Appendix for "Forecasting with Economic News"

Luca Barbaglia[a], Sergio Consoli[a], Sebastiano Manzan[b]
[a] Joint Research Centre, European Commission
[b] Zicklin School of Business, Baruch College

A Description of the FiGAS algorithm

In this section we provide additional details regarding the linguistic rules that are used in the NLP workflow discussed in Section 2 of the main paper. For a complete presentation of the FiGAS algorithm and a comparison with alternative sentiment analysis methods, we refer to Consoli et al. (2021). Our analysis is based on the *en_core_web_lg* language model available in the Python module *spaCy*[14]. The package is used to assign to each word a category for the following properties:

- POS : part-of-speech;
- TAG : tag of the part-of-speech;
- DEP : dependency parsing.

Tables 3 and 4 provides the labels used by spaCy for part-of-speech tagging and dependency parsing, respectively. In the remainder of this section, we detail the steps of our NLP workflow, namely the tense detection, location filtering, semantic rules and score propagation based on the proposed fine-grained economic dictionary.

A.1 Tense

We rely on the spaCy tense detection procedure to associate a verb (`POS == VERB`) to one of the following tenses:

a. past (i.e., `TAG == VBD` or `TAG == VBN`);

b. present (i.e., `TAG == VBP` or `TAG == VBZ` or `TAG == VBG`);

c. future (i.e., `TAG == MD`).

If no tense is detected, spaCy assigns the "NaN" value. For each text chunk, we compute a score of each of the tenses above: the score is based on the number of verbs detected with that tense, weighted by the distance of the verb to the ToI (the closer the verb to the ToI, the larger the score). We then assign the tense with the highest score to the whole chunk.

[14] `https://spacy.io/api/annotation`.

Table 3: Part-Of-Speech (POS) tagging.

TAG	POS	DESCRIPTION
CC	CONJ	conjunction, coordinating
IN	ADP	conjunction, subordinating or preposition
JJ	ADJ	adjective
JJR	ADJ	adjective, comparative
JJS	ADJ	adjective, superlative
MD	VERB	verb, modal auxiliary
NN	NOUN	noun, singular or mass
NNP	PROPN	noun, proper singular
NNPS	PROPN	noun, proper plural
NNS	NOUN	noun, plural
RBR	ADV	adverb, comparative
RBS	ADV	adverb, superlative
VB	VERB	verb, base form
VBD	VERB	verb, past tense
VBG	VERB	verb, gerund or present participle
VBN	VERB	verb, past participle
VBP	VERB	verb, non-3rd person singular present
VBZ	VERB	verb, 3rd person singular present

Table 4: spaCy dependency parsing.

DEP	DESCRIPTION
acl	clausal modifier of noun (adjectival clause)
advcl	adverbial clause modifier
advmod	adverbial modifier
amod	adjectival modifier
attr	attribute
dobj	direct object
neg	negation modifier
oprd	object predicate
pcomp	complement of preposition
pobj	object of preposition
prep	prepositional modifier
xcomp	open clausal complement

A.2 Location

The location detection builds on the ability of the spaCy library to perform named-entity recognition. For our location task, we use the following entities:

- countries, cities, states (`GPE`);
- nationalities or religious or political group (`NORP`);

- non-GPE locations, mountain ranges, bodies of water (`LOC`);
- companies, agencies, institutions, etc. (`ORG`).

We search for the desired location among the above entities with the following procedure. We first label each article with its most frequent location among the above recognized named-entities. Then, we look for the most frequent location in each chunk: if there are no entities, we then assign the article location also to the chunk. FiGAS continues considering the specific chunk only if the assigned location is one of the locations of interest to the user, which in our case corresponds to terms referring to the United States[15].

A.3 Rules

After the detection of the location and verbal tense, the FiGAS analysis continues with the evaluation of the semantic rules and the calculation of the sentiment value. In particular, the algorithm selects the text that satisfies the rules discussed below and assigns a sentiment value only to the tokens related to the ToI. Furthermore, if the chunk contains a negation or a term with a negative connotation (`DEP == neg`), the final sentiment polarity of the whole chunk is reversed (*negation handling*). Below is the list of rules that we use to parse dependence:

1. *ToI associated to an adjectival modifier*
   The ToI is associated to a lemma with an adjectival modifier dependency (`DEP == amod`) such as:

   a. an adjective (`POS == ADJ`) in the form of (i) standard adjective (`TAG == JJ`), (ii) comparative adjective (`TAG == JJR`), or (iii) superlative adjective (`TAG == JJS`)

   b. a verb (`POS == VERB`).

