International Journal of Forecasting 39 (2023) 791–808

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

The power of text-based indicators in forecasting Italian economic activity

Valentina Aprigliano [a,1], Simone Emiliozzi [a,*], Gabriele Guaitoli [b], Andrea Luciani [a], Juri Marcucci [a], Libero Monteforte [a,c]

[a] Bank of Italy, Italy
[b] University of Warwick, United Kingdom
[c] Ufficio Parlamentare di Bilancio, Italy

ARTICLE INFO

Keywords:
Forecasting
Text mining
Sentiment
Economic policy uncertainty
Big data
BMA

ABSTRACT

Can we use newspaper articles to forecast economic activity? Our answer is yes; and, to this end, we propose a high-frequency Text-based Economic Sentiment Index (TESI) and a Text-based Economic Policy Uncertainty (TEPU) for Italy. Novel survey evidence regarding Italian firms and households supports the rationale behind studying text data for the purposes of forecasting. Such indices are extracted from approximately 1.5 million articles from 4 popular newspapers, using a novel Italian economic dictionary with valence shifters. The TESI and TEPU can be updated daily for the whole economy and for specific sectors or economic topics. To test the predictive power of our indicators, we propose two forecasting exercises. Firstly, we use Bayesian Model Averaging (BMA) techniques to show that our monthly text-based indicators greatly reduce the uncertainty surrounding the short-term predictions of the main macroeconomic aggregates, especially during recessions. Secondly, we employ these indices in a weekly GDP tracker, achieving sizeable gains in forecasting accuracy, both in normal and turbulent times.

1. Introduction

Many radical transformations have been reshaping the structure of the economy in recent decades. Globalisation fostered worldwide capital and financial linkages, while the shift from manufacturing to services and digital technologies is structurally changing our economies. Together with the legacies of the Global Financial Crisis, these radical transformations have been challenging the cornerstones of macroeconomic theory and traditional statistical tools for economic analysis and forecasting (Bok et al., 2018; Ng & Wright, 2013). However, analysts today can benefit from the proliferation of non-traditional data, characterised by high frequency and high dimensionality, obtained from unconventional and previously unexplored sources. The big-data phenomenon is indeed spreading in economic analysis.

In this work, we explore one of the most particular types of this new information, text data (Gentzkow et al.,

* Correspondence to: Bank of Italy, via Nazionale, 91, 00184, Rome, Italy.
E-mail address: simone.emiliozzi@bancaditalia.it (S. Emiliozzi).
1 We gratefully acknowledge comments from the Editor Dick van Dijk, two anonymous referees, Eleni Kalamara, Andrea Nobili, Dooruj Rambaccussing, Roberta Zizza and participants of the 31st (EC)2 Conference on "High Dimensional Modeling in Time Series", the 2020 Bank of Italy and Federal Reserve Board Joint conference on "Nontraditional Data & Statistical Learning with Applications to Macroeconomics", the Bank of England virtual conference on "Modelling with Big Data & Machine Learning: Measuring Economic Instability", the 21st IWH-CIREQ-GW Macroeconometric Workshop on "Forecasting and Uncertainty", and the 2020 World Congress of the Econometric Society for their helpful comments. An earlier version of the paper circulated under the title "*Sentiment and Uncertainty Indexes for Forecasting Italian Economic Activity*". The views expressed herein are those of the authors and do not necessarily represent the views of the Bank of Italy or the Eurosystem or the Ufficio Parlamentare di Bilancio. All remaining errors are our own.

https://doi.org/10.1016/j.ijforecast.2022.02.006
0169-2070/

2019), which differs from traditional data because of its unstructured nature. In the same way as this paragraph, text data represents a source of valuable qualitative information for (or about) economic agents, encoded through the complex rules of language. Recent advances in statistical and computational methods have allowed researchers to access and process such data, in order to extract quantitative information suitable for statistical analysis and forecasting.

We use a database of articles from four Italian newspapers, focussing on economic news, to build a sentiment and an uncertainty time series along the lines put forward by Soroka (2006), Tetlock (2007), Loughran and McDonald (2011, 2014), Baker et al. (2016), Ardia et al. (2021, 2019). Then, we use these indicators to nowcast and forecast Italian economic activity.

Our study contributes to the current literature on forecasting with non-traditional data in several ways. Firstly, we provide a rationale for using text data, and in particular that originating from newspapers. Using the Survey on Inflation and Growth Expectations (SIGE) and the Special Survey on Italian Households (SSIH), both published by the Bank of Italy, we deliver new evidence on how Italian firms and households devote attention to newspaper news, citing these outlets as one of the most important sources of information for their economic decisions. Despite the recent rise in alternative media, agents pay attention to the economic news reported in newspapers, which suggests that the information contained within them is timely and significant for their economic choices. This provides a rationale for exploring how newspaper data can help in tracking the business cycle.

Secondly, we build a novel Italian economic dictionary tailored to capture the sentiment of the news regarding the Italian economy. We enrich our approach with valence-shifting terms,[2] which help to better capture the polarity of each sentence.

Finally, and most importantly, we show how the text-based sentiment and economic policy uncertainty indices (TESI and TEPU) that we propose help in tracking Italian economic activity at different frequencies. Using the Bayesian Model Averaging (BMA) approach, we offer evidence to show that, in several cases, our monthly indicators are selected in the predictive models more often than the classic soft and hard indicators commonly adopted in nowcasting. Furthermore, their inclusion tends to improve the predictive density of nowcasts and reduce forecasting errors during recessions. We also explore the weekly properties of our text-based measures, building a high-frequency tracker of Italian GDP, as in Lewis et al. (2021), Delle Monache et al. (2021) and Eraslan and Götz (2021). The text-based indices provide point forecast gains over a simple benchmark model when nowcasting Italian GDP.

Our newspaper articles database is extracted from Dow Jones Factiva, one of the largest archives used in the literature of forecasting with textual based indicators (Ardia et al., 2019; Ashwin et al., 2021; Barbaglia et al., 2020; Bybee et al., 2021; Fraiberger, 2016; Kelly et al., 2021; Shapiro et al., 2020; Thorsrud, 2018). Kalamara et al. (2020) analyse around half a million articles from three main British newspapers, showing that simple text-based indicators can improve economic forecasts, particularly during downturns. Rogers and Xu (2019) test the predictive performance of different measures of economic uncertainty in the US to forecast real and financial outcome variables, showing some predictive content of such measures. Similarly, Rambaccussing and Kwiatkowski (2020) use around 400,000 UK newspaper articles from 1990 to 2018 to nowcast inflation, unemployment and output in the UK. While they do not find evidence of performance gains for inflation, they find evidence of gains in forecasting output and unemployment. As with recent contributions from the text-mining literature, we use multiple newspapers to build the text-based sentiment and uncertainty measures. This allows us to smooth over individual-paper bias and outliers. Moreover, we employ articles written in Italian, and analyse them without translating them into English, thanks to our Italian economic dictionary. This sets us apart from most of the literature on text data, where newspaper articles are either translated into English, or analysed through techniques mostly developed for the English language and which exclude local jargon.

The rest of the paper is organized as follows. Section 2 describes our textual database, the dictionary and the methodology proposed to compute the sentiment indices. Sections 3 and 4 present, respectively, the overall Textual Economic Sentiment (TESI) and Textual Economic Policy Uncertainty (TEPU) indices and their sub topics. Section 5 contains the two empirical forecasting applications. Section 6 concludes the paper.

2. Data

2.1. Why newspaper-based indicators?

Sentiment indicators and topic analysis are becoming popular tools for forecasting and structural analysis. Newspaper-based indicators can be a timely source of information to track business cycles, as documented by recent contributions (Ardia et al., 2019; Ashwin et al., 2021; Barbaglia et al., 2020; Bybee et al., 2021; Kalamara et al., 2020; Rambaccussing & Kwiatkowski, 2020; Shapiro et al., 2020; Thorsrud, 2018).

We provide novel evidence on why newspaper-based sentiment indicators can be helpful in forecasting Italian economic activity by showing that a relevant share of firms and consumers consider newspaper news when making economic decisions.

We asked Italian firms surveyed in the December 2019 wave of the quarterly Survey on Inflation and Growth Expectations (SIGE), a representative survey regularly conducted by the Bank of Italy, to name the three most reliable sources of information used to make important business decisions, out of a list of six possible answers (newspapers, TV news, institutional or trade association

[2] Valence-shifting terms are negations, conjunctions or adverbs that can change the meaning of a sentence. In general, valence-shifting terms can switch the polarity of close words or amplify/reduce the intensity of sentiment.

publications, analysis carried out by private companies, contact with customers and suppliers, social networks).[3]

According to the survey's results, newspapers are a relevant source of information for Italian firms. Fig. 1(a) shows that newspapers are ranked the most important source for almost one-third of the firms, after *"Direct contact with clients and suppliers"* (a reasonable ranking, since supply-chain relations are crucial for businesses). Moreover, considering the second and third mention, too, newspapers are among the top three sources for almost 60% of the sample. Whatever the indicator, newspapers are the most frequently cited source among the publicly available ones, with TV news and social networks being ranked first by less than 3% of the firms and in the top-three by less than 25%.

Regarding the consumers, in the September 2021 wave of the quarterly Special Survey on Italian Households (SSIH), conducted by the Bank of Italy, we asked Italian households to rank the sources of information consulted when taking important economic and financial decisions, with the following options: newspapers, TV news, institutional or trade association publications, contact with relatives/friends/colleagues, social networks, using no source of information.[4]

Newspapers are considered a relatively important source of information for Italian households, too: they are the second most cited information channel, ex aequo with the category *"Contact with relatives, friends and colleagues"*, and surpassed only by "TV news". An interesting finding is that approximately half of the sample declared that they did not use *any* source of information before making important economic and financial decisions. This result is consistent with households' high degree of inattention to economic outcomes, documented in the rational inattention literature (Angeletos & Lian, 2016; Armantier et al., 2016; Coibion et al., 2021; Mackowiak et al., 2021; Sims, 2003). Fig. 1(b) shows that almost 10% of households in the sample use newspapers as the primary source of news to inform their economic and financial decisions, while 20% report newspapers among the top two most important sources. Social networks are ranked first by less than 5% of households and in the top-two by less than 12%.

Overall, Italian economic actors seem to pay a great deal of attention to newspaper news, which justifies why extracting quantitative information from this data source can have relevant implications for economics and forecasting.

2.2. The news corpus

We use articles extracted from the repository Factiva Analytics, using a general query to gather newspaper articles related to economics and finance.[5] For each article, Factiva's records include the title, full text, date and source, along with other metadata and an automatically generated category tagging (e.g. topic(s), name of company, geographical region). We download all articles related to the Italian economy from four popular national newspapers (using both paper and online editions): *Il Corriere della Sera, Il Sole 24 Ore, La Repubblica*, and *La Stampa*. All these newspapers have a large national circulation, are available for a reasonable number of years, and are written in Italian.[6] We report details on the data collection in Appendix A.1.

2.2.1. Data cleaning

We apply the following cleaning procedure to make the data-set ready to compute sentiment. Firstly, starting from 1.8 million articles downloaded using our general query on economic and financial issues, we drop all duplicates, that possibly arise from the dual paper-online editions, by matching titles and texts published within the same month of the "first" article. Around 4.6% of the sample is identified as duplicate and thus dropped.

