EUROPEAN CENTRAL BANK
EUROSYSTEM

Working Paper Series

Manuel Medina Magro, Lorena Saiz

What can newspaper articles reveal about the euro area economy?

No 3122

Abstract

This study introduces a novel approach to dictionary-based sentiment analysis that extracts valuable insights from economic newspaper articles in the euro area without requiring article translation. We develop sentiment indices that accurately measure economic, labour, and inflation perceptions in Germany, France, Italy, and Spain using native-language texts. The aggregation of these country-specific sentiments provides a reliable indicator for the euro area as a whole, demonstrating the effectiveness of our approach in several nowcasting and forecasting experiments. This translation-free method significantly reduces resource requirements, facilitates easy replication across various languages, and enables daily updates. By eliminating the translation bottleneck, our approach emerges as one of the most timely and cost-effective economic measures available, offering a powerful tool for monitoring and forecasting business cycles in the multilingual context of the euro area.

Keywords: textual analysis, forecasting, inflation, recession, output

JEL Codes: E32, E37, C53, C82

Non-technical summary

Our research explores a novel approach to understanding economic trends by analyzing the sentiment of newspaper articles. We have developed a method to extract valuable economic insights from news articles in the four largest economies in the euro area: Germany, France, Italy, and Spain.

Our approach focuses on gauging sentiment about output, inflation, and employment. Unlike traditional methods that rely on surveys, our technique uses natural language processing to analyze economic news articles. This allows us to create daily sentiment indices that can anticipate changes in GDP and other key economic indicators.

Instead of translating articles, we have developed a method that works in different languages, making it more efficient and scalable for analyzing the euro area as a whole. Circumventing the need to translate the articles saves time and money. Translating often involves paying for an API and/or waiting for the translations, which can potentially lead to errors or ambiguities. Furthermore, our method has better forecasting performance than other methods in the literature that require translation.

Our newspaper-based indices have shown impressive results. They have accurately predicted economic turning points, including the sovereign debt crisis and the COVID-19 pandemic. In fact, our indices outperformed traditional economic indicators in forecasting quarterly economic growth as well as the high inflationary period of 2022 and 2023. They also achieved small gains in forecast accuracy when predicting survey sentiment indicators.

The merit of our approach lies in its timeliness and cost-effectiveness. It provides daily updates on economic sentiment, offering policymakers and analysts a real-time, automated tool for understanding current and future economic conditions.

1 Introduction

Many economics works have investigated the role of sentiment shocks as business cycle drivers and concluded that these shocks can explain a significant part of business cycle fluctuations (see, for instance, Lagerborg et al. (2023)). Sentiment shocks can be seen as changes in the optimism or pessimism of agents' expectations, unrelated to economic fundamentals. Keynes referred to it as 'animal spirits' or the state of confidence and considered it one of the major determinants of investment. In practice, sentiment is not directly observable and needs to be proxied. Surveys are a very effective way to get consumer and business expectations. More recently, numerous works have explored alternative approaches to measuring sentiment based on natural language processing (NLP) techniques to extract the sentiment from news articles, social media or financial reports.

A rapidly growing literature has found that textual sentiment indicators can complement survey-based sentiment measures and be helpful for macroeconomic forecasting (see, for example, Barbaglia et al. (2022, 2025); Shapiro et al. (2022); Kalamara et al. (2022)). This literature faces limitations in conducting multilingual analysis, which is essential for deriving sentiment in linguistically diverse regions such as the euro area. Moreover, most of the dictionaries and language models have been developed for the English language. Therefore, for the euro area, given the multilingual context, it is necessary to create new dictionaries or models for non-English languages or translate texts into English (Barbaglia et al. (2024); Ashwin et al. (2024)). Large language models such as ChatGPT can help overcome this issue, but they often come with monetary costs and security concerns, and little is known about their training processes or potential biases in specific domains like economics.

In this paper, we propose an approach that circumvents the need for article translation, making it suitable for forecasting economic variables in the euro area. More specifically, we construct textual sentiment indices for the four largest euro area countries and then aggregate them to obtain sentiment indices for the euro area. Since the euro area currently comprises 20 countries and has 17 official languages, conducting textual analysis would be a titanic task due to the large language dimension. For this reason, as in Ashwin et al. (2024), our analysis focuses only on the four largest countries in the euro area (and therefore four non-English languages). Moreover, for predicting euro area macroeconomic aggregates, we found that using aggregated textual measures yields more accurate results than using other typically used macroeconomic indicators (e.g., survey indicators, industrial production).

Regarding how we construct sentiment measures, we follow Ochs (2021) and focus on specific economic topics and on directional changes in some macroeconomic aggregates reported in the news articles rather than the tone (i.e. positive, negative, or neutral) of the articles. This approach has several advantages. First, counting the number of times the words 'economy', 'inflation', or 'employment' appear in the news provides information about the relevance of these topics in the news media. Second, in combination with directional keywords, it informs about the directional change of output, inflation, or employment reported by the media. Moreover, this approach is relatively simple to compute in any language, given that the set of directional change keywords is rather limited and stable over time (e.g. increase, decrease, accelerate, decelerate). In contrast, developing new dictionaries in a specific language with a long, exhaustive set of words that have been previously classified and scored, which can be a challenging task. Also, as shown by Ashwin et al. (2024) in an application for the euro area, the choice of the text dictionary matters for the results. We chose a simple dictionary with few words for several reasons: it facilitates scalability, enhances interpretability, and ensures the transparency and replicability of our analysis.

We also built a simple measure that counts the number of times the words 'recession' or 'crisis' appear in news articles (the so-called R-word index[1]), and is well known as a good predictor of turning points (see for instance Ferrari Minesso et al. (2022) and Bybee et al. (2024)). In a similar spirit to Hassan et al. (2023), we also built a measure that counts the number of times these words appear together with the words 'risk' or 'uncertainty' (the R-risk index). This measure identifies spikes in perceived recession risk in the media, even before the recession starts.

To validate our textual indices, we investigate their correlation with survey-based sentiment measures and macroeconomic aggregates. Moreover, to highlight the value added of our straightforward approach, we compare our textual measures with those obtained using well-known English dictionaries (e.g., VADER) in the English-translated news articles. We find that while

[1]The methodology to build this index is similar to The Economist's R-word index. See Gauging the gloom, The Economist, Sep. 16th 2011.

all indices are highly correlated, our approach delivers more accurate results in forecasting GDP.

Overall, our work demonstrates that these simple measures can provide valuable insights into the state of the economy and the euro area business cycle. The directional sentiment measures can anticipate changes in the primary macroeconomic aggregates (i.e. output, inflation). At the same time, the R-word and R-risk index help to detect turning points.

The paper is structured as follows. Section 2 outlines the proposed methodology for constructing news-based indicators. Section 3 investigates the relationship between our news sentiment measures and survey-based sentiment, output, employment, and inflation. Section 4 shows that news sentiment contains information that helps forecast GDP and inflation, as well as identify turning points and nowcast survey sentiment. Section 5 compares our methodology with other alternative approaches. And Section 6 concludes.

2 Methodology

This section presents our proposed approach for constructing news-based indicators. We used articles from major print newspapers for the four largest euro area economies from the Dow Jones Factiva database. In each case, we restrict the data to articles that are tagged as either economic, corporate, or financial market news. Table 9, in Appendix A, provides more details on the news articles used in the analysis.

We construct two recession indices and three sentiment indices corresponding to key economic topics: economy, labour, and prices. The emphasis on these topics arises from the understanding that central bankers prioritize the output gap, the (un)employment gap, and the deviation of actual inflation from policy targets when making monetary policy decisions (Taylor, 1993).

Numerous studies in the literature have demonstrated that indices with good forecasting properties can be constructed using relatively simple methods. Our proposed approach consists of counting the number of articles that contain specific keywords (Ferrari Minesso et al., 2022) and combining keyword detection with directional terms (Ochs, 2021; Aguilar et al., 2021). Due to the multilingual dimension of our work, using this approach with a few essential keywords and directional words is the most efficient way to build the news indicators in terms of minimizing translation effort and computational costs. Instead of translating the corpus of articles, which would drastically increase the computational cost of producing the indices, we use language-specific dictionaries for each different language. These dictionaries were created with the assistance of native-speaking economists from the Business and Cycle Analysis Division (BCA) at the ECB, aiming to reflect the typical usage of words in newspapers.

For the recession indices, we generate the Rword index, which represents the percentage of news articles containing the words 'recession' or 'crisis'. We also produce the Rrisk index, defined as the percentage of recession-related news that includes the words 'risk' or 'uncertainty'. The Rrisk index reflects discussions in the news about recessions as a potential risk that may or may not materialize.

The sentiment indices for the economy, labour, and inflation are constructed in several steps.

First, we keep only the news articles that contain the name of the country of origin of the newspaper to filter out articles about foreign economies.

Second, we use the roots of topic-specific keywords to detect any possible variations of the words and identify the topics the news is discussing. We do this by iterating over every word in every article and checking whether they start with the roots of specific keywords (e.g. econom*). When this condition holds, we open a window that goes from 5 words before the keyword to 5 words after the keyword and check for the presence of directional terms (e.g. for increase: improv*, accelerat*, etc.; and for decrease: slow*, contract*, etc.). The sentence is classified as an increase or decrease if a directional term is found within the window. If no directional term is found, the sentence is considered neutral.

We also select a set of single words that indicate 'increase' or 'decrease' on their own, without the need to be accompanied by a topic keyword. For example, deflation means a 'decrease' in prices, even if the topic keywords 'price' or 'inflation' are not within the +/-5 word window. We call these words 'strong increase/decrease', and the presence of one of these words in the article has the same weight as an 'increase/decrease' sentence. The German dictionary has a larger proportion of strong words because the German language tends to merge several words into one word, so frequently, just one word refers to both the topic and the directionality.

Third, we classify each article as positive if it contains more sentences classified as 'increase' than 'decrease'. It is considered neutral if it has an equal number of both types of sentences, and negative otherwise.

Fourth, we calculate the daily values of the sentiment indicators as the difference between the number of positive and negative news articles compared to the total news articles that contain the name of the country. Finally, we standardize the daily indicators considering the pre-COVID period so that the volatility of each series is comparable across newspapers. Next, we calculate the (simple) average news sentiment[2] by country and then the (simple) average news sentiment for the Big-4 euro area countries as a proxy for the euro area. All the dictionaries used to build these indices are listed in Appendix B.

This procedure is repeated for each topic, using slightly different topic-specific dictionaries, resulting in these three indicators: Economic News Sentiment (ENS), Inflation News Sentiment (INS), and Labour News Sentiment (LNS). We differentiate between 'employment' and 'unemployment' keywords for the LNS index. The directional terms are the same, but the directionality of the sentence is inverted when 'unemployment' is found instead of 'employment', such that 'employment increase' and 'unemployment decrease' would both be detected as 'increase' sentences in the LNS calculations.

3 Sentiment and the business cycle

This section presents our news sentiment indicators and compares their trends with survey-based sentiment indicators and macroeconomic aggregates for validation purposes.

Starting with the Economic News Sentiment (ENS), Figure 1 displays the ENS alongside two popular surveys and quarterly real GDP growth in the euro area. The monthly frequency surveys are: i) the composite Purchasing Managers' Index for output (pmicomp) published by S&P Global which covers manufacturing and services and is one of the most watched indicators; ii) the Economic Sentiment Index (ecesi) published by the European Commission, which integrates sentiment in manufacturing, construction, retail, and services sectors and by con-

[2] Both taking the simple average and weighted averages based on the relative size of the countries report similar results.

sumers. The high co-movement between our textual sentiment measure and survey-based sentiment is evident. Moreover, all sentiment measures are highly and positively correlated with GDP. These findings align with the results of other studies (for example, Ashwin et al. (2024); Aguilar et al. (2021); Aprigliano et al. (2023)).

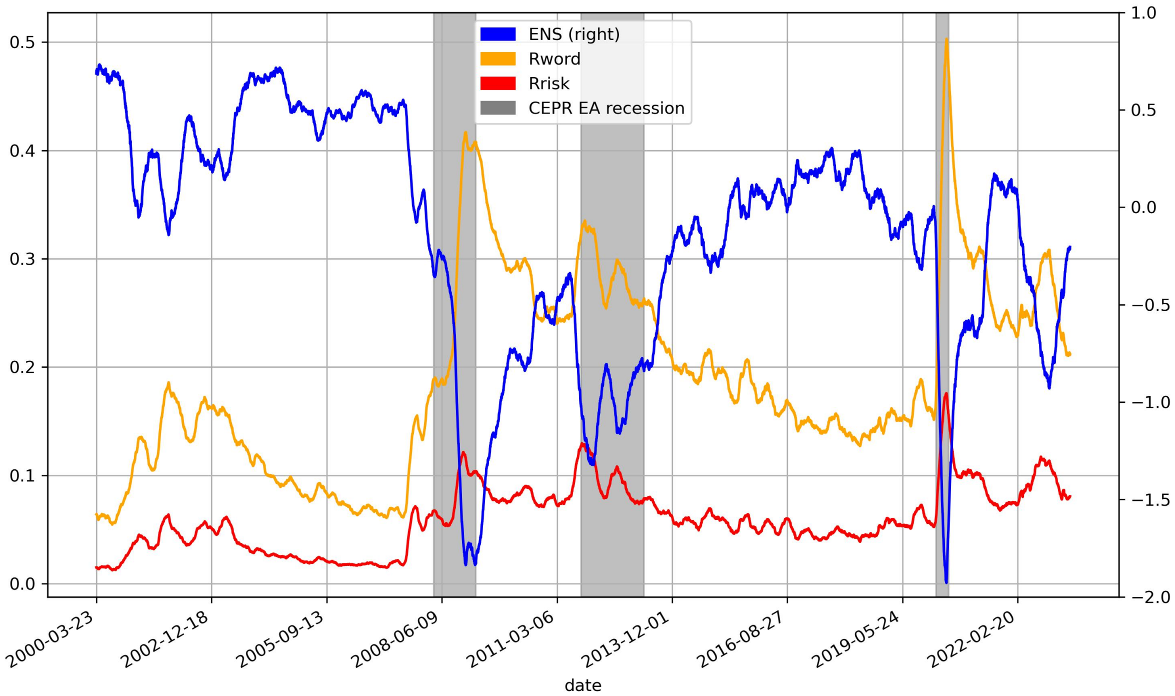
Figure 1: News and survey sentiments vs GDP growth

Notes: All the variables are standardized and for reasons of scale the pre- and post-COVID periods are shown and standardized separately.

