GDP nowcasting power of PMI news releases using ChatGPT

Authors

Gabe J. de Bondt | European Central Bank (ECB)
Yiqiao Sun | European Central Bank (ECB)

Abstract

This policy brief summarizes the findings of a novel approach to economic forecasting as outlined in the ECB Working Paper, "Enhancing GDP Nowcasts Using ChatGPT: A Novel Application to PMI News Releases". The study demonstrates how artificial intelligence (AI), specifically ChatGPT, can improve euro area GDP nowcasts by deriving sentiment activity scores from Purchasing Managers' Index (PMI) news releases. The forecast gain varies by calendar year and is on average around 20% apart from the two most recent years. Our findings raise the question of whether economic forecasters should prioritize the use of big data or instead focus on small but economically meaningful datasets — as little as a few pages of text, as in our case.

Key insights

1. ChatGPT-derived PMI text scores enhance GDP nowcasts

A PMI text score is derived by tasking ChatGPT to classify activity reported in PMI news releases. The ChatGPT-derived scores quantify the sentiment about activity expressed in the narratives and anecdotes of the news release, ranging from strongly expanding activity to strongly contracting activity. The study then integrates these ChatGPT-derived activity scores in traditional GDP nowcasts, i.e., forecasts of real GDP growth in the current quarter, to assess their predictive power. Our study shows that ChatGPT-generated sentiment activity scores significantly improve the accuracy of real GDP growth nowcasts, enhancing two hard-to-beat benchmarks:
- ECB/Eurosystem Staff Projections (*ESP*), which incorporates economic expert judgement;
- Eurostat first GDP estimate (*1st GDP*), which relies on statistical methodologies to address data gaps especially for the services sector.

The equation estimated to test whether an enhancement of the benchmark GDP nowcast with the PMI text score is fruitful reads as follows, with y equal to quarter-on-quarter real GDP growth according to the latest GDP release:

$$y_t = \beta \cdot \text{PMI text score}_t + 1.0 \cdot \text{benchmark GDP nowcast}_t + \varepsilon_t \qquad (1)$$

2. Robustness across methodologies

The findings are robust across various methodological adjustments (see upper panel of Table 1):
- excluding extraordinary economic events (e.g., the pandemic);
- considering different GDP growth quantiles;
- employing alternative prompts for deriving text scores;
- applying different weighting schemes.
This robustness underscores the usefulness of ChatGPT-based sentiment analysis as a forecasting tool.

3. Accuracy improves with incoming news

The accuracy of GDP nowcasts improves as more information becomes available throughout the quarter (see lower panel of Table 1). The most significant enhancements are observed when incorporating data from all three months of the quarter (reported in the upper panel) or in addition the first month of the next quarter, which is available just before the first GDP estimate. This underscores the importance of continuously updating forecasts with the latest PMI news releases to maximise predictive accuracy.

4. Text scores outperform PMI diffusion indices

While the PMI diffusion index has traditionally been used for nowcasting GDP (de Bondt 2012 and 2019), the study finds that ChatGPT-derived text scores provide greater predictive power. Regression analyses reveal that diffusion indices contribute little to improving GDP nowcasts and, in some cases, negatively affect forecast accuracy.

Table 1. In-sample estimates of GDP nowcast benchmarks enhanced with ChatGPT-derived PMI text scores

Sample	β (ESP)	β (1st GDP)	β (ESP)	β (1st GDP)	β (ESP)	β (1st GDP)	β (ESP)	β (1st GDP)
	All three months of the quarter							
	Basic prompt		ECB contacts' prompt		Sector split prompt		Theoretical weights	
All	0.08*	0.03	0.07	0.02	0.08**	0.04**	0.06	0.03
Exluding extremes	0.05*	0.04***	0.05	0.03*	0.06**	0.05***	0.04	0.03**
0.25 GDP growth quantile	0.08***	0.07***	0.05***	0.07***	0.02	0.04**	0.04*	0.04***
0.50 GDP growth quantile	0.10***	0.08***	0.05**	0.06**	0.07***	0.05***	0.05*	0.05***
0.75 GDP growth quantile	0.15***	0.08***	0.05**	0.04***	0.09***	0.08***	0.07***	0.06***
	Basic prompt							
	1st month of the quarter		1st and 2nd month		4 months		1st month of next quarter	
All	0.05	0.00	0.05	0.02	--	0.04**	--	0.05***
Exluding extremes	0.02	0.02	0.03	0.02	--	0.04***	--	0.04***
0.25 GDP growth quantile	0.03	0.02	0.04**	0.04**	--	0.05***	--	0.05***
0.50 GDP growth quantile	0.04	0.04**	0.04*	0.05***	--	0.05***	--	0.05***
0.75 GDP growth quantile	0.05**	0.05***	0.07***	0.06***	--	0.07***	--	0.05***

