interest rates, exchange rates, liquidity conditions, balance sheet conditions and so on. Easier financial conditions, lower interest rates and exchange rate depreciation tend to raise growth and inflation. These factors reflect shifting conditions in the financial sector that are not directly induced by changes in monetary policy.
- Financial, monetary policy drivers: These are shifts in overall financial conditions that are primarily induced or caused by shifts in monetary policy actions. Monetary policy actions may include shifts in the policy interest rate, asset purchases/sales or central bank balance sheet compositions, as well as communication or signals to shift these policies in the foreseeable future. Do not answer "monetary policy" if there are multiple competing factors causing changes in overall financial conditions, and it is not clear that monetary policy is the most important factor.
- Undeterminable: The demand drivers cannot be classified as either of the above. In this case, please provide one or two key words to summarise the type of demand drivers.

For supply drivers, choose:

- Commodity prices: e.g. changes in the prices of oil, energy, metal, agricultural products and other raw materials.
- Policy shifts: e.g. shifts in trade policy, tariffs, subsidies, structural reforms or any other government policies that affect the economy's supply side.
- Supply disruptions: e.g. global supply chain problems, bottlenecks, or factors related to global value chain.
- Others: All other drivers. This may include natural disasters, productivity shocks and so on. In this case, please provide one or two key words to summarise the type of supply drivers.

Output JSON format (rules embedded):

```
{
  "Input context": {
    "Sentiment": "<Positive|Negative|Neutral> # Copied exactly from input",
    "Main drivers": "<Demand drivers|Supply drivers >"
  },
  "Driver subtypes": {
    "Demand drivers": "<Real, private drivers|Real, fiscal policy
    ↪  drivers|Financial, private drivers|Financial, monetary policy
    ↪  drivers|Undeterminable|Null>  # Non-Null only if demand drivers are
    ↪  present",
    "Supply drivers": {
      "Type": "<Commodity prices|Policy shifts|Supply disruptions|Others|Null>
      ↪  # Non-Null only if supply drivers are present",
      "Supply-Other Details": "<Open text field for additional details if
      ↪  'Others' is selected; otherwise Null>"
    }
  },
  "confidence": {
    "subtypes": "<0.0-1.0>  # Confidence in sub-driver classification"
  },
  "explanation": "<Concise summary of reasoning behind the classifications>"
}
```

A.3.2 Inflation

You are a macroeconomist specializing in U.S. economic trends. You will be given a news article, together with a sentiment about U.S. inflation development–positive (higher inflation), negative (lower inflation) or neutral and its main drivers (demand or supply). Your task is to provide a more detailed classification of the driver types.

For demand drivers, choose:

- Real, private drivers: These are non-financial drivers such as confidence, perception of uncertainty, expected future income, and other factors reflecting shifts in private sector's willingness to spend. These drivers are unrelated to fiscal policy.

- Real, fiscal policy drivers: These are predominantly shifts in federal fiscal policy—such as new or revised legislation or an officially proposed federal budget—that directly affects aggregate demand via changes in taxation or government spending (e.g., infrastructure, transfers, subsidies). Include cases where fiscal policy is affected by administrative factors such as government shutdowns, legislative delays or stronger enforcement. The change must be substantial enough to shift national output and/or inflation. For example, you should exclude cases like allocation of existing budgets that do not expand or contract aggregate demand; state or municipal policies, unless they are federally coordinated and have meaningful national demand implications; analysis or evaluation of past fiscal policies, unless there are clear implications for new fiscal actions in response; or any government policy that does not directly impact demand, such as regulatory changes, trade policies, or administrative rules. Be conservative in assigning "fiscal policy." If there is any ambiguity or lack of direct impact on national aggregate demand, classify as "not fiscal policy."

- Financial, private drivers: These include financial conditions, credit conditions, cost of funds, interest rates, exchange rates, liquidity conditions, balance sheet conditions and so on. Easier financial conditions, lower interest rates and exchange rate depreciation tend to raise growth and inflation. These factors reflect shifting conditions in the financial sector that are not directly induced by changes in monetary policy.

- Financial, monetary policy drivers: These are shifts in overall financial conditions that are primarily induced or caused by shifts in monetary policy actions. Monetary policy actions may include shifts in the policy interest rate, asset purchases/sales or central bank balance sheet compositions, as well as communication or signals to shift these policies in the foreseeable future. Do not answer "monetary policy" if there are multiple competing factors causing changes in overall financial conditions, and it is not clear that monetary policy is the most important factor.

- Undeterminable: The demand drivers cannot be classified as either of the above. In this case, please provide one or two key words to summarise the type of demand drivers.