Secondly, we drop the articles that are (i) financial markets' purely technical news, or (ii) that contain economic words but are not proper economic news, exploiting Factiva's classification algorithm. This reduces a further 12.5% of the observations for a final dataset of slightly more than 1.5 million articles. Table 1 reports the effects of this cleaning procedure on the article count, both in total and by newspaper.

Il Sole 24 Ore turns out to be the newspaper with the most articles relevant to our analysis, as expected, given that it is specialized in economic and financial news.

The text ("corpus") undergoes a preliminary screening ("pre-processing") to eliminate stopwords and characters that are non-informative (i.e. articles, most conjunctions, non-meaningful punctuation). Finally, we stem the text to further reduce the dimensionality of the problem.[7] A complete description of the cleaning algorithm can be found in Appendix A.1.2.

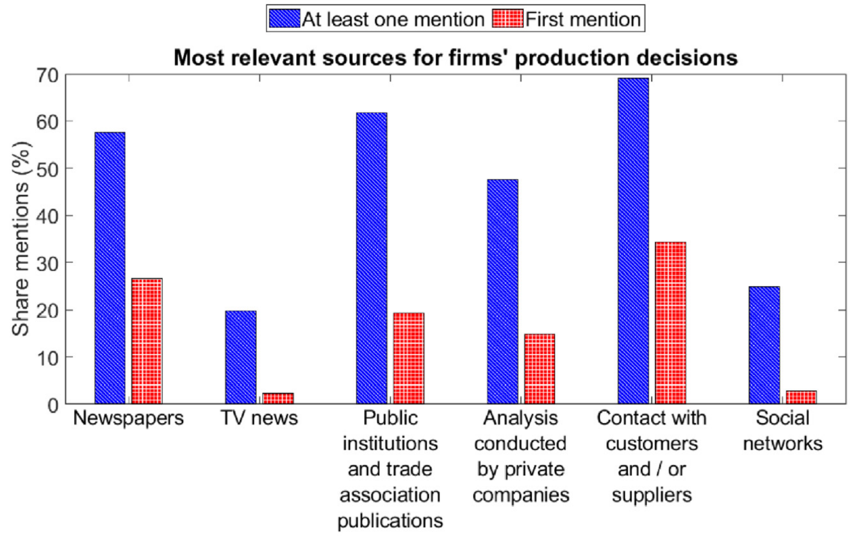
[3] The questionnaire of the December 2019 SIGE survey can be found at https://www.bancaditalia.it/pubblicazioni/indagine-inflazione/2019-indagine-inflazione/12/index.html?com.dotmarketing.htmlpage.language=1. The exact wording of the question and the available options have been reviewed and approved after a pilot study. The questionnaire was filled by almost 1,200 Italian firms during the last quarter of 2019.

[4] The questionnaire of the September 2021 SSIH survey will be published at https://www.bancaditalia.it/statistiche/tematiche/indagini-famiglie-imprese/indag-straord-famiglie-italiane/index.html?com.dotmarketing.htmlpage.language=1. The exact wording of the question and the available options were reviewed and approved after a pilot study. The order in which the respondents were prompted with the options was randomized, in contrast with the earlier firms' survey. The questionnaire was completed by almost 2,063 Italian households during the third quarter of 2021. More details about the results of the survey can be found in Rondinelli and Zanichelli (2021).

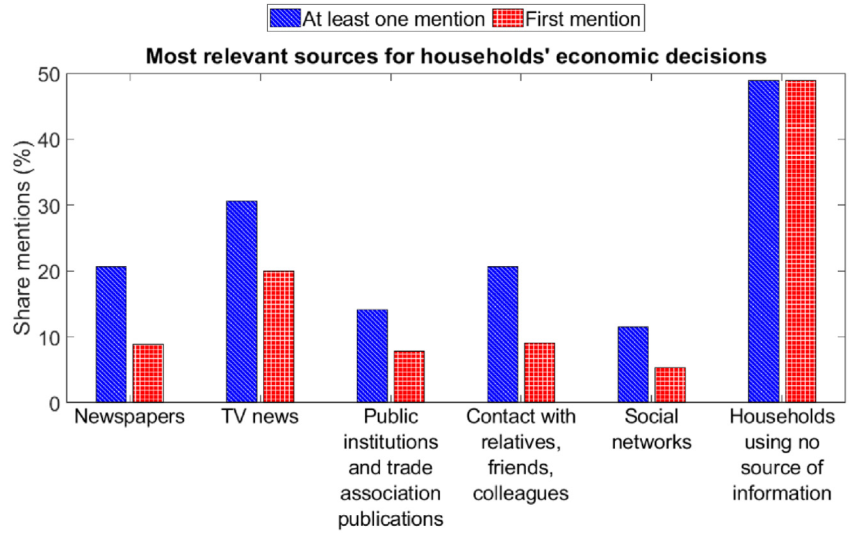
[5] For more details on Factiva Analytics, visit https://professional.dowjones.com/factiva/. For more details on the query, see the online Appendix A.1.

[6] According to *Accertamenti Diffusione Stampa* (https://www.adsnotizie.it/index.asp), *Il Corriere della Sera* and *La Repubblica* are the top two newspapers in Italy in terms of average circulation (with 216,149 and 165,748 copies sold in 2018, respectively); *La Stampa* is the fifth (131,744 copies), while *Il Sole 24 Ore* is the tenth (79,928 copies) but it is the most important Italian economic/financial newspaper.

[7] Our corpus has approximately 350,000 unique terms before stemming: too many to be directly utilized for statistical applications such as forecasting. This large number of elementary words is far above the 160,000 terms commonly used in the Italian language, due to the presence of English words, names, strictly financial terms appearing in short forms, as well as typos.



(a) Firms' information sources ranking from the December 2019 wave of SIGE survey.



(b) Households' information sources ranking from the September 2021 wave of SSIH survey.

Fig. 1. Information sources ranking provided by Italian firms and households. *Note*: the SIGE sample size includes 1199 respondent firms. The *"At least one mention"* bars (in blue) do not add up to three, even though three choices were available, as some respondents (12.5%) provided only one (6.7%) or two (5.8%) answers. The *"First mention"* bars (in red) add up to 100%. The SSIH sample size includes more than 2,063 respondent households. The *"At least one mention"* bars (in blue) do not add up to 2, despite the fact that there were two questions in the survey, as only some respondents (55.6%) provided an answer to the second question. The *"First mention"* bars (in red) add up to 100%. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.3. The Italian economic dictionary

To extract quantitative information from textual data and shrink the dimensionality of the problem, we use a "dictionary-based approach". We compile an Italian economic dictionary focussed on economic and financial terms, adapted to journalistic jargon (Fraiberger, 2016; Loughran & McDonald, 2011, 2014; Soroka, 2006; Tetlock,

Table 1
Article count and cleaning process.

	Sole24Ore	Corriere della Sera	La Repubblica	La Stampa	Total
Raw Count	712,614	450,121	388,137	259,751	1, 810, 623
No Duplicates	679,093	439,938	357,339	251,136	1, 727, 506
Only Economic Topics	624,332	375,300	291,671	219,113	1, 510, 416

Note: the table shows the count of articles at different stages of cleaning, for each newspaper and in total. “Raw Count” refers to the dataset downloaded through our query, until December 2019. “No Duplicates” corresponds to the number of articles left after we exclude all duplicates, as described in the main text. “Only Economic Topics” represents the final database we use, after excluding all articles which are part of topics that are not strictly economic, or which only refer to purely financial news reports (such as stock market closing values of the day).

2007). The dictionary is a set of words (unigrams) and phrases (n-grams) associated with a “polarity" (positive or negative meaning) and a “weight" (module that characterizes the relative importance of a word in a statement). In contrast with the literature, our dictionary is enriched with a set of valence-shifting words, specifically tailored for newspaper jargon.

Some words or n-grams carry a polarity, taking positive or negative signs according to their economic meaning. We call these *“polarity terms"*. As an example, *debito* (debt) has a negative polarity in our dictionary. Some common associations of words carry their own polarity, possibly different from the sum of the individual words. As an example, consider the bi-gram *debt growth*, which carries an overall negative meaning, despite the usually positive meaning associated to the word “growth" without qualification.[8]

A *“valence shifter" term* is a word that does not have a meaning on its own but affects the polarity sign or module of other terms. For example, consider negations (e.g. *GDP did not grow*) and amplifications (deamplifications) of a polarity word, which increase (decrease) the importance given to its meaning (e.g. *a lot of uncertainty*).

2.3.1. Dictionary construction

N-grams are included in the dictionary via a supervised process where the researcher selects relevant economic and financial terms. We extract all n-grams, with $n \in \{1, 2, \ldots, 6\}$, from the text corpus and articles' titles. Then, we discard all stopwords and non-meaningful phrases, all n-grams with fewer than 10 occurrences in the corpus; and from the remaining n-grams, we select and assign a positive, negative or neutral polarity to all phrases with meaningful economic sense. Finally, we discard all n-grams with neutral polarity, as they have no effect within our methodological approach. In this respect, our approach is different from the unsupervised algorithms such as Latent Dirichlet Allocation (LDA) (Blei et al., 2003), *Word2Vec* (Mikolov et al., 2013) or *GloVe* (Pennington et al., 2014). However, our methodology is transparent, as the n-grams included in the dictionary are predetermined and it is possible to assess the opportunity of their inclusion (or exclusion). Valence-shifting words are selected by consulting Italian dictionaries and the corpus of a selection of articles. Our final dictionary, after stemming and stopword removal, contains 433 and 190 unique polarity and valence-shifting terms, respectively.

2.4. Computing the sentiment

Our sentiment indicators are computed using the R package *sentometrics* (Ardia et al., 2021), which allows fast computations.[9] Once the text is cleaned and processed, the sentiment is calculated for each article as the sum of polarities, as modified by accounting for nearby shifters, divided by the number of words:

$$\mathrm{SS}_j = \frac{\sum_i^{\text{N words}} \text{Polarity}_i \times \text{Shifter}(i)}{\text{N words}} \tag{1}$$

To give an example of how this approach works in practice, consider the following sentence, extracted from an article published on 10 December 2019, in the online edition of *Il Sole 24 Ore*: “*tra gennaio e ottobre il calo complessivo dell'output è dell'1,2%*" (“Production fell by 1.2% overall between January and October"). After our text-processing routine, the sentence is collapsed into:

gennaio ottobre calo complessivo output
(January October fall overall production)

to which our estimation procedure assigns a sentiment score of −0.2. In order to see how this is computed, we highlight and classify the words for which there is a match in our dictionary:

$$\text{gennaio ottobre } \underbrace{\text{calo}}_{\text{Shifter }(-1)} \text{ complessivo } \underbrace{\text{output}}_{\text{Polarity }(1)}$$

The token “*output*” is assigned total polarity $1 \times (-1) = -1$ as a shifter with value (-1) (“*fall*”); it appears within five words and it has not already been used in conjunction with another polarity term. Since there are five words in the trimmed sentence, the total sentiment score of article j (SS_j) is computed as

$$\mathrm{SS}_j = \frac{1 \times (-1)}{5} = -0.2$$

Table 2 shows some examples of how our dictionary works on a selection of real newspaper headlines. While we almost always correctly classify the sign of the sentence, we can only partially capture the actual depth of the positivity/negativity. Our classification seems to be helped by the inclusion of valence-shifter terms, which allow a better signaling of some polarity terms.