Notes: This table reports the in-sample estimates of Eq. (1). *** <1%, **<5%, *<10% significance level. The sources are Tables 2, 3, 4 and 8 in de Bondt and Sun (2025). ESP = ECB/Eurosystem Staff Projections GDP nowcast. 1st GDP = Eurostat first GDP estimate. ECB contacts' prompt refers to a prompt along the activity question regularly asked in the ECB's dialogue with non-financial companies (Elding et al., 2021).

## 5. Narrative and services information explain forecast gain

The forecast gain from ChatGPT-derived sentiment scores can be attributed to the richness of qualitative information embedded in PMI news releases. These releases often include narratives, anecdotes, and sector-specific insights — particularly about services — that are not fully captured in the initial GDP estimate due to incomplete data. According to standard statistical practices, it can take up to three years for the final GDP statistics to be determined. Services, which constitute a significant component of GDP in the euro area, appear to be better represented through textual analysis, enabling a more accurate and nuanced understanding of the economic situation.

6. Sizeable out-of-sample forecast gains on average but varying by calendar years

The integration of PMI text scores into the two benchmark GDP nowcasts results in forecast gains of approximately 20% excluding the two most recent years (Figure 1). The forecast gain relative to the two hard-to-beat benchmarks is sizeable but do vary a lot across calendar years. This highlights a high potential of qualitative data to complement quantitative methods in economic forecasting, but at the same time caution is needed given a high variation in relative accuracy.

7. Minimal data requirements

Unlike traditional AI applications requiring vast datasets, this approach achieves significant forecast gains using just a few pages of text from PMI news releases. This efficiency makes it a practical tool for policymakers and analysts. Our small data exercise contrasts other studies using big data.

Figure 1. Out-of-sample performance relative to two hard-to-beat GDP nowcasts
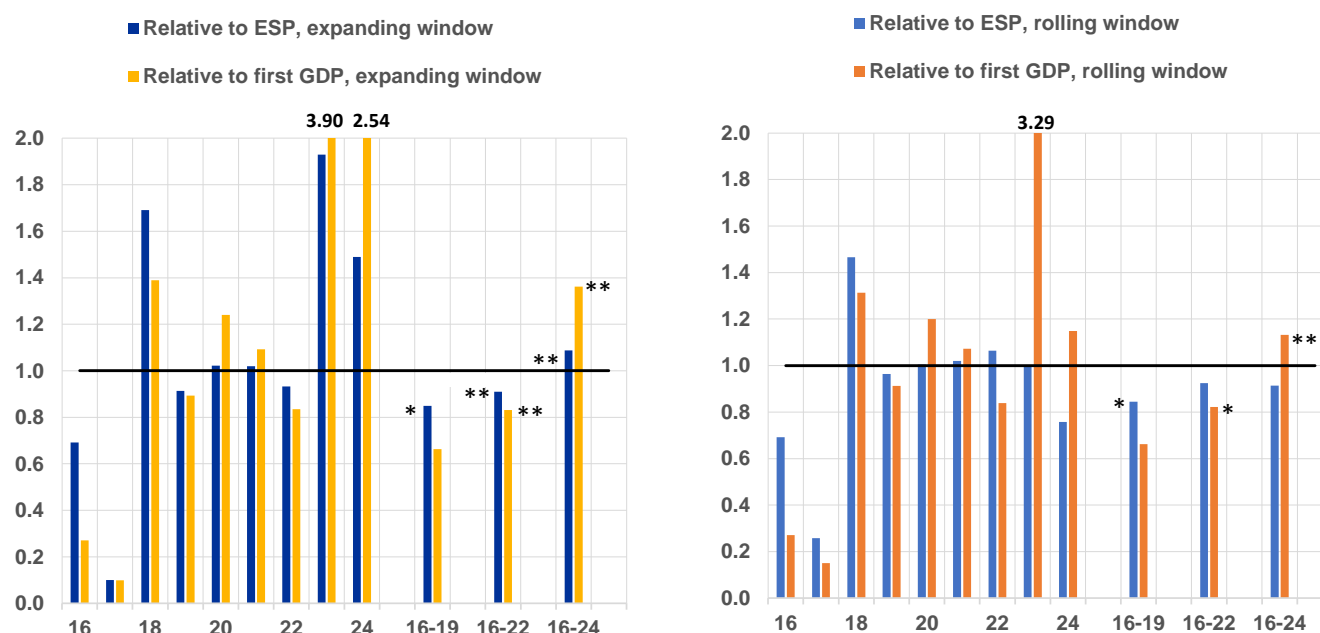
Notes: This figure plots the root mean squared error of out-of-sample GDP nowcasts enhanced with PMI text score relative to ECB/Eurosystem Staff Projections (ESP) and first GDP estimate benchmarks, with expanding window excluding extremes as well as with a rolling eight-year window. ** $p < 0.05$, * $p < 0.10$ refers to the Clark West test statistics, indicating a significant difference in the nowcasts from the model including the PMI text score versus those of the respective benchmark model. The source is Figure 2 in de Bondt and Sun (2025).