For supply drivers, choose:

- Commodity prices: e.g. changes in the prices of oil, energy, metal, agricultural products and other raw materials.

- Policy shifts: e.g. shifts in trade policy, tariffs, subsidies, structural reforms or any other government policies that affect the economy's supply side.

- Supply disruptions: e.g. global supply chain problems, bottlenecks, or factors related to global value chain.

- Others: All other drivers. This may include natural disasters, productivity shocks and so on. In this case, please provide one or two key words to summarise the type of supply drivers.

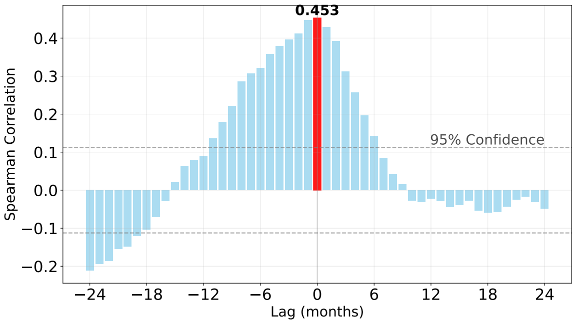
Output JSON format (rules embedded):

```
{
  "Input context": {
    "Sentiment": "<Positive|Negative|Neutral> # Copied exactly from input",
    "Main drivers": "<Demand drivers|Supply drivers >"
  },
  "Driver subtypes": {
    "Demand drivers": "<Real, private drivers|Real, fiscal policy
    ↪  drivers|Financial, private drivers|Financial, monetary policy
    ↪  drivers|Undeterminable|Null>  # Non-Null only if demand drivers are
    ↪  present",
    "Supply drivers": {
      "Type": "<Commodity prices|Policy shifts|Supply disruptions|Others|Null>
      ↪  # Non-Null only if supply drivers are present",
      "Supply-Other Details": "<Open text field for additional details if
      ↪  'Others' is selected; otherwise Null>"
    }
  },
  "confidence": {
    "subtypes": "<0.0-1.0>  # Confidence in sub-driver classification"
  },
  "explanation": "<Concise summary of reasoning behind the classifications>"
}
```