[8] One could argue that “growth” is a neutral word, unless associated with a meaningful name, but it is classified as a positive word by several generic and economic polarity dictionaries. In fact, talking about “growth" in economics usually implies an increase in output or GDP and therefore economic development or upswings in the business cycle.

[9] See Appendix A.1.4 for a schematic pipeline of the procedure.

Table 2
Sentences and associated Sentiment score.

Sentence (translated from Italian)	Sentiment
GDP has fallen	−1.000
Ex Ilva, does ArcelorMittal risk to leave Taranto?	−0.257
Oil: new cuts by OPEC, but the group argues about quotas	−0.143
Clash on plastic and sugar taxes. Renzi raises the bar	−0.083
Censis: Italians, left with no confidence, leave even the BOT.	−0.029
More jobs and output. The government's plan for a "mixed-property" ILVA	0.257
China lends a helping hand on duties; financial indices are positive	0.000
Istat, GDP at +0.2% in 2019. Growth expected to pick up in 2020	0.250
Ex Ilva, Conte against ArcelorMittal: "the plan is not yet ok: we will reject it"	0.000
Debt has grown	−0.500

Note: the table shows sample sentences and the associated Sentiment, obtained by applying Eq. (1) over the cleaned text (not reported here). The sentences are reported as translated from Italian, while the Sentiment has been computed using our Italian dictionary with the original phrases/sentences. All the sentences, except for the first and the last ones, are headlines taken from our newspaper corpus.

2.4.1. Aggregation across newspapers.

The final sentiment indicator in a given day t, the TESI_t, is obtained by aggregating the individual series from each newspaper. The series for each newspaper is given by the average score(s) of individual articles' scores (SS_j), normalised to have mean zero and unit variance. Then, to take into account breaks in the series generated by the entrance in Factiva's database of new newspapers (for example, March 2001 for the entrance of *Il Sole 24 Ore*, July 2005 for *La Repubblica*), we rescale each newspaper series in order to match the sentiment score of *Il Sole 24 Ore* in August 2007.[10] The TESI_t is obtained as a weighted average (with weights equal to the number of articles published in a given month) of each newspaper's rescaled sentiment score.

The exact procedure to calculate the aggregate TESI is as follows: (i) we compute the sentiment for each newspaper S_{jt}, with j being the specific newspaper, by averaging $SS_{a(j)}$ across all articles a of newspaper j (an article with no polarity words at all has $SS_{a(j)} = 0$); (ii) we normalize the four newspaper-specific series to have mean zero, and variance one; (iii) we offset the normalized newspaper j's sentiment by a constant α_j, so that $S_{j,(\text{Aug 2007})} = S_{(\text{Sole24Ore}),\,(\text{Aug 2007})}$ (the constant is obviously the difference between the two measures in August 2007); call this series $\tilde{S}_{jt}$;[11] (iv) we take the weighted average of all $\tilde{S}_{jt}$, using as weights the number of articles published in j at time t (N_{jt}); (v) finally, we normalize the weighted average aggregate series to have mean zero and variance one. This final series represents our TESI_t.

The same procedure is applied separately to all topics and sector series (so that to generate another series based on a subset of articles, we follow steps (i)-(v) using only the articles within that subset).

[10] We choose this month to match the sentiments, as: (i) no newspaper presents any local outlier, (ii) all newspapers are available (La Repubblica enters the sample in 2005), and (iii) it is right before the start of the Financial Crisis. Sentiment calculation is robust to changes in the rescaling period.

[11] For weekly series, we offset by the average difference among all August 2007 weeks, while for daily series, we offset by the average difference among all August 2007 days.

3. A text-based sentiment indicator for Italy

Fig. 2 depicts our overall TESI, derived from articles published between January 1997 and December 2019,[12] along with the quarterly GDP growth rate (both standardized). Vertical bars represent key historical events affecting the Italian economy. As from 1997 until early 2001 TESI is computed with articles from only two newspapers (*La Stampa* and *Il Corriere della Sera*), the series in this period must be interpreted with caution.[13]

The contemporaneous correlation between the two series over the whole sample at the quarterly frequency is 0.45: it is low at the beginning of the sample, between 1997 and 2001 (when only *La Stampa* and *Il Corriere della Sera* were available), but it starts to increase from 2002 onwards.

TESI tracks episodes important for the Italian economy. It decreases after adverse events, such as the Iraq disarmament crisis in 1998, the Dot-com bubble burst in 2000, the Iraq war in 2003, the US stock market collapse in 2007, the financial crisis after the collapse of Lehman Brothers in 2008, the Greek risk of default in 2010, the start of the Sovereign Debt Crisis in 2011, the Brexit referendum in 2016, the 2019 Italian budget law discussion in October 2018, and the Italian government crisis in August 2019. It increases rapidly after the Dot-com bubble crash in 2000, before the Great Recession, during the recovery in 2014 and 2015, and also in 2017, when global trade was robustly growing.

[12] The computations to generate TESI indices are based on articles downloaded from the Factiva database in April 2020 via the query presented in Appendix A.1. GDP data come from a vintage downloaded in January 2020. Since Factiva constantly updates its databases, if one were to download all the articles as of today for the time span considered, namely January 1997 - December 2019, one would possibly find minimal differences in the composition of the corpus. However, we are confident that such updates are a limited phenomenon and do not compromise the robustness and stability of our results, given the large number of articles we use.

[13] Figure A.1 in Figure A.1 displays the annual time series evolution of the number and share of articles by newspaper used in the TESI and TEPU computation.

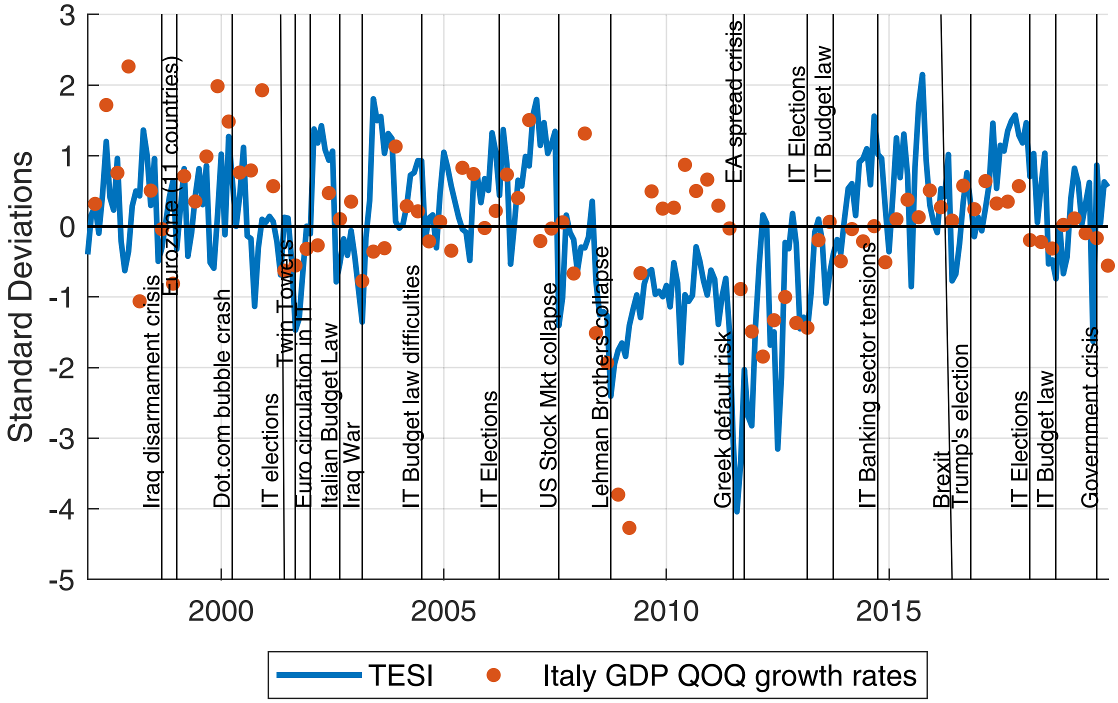


Fig. 2. Textual Economic Sentiment Indicator (TESI) and Italian GDP growth. *Note*: both TESI and GDP growth are standardized to have zero mean and unit variance over the period 1997–2019. TESI is computed daily (using articles last downloaded in April 2020) and aggregated at monthly frequency. GDP growth rates are calculated using the vintage published by Istat in January 2020. Vertical bars highlight key events for the Italian economy.

3.1. Sentiment by topic and economic sector

We use Factiva's proprietary algorithm for article tagging, which assigns each article to one or more granular subject (e.g. economic news, sales figures, economic growth/recession), to estimate sentiment indicators for several topics and economic sectors. We group the almost 300 categories provided by Factiva into 15 topics (such as economic conditions, labor market, and monetary policy) and 21 sectors (such as manufacturing, services, and retail). In the forecasting application in Section 5, we select economic-relevant topical and sectoral TESI topics as timely proxies for important economic variables (such as manufacturing output) along with the overall TESI.

Appendix A.2 discusses the construction and the properties of these topics more in detail.

3.2. Comparison with other soft indicators of economic activity

We compare TESI with measures of consumers' and firms' expectations on the economy that are commonly used in forecasting GDP: the Markit Composite Purchasing Managers' Index (PMI) for Italy, and the Italian Business (IESI) and Consumer Confidence indices produced by the national statistical institute (Istat).

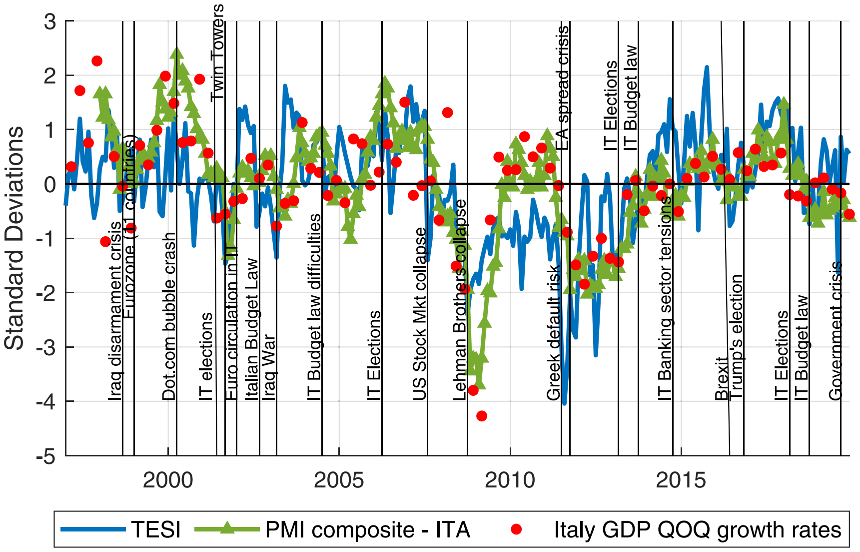
Fig. 3(a) shows TESI, the Italian composite PMI and the quarter-on-quarter (QOQ henceforth) change of the Italian GDP. The contemporaneous correlation between our text-based sentiment index and the composite PMI is 0.5 at the monthly level.[14] Fig. 3(b) displays TESI together with business and consumers' confidence indices released by Istat. Looking at the contemporaneous correlation, TESI seems to align better with the confidence of businesses rather than households.