Figure 2 shows the daily ENS together with the Rword and Rrisk indices and the official CEPR recession dates (grey areas) for the euro area. An increase in the Rword or Rrisk means that a greater share of economic articles contains the words recession, or recession and uncertainty or risk. In contrast, an increase in the ENS signals an improvement in the economic sentiment. Both recession indices are clearly countercyclical, while economic news sentiment is procyclical. The three news-based indicators displayed very negative signals before both the financial and sovereign debt crises. Although a negative signal was also observed before the onset of the COVID-19 crisis, it was not as pronounced as in the previous two crises. This may be attributed to the unforeseeable nature of the COVID-19 pandemic, as little was known about COVID or its potential consequences until it became a significant issue.

These findings suggest that negative signals provided by news-based sentiment measures can offer valuable insights into broader economic trends and serve as early warning signals for economic crises.

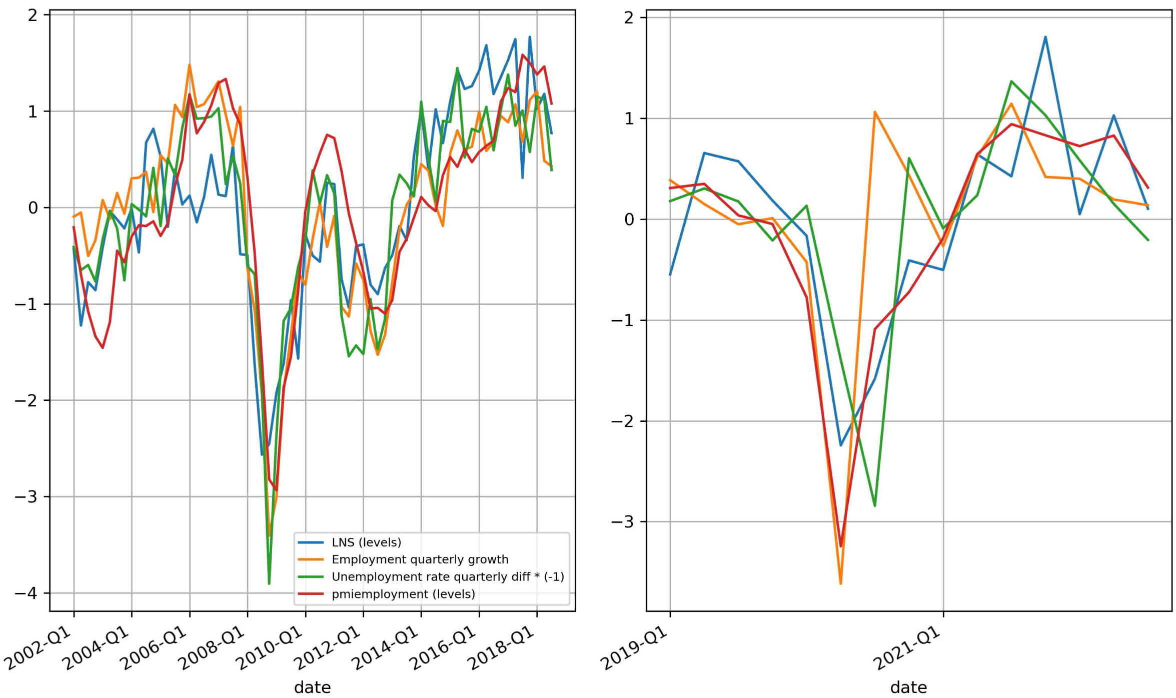
Figure 2: Euro area news indices (90-day moving average)



Regarding the labour news sentiment (LNS), Figure 3 shows the high co-movement with survey sentiment (proxied by the composite Purchasing Managers' Index for employment) as well as with quarterly employment growth and unemployment rate changes.

Before delving into the forecasting exercises, we can better understand how to use these indicators with Figure 4. Given that there is no evidence to claim that newspaper sentiment is stationary in levels (i.e. we cannot statistically reject the null hypothesis of a unit root), the higher correlation of their levels with GDP growth might be spurious due to newspaper sentiment persistence. However, when we examine the correlation between quarterly GDP growth and the quarterly growth of newspaper sentiment, we find that quarterly GDP growth has the highest correlation with a one-quarter lag of newspaper sentiment, and the correlation decays rapidly when we look at two and three-quarter lags. From these results, we can infer that newspaper sentiment has good properties for short-term forecasts but not long-term ones. As shown in the following sections, quarterly differences in newspaper sentiment are more effective than their levels for forecasting one-quarter ahead of GDP growth, and the levels are better for nowcasting whether the state of the economy is a recession or an expansionary period.

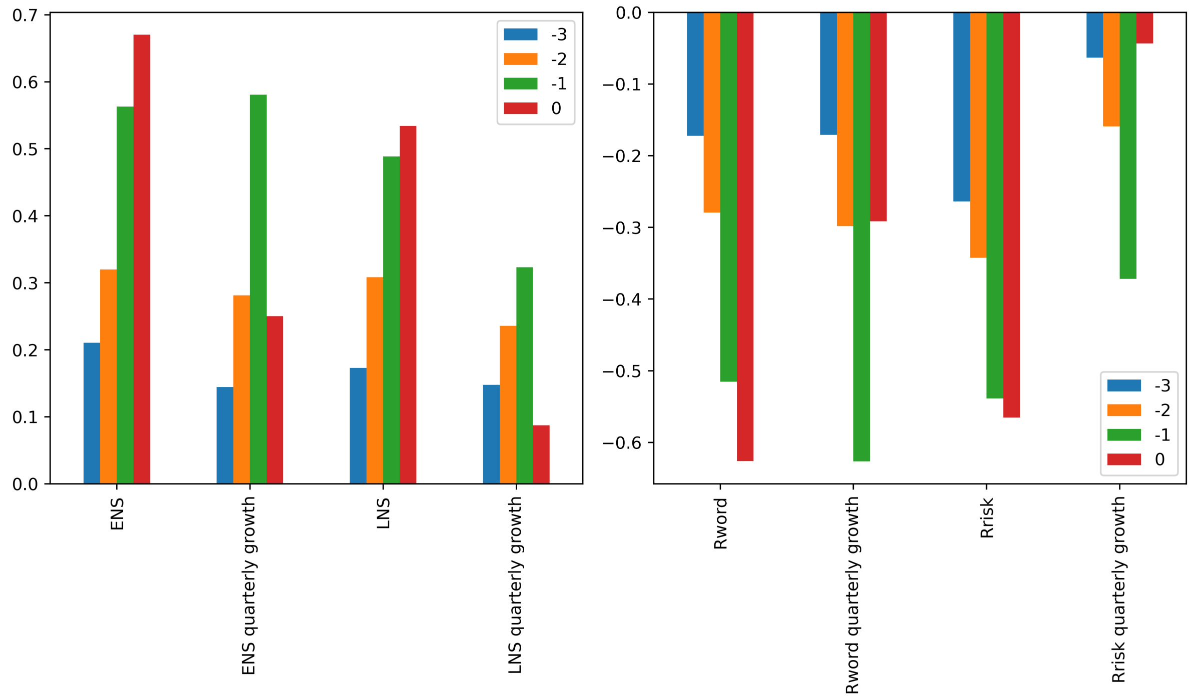
Figure 3: News and survey sentiments vs Employment and Unemployment



Notes: The sign of the quarterly difference of the unemployment rate is inverted to facilitate its comparison with the other indices. All the variables are standardized, and for reasons of scale, the pre- and post-COVID periods are shown and standardized separately.

Recently, de Bandt et al. (2023) and Angelico et al. (2022) found that textual indicators can explain inflation and inflation expectations in France and Italy. In line with these studies, we compare the euro area inflation news sentiment with actual inflation and household inflation expectations in Figure 5. Inflation News Sentiment (INS), household inflation expectations 12 months ahead from the European Commission consumers' survey, and inflation (measured as year-on-year growth of the Harmonized Index of Consumer Prices (HICP)) exhibit similar trends. In fact, we can observe that the INS anticipates changes in the inflation trend. During the Great Recession, the INS reached its trough in 2008-12 and then its trend shifted up, a behavior that was mirrored by actual inflation one year later. The last downward shift signaling the end of the high inflationary period experienced in 2022 is very similar for INS and inflation, but INS reached the turning point in 2022-07 and actual inflation in 2023-03. However, the relationship between INS and household inflation expectations is stronger contem-

Figure 4: Correlation between GDP quarterly growth and news indicators

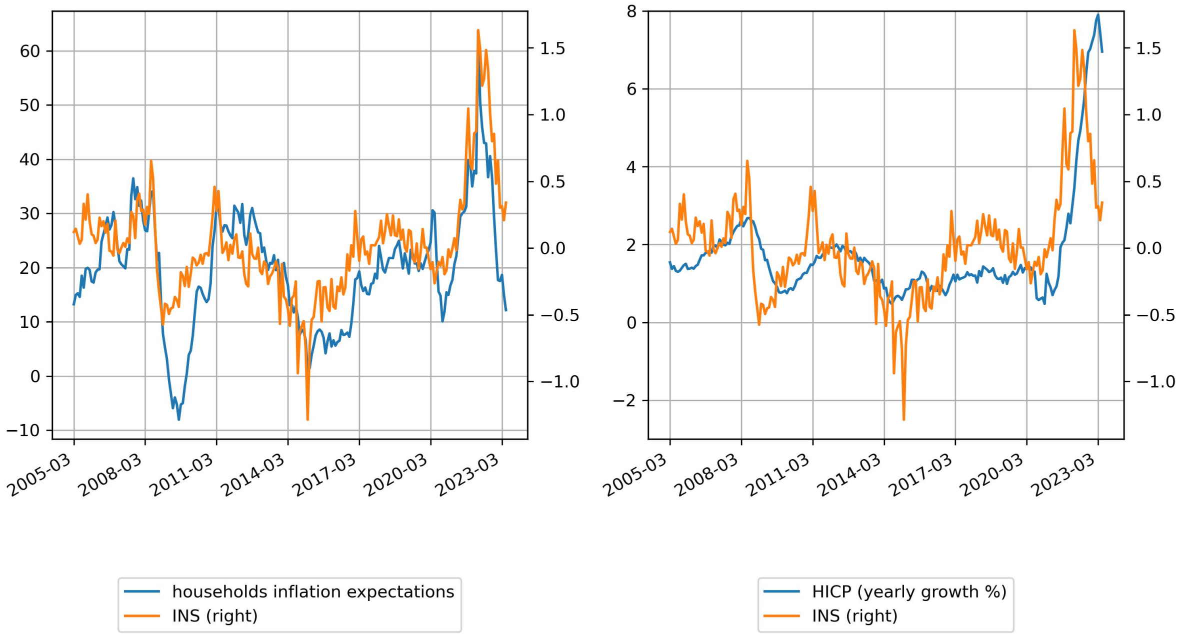


Note: The correlations are between GDP quarterly growth at time t and each news indicator at time t + lag. As indicated in the legend, only 0, 1, 2 and 3 lags are considered. The news indicators are considered both in levels and first differences or quarterly growth.

poraneously. The strong correlation (0.76) between household inflation expectations and news sentiment may suggest that a large proportion of households frequently update their expectations with those of the news. This is consistent with Carroll (2003)'s theoretical framework, which states that individuals are likely to update their inflation expectations with those of the news every period, and if they are not updated, their previous expectations obtained from the latest news that they read remain. In Section 4.3, we econometrically show that Inflation News Sentiment is a meaningful variable to explain inflation and household expectations, surpassing Consensus Inflation Forecasts in terms of explanatory power when we include the latest high inflationary period in the analysis.

Finally, examining Granger causality offers valuable insights into the predictive relationships between economic variables and newspaper sentiment, shedding light on how sentiment indicators may serve as leading indicators. Evidence from several bivariate VARs with different combinations of quarterly real GDP growth and newspaper sentiment indices indicates that GDP is Granger-caused by ENS, LNS, Rword and Rrisk at a 1% significance level during the

Figure 5: Inflation, News Sentiment and Household Expectations



pre-COVID period. This finding is in line with the results of Zheng et al. (2024), who found that the Chinese GDP was Granger-caused by most of the textual sentiment indicators they built for China. Also with Ellingsen et al. (2022), who found that it is more common that the news-based data Granger causes hard indicators in the US. Furthermore, the quarterly HICP growth and household inflation expectations in the euro area are Granger-caused by Inflation News Sentiment quarterly growth during the period 1999-2023 at a 5% significance level. These results also suggest that newspaper sentiment exhibits leading properties that will be studied in the following sections.