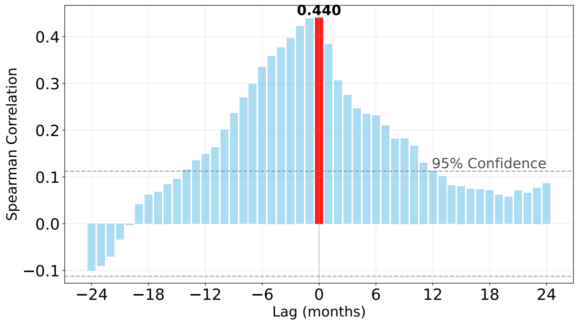
Appendix B Additional results

Figure B.1: Cross-correlograms of growth drivers and data counterparts

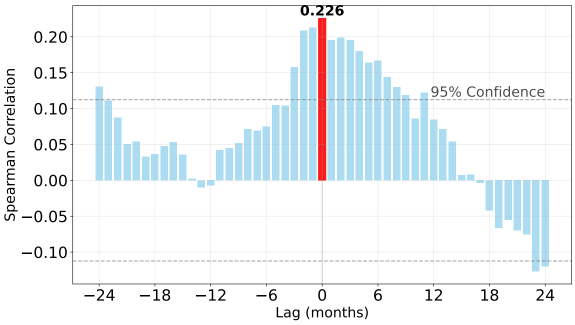
(a) Real demand drivers vs PCE



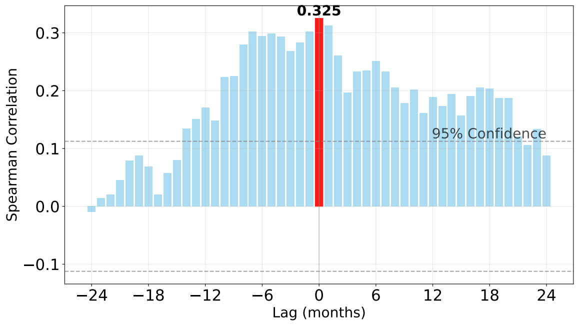
(b) Financial demand drivers vs FCI



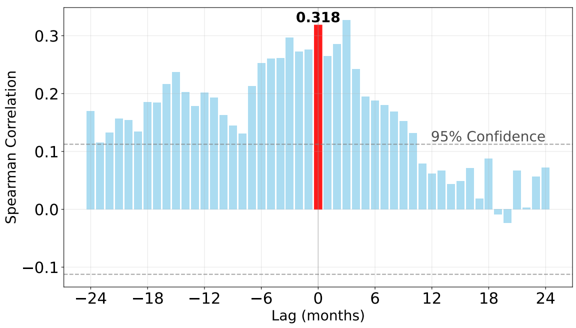
(c) Commodity prices vs WTI



(d) Supply disruptions vs GSCPI



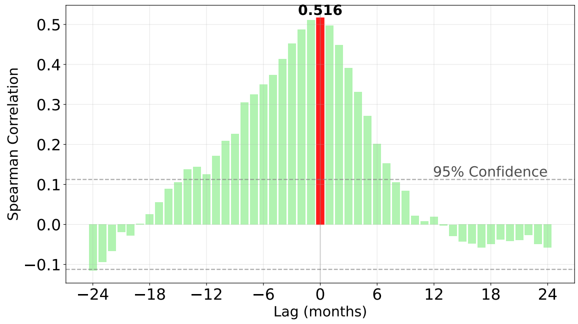
(e) Policy shifts vs TPU



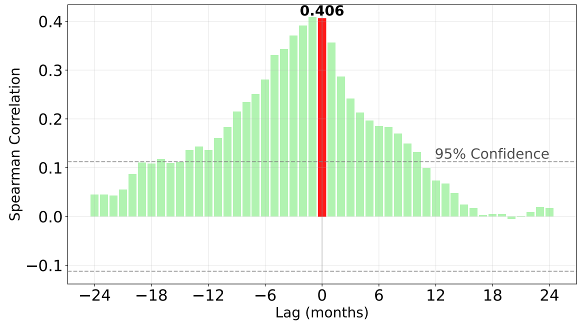
Note: Negative lags = Growth sentiment drivers lead; positive lags = Indicators lead.

Figure B.2: Cross-correlograms of inflation drivers and data counterparts

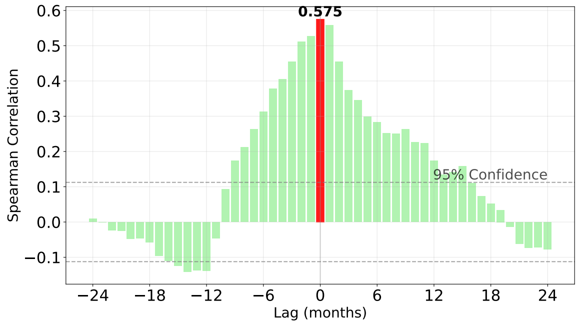
(a) Real demand drivers vs PCE



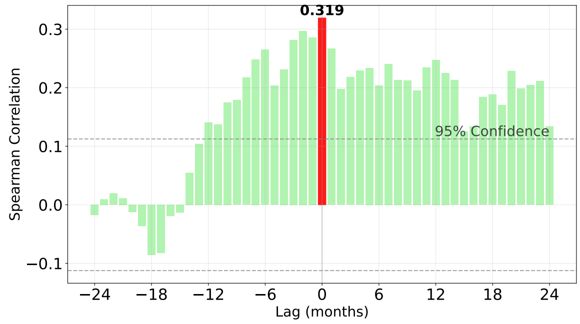
(b) Financial demand drivers vs FCI



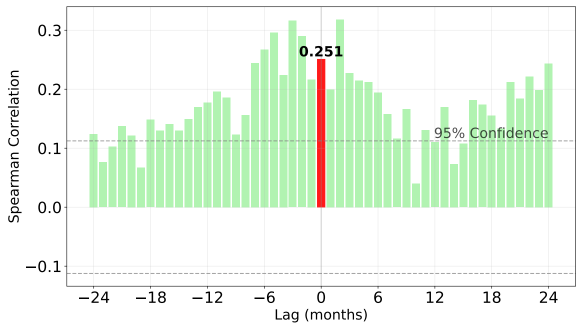
(c) Commodity prices vs WTI



(d) Supply disruptions vs GSCPI



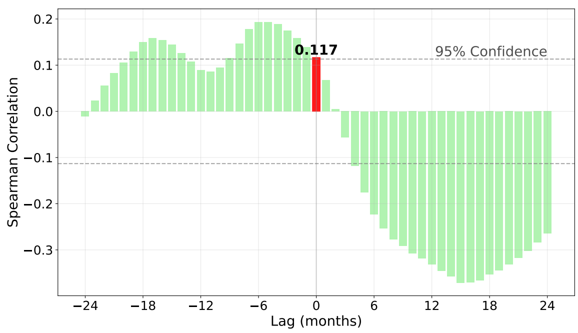
(e) Policy shifts vs TPU



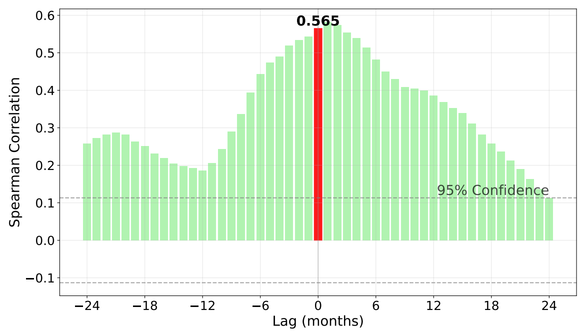
Note: Negative lags = Inflation sentiment drivers lead; positive lags = Indicators lead..