TESI has two methodological advantages over these traditional monthly confidence indicators. Firstly, it can be computed at higher frequencies. This allows for both infra-month updates, as well as exercises at higher frequencies, such as the weekly one presented in Section 5. Secondly, the monthly Markit and Istat indicators are based on data collected much earlier than their publication date. Accounting for processing and publishing time, these series are published around two weeks after the end of the period of data collection. For example, if a relevant event for the Italian economy occurred in the last week of a given month, it would enter the PMI and Istat confidence surveys only in the following month. In contrast, our sentiment can be updated continuously, which means that this event would enter TESI with a delay of only one day.[15]

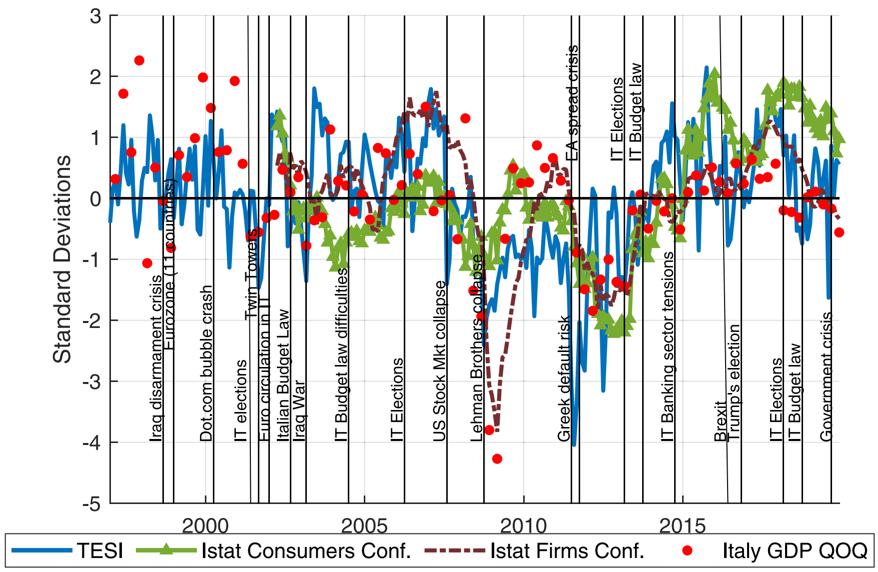
Finally, Fig. 4 shows a comparison at the monthly frequency between two text-based indices for Italy: our

[14] This high correlation is recorded also between the sentiment and the other PMI sub-components, such as PMI manufacturing, services and their forward-looking components.

[15] The possibility of a fast update of our textual indexes can be extremely useful when major and unexpected events for the economy happen abruptly. In these cases, policymakers need timely information to make the appropriate policy decisions.



(a) Comparison between TESI and PMI composite



(b) Comparison between TESI, Consumers' confidence indicator, and Business confidence indicator (IESI)

Fig. 3. Comparison between TESI and other soft indicators. *Note*: all the series in the graphs are standardized to have zero mean and unit variance.

TESI, based on newspaper articles, and Istat's Social Mood on Economy (SME) index, based on Twitter feeds.[16] The contemporaneous monthly correlation between the two indices is moderate (0.3). It is larger at the beginning and at the end of the sample (from February 2016 until June 2017), but the two series tend to move in opposite directions during the first phase of the US-China trade dispute, in July 2017 and in 2019. During the Italian government crisis in August 2019, our indicator collapsed, while SME jumped up to high positive values. Our measure has also several methodological advantages: (i) it covers a longer time span, while SME starts in February 2016; (ii) we provide decomposition of sentiment into sectors and topics, unlike the SME; (iii) our index is based on newspaper news, which we have proven, through representative surveys of firms and households, to be among the principal sources of information for economic actors. Instead, SME only relies on one social network.

16 Istat's SME index is a daily index available at https://www.istat.it/en/archivio/219600.

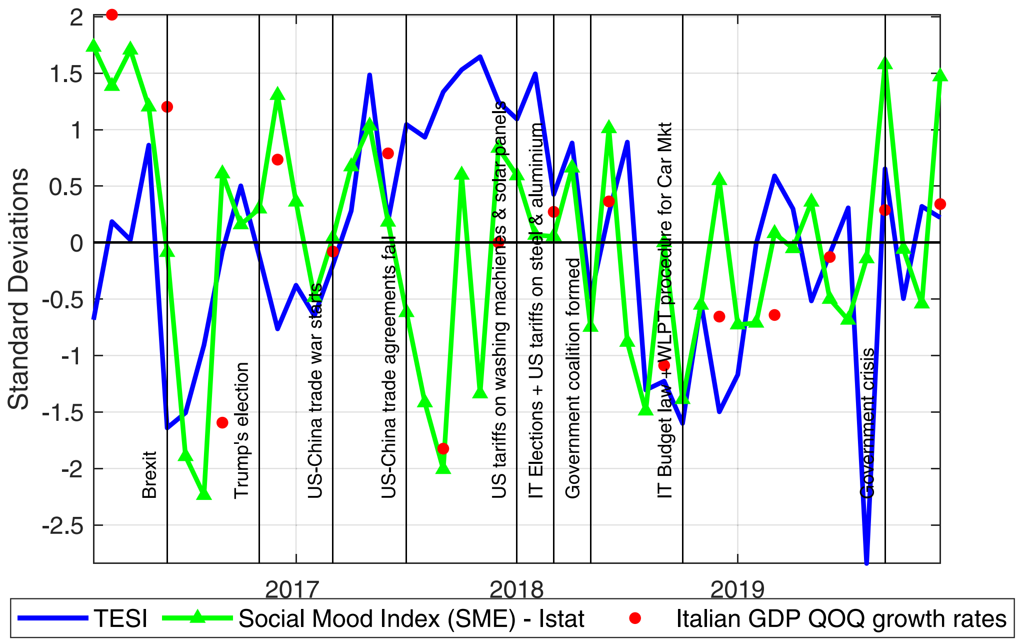


Fig. 4. Comparison between TESI and Social Mood on Economy (SME) Index. *Note*: the series in the graphs are standardized to have zero mean and unit variance. SME is available since 2016, hence the comparison is over the period February 2016 - December 2019.

4. Textual economic policy uncertainty for Italy

Measuring Economic Policy Uncertainty (henceforth, EPU) has become a key issue in recent years, since the seminal work of Baker et al. (2016) (henceforth, BBD2016; see also (Rogers & Xu, 2019) for an interesting application to forecasting). Inspired by their methodology, we propose several textual EPU indices for Italy (henceforth, TEPU) at daily, weekly and monthly frequency. The availability of the topic tags allows us to construct granular topical-uncertainty indicators, unlike (Ardizzi et al., 2019), who propose an overall Italian EPU index derived both from articles in Bloomberg and *tweets* from Twitter,[17] or (Donadelli et al., 2020), who use only the first 100 most read economic articles of the day in *Il Sole 24 Ore*. Furthermore, our TEPU series covers a longer time span than the Italian EPU index of BBD2016 and the EPU24 index proposed by Donadelli et al. (2020).

Our TEPU is computed as the share of articles in a given time period t that are “EPU" articles (that is, those which talk about uncertainty, policy and the economy):

$$\mathrm{TEPU}_t = \frac{\sum_i^{N_t} \mathbf{1}\left[\text{EPU Article}\right]_i}{N_t} \tag{2}$$

where N_t is the total number of economic and financial articles extracted from the four Italian newspapers in day t and filtered as described in Section 2, $\mathbf{1}\left[\cdot\right]_i$ is an indicator function taking value 1 if i is an EPU article, and 0 otherwise. We define “EPU Article" as any piece of news that contains at least one “uncertainty word” (*incert** or *incertezz** in Italian, with * being a wild card) and one “policy word" as defined in BBD2016 for its Italian EPU indicator.[18] Hence the $TEPU_t$ index is the share of articles jointly satisfying the (E)conomic, the (P)olicy, and the (U)certainty criteria described above.[19] Notice that this methodology slightly differs from BBD2016, but produces substantially similar results, as we discuss later.

4.1. TEPU index for Italy

The monthly Italian TEPU index is shown in Fig. 5 together with the quarterly growth rates of the Italian GDP and selected important events for Italy. The negative correlation between TEPU and GDP growth is strong. The tone of the newspaper articles becomes more uncertain after episodes associated with a slowdown in economic growth. This goes with strong spikes in our TEPU index, as for example during the *Dot.com bubble* in 2000; after the struggling process to approve the budget laws in Parliament in 2004; and, more recently, in 2018; during the Global Financial Crisis; during the Sovereign Debt Crisis in 2011; after Brexit and Trump's elections and during the Italian government's crisis in August 2019.

4.1.1. Differences from other EPU indices

With respect to Baker et al. (2016), we use a larger set of newspapers. The original EPU index uses only articles from *il Corriere della Sera* and *La Stampa*. We also add

[17] Actually, Ardizzi et al. (2019) are the first authors to suggest using Twitter to measure EPU. Altig et al. (2020) adopt the same idea to measure EPU during the COVID-19 pandemic in the US and in the UK.

[18] See the Appendix in Ardizzi et al. (2019) for adaptating the “policy word” to the Italian case. Another difference with respect to the Italian EPU index by BBD2016 is that we add the keyword “*bankitalia*”, which is often used by journalists to talk about “*Banca d'Italia*”.

[19] Most articles in our database satisfy the (E)conomic requirement by construction.