4 Are news useful to predict the euro area economy?

In this section, we employ econometric and machine learning techniques to demonstrate that the indicators presented in the previous sections are useful for nowcasting and forecasting exercises, and surpass other traditional macroeconomic indicators in terms of short-term forecasting accuracy.

4.1 Forecasting quarterly GDP growth

In this section, we investigate the predictive performance of our newspaper-based indicators in forecasting real GDP growth. The focus is on the pre-COVID period, distinguishing between recessionary and expansionary periods. The forecast accuracy of these textual indicators is compared with macroeconomic indicators traditionally used for economic forecasting.

As the modeling framework, we utilized Vector Autoregression (VAR) models that included only one lag of GDP growth along with the following indicators: industrial production (index IPI), retail sales (RS), composite purchasing managers' output index (PMI), economic sentiment indicator (ESI), recession risk (Rrisk), recession word (Rword), labour news sentiment (LNS), inflation news sentiment (INS) and economic news sentiment (ENS). The VARs are estimated using an expanding window approach for the period 2006-2024. The focus of the analysis is forecasting GDP growth one quarter ahead.

Table 1: Bivariate VAR: GDP forecast Benchmark

MCS pval	RMSE 04-19	RMSE 06-14	RMSE 14-19	RMSE 20-21	RMSE 22-24	Variable Selection
1.00	0.51	0.68	0.26	10.69	0.30	'GDP', 'Rword'
0.62	0.54	0.72	0.24	10.51	0.32	'GDP', 'ENS'
0.62	0.58	0.78	0.26	10.69	0.33	'GDP', 'ESI'
0.62	0.59	0.80	0.24	10.73	0.30	'GDP', 'Rrisk'
0.62	0.61	0.84	0.23	10.79	0.35	'GDP', 'LNS'
0.62	0.62	0.83	0.27	10.42	0.25	'GDP', 'PMI'
0.62	0.62	0.85	0.25	11.07	0.34	'GDP', 'LNS level'
0.62	0.64	0.87	0.25	11.98	0.24	'GDP', 'PMI level'
0.62	0.64	0.87	0.25	10.68	8.43	'GDP', 'INS'
0.62	0.65	0.88	0.26	10.63	0.26	'GDP' (AR model)
0.57	0.65	0.88	0.25	11.48	0.24	'GDP', 'ESI level'
0.38	0.66	0.89	0.26	10.36	0.26	'GDP', 'RS'
0.62	0.66	0.90	0.29	11.66	0.25	'GDP', 'IPI'

Notes: 'Variable level' indicates the variable in levels was introduced in the model, otherwise, the variable is in quarterly growth. To avoid parameter instability, the years 2020–2021 were excluded from the estimation when generating forecasts for 2022–2024. See Table 13 for data description. First column shows the p-values for the Model Confidence Set (MCS) test (Hansen et al. (2011), https://github.com/JLDC/model-confidence-set). A high p-value indicates small relative error. For a significance level below the p-value, the model would be excluded from the MCS. The period 2004-2019 was used for the MCS test.

To evaluate the accuracy of the forecasting models, we focus on the Root Mean Squared Error (RMSE) and Model Confidence Set (MCS) tests. Tables 1 and 2 show the RMSE of bivariate and multivariate VAR models over different periods: the 2006-2014 period which included both the financial and sovereign debt crises; the 2014-2019 period, representing a rela-

Table 2: Multivariate VAR: GDP forecast Benchmark

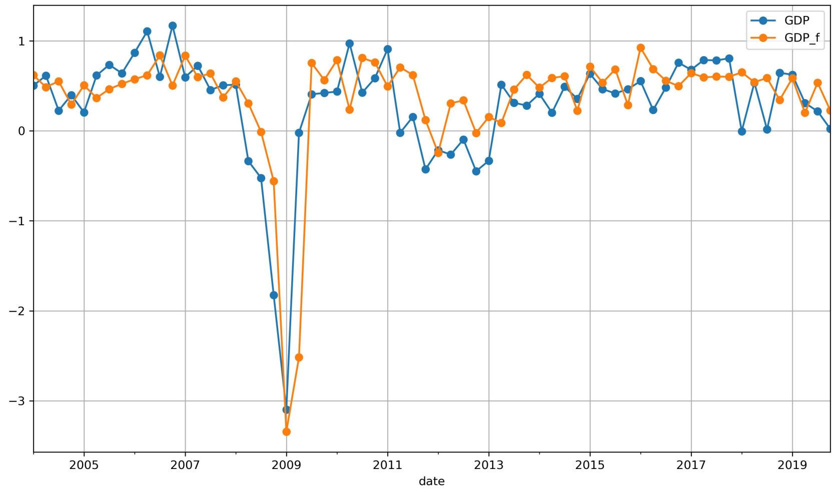
MCS pval	RMSE 04-19	RMSE 06-14	RMSE 14-19	RMSE 20-21	RMSE 22-24	Variable Selection
1	0.47	0.61	0.28	11.27	0.37	'GDP', 'Rword', 'Rrisk', 'LNS level'
0.56	0.48	0.63	0.29	10.66	0.28	'GDP', 'Rword', 'Rrisk', 'ENS'
0.56	0.50	0.65	0.28	9.98	0.33	'GDP', 'Rword', 'Rrisk', 'ENS', 'PMI', 'ESI'
0.45	0.53	0.69	0.29	11.59	0.30	'GDP', 'Rword', 'Rrisk', 'ENS', 'IPI', 'RS'
0.56	0.59	0.79	0.26	10.17	0.38	'GDP', 'PMI', 'ESI'
0.45	0.59	0.78	0.28	11.13	0.37	'GDP', 'Rword', 'Rrisk', 'ENS', 'IPI', 'RS', 'PMI', 'ESI'
0.45	0.64	0.86	0.27	11.57	0.27	'GDP', 'PMI level', 'ESI level'
0.45	0.64	0.86	0.25	11.66	0.44	'GDP', 'IPI', 'RS', 'PMI', 'ESI'
	0.65	0.88	0.26	10.63	0.26	'GDP' (AR model)
0.45	0.67	0.91	0.29	11.49	0.25	'GDP', 'IPI', 'RS'

Notes: 'Variable level' indicates the variable in levels was introduced in the model, otherwise, the variable is in quarterly growth. To avoid parameter instability, the years 2020–2021 were excluded from the estimation when generating forecasts for 2022–2024. See Table 13 for data description. First column shows the p-values for the Model Confidence Set (MCS) test (Hansen et al. (2011), https://github.com/JLDC/model-confidence-set). A high p-value indicates small relative error. For a significance level below the p-value, the model would be excluded from the MCS. This MCS test was performed only for multivariate VAR models (hence, the AR model is excluded from the test). The period 2004-2019 was used for the MCS test.

tively stable expansionary period; the COVID-19 pandemic period, covering the years 2020-2021; and the post-COVID period. In Table 1 the bivariate VARs using textual indicators and GDP had the lowest RMSE during the pre-COVID period. Table 2 shows that multivariate VAR models including GDP and news sentiment indices had the lowest RMSE during the pre-COVID period (Figure 6 shows the one-quarter ahead GDP forecasts of this model). The gains in forecast accuracy when including Rword and Rrisk are more noticeable during recessions, while forecasts using ENS and LNS as predictors perform better during stable periods. Overall, models including newspaper sentiment indexes have the lowest forecast errors for 2004-2019 and the highest relative p-values from the MCS test (MCS, Hansen et al. (2011)), which indicates greater relative forecasting performance.

We also estimated a LASSO regression for quarter-on-quarter GDP using the same indicators as in the previous VAR exercise. We include GDP at t+1 and the indicators at t to determine which indicator is more informative in explaining GDP growth in the next quarter. Figure 7 shows the values of the coefficients of each indicator for different values of the penalization parameter (alpha). As alpha increases, the LASSO becomes more restrictive, and the shrinkage of the parameters to zero strengthens. We find that the Rword and the first lag of GDP are the regressors whose coefficients survive stronger penalization levels of the LASSO. These results suggest that Rword was the most informative variable to explain future GDP growth values, which is consistent with the results found in Table 1. Both of these results indicate

Figure 6: Best VAR model for 1-quarter ahead GDP forecast

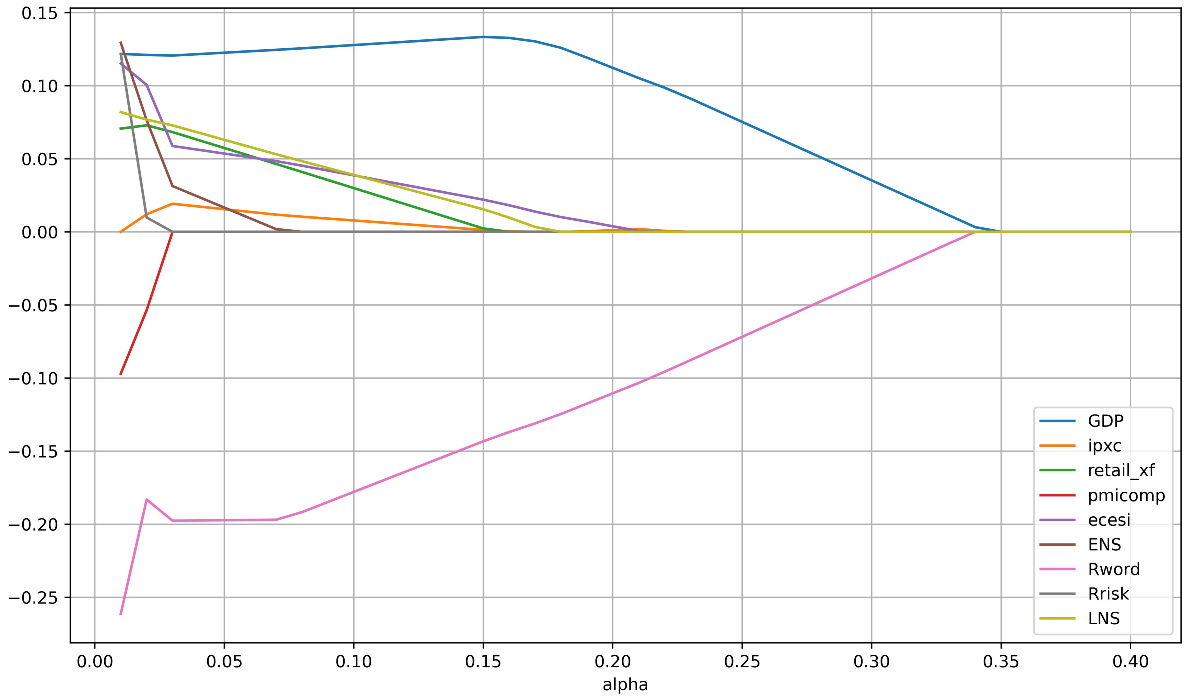


that newspaper sentiment exhibits leading properties. In addition, the fact that Rword, the simplest textual index, outperforms other more complex textual indicators, such as ENS and LNS, suggests that increasing the complexity of the indicators does not always lead to better forecasting properties.

To evaluate the forecasting importance of newspaper sentiment over time, we adapted Koop & Korobilis (2012)'s dynamic model averaging (DMA) for GDP. This model operates within a Bayesian framework, estimating time-varying probabilities of inclusion for multiple bivariate models and employing these probabilities as dynamic weights. Each bivariate model comprises the target variable (GDP) and a single predictor, allowing the time-varying inclusion probability of each model to serve as an indicator of the forecasting importance of the corresponding predictor. Through this approach, we analyze the evolution of the weights assigned to each forecasting model over time, those associated with newspaper sentiment, and benchmark these probabilities against those attributed to other macroeconomic indicators. This comparative analysis provides information on the temporal relevance of newspaper sentiment in relation to traditional economic predictors.