Policy implications

The findings underscore the transformative potential of AI in economics, paving the way for innovative and more effective forecasting tools. The demonstrated success of ChatGPT-derived PMI text scores highlights the potential of qualitative data to enhance economic assessments. Our approach complements studies applying AI to other qualitative economic texts, such as central bank communications (Hansen and Kazinnik, 2024; Smales, 2023) and reports (Díaz Sobrino et al., 2022, Sharpe et al., 2023; Du et al., 2024). Policymakers and institutions should integrate AI-driven sentiment analysis into their forecasting toolkits. While this study focuses on the euro area, the methodology can be easily adapted to other regions or countries.

References

de Bondt, G.J. (2012) Nowcasting: trust the Purchasing Managers' Index or wait for the flash GDP estimate? in: G.T. Papanikos (ed.), Economic Essays, Athens Institute for Education and Research, Chapter 8: 83-97.

de Bondt, G.J. (2019) A PMI-based real GDP tracker for the euro area, Journal of Business Cycle Research, 15(2): 147-170.

de Bondt, G.J. and Y. Sun (2025) Enhancing GDP nowcasts with ChatGPT: a novel application of PMI news releases, ECB Working Paper, No. 3063.

Díaz Sobrino, N., Ghirelli, C., Hurtado, S., Pérez, J. J., and Urtasun, A. (2022). The narrative about the economy as a shadow forecast: an analysis using Bank of Spain quarterly reports, Applied Economics, 54(25): 2874–2887.

Du, S., Guo, K., Haberkorn, F., Kessler, A., Kitschelt, I., Lee, S. J., Monken, A., Saez, D., Shipman, K. and Thakur, S. (2024) Do anecdotes matter? Exploring the Beige Book through textual analysis from 1970 to 2023, Irving Fisher Committee on Central Bank Statistics (IFC) Bulletin, No. 57.

Elding, C., Morris, R. and Slavik, M. (2021) The ECB's dialogue with non-financial companies. ECB Economic Bulletin, Issue 1.

Hansen, A.L. and Kazinnik, S. (2023) Can ChatGPT Decipher Fedspeak?, August 5. Available at SSRN: https://ssrn.com/abstract=4399406.

Sharpe, S., Sinha, N. and Hollrah, C. (2023) The power of narrative sentiment in economic forecasts, International Journal of Forecasting, 39(3): 1097–1121.

Smales, L.A. (2023) Classification of RBA monetary policy announcements using ChatGPT, Finance Research Letters, 58:104514.

About the author(s)

Gabe J. de Bondt is employed as Senior Lead Economist in the Business Cycle Analysis Division of the European Central Bank (ECB). Prior to joining the ECB in 2000, he was Economist at De Nederlandsche Bank. He holds a PhD from the University of Amsterdam. His research interests are in the field of business cycle (leading) indicators and surveys, monetary economics, and financial markets.

Yiqiao Sun is currently Economist in the Forecasting and Policy Modelling Division of the European Central Bank. Previously, she also worked in the Prices and Costs Division and Business Cycle Analysis Division. She holds a PhD from WHU-Otto Beisheim School, Germany. Her research interests are business cycles, monetary economics and economic forecasting.