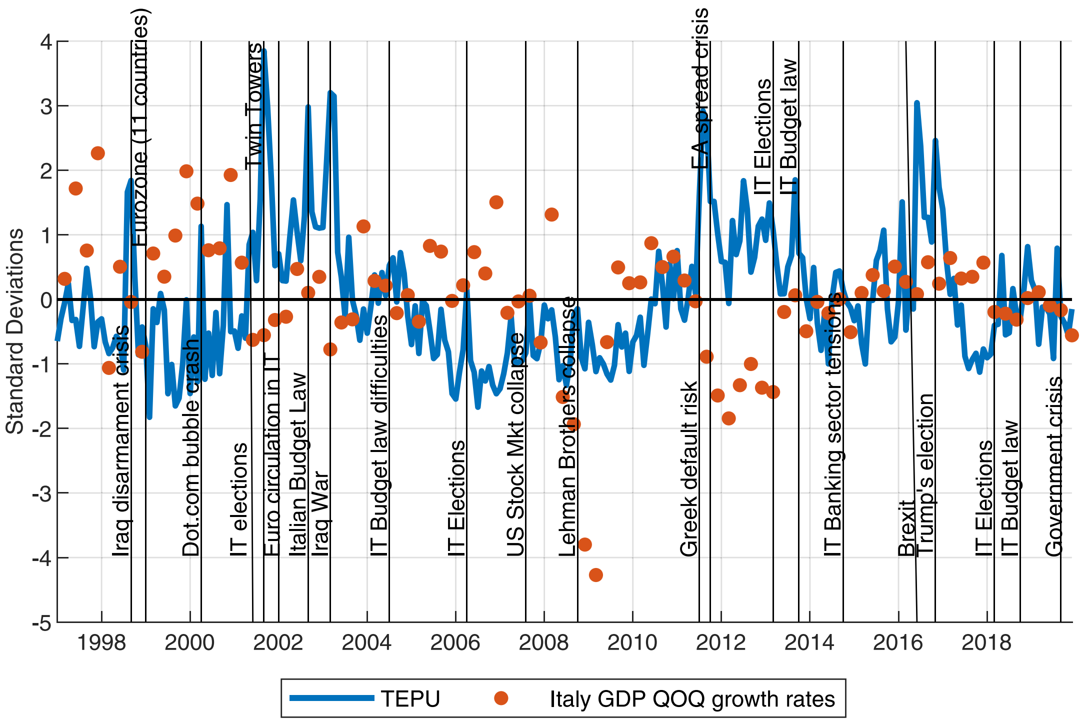


Fig. 5. Textual Economic Policy Uncertainty (TEPU) and Italian GDP growth. *Note*: the TEPU is calculated using daily articles in *real time*, while GDP growth rate figures are calculated using the vintage published by Istat in January 2020. These figures include the whole sample of September 1996 - December 2019 and all sources. Vertical bars highlight key events for the Italian economy.

to the (P)olicy criterion the term "*bankitalia*", often used by journalists to mean "*Banca d'Italia*". Furthermore, we exclude all the articles related to sport, entertainment, or weather. Moreover, while BBD2016 use as a denominator all the articles published in each newspaper on a given day, we use only those articles focussing on economic and financial issues. While a change in their EPU index can be driven by an increase in the coverage of economic news over the total amount of news in a given day, TEPU mostly captures the intensive margin within economic news. Regarding the aggregation in a single EPU index of information emanating from different newspapers, BBD2016 build their EPU index by (i) dividing the raw counts of articles satisfying the EPU criteria with the total number of articles in the same newspaper in each month; (ii) dividing each newspaper series by their standard deviation prior to 2011; (iii) taking averages of the two standardized series; (iv) normalizing the averaged series so that it has mean 100 prior to 2011. Instead, TEPU is constructed by pooling together the total number of EPU-related articles available at each date from the four Italian newspapers and applying Eq. (2). Despite these differences, our overall TEPU index for Italy is strongly correlated with the monthly BBD2016 EPU index: the contemporaneous correlation is close to 0.8, as shown in Fig. 6.

Donadelli et al. (2020) construct their Economic Policy Uncertainty index (EPU24) using only articles from *Il Sole 24 Ore* and choosing the 100 most popular economic articles in each month. Their indices cover the period 2012–2019. Instead, TEPU indices presented in this work are based on a larger set of articles, on a longer time period, and with details from components for topics and sectors. On the other hand, we use *all* available economic articles, regardless readership, which could be a proxy of attention/relevance for a piece of news. The monthly correlation in the period 2012–2019 of TEPU and EPU24 is slightly above 0.4.

TEPU indices by topic and sector. As with sentiment, we use Factiva's proprietary algorithm to calculate TEPU indices by topic and sector. See Appendix A.3.2 for a more detailed discussion.

5. Empirical applications

5.1. Monthly BMA model

We conduct a pseudo real-time forecasting experiment to assess the ability of TESI and TEPU to track economic activity.[20] We adopt the BMA approach described in Bencivelli et al. (2017) to nowcast and forecast one step ahead

[20] The asynchronous availability of macroeconomic indicators implies that the information set in real-time typically displays the *ragged-edge* shape at the end of the sample. This generates missing observations for variables that are released with a delay with respect their reference period (i.e. industrial production in month t is available with a 45-day delay). In order to reproduce the (pseudo) real-time information setting in our out-of-sample simulations, we take the latest available dataset, updated until December 2019, and we cut it recursively month by month, in order to mimic the same pattern of missing values at the end of the sample faced in real time.

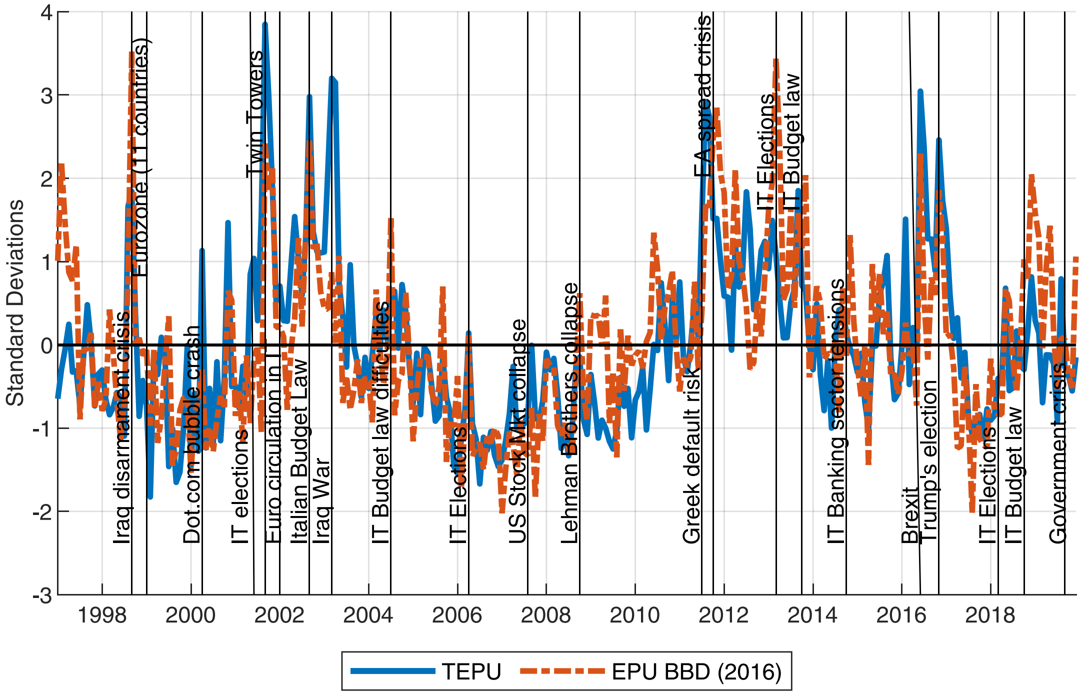


Fig. 6. Comparison between TEPU and the (Baker et al., 2016) EPU index for Italy. Note: the EPU by Baker et al. (2016) correlates 0.76 with TEPU measure for Italy and 0.83 with the TEPU measure for economic conditions. TEPU and TEPU for economic conditions have a correlation of 0.92.

the quarterly growth rates of GDP, as well as its main components, such as value added in the service sector (VAS), which accounts for about 70% of total economic activity in Italy, household consumption (HHC), and gross fixed investments (GFI).

The BMA modelling approach has two non-negligible advantages over a frequentist one when using a "large" set of correlated predictors. Firstly, it allows for enhanced transparency and interpretation of the selection of regressors included in the predictive models. Secondly, the BMA's predictive densities jointly incorporate within each model the uncertainty stemming from both model selection and coefficient estimation. In fact, the BMA approach estimates and combines forecasts of a large number of different models, weighting them by the likelihood of being the true data-generating process. These characteristics of the BMA posterior statistics are extremely useful in evaluating the extent to which specific series (TESI and TEPU in our case) contribute to improving both point and density forecasts. The Markov Chain Monte Carlo Model Composition algorithm[21] (MC^3 for short) is used to construct the confidence bands of the regression coefficients, the Posterior Inclusion Probabilities (PIPs) of each regressor, and to assess the ability of the text-based indices to shrink the uncertainty surrounding the forecasts.

BMA averages across the S randomly drawn models

$$y = \alpha_s \mathbf{1}_T + X_s\beta_s + \epsilon, \quad \epsilon \sim N(0, \omega^{-1} I_T) \tag{3}$$

where X_s is a $T \times K_s$ matrix with a subset of K_s regressors, weighted by their posterior probabilities (assuming an equal prior probability for all models). Model M^s is drawn from among all the $R = 2^n$ possible models, where n is the number of regressors. At each iteration, model M^{s+1} includes a variable in M^s or removes it. Based on M^s, the marginal likelihood $p(y|M^s)$ is estimated, and the acceptance probability is computed as

$$p(M^s, M^{s+1}) = min\left[\frac{p(y|M^{s+1})}{p(y|M^s)}, 1\right] \tag{4}$$

If model M^s is accepted, its h-step-ahead forecast is[22]

$$\hat{y}^s_{t+h} = E(y_{t+h}|y_t, M^s) \tag{5}$$

and the final h-step-ahead forecast is

$$\hat{y}_{t+h} = \frac{1}{S}\sum_{s=1}^{S} \hat{y}^s_{t+h} \tag{6}$$

Following Fernández et al. (2001a), we assume the non-informative prior for both the error precision[23]

$$p(\omega) \propto \frac{1}{\omega} \tag{7}$$

[21] Introduced in Madigan and York (1995). As in Fernández et al. (2001b), to assess the convergence of the chains we compute the correlation between the actual Bayes factor in Eq. (4) and the empirical relative frequency of model visits. This correlation is approximately 90%, which reasonably guarantees the convergence of the chains of the MCMC Model Composition algorithm (MC^3).

[22] In our empirical application for nowcast, we set $h = 0$, while for 1-step-ahead forecasts $h = 1$.

[23] Our main results are obtained following the literature on BMA estimation, and hence setting the ω parameter to a value close to the historical variance of each target variable. Our results are quite insensitive to small perturbations of the ω parameter.

Table 3
Variables included in the BMA exercise: Baseline vs text-based specification.

N	Label	Description	Treatment	GDP		HHC		GFI		VAS	
				Base	Text	Base	Text	Base	Text	Base	Text
1	ITCNFCONR	ITA household confidence index	none	✓	✓	✓	✓	✓	✓	✓	✓
2	ITCNFBUSQ	ITA business confidence indicator	none	✓	✓	✓	✓	✓	✓	✓	✓
3	ITTOTPRDR	ITA business survey: production	none	✓	✓			✓	✓	✓	✓
4	ITEUSVCIQ	ITA services: confidence SA	none	✓	✓			✓	✓	✓	✓
5	ITIPMAN.G	ITA industrial prod.-manufacturing	deltalog(1)	✓	✓	✓	✓	✓	✓	✓	✓
6	EMPMIM..Q	PMI Manufacturing - EA	none	✓	✓	✓	✓	✓	✓	✓	✓
7	ITPMIM..Q	PMI Manufacturing - IT	none	✓	✓	✓	✓	✓	✓	✓	✓
8	ITPMIS..Q	PMI Services - IT	none	✓	✓	✓	✓	✓	✓	✓	✓
9	@:ITMSCIP	Weighted ave. Std. Dev. of the EPS forecast for the t+1 Fiscal Year	deltalog(12)	✓	✓			✓	✓	✓	✓
10	AUTOD	Car registrations in Italy	deltalog(1)			✓	✓				
11	fct_sent	TESI	zscore MA(3)		✓		✓		✓		✓
12	fct_epu	TEPU	zscore MA(3)		✓		✓		✓		✓
13	fct_sent_man	TESI - Manufacturing	zscore MA(3)		✓		✓		✓		✓
14	fct_sent_ser	TESI - Services	zscore MA(3)		✓		✓		✓		✓
15	fct_epu_man	TEPU - Manufacturing	zscore MA(3)		✓		✓		✓		✓
16	fct_epu_ser	TEPU - Services	zscore MA(3)		✓		✓		✓		✓
17	fct_sent_lab	TESI - Labor market	zscore MA(3)		✓		✓		✓		✓
18	fct_sent_ret	TESI - Retail sales	zscore MA(3)		✓		✓		✓		✓

Note: Base is the set of data entering the Baseline model without textual data. Text is the set of variables included in the estimation of the text-based model, which includes text-based indices.

and the intercept as $p(\alpha) \propto 1$.