Figure 8 shows the evolution of the weights for the GDP forecasting models. Each line represents the weight associated with a forecasting model using the first log differences of GDP and one of the indicators. The European Commission's economic sentiment indicator (ecesi) was the most important indicator before the financial crisis. However, after the model learned

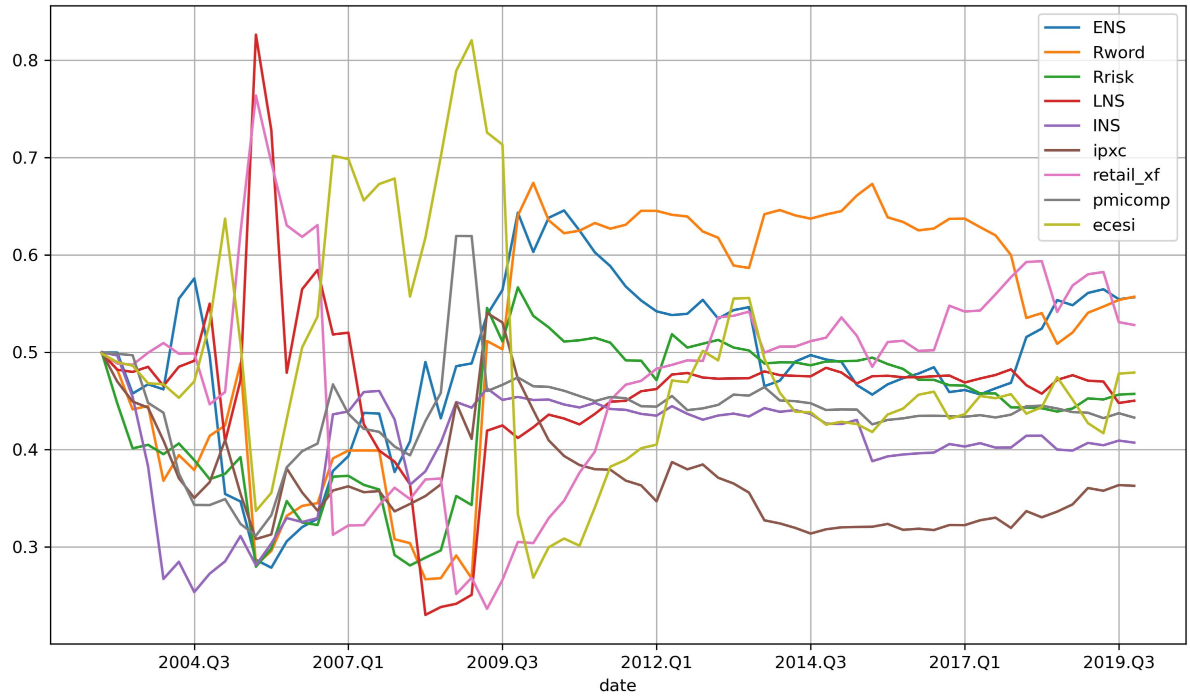
Figure 7: LASSO regression coefficients with an increasing penalization parameter



from the financial crisis, the recession word (Rword) and Economic News Sentiment indicator (ENS) became two of the most important GDP predictors. This result is consistent with Table 1, which showed that forecasting models with newspaper sentiment performed particularly better during the 2006-2014 period. These results are in line with the results of Ellingsen et al. (2022) for the US economy. They found that news-based predictors consistently get a higher weight than models containing hard economic indicators when optimally combining forecasts recursively.

In summary, this section shows that the newspaper sentiment indicators have good forecasting properties to track the GDP of the euro area one quarter ahead. Our results are in line with those of many other authors who found news to be relevant for forecasting the business cycle in many different countries (Ashwin et al., 2024; Aprigliano et al., 2023; Aguilar et al., 2021; Shapiro et al., 2022; Algaba et al., 2023; Barbaglia et al., 2022, 2024; Thorsrud, 2016; Bybee et al., 2024; Fraiberger, 2016; Rambaccussing & Kwiatkowski, 2020; Kalamara et al., 2022).

Figure 8: Importance of indicators for GDP forecasting over time: Estimated weights using Dynamic Model Averaging



4.2 Nowcasting quarterly GDP growth

Gilbert et al. (2017) emphasized the critical role of timing in macroeconomic news announcements, including the release of new data and revisions to existing macroeconomic indicators. Given that news sentiment data can be updated daily, this study investigates its capacity to nowcast quarterly GDP growth using the most up-to-date information available throughout the quarter.

Following Ashwin et al. (2024), we compare the GDP nowcasts derived from an OLS linear regression model that incorporates both text metrics and the PMI indicator to the model tha includes only the PMI indicator. Moreover, we also compare these nowcasts with the official GDP nowcasts from the ECB macroeconomic projections. The nowcasting models use the indicators available on each day d of quarter q to produce daily nowcasts of GDP growth in quarter q. To generate the nowcast for any given day d within the quarter, we used the most recent PMI data available on that day, in conjunction with the cumulative average of the news sentiment indices from the start of the quarter up to day d[3]. Starting as of 1 April 2006, the text and benchmark models are re-trained each day over an expanding window. In this exercise, we ensure that for a given day, data vintages available in real-time are used. In addition to comparing with the ECB nowcasts and the PMI model, we also compared our textual indices, which do not require translation, with the indices produced in Ashwin et al. (2024), which were produced with the translated articles.

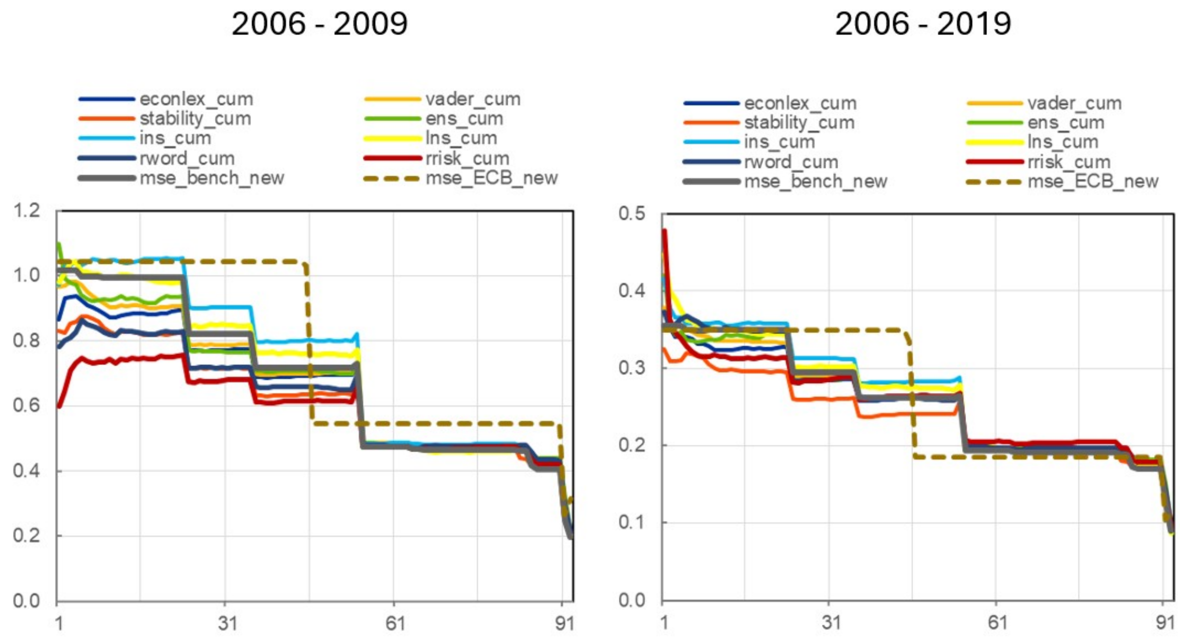
We conducted this experiment in three distinct periods: first, during the financial crisis; second, in the pre-COVID period; and finally, during the COVID-19 pandemic crisis. Although both survey and news sentiments proved helpful for nowcasting GDP in the pre-COVID period, neither the PMI model nor the text-based sentiment models outperformed the ECB nowcasts during the COVID-19 pandemic crisis. Two primary factors may explain this underperformance.

First, the relative volatility between macroeconomic indicators and GDP during the COVID-19 pandemic crisis differed significantly from the pre-COVID period. The unprecedented scale of the economic contraction during the COVID-19 pandemic provided no historical precedent for model training, resulting in poor forecast performance across all approaches. Second, survey and news sentiment indices generally have a more limited range compared to GDP, which can experience much larger fluctuations. This limitation becomes particularly relevant in crises of unprecedented magnitude, such as the COVID-19 pandemic, where sentiment measures may capture turning points but fail to fully reflect the scale of extreme GDP contractions. However, in more typical downturns, such as those observed in the pre-COVID period, sentiment indices can still provide valuable insights into the magnitude of economic declines.

Figure 9 presents the results of this experiment. Consistent with Ashwin et al. (2024), our findings indicate that newspaper sentiment is particularly valuable for nowcasting GDP in the first half of the quarter, when conventional macroeconomic indicators are not yet available. Furthermore, the forecast improvement relative to the official ECB nowcasts and the PMI model is more pronounced when focusing exclusively on the financial crisis period compared to the broader 2006–2019 sample. This suggests that the predictive advantage of newspaper-based nowcasts is primarily driven by recessionary periods, aligning with our previous fore-

[3]We also tested using moving averages and several rolling window sizes, obtaining similar results.

Figure 9: Nowcasting GDP: Mean Squared Error (MSE) across days of the quarter



Notes: We use ECB nowcasts (mse_ECB_new) and nowcasts from a model incorporating real-time vintages of PMI (mse_bench_new) as benchmarks. For textual data, we use the monthly cumulative value of the index as a predictor on each day of the quarter. In addition to our indexes (ENS, INS, LSN, Rword, and Rrisk), we also include three textual sentiment indexes from Ashwin et al. (2024): econlex, stability and VADER.

casting results. Furthermore, we find that sentiment indices constructed without article translation perform comparably to those constructed with article translation. In particular, the Rrisk index emerges as the most effective predictor during the financial crisis and the second-best over the entire 2006–2019 period. These results underscore that high-quality nowcasting can be achieved without translating articles, and that increasing the complexity of sentiment indices does not necessarily enhance nowcasting performance, since Rrisk is one of the simplest indices.

4.3 Nowcasting daily recession probabilities

Based on previous results, this section investigates whether our news-based indices can help forecast turning points, particularly in signaling recessions. Although most works focus on predicting recessions at quarterly or monthly frequencies, we study daily probabilities of recession, exploiting our daily news indices.

Ferrari Minesso et al. (2022) proposed to use the R-word index in a probit model to forecast recessions in the United States (US). They found that the R-word index outperforms the yield curve, but adding the yield curve into the model increases the forecast accuracy for horizons of one to twelve months ahead. In this article, we focus on nowcasting daily recession probabilities. We find that the term spread is not a good predictor for the last three crises in the euro area, and newspaper indices outperformed it.

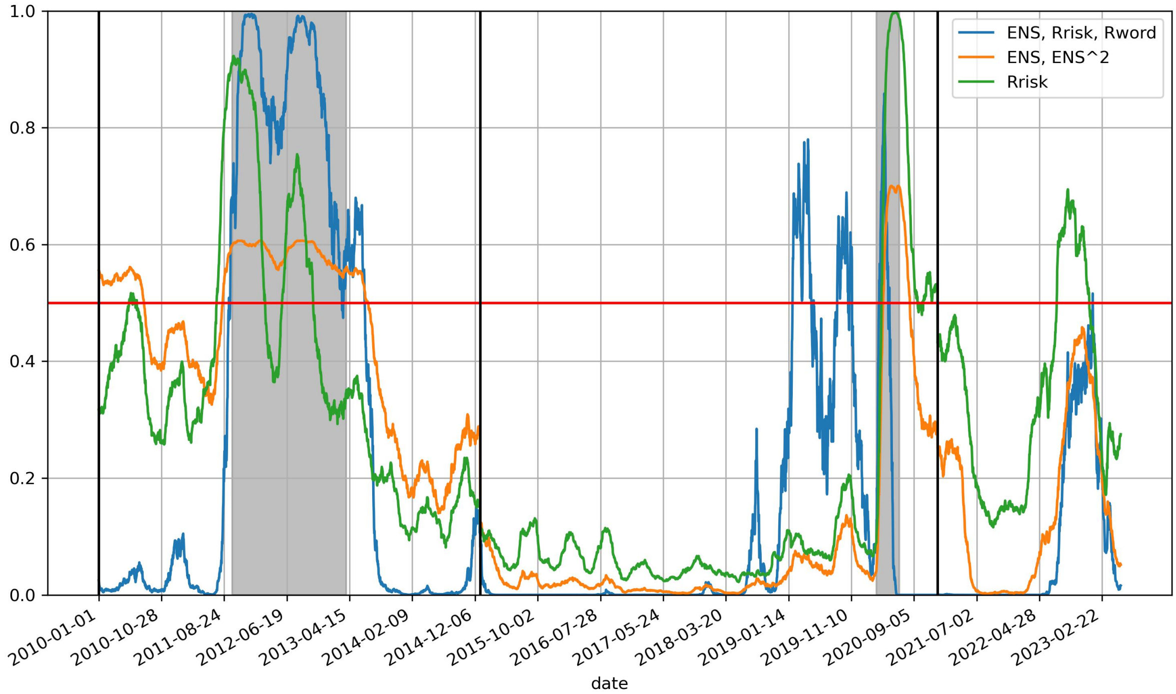
Estrella & Hardouvelis (1991), Chauvet & Potter (2002), Wright (2006) found that simple probit models are more precise than alternative continuous models to predict recessions. Following Ferrari Minesso et al. (2022)'s approach, we estimate several probit models using different combinations of newspaper indices and the term spread. The main difference is that the model is estimated at a daily frequency (instead of monthly) to fully exploit the high frequency of the newspaper indices. Furthermore, these authors used the NBER[2] binary recession indicator built with the US recession dates. Instead, We use the CEPR[3] recession dates for the euro area.

Typically, forecasting models are estimated using monthly or quarterly data to project several periods into the future. Although extending the forecast horizon enables the identification of potential risks before their materialization, it may also lead to a decrease in accuracy. On the other hand, high-frequency data have the advantage of facilitating more precise nowcasting models. Although monthly or quarterly forecasting models can provide an approximation of the timing of a recession, it is only through nowcasting models that are estimated daily, utilizing high-frequency data, that we can obtain exact information about when a recession begins.