   Example (a.) We observe a *stronger* industrial production in the US than in other countries.
   Example (b.) We observe a *rising* industrial production in the US.

2. *ToI associated to a verb in the form of an open clausal complement or an adverbial clause modifier*
   The ToI is related to a verb that is dependent on:

   a. an open clausal complement (i.e., adjective is a predicative or clausal complement without its own subject; `DEP == xcomp`); or

[15]The *locations and organizations* used for restricting the searches to US articles only are: America, United States, Columbia, land of liberty, new world, U.S., U.S.A., USA, US, land of opportunity, the states, Fed, Federal Reserve Board, Federal Reserve, Census Bureau, Bureau of Economic Analysis, Treasury Department, Department of Commerce, Bureau of Labor Statistics, Bureau of Labour, Department of Labor, Open Market Committee, BEA, Bureau of Economic Analysis, BIS, Bureau of Statistics, Board of Governors, Congressional Budget Office, CBO, Internal Revenue Service, IRS.

b. an adverbial clause modifier (i.e., an adverbial clause modifier is a clause which modifies a verb or other predicate -adjective, etc.-, as a modifier not as a core complement; `DEP == advcl`).

Example (a.) We expect output to *grow*.
Example (b.) Output dipped, *leaving* it lower than last month.

3. *ToI associated to a verb followed by an adjective in the form of an adjectival complement*
The ToI is associated to a verb that is followed by an adjectival complement (i.e., a phrase that modifies an adjective and offers more information about it; `DEP == acomp`) such as:

   a. standard adjective (`TAG == JJ`);

   b. comparative adjective (`TAG == JJR`);

   c. superlative adjective (`TAG == JJS`).

Example (a.) The economic outlook looks *good*.

4. *ToI associated to a verb followed by an adjective in the form of an object predicate*
The ToI is associated to a verb followed by an object predicate (i.e. an adjective, noun phrase, or prepositional phrase that qualifies, describes, or renames the object that appears before it; `DEP == oprd`) such as:

   a. standard adjective (`TAG == JJ`);

   b. comparative adjective (`TAG == JJR`);

   c. superlative adjective (`TAG == JJS`).

Example (b.) The FED kept the rates *lower* than expected.

5. *ToI associated to a verb followed by an adverbial modifier*
The ToI is associated to a verb followed by an adverbial modifier (i.e. a non-clausal adverb or adverbial phrase that serves to modify a predicate or a modifier word; `DEP == advmod`) defined as:

   a. a comparative adverb (`TAG == RBR`);

   b. a superlative adverb (`TAG == RBS`).

Example (a.) UK retail sales fared *better* than expected.

6. *ToI associated to a verb followed by a noun in the form of direct object or attribute*
ToI associated to a verb followed by a noun in the form of:

   a. a direct object (noun phrase which is the (accusative) object of the verb. `DEP == dobj`);

   b. an attribute (it qualifies a noun about its characteristics. `DEP == attr`).

Example (a.) The economy suffered a *slowdown*.
Example (b.) Green bonds have been a *good* investment.

7. *ToI associated to a verb followed by a prepositional modifier*
ToI associated to a verb followed by a prepositional modifier (i.e. words, phrases, and clauses that modify or describe a prepositional phrase. `DEP == prep`) and by a:

   a. a noun in the form of an object of preposition (i.e., a noun that follows a preposition and completes its meaning. `DEP == pobj`);

   b. a verb in the form of a complement of preposition (i.e., a clause that directly follows the preposition and completes the meaning of the prepositional phrase. `DEP == pcomp`).

Example (a.) Markets are said to be in *robust* form.
Example (b.) Forecasts are bad, with a falling output.

8. *ToI associated to an adjectival clause*
ToI associated to an adjectival clause (i.e., a finite and non-finite clause that modifies a nominal. `DEP == acl`).

   - *Example* There are many online tools providing investing platforms.

A.4 Sentiment polarity propagation

The following step is to compute the sentiment of the text. In case there are terms wiht non-zero sentiment (excluding the ToI), we proceed by:

i. summing the sentiment of the term closer to the ToI with the sentiment of the remaining terms weighted by the tone of the closest term;

ii. the sentiment value of the ToI is not included in the computation and it is only used to invert the sign of the resulting sentiment in case the ToI has negative sentiment.