Table 3 illustrates the detailed list of variables included in each model and their pre-treatment. The information set includes eight text-based indices distinguished by topic (TESI and TEPU for economic conditions, service and manufacturing activity, and sentiment for the labor market and retail sales) and a selection of standard indicators commonly adopted to track the short-term economic outlook, such as the industrial production, Purchasing Managers' Indices (PMI), business and consumers' confidence indices, investors' expected earnings and, only for the HHC model, new cars registrations.[24]

The standard deviation of investors' earning expectations can be interpreted as a non-textual uncertainty measure of Italian firms' future profitability, and has an appreciable negative correlation with the quarterly variation of the GDP. All the variables are demeaned.

We start by estimating the models with data from January 2001 to December 2010 and then produce the first nowcast, relative to 2011.Q1, and the first one-step-ahead forecast, relative to 2011.Q2. We continue recursively, by adding a new month to the sample and once again estimating the BMA, proceeding in this fashion until the end of the sample (nowcasting the 2019.Q4 in December 2019).[25]

Fig. 7 shows the nowcasts made with a model including our text-based indices (TB-model) with respect to a baseline model, which excludes them from the information set. BMA estimates are shown with the 25th and 75th percentile bands of uncertainty and with the QOQ growth of the targets.

The TB-model turns out to grasp more effectively the depth of the downturn during the sovereign-debt crisis for all targets. Unlike the baseline, the TB-model's nowcast is significantly negative for GFI in 2012.Q1, when the investments dropped by almost 6%. The TB-model is also better at tracking the swing of investments in 2017 and 2018.[26] It follows the contraction of household consumption between 2011 and 2014 significantly better than the baseline, but stops outperforming in 2016, when it starts to overestimate the slowdown in HHC growth.

In the full sample, the two models' point forecasts display an equal predictive accuracy as measured by their relative RMSFE (the ratio between the TB-model's and the baseline's RMSFE; see Table 4 and Appendix A.5). However, the TB-model tends to have a lower RMSFE during the most turbulent period in the sample, in particular for household consumption.[27]

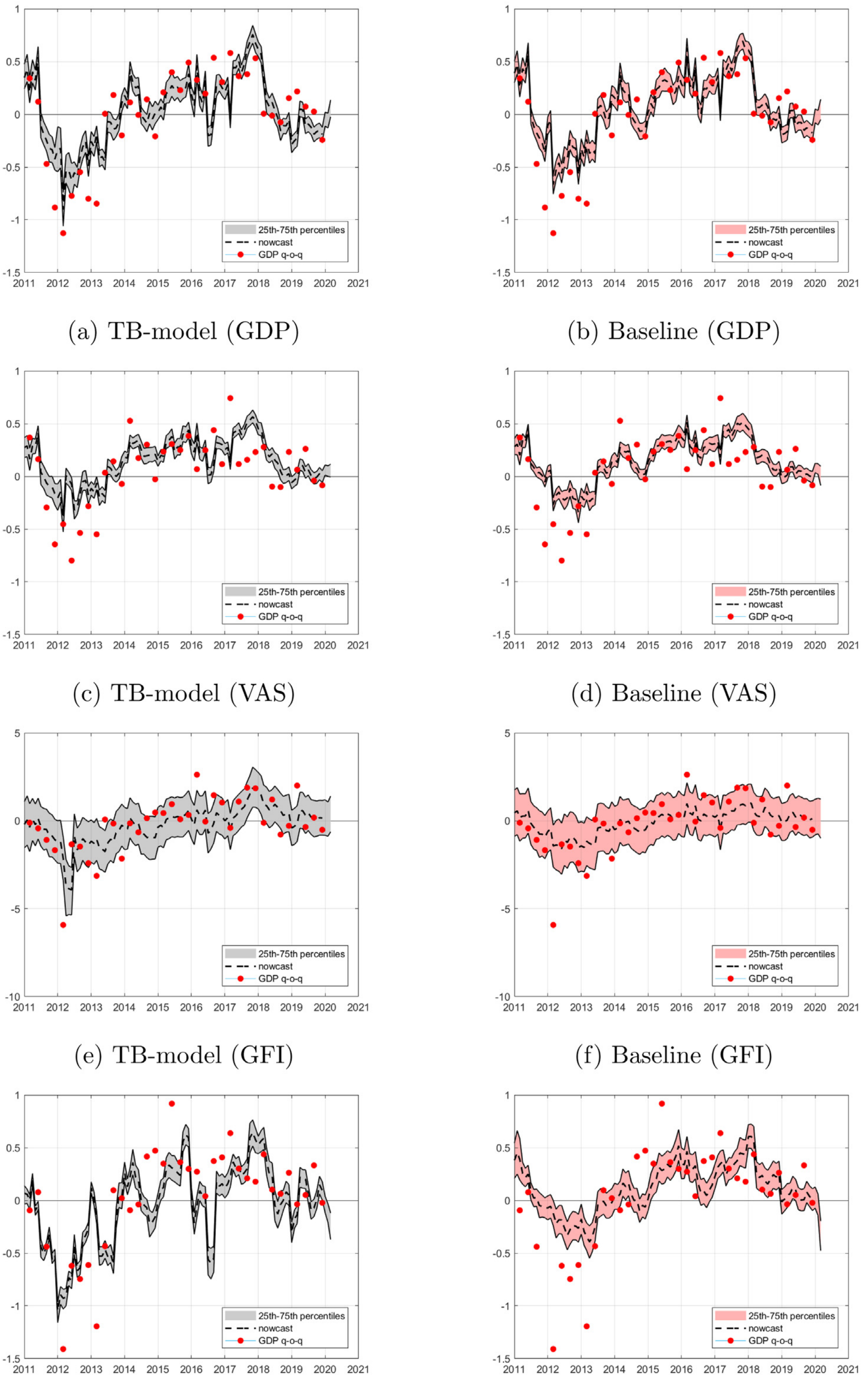
The most interesting results derive from the density forecast comparison, based on the weighted likelihood ratio test by Amisano and Giacomini (2007) (see Table 5 and Appendix A.5). TB-model outperforms the baseline, particularly during the sovereign-debt crisis, when the higher

24 The HHC model does not include business confidence for service and manufacturing sectors and investors' expected earnings, as they are more related to the supply side of the economy, such as GDP and value added, and less concerned with the demand side.

25 The one-step-ahead forecast for 2020.Q1 was not used in the evaluation of the forecast in this forecasting exercise, as the official National Accounts for 2020.Q1 had not yet been released.

26 This swing was likely due to the uncertainty regarding the renewal of tax incentives, as frequently reported by the media at the time.

27 The negligible difference between the RMSFEs of the two competing models is not significant, according to the Diebold–Mariano test (Diebold & Mariano, 1995).



(a) TB-model (GDP) (b) Baseline (GDP)

(c) TB-model (VAS) (d) Baseline (VAS)

(e) TB-model (GFI) (f) Baseline (GFI)

(g) TB-model (HHC) (h) Baseline (HHC)

Fig. 7. Nowcasts of the quarterly growth of GDP and its components (percentage changes) with BMA. *Note*: the figures depict the BMA estimates with the 25th and 75th percentile bands of uncertainty. "TB-model" includes the text-based uncertainty and sentiment indicators; "Baseline" excludes them. VAS = value-added in-service sector, GFI = gross fixed investments and HHC = household consumption.

Table 4
Relative Root Mean Squared Forecast Error (RMSFE) for nowcasts (n) and 1-step-ahead forecasts (f).

	2011.1 - 2014.12		2015.1 - 2019.12		Whole sample	
	n	f	n	f	n	f
GDP	0.93	0.91	1.17	1.16	1.00	1.00
VAS	0.97	1.21	1.08	1.08	1.00	1.00
GFI	1.03	0.94	1.13	1.08	1.03	1.00
HHC	0.83	0.79	1.46	1.29	0.99	1.00

Note: the relative RMSFE is the ratio between the BMA text-based model's RMSFE and the baseline model's RMSFE, estimated only with traditional macroeconomic data. Using a Diebold–Mariano (DM) test we do not find statistically significant improvements in predictive accuracy in the nowcasts and the 1-step-ahead forecasts for the samples and macroeconomic variables considered.

Table 5
Weighted Likelihood Ratio Test for nowcasts (n) and 1-step-ahead forecasts (f).

	2011.1 - 2014.12		2015.1 - 2019.12		Whole sample	
	n	f	n	f	n	f
GDP	**9.1*****	**8.6*****	−15.6***	−22.4***	**6.1*****	**6.6*****
VAS	**5.3*****	**6.6*****	−7.1***	−10.3***	**3.7*****	**6.4*****
GFI	**3.6*****	**23.7*****	**6.9*****	**10.5*****	**4.6*****	**24.5*****
HHC	**14.6*****	**12.2*****	−9.4***	−11.8***	**11.4*****	**11.8*****

Note: all results are significant at 1% level. Positive values of the Weighted Likelihood Ratio Test (in boldface) imply that the TB-model produces more accurate density forecasts than the baseline.

volatility makes forecasting very difficult. Our TESI and TEPU indices squeeze the forecast uncertainty of the GDP and its main components, both in normal and turbulent times.

Finally, we evaluate the Posterior Inclusion Probabilities (PIPs), which are a measure of the relative importance of each regressor in predicting the dependent variables. Exploiting the pseudo real-time design of our forecasting exercise, we can evaluate their temporal evolution. Results presented in Fig. 8 refer to nowcasting: using the PIPs metric, TESI is an outstanding competitor of the popular Purchasing Managers's Indices (PMIs) produced by Markit (manufacturing and services PMI are displayed in Fig. 8 with the acronyms PMI-man and PMI-services, respectively) and by Istat (Italian Economic Sentiment Indicator, IESI), which are both released monthly.

TESI is selected in the forecasting models more frequently than Istat's IESI and TEPU indices when nowcasting GDP's quarterly growth. Furthermore, for all the target variables in the forecasting exercise, the sectoral "services activity" TESI outperforms both its PMI and Istat counterpart throughout the sample: after the trough recorded in the sovereign-debt crisis, it is selected by 80% of the forecasting models for nowcasting GDP. Also, the "manufacturing" TESI PIPs are comparable in magnitude to its PMI and Istat counterparts. In the HHC model, the overall TESI markedly outweighs the Istat consumers' confidence index in the aftermath of the crisis and in the last two years, when it is persistently selected for nowcasting.