Figure B.3: Cross-correlograms of inflation sentiment drivers and Shapiro's counterparts

(a) Demand components



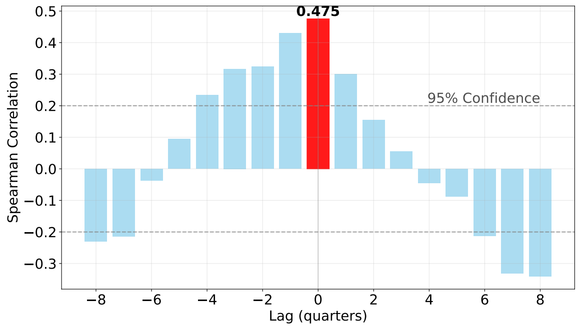
(b) Supply components



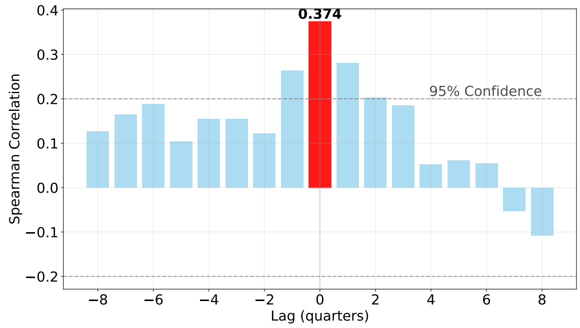
Note: Negative lags = Inflation sentiment drivers lead; positive lags = Shapiro's inflation shocks lead.

Figure B.4: Cross-correlograms of macroeconomic sentiment drivers and Eickmeier-Hofmann counterparts