A.5 Fine-grained economic and financial dictionary

Several applications in economics and finance compute sentiment measures based on the dictionary proposed by Loughran and McDonald (2011) (LMD). The LMD dictionary includes approximately 4,000 words extracted from 10K filing forms that are classified in one of the following categories: negative, positive, uncertainty, litigious, strong, weak modals, and constraining. A limitation of the LMD dictionary is that it does not provide a continuous measure (e.g., between $\pm$ 1) of the strength of the sentiment carried by a word.

To overcome this limitation, we enriched the LMD dictionary by extending the list of words and by assigning a numerical value to the sentiment. More specifically, we removed a small fraction of the 10K-specific terms not having polarity (e.g., juris, obligor, pledgor), and added to the dictionary a list of words that were frequently occurring in the news text presented in Section 3 of the main paper. Next, we annotated each word in the expanded list by employing a total of 15 human annotators. In particular, we collected 5 annotations from field experts[16], and additional 10 from US based professional annotators. The annotation task consisted of assigning a score to each term based on the annotator's perception of its negative, neutral or positive tone in the context of an article that discusses economic and financial issues. We required annotators to provide a score in the interval $[-1, 1]$, with precision equal to one decimal. We also reminded annotators that a negative values $[-1, 0)$ would indicate that the word is, for the average person, used with a negative connotation in an economic context, whereas positive scores $(0, 1]$ would characterize words that are commonly used to indicate a favourable economic situation. The annotator was given the option to assign the value 0 if he/she believed that the word was neither negative nor positive. The magnitude of the score represents the sentiment intensity of the word as perceived by the annotator: for instance, a score close to 1 indicates an extremely positive economic situation, while a small positive value would be associated with a moderately favourable outcome. In case the annotator was doubtful about the tonality of a word, we provided an example of the use of the word in a sentence containing that specific term. These examples were taken at random from the data set presented in Section 3.

The left panel of Figure 6 compares the mean scores assigned by the US-based and field expert annotators. In general, there seems to be an overall agreement between the two annotator groups, as demonstrated by the positive correlation between the average scores. Moreover, the positivity or negativity of the resulting polarity scores have been verified with the categories of the Loughran and McDonald (2011) dictionary, which are identified by colors in the Figure. In the right panel of Figure 6 we compare the median score across the 15 annotators with the score in the SentiWordNet dictionary proposed by Baccianella et al. (2010). Overall, there seems to be a weak positive corre-

[16] These experts are the authors of this paper and two economist colleagues at the Joint Research Centre.

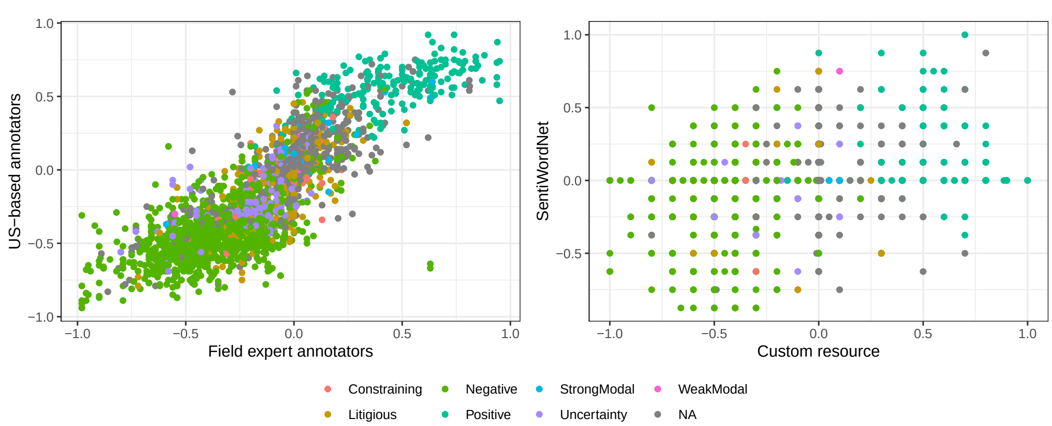

Figure 6: Mean scores US-based annotators versus field experts annotations (left panel), and SentiWordNet scores versus median score of the proposed resource (right panel). Colors indicates the categories identified by Loughran and McDonald (2011).

lation between the two resources, suggesting that often there is disagreement between the general purpose and the field-specific dictionaries. The resulting dictionary contains 3,181 terms with a score in the interval ± 1. We plan to further extend the dictionary to represent a comprehensive resource to analyze text in economics and finance.