The overall "manufacturing" TEPU index is selected as many times as a strong competitor, such as the standard deviation of investors' expected earnings, which is a robust predictor for the short-term evolution of Italian economic activity. The "TEPU Services" index is selected more frequently than the investors' expected earnings in both VAS and GFI models; for the latter, the PIP of the uncertainty measure extracted from the newspapers reaches almost 50% during the recovery phase, notwithstanding its high variability.[28]

5.2. A text-based weekly economic tracker

Following Lewis et al. (2021), Delle Monache et al. (2021) and Eraslan and Götz (2021) we build a weekly indicator of economic activity for Italy with two innovations: (i) a new dataset, focussed on sentiment (TESI) and uncertainty indices (TEPU); (ii) the use of the Italian GDP year-on-year growth rate. Including the weekly TESI and TEPU series in a dataset of macroeconomic variables commonly used in nowcasting (see Table 6 for the full list of weekly and monthly variables), we obtain large point-forecast gains in the prediction of the quarterly Italian GDP year-on-year growth rate. With respect to a benchmark weekly tracker that excludes textual data, the forecast accuracy improvement is large across the whole out-of-sample period of 335 weeks (2011–2019), with a statistically significant gain consisting of a 15% reduction in the relative RMSFE. This result is confirmed either using an expanding or a rolling window.

We build two weekly datasets, where monthly variables (e.g. PMIs and Istat Confidence indices) have been disaggregated at weekly frequency via Chow-Lin methodology (Chow & Lin, 1971), using the available indicators at higher frequencies. The first dataset (X^{Bench}), used to calculate the benchmark weekly index, contains traditional data used in nowcasting, such as electricity consumption and PMIs; the second dataset (X^{Text}) expands the first

[28] We provide additional details on the BMA estimated coefficients and the confidence band construction in Appendix A.5.

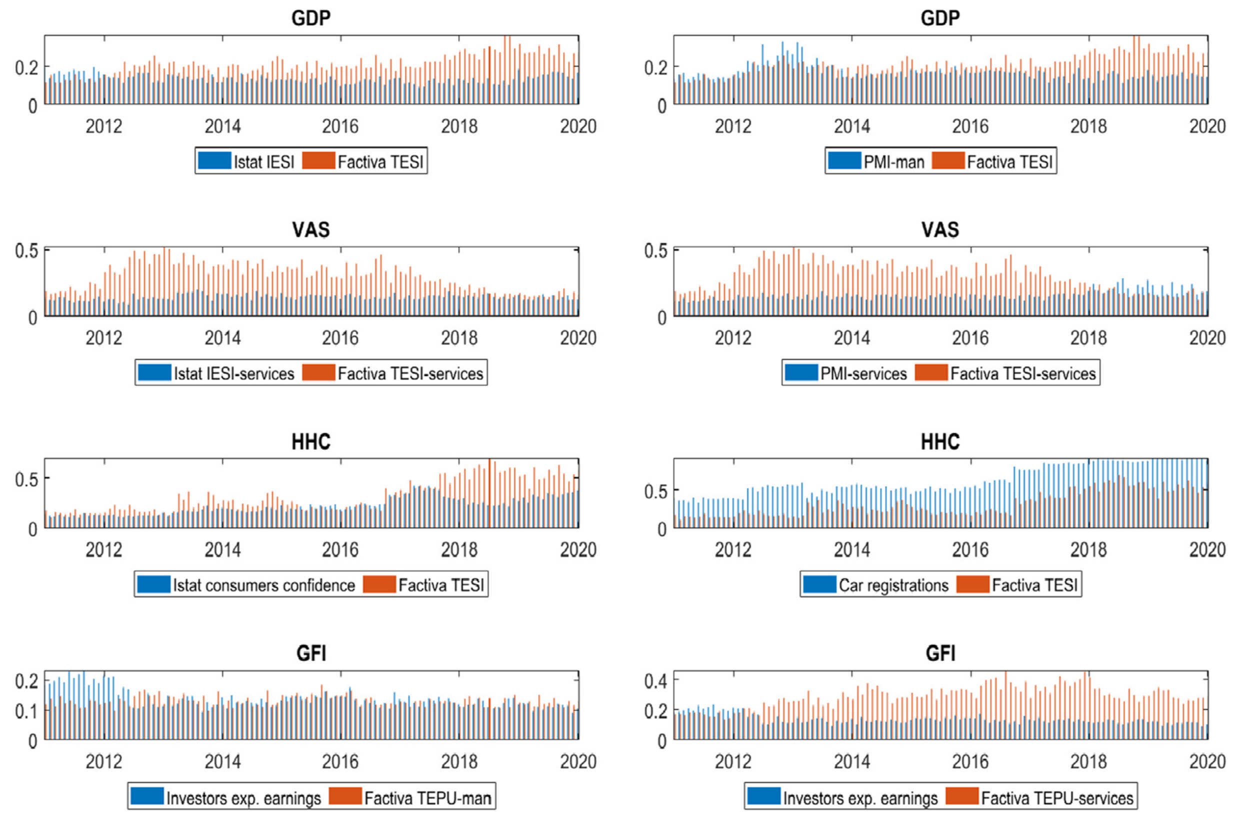


Fig. 8. Posterior Inclusion Probabilities (PIPs) - nowcasting. *Note*: PIPs are useful for evaluating the relative importance of a regressor in the BMA predictive regressions for GDP, VAS, HHC, GFI. PMI-man and PMI-services are the Purchasing Managers' Indices (PMIs) for manufacturing and services, respectively; Istat IESI is the Economic Sentiment Indicator produced monthly by the Italian Statistical Office, Istat. Sources: Markit, Istat, IBES MSCI and Italian Association of Automotive Industry.

Table 6
Variables used in the estimation of the weekly index of economic activity.

Variable	Frequency	Baseline	Text-Based
Electricity consumption	Weekly	✓	✓
Earning forecast std	Weekly	✓	✓
Istat IESI	Monthly	✓	✓
PMI manufacturing	Monthly	✓	✓
PMI services	Monthly	✓	✓
TESI	Weekly		✓
TEPU	Weekly		✓
TESI manufacturing	Weekly		✓
TEPU manufacturing	Weekly		✓
TESI services	Weekly		✓
TEPU services	Weekly		✓
TESI labor	Weekly		✓
TESI retail	Weekly		✓
TESI leisure	Weekly		✓

dataset to include sentiment (TESI) and uncertainty indicators (TEPU) for Italy. Table 6 provides details of the variables included in the baseline and in the text-based model for the weekly index of economic activity.

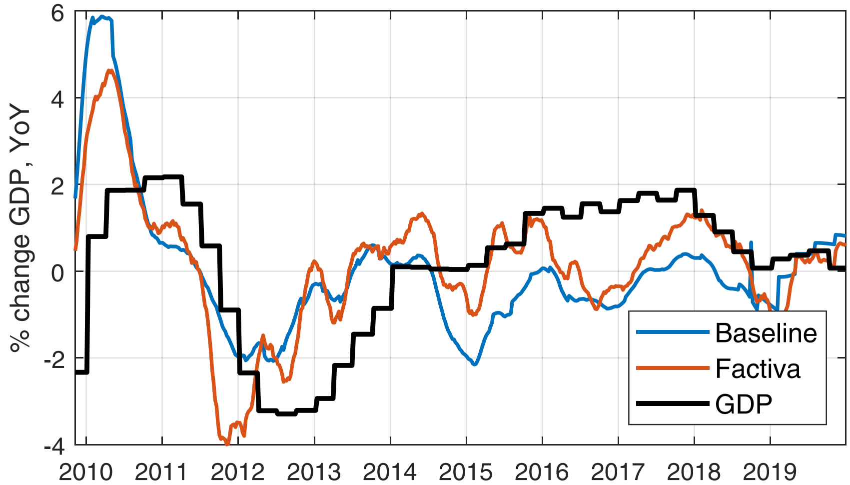
We perform a horse race between two weekly economic activity indices, $Z^{Bench}_{MA,t}$ and $Z^{Text}_{MA,t}$, derived from the two datasets as follows:

1. We extract for each dataset the first principal component following the literature on weekly economic trackers;
2. We take a 13-period ($\approx$ one quarter) moving average of the first principal component.

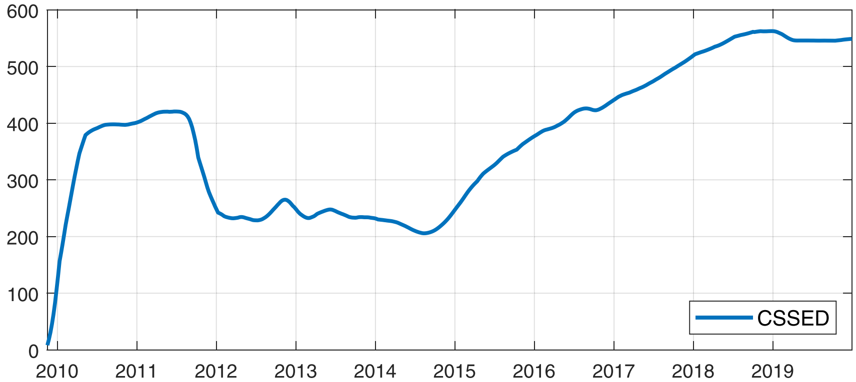
As $Z^{Bench}_{\mathrm{MA},t}$ and $Z^{Text}_{\mathrm{MA},t}$ are standardized principal components, we scale them to the quarterly year-on-year GDP growth rate, $\Delta Y_{yoy,t}$, via the following regression:

$$\Delta Y_{yoy,t} = \alpha^i_t + \beta^i_t Z^i_{\mathrm{MA},t} + \varepsilon^i_t \\ t = 1, \dots, T \quad and \quad i = \{Bench, Text\} \tag{8}$$

Using the estimated coefficients in (8), we compute the value of the two weekly GDP trackers at week t, W^i_t for



(a) Italian weekly real-time index of economic activity



(b) CSSED on the nowcasting errors comparing the models with and without text-based indices

Fig. 9. Real-time weekly indicators of the GDP - rolling sample results. *Note*: these results refer to the weekly GDP tracker presented in Section 5.2. Panel (a) shows the GDP QOQ growth rate (black line), the baseline tracker (blue line) and the tracker augmented with TESI and TEPU indices (red line), as calculated in real time by the rolling window model. Panel (b) shows the CSSED of the two models. A positive value indicates that - from the start of the out-of-sample exercise to that point in time - the text-based model performs better than the baseline. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$i = \{Bench, Text\}$, as:

$$W_t^i = \widehat{\alpha}_t^i + \widehat{\beta}_t^i Z_{\mathrm{MA},t}^i \quad t = 1, \ldots, T \quad and \quad i = \{Bench, Text\}$$

To train the model, we use only available data for both the GDP and the principal component up to week t. We compute the corresponding nowcast error as:

$$error_t^i = \Delta Y_{yoy,t} - W_t^i.$$

We train our model using weekly data from January 2001 to December 2010; thus, the out-of-sample exercise goes from January 2011 until December 2019. Forecast accuracy measures are computed on out-of-sample data using both an expanding and a rolling (335 weeks) window.