To estimate the model, we used 90-day moving averages of our daily sentiment indicators to mitigate the inherent noise of the high-frequency data. We chose this size for the rolling window because it was the most efficient in terms of model accuracy, and it can be interpreted as a quarterly trend that is observable on a daily basis. Figure 10 presents the out-of-sample forecasts of a probit model that estimates the likelihood of being in a recession based on these quarterly trends of our daily sentiment indicators. To avoid overfitting, the model is trained

[2] https://fred.stlouisfed.org/series/USREC
[3] https://eabcn.org/dc/recession-indicators

Figure 10: Daily recession probabilities



Notes: Black vertical lines indicate dates of re-estimation of the parameters. Last observation: 22/05/2023

up to the previous crisis. The financial crisis (2008) was used as the initial training sample, so its out-of-sample accuracy was assessed with the sovereign debt crisis and the COVID-19 recession. The orange and blue lines depict the first and second optimal models derived from the benchmark, listed in Table 3.

The AUC is a widely used metric to evaluate the performance of binary classification models, such as the probit model in this case. It measures the model's ability to distinguish between positive and negative instances, which, in this context, means whether the model can correctly predict the occurrence or non-occurrence of a recession. It ranges from 0.5, which corresponds to a random guessing model, to 1, corresponding to a perfect classification model.

It should be noted that the daily CEPR recession dates were obtained from the original quarterly turning points. This was achieved by assigning recession days to each recession quarter in the CEPR variable, except for the COVID-19 pandemic crisis. In this case, we dated the beginning of the crisis to mid-March, which marked the onset of the COVID-19 pandemic and the implementation of lockdowns throughout Europe.

Table 3 presents various specifications for the model from which several conclusions can be drawn.

First, the term spread or the slope of the yield curve (i.e. the difference between long- and short-term bond yields) was ineffective in nowcasting recent crises. It is uninformative in univariate models and does not improve the performance of models that only include textual information as explanatory variables. The term 'spread' has traditionally been considered a key indicator for predicting recessions. For example, Kessel (1971), Fama (1986), Estrella & Hardouvelis (1991), Estrella & Mishkin (1996, 1998), Chauvet & Potter (2002, 2005), Estrella et al. (2003), Duarte et al. (2005) and Benzoni et al. (2018) found a strong relationship between the slope of the yield curve and economic recessions. However, the central banks' asset purchase programs distorted the relationship between the slope of the yield curve and the real economy, reducing its usefulness as a recession predictor.[4]

Second, textual indices have proven beneficial for nowcasting recessions. Some forecast accuracy gains are achieved by combining multiple newspaper sentiment indices and their non-linear effects. More specifically, the best model comprises the Economic News Sentiment (ENS) and its square in the set of explanatory variables. It achieves an AUC of 0.93, an excellent score for this statistic. The estimated parameters linked to ENS and its squared value are both negative and very significant, indicating that an increase in newspaper confidence in the economy decreases the likelihood of being in a recession, and this effect magnifies for extreme values. The significant gain of adding ENS squared values as an explanatory variable in the probit indicates that the likelihood of being in a recession is more affected by newspaper sentiment when the Economic News Sentiment (ENS) index oscillates among the tail values of its distribution. This result is consistent with the findings of Labonne & Thorsrud (2023) and Adämmer et al. (2024), which indicate that newspapers are particularly informative about the lower left tail of the GDP distribution and deliver significantly better out-of-sample density forecasts than commonly used alternatives.

In Figure 10, it is evident that the model can completely capture both the sovereign debt crisis and the COVID-19 pandemic crisis. For the sovereign debt crisis, the chart shows that the probability of being in a recession is elevated before the actual recession starts. This finding

[4] See: Chauvet & Potter (2002, 2005); Christiansen (2013); Bauer & Mertens (2018a,b); Engstrom & Sharpe (2019) and Gräb & Tizck (2020)

Table 3: Daily recession probability model benchmark

Variable selection	AUC
ENS + ENS^2	0.93
Rrisk + Rword + ENS	0.92
Rrisk + ENS	0.78
Rword + Rrisk	0.77
Rrisk	0.75
Rrisk + $Rrisk^2$	0.75
ENS	0.60
ENS + ENS^2 + Term Spread	0.59
Rword	0.54
Term Spread	0.5

Note: This table shows the area under the ROC curve statistic for several probit models with different variable selection. It ranges from 0.5, which corresponds to a random guessing model, to 1, corresponding to a perfect classification model.

demonstrates the potential of these indicators to provide an early signal before the crisis hits. Interestingly, the probability of recession remains high in the aftermath of the sovereign debt crisis, consistent with the observed 'stickiness' of the R-word index, as discussed by Doms & Morin (2004). Hence, these indicators are probably helpful to accurately identify the beginning of a recession but not the end. Doms & Morin (2004) showed that consumers tend to update their expectations about the economy more frequently during periods of high news coverage compared to low news coverage. High news coverage of the economy is primarily concentrated during recessions and their aftermath, indicating a countercyclical 'stickiness' in expectations.

For instance, right after the end of a recession, a significant number of articles may still refer to the recession, not because it is happening but because it is just ending. This leads to a 'stickiness' in the newspaper sentiment, mainly concentrated right after crisis periods.

It is worth noting that in 2019, two false signals of a recession were observed. At that time, the future of the euro area economy was uncertain, and various challenges were posed by geopolitical risks, trade tensions, and the ongoing Brexit process, all of which raised concerns in the newspapers. We could also observe a false signal at the end of 2022, when the sentiment fell due to the Russia-Ukraine war, which triggered a negative energy supply shock and a surge in inflation. Although no recession occurred, GDP growth stagnated for several consecutive quarters. Sometimes, newspapers may be incorrect about impending recessions or confuse recessions with periods of stagnation. However, in these cases, the false signals have a lower magnitude and do not last as long as the true signals preceding a recession. In the case of the best performing probit model, none of these false signals mentioned surpassed the 50% likelihood of being in a recession.

4.4 Do inflation news sentiment explain actual inflation and household expectations?

In previous sections, we focused on the relationship between news sentiment and economic activity. This section examines whether news sentiment is a useful predictor of inflation and household inflation expectations.

We investigated the ability of news sentiment to predict quarter-on-quarter HICP growth rates in an out-of-sample forecasting exercise with a VAR model. The choice of parsimonious models reflects the study's focus on assessing the core characteristics of news sentiment indicators rather than identifying the optimal modeling framework. Consequently, traditional Ordinary Least Squares (OLS) regression is used to examine the coincident relationship, while Vector Autoregression (VAR) models are utilized to assess leading behavior.

For household inflation expectations, we follow Angelico et al. (2022) and de Bandt et al. (2023), who study whether inflation textual indicators have some additional explanatory power to explain Italian (Angelico et al., 2022) and French (de Bandt et al., 2023) household inflation expectations compared to standard variables in the literature (inflation-linked swaps and professional forecasts). We perform this analysis for households' inflation expectations in the euro area. We follow de Bandt et al. (2023), adding oil prices as an explanatory variable and cleaning the other explanatory variables from oil price effects before including them in the model. Since oil prices are strongly correlated with the other explanatory variables, introducing a multicollinearity problem, which is addressed by regressing the other explanatory variables on oil prices and using the residuals from these regressions, along with oil prices, in the regression.

For inflation, similar to de Bandt et al. (2023), we study whether inflation news sentiment (INS) has additional explanatory power compared to survey-based and professional forecast measures using the simplest form of the hybrid Philips curve (Galı & Gertler, 1999).

We find that INS is significant for explaining households' inflation expectations and quarterly HICP growth, both when considering a low inflation volatility period (2005-2019) and when including the latest high inflation period (up to 2023Q1).

During the 2000-2019 period, euro area inflation only moved between 0.5 and 3% of HICP year-on-year growth; but in 2022-Q4, euro area inflation reached a 20-year high of nearly 10%. In the first period, there was little variation in euro area inflation, with an AR(1) being hard to beat. However, when we include the last high inflationary period in the sample of analysis, we find that Inflation News Sentiment becomes crucial to forecast inflation (see Table 6 and Figure 12).

4.4.1 Inflation Expectations

To study the importance of inflation news sentiment in explaining household inflation expectations, we run several versions of de Bandt et al. (2023) households' inflation expectation equation:

$$\pi^e_{t,t+12} = \beta_0 + \rho\pi^e_{t-1,t+11} + \beta_1\pi_{t-1} + \beta_2 ILS^{1Y}_t + \beta_3 CF^{1Y}_{t-1} + \beta_4 INS_t + \beta_5 Oil\_prices_t + \epsilon_t \quad (1)$$

Where $\pi^e_{t,t+12}$ is the survey-based households' inflation expectation in month t as published by the European Commission, π_t is HICP year-on-year growth rate, ILS^{1Y}_t is market-based 1-year ahead inflation expectations as extracted from inflation-linked swaps, and oil prices is the yearly growth of crude oil prices, $CF^{1y-ahead}_{t-1}$ is 1-year ahead inflation expectations of professional forecasters in the Consensus Forecast survey, and INS is Inflation news sentiment.

Table 4 shows the regression of equation 1. Comparing columns 2, 3, 4 and 5 we observe that INS is more significant for explaining household inflation expectations when compared to inflation-linked swaps, Consensus Forecasts, and actual past inflation. In addition, its inclusion in the baseline regression (column 1) is the one that further improves the adjusted R2;

and when the three of them are included, only INS is significant at 5%. These results hold when we re-estimate the equation only until 2019 (Table 11). Angelico et al. (2022), with a Twitter inflation sentiment index, and de Bandt et al. (2023), with both news and Twitter sentiment indices, found similar results for Italy and France. Our results show that inflation news sentiment also helps to explain household inflation expectations in a multinational framework.

Table 4: Households' Inflation Expectations Regressions (2005–2023Q1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\pi^e_{t-1,t+11}$ (households)	0.883***	0.889***	0.873***	0.877***	0.879***	0.876***	0.876***
	(0.023)	(0.023)	(0.024)	(0.021)	(0.025)	(0.022)	(0.024)
Δ_{12}oil_prices$_t$	3.810***	3.737***	3.953***	3.846***	4.008***	3.871***	3.963***
	(0.667)	(0.659)	(0.677)	(0.620)	(0.741)	(0.634)	(0.682)
ILS^{1Y}_t (adjusted)		3.639***				1.699	1.104
		(1.363)				(1.361)	(1.421)
π_{t-1} (adjusted)			2.153			0.481	0.774
			(1.305)			(1.292)	(1.368)
INS_t (adjusted)				7.529***		7.003***	7.829***
				(1.236)		(1.315)	(1.386)
CF^{1Y}_{t-1} (adjusted)					-0.674		-1.221*
					(0.760)		(0.726)
Intercept	1.846***	1.753***	2.040***	1.965***	1.914***	1.990***	1.973***
	(0.471)	(0.466)	(0.487)	(0.443)	(0.514)	(0.458)	(0.483)
Obs	223	223	221	219	199	217	195
R^2	0.911	0.914	0.912	0.924	0.913	0.925	0.929
Adj. R^2	0.910	0.913	0.911	0.923	0.912	0.923	0.927
AIC	1160.812	1155.669	1151.541	1107.065	1050.690	1100.738	997.165
BIC	1171.033	1169.297	1165.134	1120.621	1063.863	1121.017	1020.076

Note: Adjusted variables are the residuals from regressions of the respective original variables on oil prices.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.4.2 Inflation

To study the importance of inflation news sentiment in explaining actual inflation, we run several versions of de Bandt et al. (2023) inflation equation:

$$\pi_t = \beta_0 + \rho\pi_{t-1} + \beta_1 Output\_gap_t + \beta_2 \pi^{Exp}_{t+1} + \epsilon_t \tag{2}$$

Where π_t is inflation in t (measured as quarter-on-quarter growth of seasonally adjusted HICP excluding energy), output gap is proxied with the principal component of the cyclical components of the capacity utilization rate, the unemployment rate, the business climate indicator and the industrial production index obtained with the Hodrick-Prescott filter; and π_{t+1}^{Exp} is either the Consensus Forecast expectation for inflation in the year ahead or the Inflation News Sentiment index.

We differ from de Bandt et al. (2023) in the way we account for the effect of oil prices. They include real oil prices into the regressions to account for its effect on HICP, and we use the version of the HICP that excludes energy.

Table 5 shows the regression of equation 2. When we ran the regression until 2023, we found that inflation news sentiment was significant in explaining inflation and added more explanatory power than Consensus Forecasts. However, if we run the regression for 2007-2019 (Table 12), the Consensus Forecast added more explanatory power. Inflation News Sentiment performs better than the Consensus Forecast only if we consider the 2022-2023 high inflation period in the regression. However, for the 2007-2019 period, the Inflation News Sentiment still adds explanatory power compared to the baseline model (column 5 vs 3 Table 12).

Table 5: Inflation (HICP quarter-on-quarter growth rate) (2007–2023Q1)

	(1)	(2)	(3)	(4)	(5)
π_{t-1}	0.965***	0.934***	0.855***	0.615***	0.713***
	(0.066)	(0.068)	(0.145)	(0.115)	(0.080)
$Cycle_{t-1}$		-0.023	-0.022	-0.016	-0.009
		(0.015)	(0.015)	(0.014)	(0.014)
CF_{t-1}^{1Y}			0.162		
			(0.264)		
CF_{t}^{1Y}				0.698***	
				(0.211)	
INS_t					0.333***
					(0.079)
Intercept	0.036	0.049	0.016	-0.105*	0.145***
	(0.038)	(0.039)	(0.067)	(0.059)	(0.041)
Obs	64	64	64	64	64
R^2	0.776	0.784	0.786	0.818	0.833
Adj. R^2	0.772	0.777	0.775	0.809	0.825
AIC	-20.942	-21.448	-19.849	-30.210	-35.972
BIC	-16.624	-14.971	-11.214	-21.575	-27.336

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

From these results, we can conclude that Inflation News Sentiment is more beneficial than Consensus Forecasts to explain inflation for the euro area during turbulent times. For periods of low inflation volatility, Consensus Forecasts are more accurate but less timely explanatory variables for inflation.