We now discuss an example that illustrates the FiGAS approach to sentiment analysis. Let's consider the following sentence: *the US manufacturing sector has carried the brunt of the global economic slowdown over the past few months.* The FiGAS methodology would proceed in the following steps to compute sentiment for, e.g., the ToI *manufacturing sector*:

- The ToI manufacturing sector is detected by looping on the part-of-speech tags;
- The ToI depends on *to carry* (`POS == VERB`), which is neutral (i.e., sentiment equal to 0) based on the sentiment dictionary;
- The dependence analysis further shows that the `VERB` *to carry* is linked to the `DOBJ` (*direct object*) *brunt* which is a `NOUN` and has negative sentiment equal to -0.125 according to the dictionary;
- Polarity is propagated as follows *brunt* $\rightarrow$ *to carry* $\rightarrow$ manufacturing sector and the final sentiment score of the ToI *manufacturing sector* in this sentence is -0.125

B Verbal tense

As discussed in the previous section and in the main paper, we designed the FiGAS methodology to detect the tense of the verb that is related to the ToI, if any. The objective of this feature is to add a temporal dimension to the sentiment measures that could be useful when forecasting macroeconomic variables. More specifically, we calculate sentiment separately for the chunks of text in which we find that the verbal form is either *past*, *future*, or *present*. Instead, we denote by *NaN* the cases when the ToI is not dependent on a verb. Combining the six sentiment measures with the four verbal tenses provides a total of 24 combinations of sentiment-tense measures. While in the paper we provide the results for the aggregate (by tense) sentiment indicators, the aim of this section is to discuss the potential benefits of separating the verbal tenses. We evaluate the statistical significance of the sentiment-tense measures based on the double lasso methodology presented in Section 2 of the paper.

Figure 7 shows the p-value adjusted for multiple testing of the sentiment-tense coefficient η_h in Equation (1). For GDPC1 we find that at longer forecasting horizons the sentiment about the *economy* is significant at the future tense, while the *inflation* is relevant when it is not associated to a verbal form. When the goal is to nowcast GDP, we find that the *economy* in the past tense is typically significant at 10%, while *unemployment* in the present tense shows relevance at some horizons during quarter t. Moving to the monthly variables, the results for INDP show that at the 4-month horizon the only significant sentiment indicator is the *economy* at past tense. When