Table 7 shows the relative RMSFE as the ratio between the two weekly activity trackers (W_t^{Text}) and (W_t^{Bench}), respectively. The tracker (W_t^{Text}), which uses both TESI and TEPU indices, has a better performance in the weekly nowcast of the year-on-year GDP growth, as the relative RMSFE is smaller than 1 on the whole out-of-sample period. For the rolling sample, we run a Diebold–Mariano test[29] (Diebold & Mariano, 1995) and find that gains are significant at the 5% level, apart from the negative GDP growth sub-sample. Fig. 9(a) shows the point forecast gains documented in Table 7 for the rolling case over the whole out-of-sample horizon. Fig. 9(b) displays the Cumulated Sum of Squared Errors Difference (CSSED)[30] proposed by Welch and Goyal (2007): the forecasting gains of the W_t^{Text} tracker are large and stable throughout the whole out-of-sample evaluation period, when a rolling scheme is used. The overall relative RMSFE gains are explained by the fact that the TB-model consistently

[29] The DM test is not consistent with nested models if the sample is expanding.

[30] The CSSED (Cumulated Sum of Squared Errors Difference) is calculated as $CSSED_\tau = \sum_{\tau=R}^{T}(e^2_{\tau,(\mathrm{baseline})} - e^2_{\tau,(\mathrm{TB})})$, so that positive values show that the TB model outperforms the benchmark in the out-of-sample between time R and T.

Table 7
Relative RMSFE of weekly GDP trackers.

	Expanding	Rolling (335 weeks)
All sample	0.85	0.84^{**}
Negative GDP growth periods	0.87	0.96
Positive GDP growth periods	0.83	0.75^{***}

Note: significance level: * = 10%; ** = 5%; *** = 1%. The table presents the relative RMSFE of the text-based weekly GDP tracker, relative to an equivalent model not incorporating text-based indices. The null hypothesis H_0 : $RMSFE = 1$ has been tested with a Diebold–Mariano (DM) test, with parameters $\tau = 12$. As the DM test is inconsistent with nested models with expanding windows, no testing is reported for the first column.

outperforms the benchmark between 2010 and 2011, and between late 2014 and the end of the sample.

6. Conclusions

We use daily newspaper articles and text-mining techniques to construct sentiment (TESI) and economic policy uncertainty (TEPU) indicators. They match up well with the Italian business cycle phases and with important events shaping them. These high-frequency indicators can be computed in real time, making them suitable candidates for tracking short-run developments in Italian economic activity.

We provide survey evidence that newspaper articles are a relevant source of information for both firms and households for their economic and financial decisions.

We compute the text-based indicators for 36 economic and sector-specific topics, using a dictionary-based approach on a very large corpus of newspaper articles from January 1997 to December 2019. Our dictionary is tailored for the Italian language and for economic news. Unlike most of the literature, our dictionary contains valence-shifting words, which help to interpret the overall meaning of each sentence.

Our indicators prove to be effective in the short-term forecasting of economic activity, both at monthly and weekly frequency. A monthly BMA application shows that our text-based sentiment and uncertainty indicators improve the accuracy of point forecasts during recessions, even though they provide few gains over the whole sample. In addition, we find statistically significant gains when focussing on density forecasts. When used at the weekly frequency, our indicators provide sizeable and statistically significant gains in nowcasting.

Several interesting questions remain to be answered. Disentangling information about the past and the future, as implemented in Angelico et al. (2022) or Byrne et al. (2021), can sharpen the accuracy of text-based indicators. In addition, the application of more sophisticated machine-learning techniques, to refine textual-based indicators and extract robust signals for forecasting, will be a fruitful avenue for future research.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijforecast.2022.02.006.

References

Altig, Dave, Baker, Scott, Barrero, Jose Maria, Bloom, Nicholas, Bunn, Philip, Chen, Scarlet, Davis, Steven J., Leather, Julia, Meyer, Brent, Mihaylov, Emil, Mizen, Paul, Parker, Nicholas, Renault, Thomas, Smietanka, Pawel, & Thwaites, Gregory (2020). Economic uncertainty before and during the COVID-19 pandemic. *Journal of Public Economics*, *191*, 104274.

Amisano, Gianni, & Giacomini, Raffaella (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics*, *25*(2), 177–190.

Angeletos, G. Marios, & Lian, Chen (2016). Incomplete information in macroeconomics: accommodating frictions in coordination. *Vol. 2*, In *Handbook of macroeconomics* (pp. 1065–1240). Elsevier.

Angelico, Cristina, Marcucci, Juri, Miccoli, Marcello, & Quarta, Filippo (2022). Can we measure inflation expectations using Twitter?. *Journal of Econometrics*, *228*(2), 259–277.

Ardia, David, Bluteau, Keven, Borms, Samuel, & Boudt, Kris (2021). The R package sentometrics to compute, aggregate and predict with textual sentiment. *Journal of Statistical Software*, *99*(2), 1–40.

Ardia, David, Bluteau, Keven, & Boudt, Kris (2019). Questioning the news about economic growth: sparse forecasting using thousands of news-based sentiment values. *International Journal of Forecasting*, *35*(4), 1370–1386.

Ardizzi, Guerino, Emiliozzi, Simone, Marcucci, Juri, & Monteforte, Libero (2019). News and consumer card payments. *No. 1233*, In *Bank of Italy temi di discussione (Working paper)*.

Armantier, Olivier, Nelson, Scott, Topa, Giorgio, Van der Klaauw, Wilbert, & Zafar, Basit (2016). The price is right: updating inflation expectations in a randomized price information experiment. *Review of Economics and Statistics*, *98*(3), 503–523.

Ashwin, Julian, Eleni, Kalamara, & Saiz, Lorena (2021). Nowcasting euro area GDP with news sentiment: a tale of two crises. In *ECB working paper*.

Baker, Scott R., Bloom, Nicholas, & Davis, Steven J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, *131*(4), 1593–1636.

Barbaglia, Luca, Consoli, Sergio, & Manzan, Sebastiano (2020). Forecasting with economic news. Available at SSRN.

Bencivelli, Lorenzo, Marcellino, Massimiliano, & Moretti, Gianluca (2017). Forecasting economic activity by Bayesian bridge model averaging. *Empirical Economics*, *53*(1), 21–40.

Blei, David M., Ng, Andrew Y., & Jordan, Michael I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*(Jan), 993–1022.

Bok, Brandyn, Caratelli, Daniele, Giannone, Domenico, Sbordone, Argia M, & Tambalotti, Andrea (2018). Macroeconomic nowcasting and forecasting with big data. *Annual Review of Economics*, *10*, 615–643.

Bybee, Leland, Kelly, Bryan T., Manela, Asaf, & Xiu, Dacheng (2021). Business news and business cycles. In *Working paper 29344*, National Bureau of Economic Research.

Byrne, David, Goodhead, Robert, McMahon, Michael, & Parle, Conor (2021). *The Central Bank Crystal Ball: Temporal information in monetary policy communication*. University of Oxford.

Chow, Gregory C., & Lin, An-loh (1971). Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The Review of Economics and Statistics*, 372–375.

Coibion, Olivier, Gorodnichenko, Yuriy, & Weber, Michael (2021). Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy, forthcoming*, 1325–1369.

Delle Monache, Davide, Emiliozzi, Simone, & Nobili, Andrea (2021). Tracking economic growth during the covid-19: a weekly indicator for italy. *Bank of Italy, Covid-19 Note*.

Diebold, Francis X., & Mariano, Roberto S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *13*(3), 253–263.

Donadelli, Michael, Gufler, Ivan, & Pellizzari, Paolo (2020). The macro and asset pricing implications of rising Italian uncertainty: evidence from a novel news-based macroeconomic policy uncertainty index. *Economics Letters*, *197*, 109606.
Eraslan, Sercan, & Götz, Thomas (2021). An unconventional weekly economic activity index for Germany. *Economics Letters*, *204*, 109881.
Fernández, Carmen, Ley, Eduardo, & Steel, Mark F. J. (2001a). Model uncertainty in cross-country growth regressions. *Journal of Applied Econometrics*, *16*(5), 563–576.
Fernández, Carmen, Ley, Eduardo, & Steel, Mark F. J. (2001b). Benchmark priors for bayesian model averaging. *Journal of Econometrics*, *100*(2), 381–427.
Fraiberger, Samuel (2016). News sentiment and cross-country fluctuations. In *Proceedings of the first workshop on NLP and computational social science* (pp. 125–131).
Gentzkow, Matthew, Kelly, Bryan, & Taddy, Matt (2019). Text as data. *Journal of Economic Literature*, *57*(3), 535–574.
Kalamara, Eleni, Turrell, Arthur, Redl, Chris, Kapetanios, George, & Kapadia, Sujit (2020). Making text count: economic forecasting using newspaper text. *Staff working paper 865*, Bank of England.
Kelly, Bryan, Manela, Asaf, & Moreira, Alan (2021). Text selection. *Journal of Business & Economic Statistics*, *39*(4), 859–879.
Lewis, Daniel J., Mertens, Karel, Stock, James H., & Trivedi, Mihir (2021). Measuring real activity using a weekly economic index. *Journal of Applied Econometrics*.
Loughran, Tim, & McDonald, Bill (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, *66*(1), 35–65.
Loughran, Tim, & McDonald, Bill (2014). Measuring readability in financial disclosures. *The Journal of Finance*, *69*(4), 1643–1671.
Mackowiak, Bartosz, Matejka, Filip, & Wiederholt, Mirko (2021). Rational inattention: a review. In *ECB working paper*.
Madigan, David, & York, Jeremy (1995). Bayesian graphical models for discrete data. *International Statistical Review*, *63*, 215–232.
Mikolov, Tomas, Sutskever, Ilya, Chen, Kai, Corrado, Greg S., & Dean, Jeff (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, *26*, 3111–3119.
Ng, Serena, & Wright, Jonathan H. (2013). Facts and challenges from the great recession for forecasting and macroeconomic modeling. *Journal of Economic Literature*, *51*(4), 1120–1154.
Pennington, Jeffrey, Socher, Richard, & Manning, Christopher D. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532–1543).
Rambaccussing, Dooruj, & Kwiatkowski, Andrzej (2020). Forecasting with news sentiment: evidence with uk newspapers. *International Journal of Forecasting*, *36*(4), 1501 – 1516.
Rogers, John, & Xu, Jiawen (2019). How well does economic uncertainty forecast economic activity?. *Finance and economics discussion series 2019-085*, Board of Governors of the Federal Reserve System.
Rondinelli, Concetta, & Zanichelli, Francesca (2021). The main results of the sixth wave of the special survey of italian households. *Bank of Italy, Covid-19 Note*.
Shapiro, Adam Hale, Sudhof, Moritz, & Wilson, Daniel (2020). Measuring news sentiment. *Journal of Econometrics*.
Sims, Christopher A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, *50*(3), 665–690.
Soroka, Stuart N. (2006). Good news and bad news: asymmetric responses to economic information. *The Journal of Politics*, *68*(2), 372–385.
Tetlock, Paul C. (2007). Giving content to investor sentiment: the role of media in the stock market. *The Journal of Finance*, *62*(3), 1139–1168.
Thorsrud, Leif Anders (2018). Words are the new numbers: a newsy coincident index of the business cycle. *Journal of Business & Economic Statistics*, 1–17.
Welch, Ivo, & Goyal, Amit (2007). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, *21*(4), 1455–1508.