Furthermore, we explore the leading properties of the INS index in a VAR setting, comparing its out-of-sample one-quarter ahead inflation forecast performance with Consensus Forecasts for inflation, household expectations, and a simple AR model for inflation. Table 6 shows the results of this exercise.

Although sophisticated inflation forecasting models, which incorporate stochastic volatility and trend-cycle decomposition, have gained widespread adoption (Stock & Watson, 2007), a consensus on the most effective approach remains elusive in the current state of the literature. Given that our primary objective is to assess the predictive capabilities of newspaper sentiment in the euro area, we opted for a more parsimonious framework employing traditional econometric techniques such as standard VAR models.

Table 6: VAR: Quarterly HICP Growth Forecast Benchmark

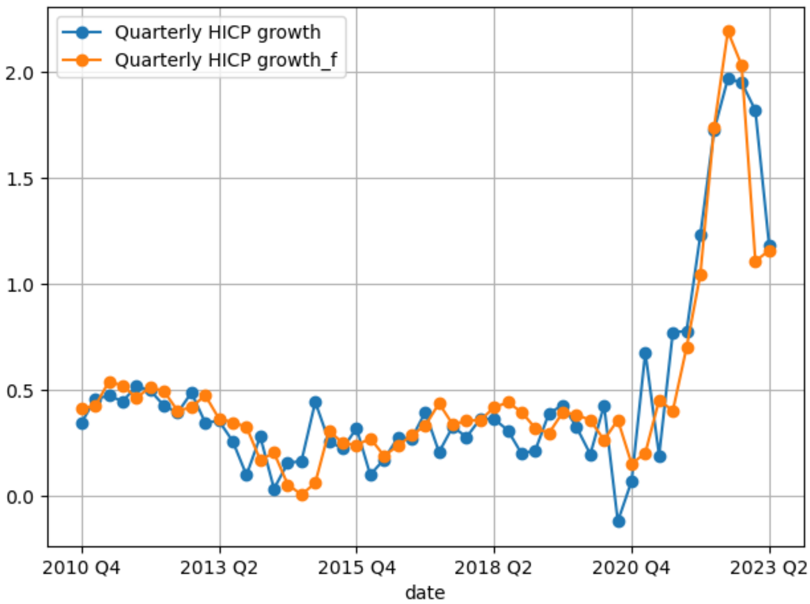
MCS p-val	RMSE 11–23	RMSE 11–19	RMSE 20–23	Variable selection
1	0.19	0.12	0.31	'HICP', 'INS', 'INS^2'
0.46	0.20	0.12	0.32	'HICP', 'CF', 'INS', 'INS^2'
0.46	0.21	0.11	0.36	'HICP', 'INS'
0.46	0.21	0.13	0.35	'HICP', 'CF', 'INS', 'INS^2', 'ILS', 'HIE'
0.46	0.22	0.12	0.37	'HICP', 'CF', 'INS'
0.46	0.23	0.13	0.38	'HICP', 'CF', 'INS', 'ILS', 'HIE'
0.02	0.24	0.11	0.42	'HICP', 'HIE'
0.02	0.25	0.11	0.44	'HICP'
0.02	0.25	0.11	0.44	'HICP', 'ILS'
0.02	0.26	0.12	0.46	'HICP', 'CF'

Note: See Table 14 for variable descriptions. All variables are in levels except for HICP and CF, which are in quarterly growth rates. The first column reports the p-values for the Model Confidence Set (MCS) test (Hansen et al. (2011), https://github.com/JLDC/model-confidence-set). A higher p-value indicates a smaller relative error. For a significance level below the p-value, the model is excluded from the MCS. The period 2020–2023 was used for the MCS test.

Table 6 shows that VAR models with INS have the lowest error for the whole sample. Since inflation did not experience much variation in 2011-2019, a simple AR model remained unbeaten compared to VAR specifications; however, the error of a VAR model containing quarterly HICP growth, INS, and INS squared is 24 % lower than the AR model when we consider the 2020-2023 sample. This VAR model, whose results are shown in Figure 11, would have been a better option than households' inflation expectations or Consensus Forecast to predict the high inflationary period observed in 2020-2023. This finding is in line with Eugster & Uhl (2024), who also find that newspaper sentiment outperforms other indicators in shorter forecast horizons for US inflation. Additionally, incorporating the quadratic term of inflation news sentiment serves as a parsimonious approach to capture the nonlinear dynamics observed during the 2020-2023 period of heightened inflation. This approach provides a simpler alternative to more complex econometric methods, such as quantile regression analysis, for modeling nonlinearities in the relationship between inflation sentiment and economic outcomes. These complex methods were previously used by Adämmer et al. (2024), who also identified the importance of nonlinearities in predicting tail risks. Furthermore, results from the Model Confidence Set (MCS) test indicate that any model without the Inflation News Sentiment index would have been excluded from the model confidence set in the period 2020-2023. This underscores how helpful this indicator would have been in forecasting the high inflationary period caused by the negative energy supply shock that followed the Ukraine-Russia war.

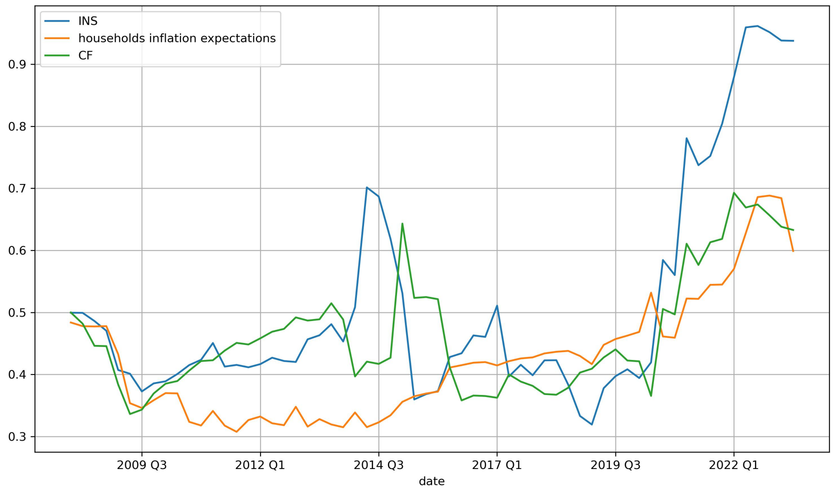
The persistent nature of inflation poses a formidable challenge in forecasting its trajectory. However, our findings indicate that INS (Inflation News Sentiment) can serve as a valuable tool for this purpose, particularly during periods of heightened inflation volatility, as witnessed during 2020-2023.

Figure 11: Best VAR model for 1-quarter ahead inflation forecast



Finally, we follow the Dynamic Model Averaging (DMA) approach by Koop & Korobilis (2012) and estimate the dynamic weights of the HICP forecasting models, including different indicators to track inflation in the next quarter. Figure 12 shows that no indicator was particularly relevant for forecasting inflation in 2009-2019, which is mainly due to two facts: the high persistence of inflation, which makes a simple AR(1) a model hard to beat, and the low volatility of inflation in this period. However, all indicators gained importance after 2020, mainly because inflation exhibited sufficient variation. Inflation news sentiment is the most important indicator for capturing the latest high-inflationary period experienced in the euro area in 2022-2023. These results further highlight the usefulness of newspaper sentiment in forecasting during periods of economic turbulence.

Figure 12: Indicators importance over time. Estimated weights from HICP forecasts with Dynamic Model Averaging.



4.5 Nowcasting survey-based sentiment

Empirical studies suggest that text-based sentiment indicators can serve as leading indicators for survey-based sentiment measures (e.g. Algaba et al. (2023); Shapiro et al. (2022)). The predictive capacity of text-based sentiment indicators stems from their ability to capture real-time economic perceptions across various sources, which reflect evolving economic narratives ahead of traditional surveys. In this section, we investigate whether our news-based sentiment indicators convey information that influences or helps predict survey-based sentiment. The focus is on out-of-sample forecasting and nowcasting monthly euro area Purchasing Managers' index (PMI) for composite output. We use a MIxed DAta Sampling (MIDAS) regressions, which are frequently used for nowcasting due to their ease of implementation and relatively good performance. We estimate MIDAS models with an AR(2) component and 90 lags of the daily sentiment metrics. The lag polynomials are Legendre polynomials of order 2. We have also tested with alternative specifications (e.g. Almon polynomials) and the results are similar. The benchmark model is a simple AR(2). Although it is possible to get PMI forecasts at multiple forecast horizons on a daily basis, we will focus on four particular situations considering the days before the release of the flash PMI: i) 45 days; ii) 30 days; iii) 15 days; and iv) 1 day. Table 7 presents the relative root mean squared forecast error (RMSFE) in relation to AR(2).

Table 7: Forecast accuracy for forecasting PMI: Relative RMSFE with respect to AR(2)

Indicator	Forecast horizon (days)	2005-2023	2005-09	2010-19	2020-23
ENS	45	1.00	1.01	1.02	1.00
	30	0.99	0.94	1.05	0.99
	15	0.97	0.93	1.04	0.97
	1	0.95	0.93	1.04	0.94
INS	45	0.99	0.84	1.11	0.99
	30	1.00	0.85	1.09	1.00
	15	1.01	0.93	1.06	1.01
	1	1.02	0.96	1.03	1.02
LNS	45	1.02	1.01	1.04	1.02
	30	0.99	1.01	1.06	0.98
	15	0.99	1.00	1.03	0.98
	1	1.01	1.01	1.05	1.01

Even though in general the forecast accuracy gains are relatively small, there is some advantage when forecasting PMI during recessions such as for the global financial crisis in the models using ENS or INS at almost all forecast horizons. Still, the fact that the forecast accuracy gains are small suggests that the textual indicators contain different information than the PMI.

5 Robustness checks

To assess the performance of our method, we compare it with other methods typically used in the literature. We consider the following lexicons: LM (Loughran & McDonald (2011), updated in 2014), HL (Hu & Liu (2004)), VADER, and Shapiro's News Lexicon augmented with

VADER and LM dictionaries (Shapiro et al. (2022)). In addition, we include Barbaglia et al. (2024)'s economic index, calculated as the average of French, Spanish, Italian, and German indices as a proxy for the euro area. We then evaluate their capacity to predict next quarter's euro area GDP growth.

Our method is the only one that can be applied to non-English text; hence, the other methods were applied to the corpus of translated articles. To generate the ENS versions of LM, HL, Shapiro and VADER methods, we identify the polarity of the economic sentences in a +/-20 word window around the appearances of the word 'economy' on the news. The only difference with respect to the methodology previously described in Section 2 is that in the economic sentences, instead of checking for the presence of directional terms, we apply the alternative sentiment analysis algorithm directly to these sentences and classify them according to the sign of the sentiment score produced by the corresponding method.

To see the gain of not translating the articles with our method, we also computed an alternative ENS using a version of our method with English dictionaries applied to the translated articles.

Figure 13: Economic News Sentiment: alternative approaches

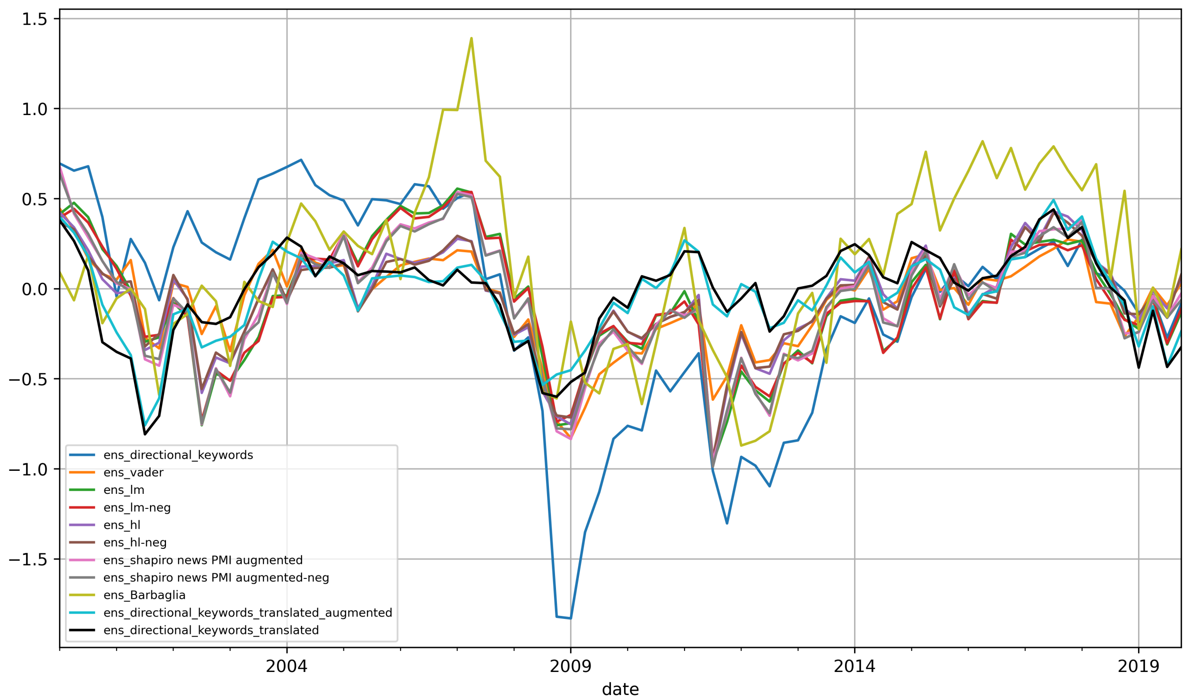


Figure 13 shows all the ENS indicators generated by the different methods. All methods generate similar indices, particularly those applied to English text. Our directional keyword approach applied to the original articles (without translation) reports the most negative signal during the financial and sovereign debt crises.