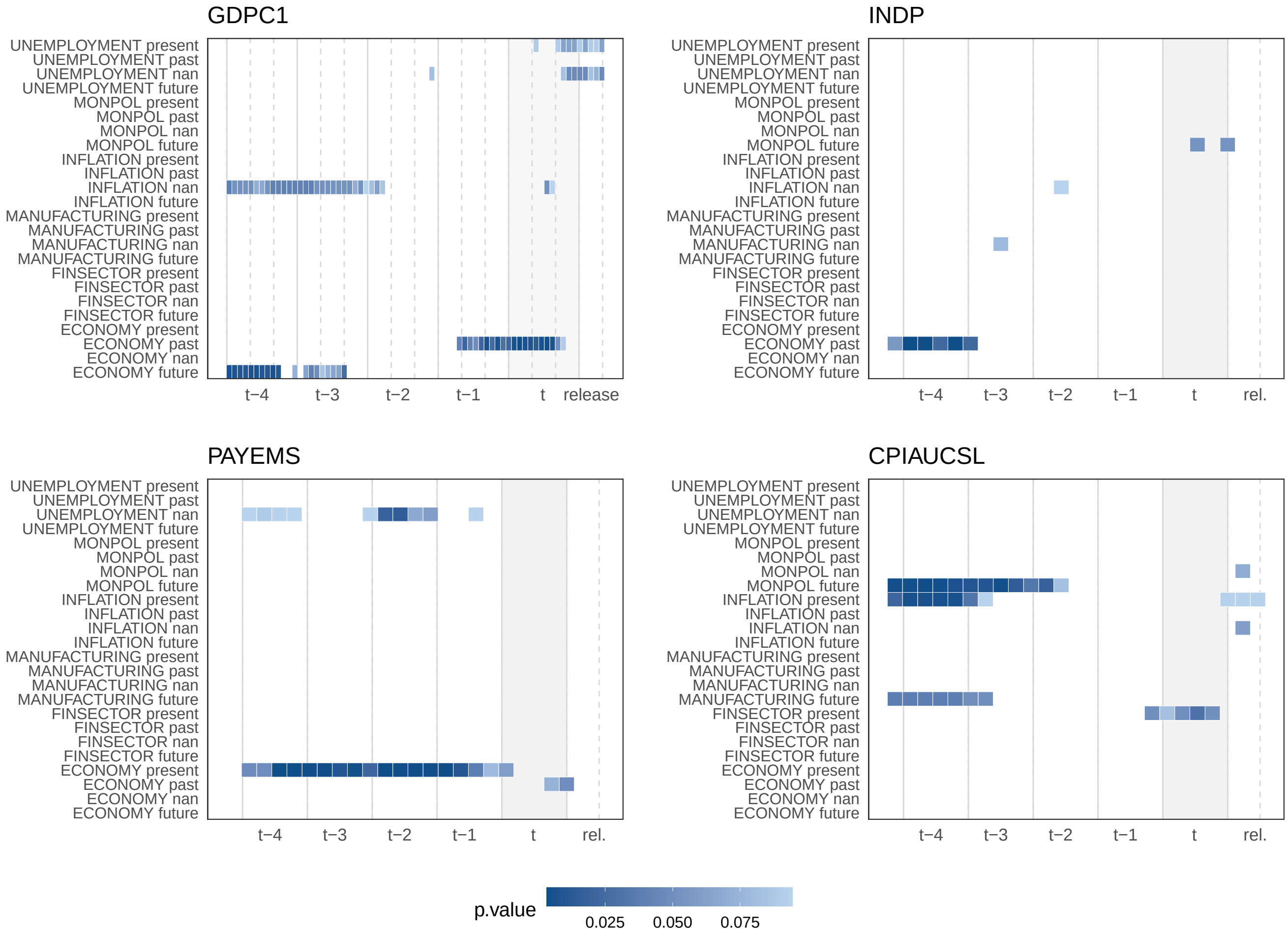

Figure 7: Statistical significance of the sentiment measures for all verbal tenses based on the double lasso penalized regression at each horizon h. The colors depend on the p-value for those variables that are significant at least at 10%: the darker the tile's color, the smaller the p-value. The grey-shaded area corresponds to the nowcasting horizons.

predicting PAYEMS we find that *economy-present* is a statistically significant predictor across horizons from $t-4$ to $t-1$. The *unemployment* indicator is also significant in predicting PAYEMS at the two and four months horizon, in particular when the verb is not detected in the sentence. Finally, the results for CPI indicate that several sentiments in the future tense are significant, in particular *monetary policy* and *manufacturing*, while *inflation* is significant at the present tense. However, at shorter horizons the present tense seems more relevant in predicting inflation for the *inflation* and *financial sector* measures. Overall, when the interest is nowcasting the macroeconomic variables the results seem to suggest that sentiment measures in the past and present tenses are more often significant while the future tense and nan seem more relevant at longer horizons.

We conduct the same analysis also at the quantile level and Figure 8 reports the adjusted p-value of the sentiment coefficient in the quantile regression model in Equation (3). Similarly to the case discussed in the paper, also splitting sentiment by verbal tense produces lower statistical significance of the measures at the median and higher significance at the tails of the distribution. In general, the results seem to suggest that the set of ToIs used to construct the sentiment indicator

Figure 8: Significance of the sentiment measures for all verbal tenses based on the double lasso penalized quantile regression at each horizon h for quantiles 0.1 (left column), 0.5 (center) and 0.9 (right). The colors depend on the p-value for those variables that are significant at least at 10%: the darker the tile's color, the smaller the p-value. The grey-shaded area corresponds to the nowcasting horizons.

is a more crucial choice relative to the verbal tense. In many instances we find that several tenses for the same sentiment indicator are significant, such as in the case of predicting PAYEMS based on the *economy* sentiment at the lowest quantile. Another example of this is the significance of the *financial sector* sentiment in predicting PAYEMS. However, there are other situations in which only one tense emerges as the most relevant, such as the *economy-present* sentiment to predict the PAYEMS median and the *unemployment-nan* to predict the top quantile of GDP. Interestingly, the future tense is significant when predicting GDP on the left tail when used in combination with the *inflation*, *financial sector*, and *economy* sentiment indicators.