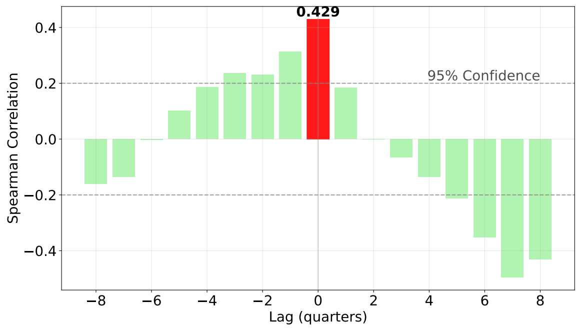
(a) Demand components of growth



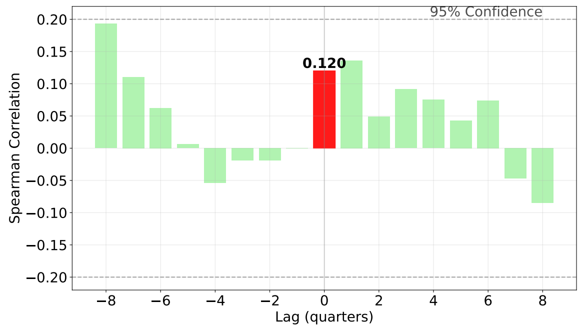
(b) Supply components of growth



(c) Demand components of inflation



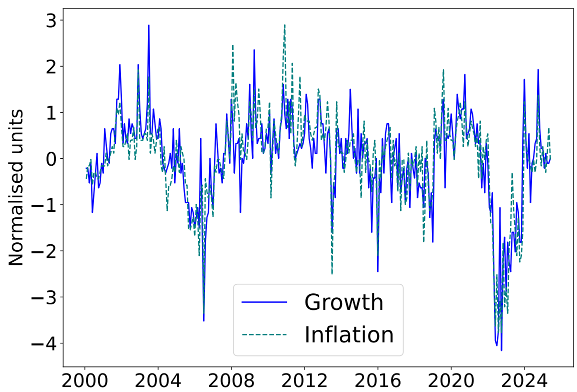
(d) Supply components of inflation



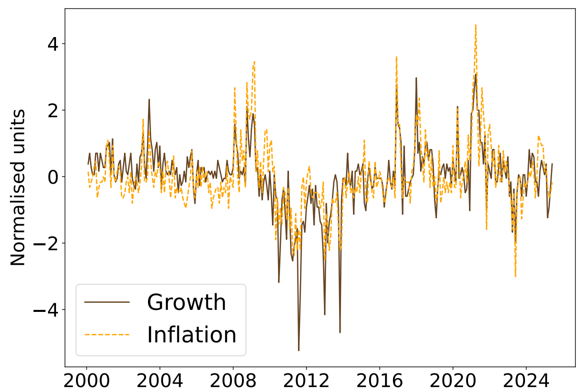
Note: Negative lags = Macroeconomic sentiment drivers lead; positive lags = Eickmeier-Hofmann's shocks lead. Eickmeier-Hofmann's shocks are from Eickmeier and Hofmann (2023).

Figure B.5: Policy drivers of demand sentiment

(a) Monetary policy drivers



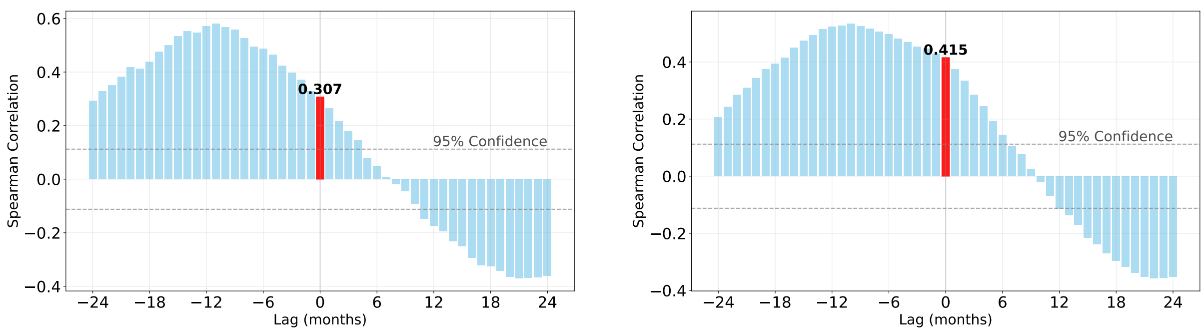
(b) Fiscal policy drivers



Note: All series are z-scores.

Figure B.6: Cross-correlograms of monetary policy sentiment and interest rates

(a) Effective fed funds (b) UST 2-year yield



Note: Negative lags = Monetary policy sentiment leads; positive lags = Interest rates lead.

Previous volumes in this series

1293 October 2025	International Risk Sharing and Wealth Allocation with Higher Order Cumulants	Giancarlo Corsetti, Anna Lipińska and Giovanni Lombardo
1292 September 2025	Macroeconomic impact of weather disasters: a global and sectoral analysis	Torsten Ehlers, Jon Frost, Carlos Madeira and Ilhyock Shim
1291 September 2025	Harnessing artificial intelligence for monitoring financial markets	Matteo Aquilina, Douglas Araujo, Gaston Gelos, Taejin Park and Fernando Pérez-Cruz
1290 September 2025	The speed premium: high-frequency trading and the cost of capital	Matteo Aquilina, Gbenga Ibikunle, Khaladdin Rzayev and Xuesi Wang
1289 September 2025	The liquidity state dependence of monetary policy transmission	Oliver Ashtari-Tafti, Rodrigo Guimaraes, Gabor Pinter and Jean-Charles Wijnandts
1288 September 2025	The capital puzzle	Eduardo Amaral
1287 September 2025	Banks' regulatory risk tolerance	Mikael Juselius, Aurea Ponte Marques and Nikola Tarashev
1286 September 2025	How Do Quantitative Easing and Tightening Affect Firms?	Egemen Eren, Denis Gorea and Daojing Zhai
1285 August 2025	R* in East Asia: business, financial cycles, and spillovers	Pierre L Siklos, Dora Xia and Hongyi Chen
1284 August 2025	What is needed for convergence? The role of finance and capital	Bryan Hardy and Can Sever
1283 August 2025	Comparing search and intermediation frictions across markets	Gabor Pinter, Semih Üslü and Jean-Charles Wijnandts
1282 August 2025	Market whiplash after the 2025 tariff shock: an event-targeted VAR approach	Gabor Pinter, Frank Smets and Semih Üslü
1281 July 2025	Integrating balance sheet policy into monetary policy conditions	Benoit Mojon, Phurichai Rungcharoenkitkul and Dora Xia
1280 July 2025	CBDC and banks: disintermediating fast and slow	Rhys Bidder, Timothy Jackson and Matthias Rottner

All volumes are available on our website www.bis.org.