Table 8 shows the results of out-of-sample forecasting exercise with several VARs in an expanding window, each with GDP and its first lag and one of the news indicators. Taking into account only the algorithms applied to the translated text, Loughran & McDonald (2011) generates the most informative indicator to explain the GDP growth in the next quarter. However, when we apply our directional keyword approach to the articles in their original languages instead of the translated version, our indicator goes from being the second least informative index to the most informative one. This result suggests that there exist gains in applying the sentiment analysis techniques directly in the original language, not only in terms of computational efficiency; but also in terms of better forecast errors.

Table 8: VAR GDP forecast method Benchmark (2000-2019)

RMSE 04-19	**variable selection**
0.538	'GDP', 'ens_directional_keywords'
0.609	'GDP', 'ens_lm-neg'
0.614	'GDP', 'ens_lm'
0.619	'GDP', 'ens_shapiro news PMI augmented-neg'
0.622	'GDP', 'ens_shapiro news PMI augmented'
0.627	'GDP', 'ens_vader'
0.632	'GDP', 'ens_hl-neg'
0.635	'GDP', 'ens_hl'
0.636	GDP', 'ens_directional_keywords_translated'
0.643	GDP', 'Economy index (Barbaglia et al., 2024)'

6 Conclusion

The proliferation of textual data presents a unique opportunity to construct sentiment indicators that enrich traditional macroeconomic and survey-based measures. This study, motivated by the growing need for timely and reliable economic insights following the COVID-19 pandemic, aims to evaluate the value of sentiment indicators gleaned from news articles. Other studies have previously shown that newspaper sentiment is helpful in forecasting the economy of several European countries (Aguilar et al., 2021; Aprigliano et al., 2023; Barbaglia et al., 2024; Ashwin et al., 2024). Following this line of research, we demonstrate in this article that newspaper sentiment has also good properties for tracking inflation and output, and predicting turning points for the euro area as a whole. Furthermore, we demonstrate that translating articles into English, the most common approach in the literature, is not necessary to achieve good-quality indexes in a multilingual framework.

Newspaper sentiment outperforms traditional macroeconomic and survey-based indicators when comparing their out-of-sample one-quarter-ahead GDP forecast performance. Similarly to Ashwin et al. (2024), we also found newspaper sentiment to be helpful in nowcasting GDP, particularly in the first half of the quarter, when conventional macroeconomic indicators are not yet available. In addition, inflation newspaper sentiment tracked household inflation expectations very well. Additionally, it demonstrated forward-looking properties in forecasting inflation during the high-inflationary period (2022-23) experienced in the euro area due to the Russia-Ukraine war and the energy shortage it triggered.

Although many authors have previously used newspaper sentiment to gauge economic growth in several countries, the literature lacks applications that take advantage of the high frequency availability of newspaper data. To address this gap, we demonstrated that a simple probit model, explaining CEPR euro area recession dates as a function of quarterly moving averages of daily newspaper sentiment, can accurately report daily probabilities of recession. The out-of-sample performance of these models reached an area under the ROC curve (AUC) of 0.93 when we evaluated the models with the sovereign debt and COVID-19 pandemic crises. The excellent value of the AUC highlights the utility of newspaper sentiment in recognizing the exact moment when a recession begins, right at the very moment it starts. Although some false recession signals were observed during 2019 and 2022, these were not of the same magnitude and did not last as long as those observed during the actual recessions. In our sample, when newspapers were concerned about a recession, it was most likely that the economy was actually in a recession. However, it is possible that newspapers may incorrectly assess the likelihood of being in a recession. However, even in this case, the tool would still be helpful, as it would warn policymakers about overly negative economic newspaper sentiment.

Last, but not least, the textual analysis methods employed to build the indicators in this article were rather simple. We based our approach on small dictionaries with a limited number of words that can be easily translated into other languages, thereby circumventing the need to translate the articles and reducing the computational cost and time required to produce and update the indices. We also found that our indices, computed from the non-translated articles, had comparable nowcast performance throughout the quarter to the news sentiment indices produced from the translated articles by Ashwin et al. (2024). Also, avoiding machine learning techniques makes the method consistent, as its results are not dependent on any training sample, and makes the algorithms' inner workings easier to understand. The low computational cost of the method eases the automation of index production, making these indicators one of the most timely economic measures available. The simplicity of the techniques used in this paper endows these newspaper sentiment indices with interpretability, transparency, and replicability.

Appendix A: Newspapers

Table 9: News Sources

Country	Newspaper	Total articles
France	**Les Échos**	378,311
	Le Figaro	242,019
	Le Monde	121,558
Germany	**Süddeutsche Zeitung**	348,905
	Die Welt	145,262
	Der Tagesspiegel	70,856
Italy	**Corriere della Sera**	301,737
	La Repubblica	216,566
	Il Sole 24 Ore	546,685
	La Stampa	152,806
Spain	**El Mundo**	139,646
	El País	275,065
	La Vanguardia	72,947
Total		3,012,363

Appendix B: Dictionaries

This section shows the dictionaries used to build the news sentiment indices. First, to create the indicators at the country level, only the articles that contain the name of the country are filtered. Then, we classify every sentence where one of the keywords appears as 'increase'/'decrease' if an upswing/downswing word is within a +/-5 word range of the keyword. Additionally, the strong upswing/downswing words are terms that do not need to go in combination with a keyword to classify the sentence as 'increase' or 'decrease'. Finally, each article is classified as positive or negative, depending on whether it has more 'increase' or 'decrease' sentences. For the case of Labor News Sentiment, we also consider keywords with inverted polarity, which, if found with an upswing/downswing word within the +/-5 word range, are classified as 'decrease'/'increase'. This adjustment is made to accommodate situations where, within the LNS index construction, sentences expressing opposite concepts, such as 'the employment increased' and 'the unemployment decreased', are both assigned positive values.

Inflation News Sentiment (INS)

Spanish:

Country : españ

Keywords : inflac, precio,ipc, iapc, cpi, hicp

Upswing words list : aceler, crec, expansi, increment, aument, sub

Downswing words list : descen, disminu, redu, ralentiz, decrec, desaceler, contracci, caída, cay, cai, baj, mejor

Strong upswing words : estanfla, stagflati

Strong downswing words : defla, desinflac

French:

Country : france

Keywords : inflation , prix , ipc , ipca , iapc , ipch , cpi , hicp

Upswing words list : accélér , croiss , augm , hausse , monte

Downswing words list : diminu , reduc , réduc , baisse , ralent , decéler , décroit , rédui , contract , amélior

Strong upswing words : stagflati

Strong downswing words : déflation , désinflation

Italian:

Country : italia

Keywords : inflaz , prezz , ipc , iapc , ipca , cpi , hicp

Upswing words list : acceler , cresc , espans , aument , increment

Downswing words list : dimin , calo , ridu , decres , scend , miglior

Strong upswing words : stagflaz , stagflat

Strong downswing words : deflaz , disinflaz

German:

Country : deutsch

Keywords : preis , inflat , vpi , hvpi , cpi , hicp

Upswing words list : beschleunigung , wachstum , zunahme , erweitert , wachsen , steigen

Downswing words list : zerbesserung , besser , verbess , verring , ermäß , verlangsa , sinkt , rückgang , abnehmen , schwinden , sink , fallen , abklingen , nachlassen , vermindern , niedr

Strong upswing words : stagflati , preisanstieg , inflationsanstieg , preissteigerung , inflationssteigerung

Strong downswing words : geldentwertung , deflat , desinflat , preissturz , preissenkung , teuerung

Economic News Sentiment (ENS)

Spanish:

Country : españ

Keywords : econom

Upswing words list : aceler , crec , expansi , increment , aument , mejor

Downswing words list : descen , disminu , redu , ralentiz , decrec , desaceler , contracci , negativ

Strong upswing words : recuperac

Strong downswing words : recesi , crisis

French:

Country : france

Keywords : econom , économ

Upswing words list : accélér , croissance , augment , hausse , amélior , expansion

Downswing words list : diminu , reduc , réduc , baisse , ralent , decéler , décroit , rédui , contract

Strong upswing words : repris , récupér

Strong downswing words : récession , crise

Italian:

Country : italia

Keywords : econom

Upswing words list : acceler , cresc , espans , aument , miglior , increment

Downswing words list : dimin , calo , ridu , decres , scend

Strong upswing words : recuper , ripre

Strong downswing words : recess , crisi , rallent , contra , declino

German:

Country : deutsch

Keywords : wirtsch , ökonomisch

Upswing words list : beschleunigung , wachstum , zunahme , zerbesserung , besser , verbess , erweitert , wachsen , steigen

Downswing words list : verring , ermäß , verlangsa , sinkt , rückgang , einbruch , abnehmen , schwinden , sink , fallen , abklingen , nachlassen , vermindern

Strong upswing words : erholung , aufschwung , wiedererlangung , zurückgewinnung , wiederverwertung , wiederbeschaffung , wiedergewinnung , bergung , wiederherstellung , wirtschaftswachstum

Strong downswing words : krise , krisis , rezession , konjunkturrückgang , wendepunkt , wirtschaftskrise , konjunktureinbruch , abschwung

Labor News Sentiment (LNS)

Spanish:

Country : españ

Keywords : empleo , mano de obra , trabaj , contrat

Keywords with inverted polarity : desempleo , paro

Upswing words list : aceler , auge , expand , rápido , gana , alto , mejor , aument , elev , fuert , fortal

Downswing words list : colaps , contrae , deceler , dismin , cae , cay , perder , pierd , pérdida , bajo , moderado , lento , ablandar , dominar , débil , debil

French:

Country : france

Keywords : emploi , travail , emploi , embauch

Keywords with inverted polarity : chômage

Upswing words list : accélérer , boom , dévelop , rapid , gagn , haut , amélior , augment fort , forc , renforc

Downswing words list : effondr , contract , ralent , dimin , autom , perdr , perte , faible , modéré , lent , adouc , maîtriser , affaiblir , faiblesse

Italian:

Country : italia

Keywords : occupazione , lavor , assumere

Keywords with inverted polarity : disoccupazione

Upswing words list : acceler , boom , espand , veloc , guada , alto , miglio , aument , salita , forte , forza , rafforzar

Downswing words list : croll , contra , fredd , deceler , dimin , autun , perd , basso , modera , lent , ammorbi , sottomett , debol , indebol , debol

German:

Country : deutsch

Keywords : anstellung , arbeit , beschäftigung , arbeitsverhältnis

Keywords with inverted polarity : arbeitslosigkeit

Upswing words list : beschleun , besser , verbess , boom , wachs , erweit , steig , expandieren , schnell , gewinnen , hoch , verbessern , zunahm , erheben , stark , stärke , stärken

Downswing words list : zusammenbr , rück , abnehmen , schwinden , sink , abklingen , nachlassen , vermindern , niedr , ermäß , verlangsa , verring , zerb , vertrag , verlangsamen , verringern , fallen , verlieren , verlust , mäßig , langsam , erweichen , unterwerfen , schwach , schwäch

Appendix C: Correlations analysis

Table 10: Correlations of news indicators with GDP quarterly growth at different lags (-) and leads (+). Sample 2002-2019.