C Comparison with the News Sentiment Index

Measuring economic sentiment from news articles is currently an active research area (Thorsrud, 2016; Kalamara et al., 2018; Kelly et al., 2018; Bybee et al., 2019; Shapiro et al., 2020, among others). The main difference between the measures proposed in these papers is represented by the methodology used to extract sentiment from text. It is then an empirical question whether these measures capture genuine predictability in macroeconomic variables. In order to compare the performance of the FiGAS measures with these alternative sentiment indicators, we consider the News Sentiment Index (NSI) proposed by Shapiro et al. (2020) that is available for download at the Federal Reserve of San Francisco website[17]. Similarly to our sentiment, the NSI time series starts in January 1980 and it is available at the daily frequency. Table 5 shows the correlations between the NSI and the six FiGAS sentiment measures at the daily, weekly, monthly, and quarterly frequencies. We consider frequencies higher than daily in order to assess the co-movement between the sentiment measures once the high-frequency fluctuations are averaged out. The results show that the correlations at the daily frequency range between 0.158 and 0.262, except in the case of *inflation* where the correlation with the NSI is close to zero (this finding also extends to the other frequencies). Aggregating to higher frequencies, by averaging the values within the period, has the effect of increasing the level of the correlations. At the quarterly frequencies, we find that the correlations between the NSI and the FiGAS sentiments are around 0.39 for *monetary policy* and *manufacturing* and 0.66 for the *financial sector*. Overall, it seems that the NSI has a high degree of comovement with the broader FiGAS sentiments related to *economy*, *unemployment* and *financial sector* that is probably due to the more frequent coverage of these topics in the news.

In order to evaluate the predictive content of the NSI and FiGAS sentiment measures, we perform an out-of-sample exercise as in the paper. We produce point forecasts from January 2002 until 2019 at horizons starting from one week before the official release of the macroeconomic variables. At each forecast date we select the baseline specification of the ARX model in which we use LASSO to select the most relevant regressors among lagged values of the dependent variables and current

[17]The NSI data is available at `https://www.frbsf.org/economic-research/indicators-data/daily-news-sentiment-index/`.

	Day	Week	Month	Quarter
ECONOMY	0.262	0.439	0.525	0.581
FINANCIAL SECTOR	0.251	0.430	0.559	0.660
MANUFACTURING	0.084	0.197	0.330	0.392
INFLATION	-0.011	-0.023	-0.042	-0.057
MONETARY POLICY	0.158	0.263	0.319	0.395
UNEMPLOYMENT	0.174	0.357	0.475	0.569

Table 5: Correlation coefficients between the NSI and the FiGAS sentiment measures. The correlations are calculated at the daily frequency and also at lower frequencies by averaging the daily values within the period.

and past values of the ADS, CFNAI, and NFCI indexes. We then augment the baseline specification with the NSI and FiGAS measures and produce forecasts of the macroeconomic variables at several horizons. To evaluate the relative accuracy of the forecasts we use the aSPA multi-horizon test proposed by Quaedvlieg (2021) as discussed in the main paper. In this exercise we consider the ARXS model that uses the NSI as predictor as the null forecasts, while the FiGAS measures provide the alternative forecasts. The results of the out-of-sample comparison are provided in Table 6 with p-values smaller than 10% indicating that FiGAS forecasts are significantly more accurate relative to NSI across the horizons considered.

The results indicate that several FiGAS-based forecasts outperform the NSI-based for INDP and PAYEMS, as in the case of the sentiments for *economy*, *manufacturing* and *monetary policy*. When forecasting INDP also the *inflation* and *unemployment* indicators contribute to provide more accurate forecasts, while *financial sector* is significant in the case of PAYEMS. For GDPC1 we find that the NSI-based forecasts outperform several alternative forecasts, in particular for *unemployment* and *financial sector* while the *average* of the FiGAS forecast outperforms NSI.

Sentiment	GDPC1	INDP	PAYEMS	CPIAUCSL
AVERAGE	**0.031**	**0.004**	**0.009**	**0.018**
LASSO	0.871	0.938	**0.010**	0.929
ECONOMY	0.587	**0.072**	**0.009**	0.856
FINANCIAL SECTOR	0.957	0.182	**0.027**	**0.022**
INFLATION	0.405	**0.084**	0.163	0.670
MANUFACTURING	0.691	**0.009**	**0.064**	0.337
MONETARY POLICY	0.887	**0.019**	**0.062**	0.509
UNEMPLOYMENT	0.901	**0.019**	0.781	0.831

Table 6: One-sided p-values of the aSPA multi-horizon test proposed by Quaedvlieg (2021). The null hypothesis of the test is that the point forecasts of the ARXS model with NSI sentiment are more accurate relative to the forecasts of the ARXS that uses the FiGAS measures across the horizons considered.

When forecasting CPI inflation we find that the NSI and FiGAS sentiments are, in most cases, equally accurate, except in the case of the forecasts based on the *financial sector* sentiment that was also found in the paper to be a useful predictor of the inflation variable.

D Fluctuations in forecasting performance

The out-of-sample results provided in the paper average the performance of the model forecasts over time, leaving open the question of whether any forecasting advantage derives from occasional episodes of predictability of the news sentiment, as opposed to a uniformly better performance over time. We use the fluctuation test proposed by Giacomini and Rossi (2010) to evaluate the possibility of time variation in the forecast performance. The fluctuation test consists of performing a predictability test on rolling windows of the out-of-sample period and using appropriate critical values to account for the non-standard setting. The null hypothesis of the test is that the ARX forecasts are more accurate relative to the ARXS forecasts and negative values of the test statistic indicate rejections of the null hypothesis. We calculate the fluctuation at three representative horizons and for a rolling window of 8 quarters for GDP and 24 months for the monthly macroeconomic variables.

Figure 9 shows the results of the fluctuation test together with the one sided critical values at 10% from Giacomini and Rossi (2010). In the figure, we centered the test statistic in the middle of the rolling window so that the statistic at time t represents the value of the test calculated on the window between $t - w/2$ and $t + w/2$, where w represents the size of the window. Overall, the findings suggest that sentiment derived from news delivers significantly more accurate forecasts during certain periods, but not uniformly over time. For instance, in the case of GDPC1 at the one and 32 week horizons some sentiment indicators provide significant predictability around 2004, during the Great Recession, and during 2015-2016. At the longest horizons the results are less remarkable, but there are still some periods with rejections of the null hypothesis of equal accuracy. Overall, it seems that the three periods discussed earlier are typically significant across the four variables considered, although not all periods for all variables. In addition, it is interesting to notice that in many occasions the predictability derives from only one or a few sentiment indicators, and only in few cases from a generalized increase of accuracy for all measures. This confirms that the six indicators capture different aspects of the economic activity which are relevant to predict some variables better than others.

E R^2 and $RMSE$ of the models

Figure 10 and 11 provide a descriptive analysis of the performance of the forecasting models at horizons that range from 1 day before release to four calendar periods (quarters for GDPC1 and months for the remaining variables). Figure 10 shows the R^2 goodness-of-fit statistic for the benchmark ARX forecasting model and the ARXS specification that augments ARX with each of

Figure 9: Fluctuation test proposed by Giacomini and Rossi (2010) with a rolling window size of 7 quarters for GDPC1 and 22 months for the monthly variables (corresponding to a 10% window). The test statistic is aligned to the middle of the rolling window.

the six sentiment indicators. The findings mirror to a large extent the findings of Figure 4, but allow to quantify in a qualitative manner the role of news sentiment measures. For GDPC1 we find that the R^2 of the model based on the sentiment indicators best performs during quarter $t-2$ when it is around 10.9% higher relative to the ARX benchmark. The macroeconomic releases from quarter $t-1$ onward contribute to increase the performance of the ARX forecast while reducing the beneficial role of the news sentiment. For the monthly growth rate of the Industrial Production Index, we find that considering news increases the fit by less than 1.3%, although we showed in Figure 4 that some indicators are statistically significant. The benefit of using the *economy* indicator for PAYEMS is evident in the Figure, since the R^2 of the ARXS specification increases between 2 and 3% at all horizons from $t-1$ to $t-4$. Finally, during months $t-2$ and $t-1$ the *financial sector* indicator contributes to increase significantly the goodness-of-fit, reaching a maximum increase in R^2 of 11% relative to the ARX model.