	-3	-2	-1	0	1	2	3
GDP quarterly growth	0.17	0.35	0.60	1.00	0.60	0.35	0.17
ENS	0.21	0.32	0.56	0.67	0.68	0.59	0.44
ENS quarterly diff	0.14	0.28	0.58	0.25	-0.01	-0.23	-0.38
LNS	0.17	0.31	0.49	0.53	0.56	0.50	0.41
LNS quarterly diff	0.15	0.24	0.32	0.09	0.03	-0.10	-0.16
Rword	-0.17	-0.28	-0.52	-0.63	-0.63	-0.58	-0.47
Rword quarterly diff	-0.17	-0.30	-0.63	-0.29	0.00	0.17	0.31
Rrisk	-0.26	-0.34	-0.54	-0.57	-0.53	-0.43	-0.30
Rrisk quarterly diff	-0.06	-0.16	-0.37	-0.04	0.08	0.20	0.26

Let the headers of the columns be "i" and the variable of the row be "X". Then, each value of the table is Corr(GDP_t, X_{t+i})

Appendix D: Pre-COVID inflation regressions

Table 11: Household expectation regressions (2005-2019)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\pi^e_{t-1,t+11}(households)$	0.916***	0.923***	0.907***	0.907***	0.905***	0.904***	0.899***
	(0.021)	(0.020)	(0.021)	(0.019)	(0.022)	(0.020)	(0.021)
$\Delta_{12}oil\_prices_t$	3.596***	3.518***	3.674***	3.717***	3.966***	3.751***	3.983***
	(0.647)	(0.634)	(0.649)	(0.603)	(0.735)	(0.610)	(0.671)
$ILS^{1Y}_t(adjusted)$		3.419***				1.679	1.587
		(1.176)				(1.173)	(1.189)
$\pi_{t-1}(adjusted)$			4.420**			2.554	2.390
			(1.733)			(1.665)	(1.755)
$INS_t(adjusted)$				5.994***		5.172***	5.604***
				(1.132)		(1.199)	(1.255)
$CF^{1Y}_{t-1}(adjusted)$					2.660		1.711
					(1.752)		(1.611)
Intercept	1.231***	1.127***	1.406***	1.392***	1.425***	1.442***	1.560***
	(0.392)	(0.385)	(0.396)	(0.366)	(0.420)	(0.375)	(0.389)
Obs	178	178	176	178	154	176	154
R^2	0.938	0.941	0.940	0.946	0.941	0.948	0.952
Adj. R^2	0.937	0.940	0.939	0.945	0.940	0.946	0.950
AIC	821.551	815.110	809.699	796.960	716.905	788.658	691.205
BIC	831.097	827.837	822.381	809.688	729.053	807.681	712.463

Note: Adjusted variables are the residuals from regressions of the respective original variables on oil prices.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Inflation (HICP q on q growth rate) (2007-2019)

	(1)	(2)	(3)	(4)	(5)
π_{t-1}	0.693***	0.537***	0.376**	0.195	0.441***
	(0.101)	(0.120)	(0.163)	(0.137)	(0.127)
$Cycle_{t-1}$		-0.022**	-0.020*	-0.014	-0.018*
		(0.010)	(0.010)	(0.009)	(0.010)
CF^{1Y}_{t-1}			0.409		
			(0.283)		
CF^{1Y}_{t}				0.943***	
				(0.241)	
INS_t					0.133*
					(0.069)
Intercept	0.102***	0.155***	0.055	-0.082	0.203***
	(0.038)	(0.044)	(0.082)	(0.071)	(0.049)
Obs	51	51	51	51	51
R^2	0.489	0.537	0.557	0.651	0.571
Adj. R^2	0.479	0.518	0.528	0.629	0.544
AIC	-78.722	-81.693	-81.912	-94.121	-83.618
BIC	-74.858	-75.897	-74.185	-86.393	-75.891

Appendix E: Data

Table 13: Data Sources and Frequency

Variable	Abbreviation	Frequency	Source
Economic News Sentiment	ENS	Daily	Factiva-based calculations
Labor News Sentiment	LNS	Daily	Factiva-based calculations
Inflation News Sentiment	INS	Daily	Factiva-based calculations
Share of articles mentioning "recession" or "crisis"	Rword	Daily	Factiva-based calculations
Share of articles mentioning ("recession" or "crisis") and ("uncertainty" or "risk")	Rrisk	Daily	Factiva-based calculations
Term spread (10-year minus 1-year government bond yield)	Term Spread	Daily	European Central Bank
Economic Sentiment Indicator (seasonally adjusted)	ESI	Monthly	European Commission
Composite Purchasing Managers' Index output	PMI/pmicomp	Monthly	S&P Global
Industrial Production Index (excluding construction)	IPI/ipxc	Monthly	Eurostat
Retail Sales Turnover Index (deflated, excluding fuel)	RS/retail$_x f$	Monthly	Eurostat
Real Gross Domestic Product (Euro/ECU Series) for Euro Area (19 Countries) (CLVMEURSCAB1GQEA19)	GDP	Quarterly	Eurostat
Composite Purchasing Managers' Index employment	pmiemp	Monthly	S&P Global
Unemployment rate	Unemployment	Quarterly	Eurostat
Total Employment (number of persons, all activities)	Employment	Quarterly	Eurostat

Table 14: Data Sources and Frequency. Inflation-related variables

Variable	Abbreviation	Frequency	Source
Inflation News Sentiment	INS	Daily	Factiva-based calculations
Survey-based households' inflation expectations	HIE	Monthly	European Commission
HICP - All-items excluding energy, Euro area (changing composition), Monthly	HICP	Monthly	ECB
Market-based 1-year ahead inflation expectations as extracted from inflation-linked swaps	ILS	daily	ECB
1-year ahead inflation expectations of professional forecasters	CF	Monthly	Consensus Forecast
Crude Oil BFO M2 Europe FOB $ BBl Commodity - Historical close, average of observations through period, Euro area (changing composition), Quarterly	Oil prices	Quarterly	ECB

References

Adämmer, P., Prüser, J., & Schüssler, R. A. (2024). Forecasting macroeconomic tail risk in real time: Do textual data add value? *International Journal of Forecasting*.

Aguilar, P., Ghirelli, C., Pacce, M., & Urtasun, A. (2021). Can news help measure economic sentiment? an application in covid-19 times. *Economics Letters*, *199*, 109730.

Algaba, A., Borms, S., Boudt, K., & Verbeken, B. (2023). Daily news sentiment and monthly surveys: A mixed-frequency dynamic factor model for nowcasting consumer confidence. *International Journal of Forecasting*, *39*(1), 266–278.

Angelico, C., Marcucci, J., Miccoli, M., & Quarta, F. (2022). Can we measure inflation expectations using Twitter? *Journal of Econometrics*, *228*(2), 259–277.

Aprigliano, V., Emiliozzi, S., Guaitoli, G., Luciani, A., Marcucci, J., & Monteforte, L. (2023). The power of text-based indicators in forecasting Italian economic activity. *International Journal of Forecasting*, *39*(2), 791-808.

Ashwin, J., Kalamara, E., & Saiz, L. (2024). Nowcasting euro area GDP with news sentiment: A tale of two crises. *Journal of Applied Econometrics*, *39*(5), 887-905.

Barbaglia, L., Consoli, S., & Manzan, S. (2022). Forecasting with economic news. *Journal of Business & Economic Statistics*, 1–12.

Barbaglia, L., Consoli, S., & Manzan, S. (2024). Forecasting GDP in Europe with textual data. *Journal of Applied Econometrics*, *39*(2), 338-355.

Barbaglia, L., Consoli, S., Manzan, S., Tiozzo Pezzoli, L., & Tosetti, E. (2025). Sentiment analysis of economic text: A lexicon-based approach. *Economic Inquiry*, *63*(1), 125-143.

Bauer, M. D., & Mertens, T. M. (2018a). *Economic forecasts with the yield curve* (FRBSF Economic Letter No. 2018-07). Federal Reserve Bank of San Francisco.

Bauer, M. D., & Mertens, T. M. (2018b). *Information in the yield curve about future recessions* (FRBSF Economic Letter No. 2018-20). Federal Reserve Bank of San Francisco.

Benzoni, L., Chyruk, O., & Kelley, D. (2018). *Why does the yield-curve slope predict recessions?* Available at SSRN. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3271363

Bybee, L., Kelly, B., Manela, A., & Xiu, D. (2024). Business news and business cycles. *The Journal of Finance*, *79*(5), 3105-3147.

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

Chauvet, M., & Potter, S. (2002). Predicting a recession: evidence from the yield curve in the presence of structural breaks. *Economics Letters*, *77*(2), 245–253.

Chauvet, M., & Potter, S. (2005). Forecasting recessions using the yield curve. *Journal of forecasting*, *24*(2), 77–103.

Christiansen, C. (2013). Predicting severe simultaneous recessions using yield spreads as leading indicators. *Journal of International Money and Finance*, *32*, 1032–1043.

de Bandt, O., Bricongne, J.-C., Denes, J., Dhenin, A., de Gaye, A., & Robert, P.-A. (2023). *Using the press to construct a new indicator of inflation perceptions in France* (Working paper series No. 921). Banque de France.

Doms, M. E., & Morin, N. J. (2004). *Consumer sentiment, the economy, and the news media* (Working Paper No. 2004-09). Federal Reserve Bank of San Francisco.

Duarte, A., Venetis, I. A., & Paya, I. (2005). Predicting real growth and the probability of recession in the euro area using the yield spread. *International Journal of Forecasting*, *21*(2), 261–277.

Ellingsen, J., Larsen, V. H., & Thorsrud, L. A. (2022). News media versus fred-md for macroeconomic forecasting. *Journal of Applied Econometrics*, *37*(1), 63–81.

Engstrom, E. C., & Sharpe, S. A. (2019). The near-term forward yield spread as a leading indicator: A less distorted mirror. *Financial Analysts Journal*, *75*(4), 37–49.

Estrella, A., & Hardouvelis, G. A. (1991). The term structure as a predictor of real economic activity. *The Journal of Finance*, *46*(2), 555–576.

Estrella, A., & Mishkin, F. S. (1996). The yield curve as a predictor of US recessions. *Current Issues in Economics and Finance*, *2*(7).

Estrella, A., & Mishkin, F. S. (1998). Predicting US recessions: Financial variables as leading indicators. *Review of Economics and Statistics*, *80*(1), 45–61.

Estrella, A., Rodrigues, A. P., & Schich, S. (2003). How stable is the predictive power of the yield curve? Evidence from Germany and the United States. *Review of Economics and Statistics*, *85*(3), 629–644.

Eugster, P., & Uhl, M. W. (2024). Forecasting inflation using sentiment. *Economics Letters*, *236*, 111575.

Fama, E. F. (1986). Term premiums and default premiums in money markets. *Journal of Financial Economics*, *17*(1), 175–196.

Ferrari Minesso, M., Lebastard, L., & Le Mezo, H. (2022). Text-based recession probabilities. *IMF Economic Review*, 1–24.

Fraiberger, S. P. (2016). *News sentiment and cross-country fluctuations.* Available at SSRN. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2730429

Galı, J., & Gertler, M. (1999). Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics*, *44*(2), 195–222.

Gilbert, T., Scotti, C., Strasser, G., & Vega, C. (2017). Is the intrinsic value of a macroeconomic news announcement related to its asset price impact? *Journal of Monetary Economics*, *92*, 78–95.

Gräb, J., & Tizck, S. (2020). Estimating recession probabilities with a less distorted term spread. *ECB Mimeo*.

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica*, *79*(2), 453–497.

Hassan, T. A., Schreger, J., Schwedeler, M., & Tahoun, A. (2023). Sources and transmission of country risk. *The Review of Economic Studies*, *91*(4), 2307-2346.

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth acm sigkdd international conference on knowledge discovery and data mining* (pp. 168–177).

Kalamara, E., Turrell, A., Redl, C., Kapetanios, G., & Kapadia, S. (2022). Making text count: economic forecasting using newspaper text. *Journal of Applied Econometrics*, *37*(5), 896–919.

Kessel, R. A. (1971). The cyclical behavior of the term structure of interest rates. In *Essays on interest rates, volume 2* (pp. 337–390). NBER.

Koop, G., & Korobilis, D. (2012). Forecasting inflation using dynamic model averaging. *International Economic Review*, *53*(3), 867–886.

Labonne, P., & Thorsrud, L. A. (2023). *Risky news and credit market sentiment* (Working Paper No. 14/2023). BI Norwegian Business School.

Lagerborg, A., Pappa, E., & Ravn, M. O. (2023). Sentimental business cycles. *Review of Economic Studies*, *90*, 1358–1393.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, *66*(1), 35–65.

Ochs, A. C. (2021). *A new monetary policy shock with text analysis* (Working Paper No. 2148). University of Cambridge.

Rambaccussing, D., & Kwiatkowski, A. (2020). Forecasting with news sentiment: Evidence with UK newspapers. *International Journal of Forecasting*, *36*(4), 1501–1516.

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

Stock, J. H., & Watson, M. W. (2007). Why has us inflation become harder to forecast? *Journal of Money, Credit and Banking*, *39*, 3–33.

Taylor, J. B. (1993). Discretion versus policy rules in practice. In *Carnegie-rochester conference series on public policy* (Vol. 39, pp. 195–214).

Thorsrud, L. A. (2016). *Nowcasting using news topics: Big data versus big bank* (Working Paper No. 6/2016). BI Norwegian Business School.

Wright, J. H. (2006). *The yield curve and predicting recessions* (Working Paper No. 2006-07). Federal Reserve Board.

Zheng, T., Fan, X., Jin, W., & Fang, K. (2024). Words or numbers? Macroeconomic nowcasting with textual and macroeconomic data. *International Journal of Forecasting*, *40*(2), 746-761.

Acknowledgements

We would like to thank seminar participants at the European Central Bank and at the University of Alicante for valuable comments and suggestions. We are also grateful for the insightful comments and discussions with Gabriel Pérez-Quirós, Julien Pinter, Gerhard Rünstler, Georg Strasser, Beatrice Pierluigi and João Sousa. The work was conducted during the year that Manuel Medina Magro spent at the European Central Bank.
The views expressed in this paper are solely those of the authors and do not necessarily represent the views of the European Central Bank.

Manuel Medina Magro
University of Alicante, Alicante, Spain; email: manuel.medina@ua.es

Lorena Saiz
European Central Bank, Frankfurt am Main, Germany; email: lorena.saiz@ecb.europa.eu

Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu

PDF ISBN 978-92-899-7455-4 ISSN 1725-2806 doi: 10.2866/0039070 QB-01-25-226-EN-N