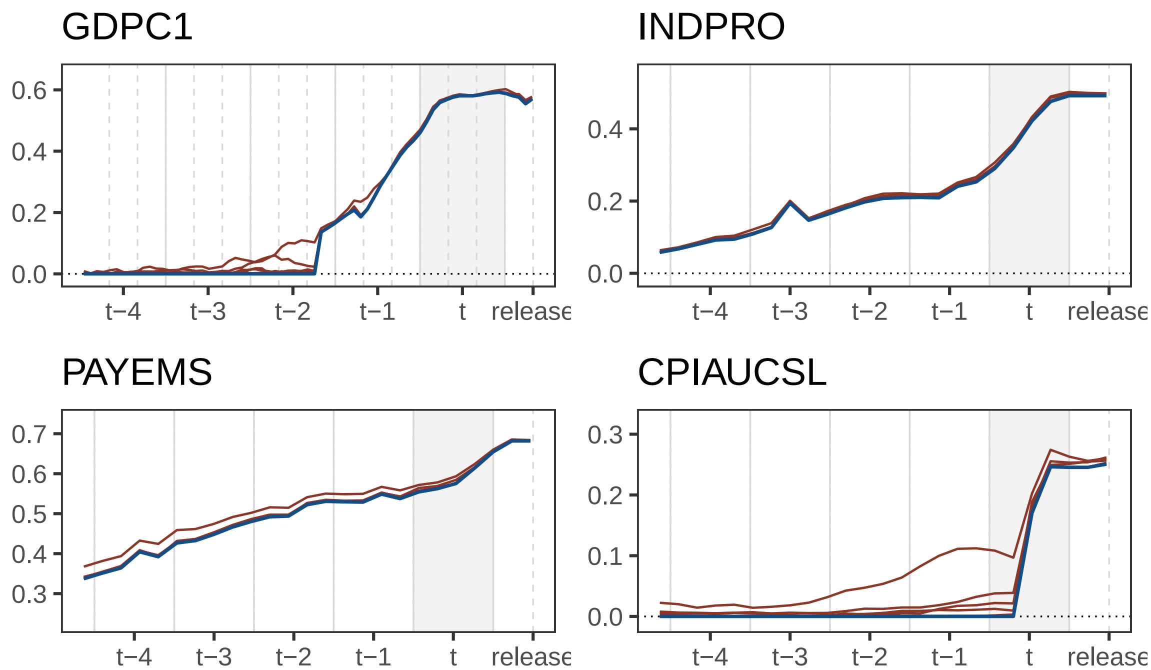

Figure 10: R^2 of ARX (blue line) and the six ARXS (red lines) models at the weekly horizons considered. The x-axis refers to quarters for GDPC1 and months for the remaining variables. The grey area represents the reference period being forecast.

The same conclusions can be drawn by looking at the RMSE of the models. For the GDP growth rate and CPI inflation the sentiment measures contribute to reduce the RMSE up to 5.5%, while the reductions for PAYEMS and INDPRO are smaller being up to 2.7% and 1.2%, respectively.

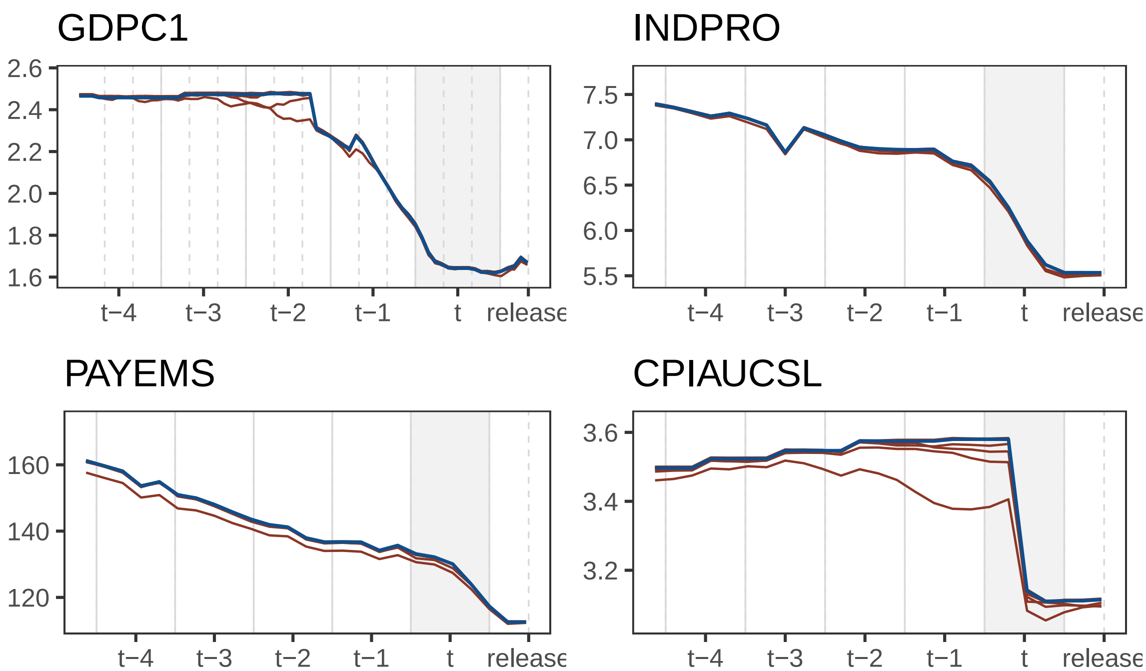

Figure 11: $RMSE$ of ARX (blue line) and the six ARXS (red lines) models at the weekly horizons considered. The x-axis refers to quarters for GDPC1 and months for the remaining variables. The grey area represents the reference period being forecast. The scale of the variables is percentage quarterly or monthly growth for GDPC1, INDPRO, and CPIAUCSL, while it is thousand of persons in the case of PAYEMS.