THE NARRATIVE ABOUT THE ECONOMY AS A SHADOW FORECAST: AN ANALYSIS USING BANCO DE ESPAÑA QUARTERLY REPORTS

2020

BANCO DE ESPAÑA
Eurosistema

Documentos de Trabajo
N.º 2042

Nélida Díaz Sobrino, Corinna Ghirelli, Samuel Hurtado,
Javier J. Pérez and Alberto Urtasun

THE NARRATIVE ABOUT THE ECONOMY AS A SHADOW FORECAST: AN ANALYSIS USING BANCO DE ESPAÑA QUARTERLY REPORTS

THE NARRATIVE ABOUT THE ECONOMY AS A SHADOW FORECAST: AN ANALYSIS USING BANCO DE ESPAÑA QUARTERLY REPORTS (*)

Nélida Díaz Sobrino (**)

UNIVERSIDAD NEBRIJA

Corinna Ghirelli (***), Samuel Hurtado (****), Javier J. Pérez (*****)
and Alberto Urtasun (******)

BANCO DE ESPAÑA

(*) We thank Laurent Ferrara, Gergely Gánics, Ángel Gavilán, José González, Malte Knueppel, Andrew Martínez, Enrique Moral, Diego Torres, and Ernesto Villanueva for their comments and suggestions. We also thank participants at the internal seminar of the Banco de España (September 2020) and the International Symposium on Forecasting (October 2020) for comments. The views expressed in this paper are the authors' and do not necessarily reflect those of the Banco de España or the Eurosystem.
(**) Universidad Nebrija, Calle de Santa Cruz de Marcenado, 27, 28015 Madrid, Spain.
(***) Corresponding author: Economic Developments Department, Banco de España, Calle Alcalá 48, 28014 Madrid, Spain. E-mail: corinna.ghirelli@bde.es.
(****) Economic Developments Department, Banco de España, Calle Alcalá 48, 28014 Madrid, Spain. E-mail: samuel.hurtado@bde.es.
(*****) International Economics and Euro Area Department, Banco de España, Calle Alcalá 48, 28014 Madrid, Spain. E-mail: javierperez@bde.es.
(******) Economic Developments Department, Banco de España, Calle Alcalá 48, 28014 Madrid, Spain. E-mail: aurtasun@bde.es.

Documentos de Trabajo. N.º 2042

2020

The Working Paper Series seeks to disseminate original research in economics and finance. All papers have been anonymously refereed. By publishing these papers, the Banco de España aims to contribute to economic analysis and, in particular, to knowledge of the Spanish economy and its international environment.

The opinions and analyses in the Working Paper Series are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem.

The Banco de España disseminates its main reports and most of its publications via the Internet at the following website: http://www.bde.es.

ISSN: 1579-8666 (on line)

Abstract

The aim of this paper is to construct a text-based indicator that reflects the sentiment of the Banco de España economic outlook reports. Our sentiment indicator mimics very closely the first release of the GDP growth rate, which is published after the publication of the reports, and the Banco de España quarterly forecasts of the GDP growth rate. This means that the qualitative narrative contained in the reports contains similar information to the one conveyed by the quantitative forecasts. In addition, the narrative complements the quantitative projections by discussing information which is not directly reflected in the point forecasts.

Keywords: textual analysis, sentiment analysis, GDP growth rate, forecasting, central bank reports.

JEL classification: C53, C55, E37, E66, E58.

Resumen

El objetivo de este trabajo es construir un indicador que mide el sentimiento sobre la coyuntura económica que se desprende del texto de los ejemplares del *Boletín Económico* publicados por el Banco de España. Este indicador de sentimiento evoluciona de manera muy parecida a la primera estimación del PIB y a las previsiones realizadas por el Banco de España. Esto implica que la narrativa cualitativa de los informes contiene información parecida a la de las proyecciones cuantitativas. Además, la narrativa complementa las previsiones cuantitativas, porque incorpora también información adicional, que no se utiliza en sí para estimar las previsiones.

Palabras clave: análisis textual, análisis de sentimiento, tasa de crecimiento del PIB, previsiones, informes de los bancos centrales.

Códigos JEL: C53, C55, E37, E66, E58.

1 Introduction

Central bank forecasts are usually published along with some explanatory text. This text may contain additional qualitative information that enriches the point forecast, e.g., a discussion of the uncertainty underlying the point forecast or an assessment of the key factors driving the forecast. Hence, it is the entire package that conveys the full forecast exercise. Each quarter since 1999, the Bank of Spain has published an economic outlook of the Spanish economy. The content of the report has changed over time. Until 2006, the report amounted to a qualitative description of the state of the economy and a qualitative risk assessment, i.e., it contained only text. More recently, it also started to provide quantitative projections for the main macroeconomic variables: once per year from 2007, and in each quarter from 2014 onwards. However, although not published, the Bank keeps a record of the full series of quantitative projections produced since 1999.

In this paper, we build an index that reflects the sentiment of the Bank of Spain's economic outlook reports by means of sentiment analysis. Sentiment analysis in economics is growing fast, and a number of recent papers have studied the information contained in other central banks reports by means of similar exercises. Clements and Reade (2016) study text published in the Bank of England Inflation Reports, which contains numerical forecasts of inflation and output growth. They find that the text-based sentiment indicator can improve the output growth numeric forecast. Moreover, changes in text-based sentiment can predict changes in the short-horizon numeric forecast. Similarly, Jones et al. (2019) focus on the narratives related to the output growth forecasts published in the Bank of England quarterly Inflation Reports, convert them into numeric scores, and compare these to real-time output growth data as well as to the corresponding quantitative projections published by the bank. They find that the narratives correctly reflect the overall state of the UK economy. Moreover, the text conveyed clearer signs of the incoming Great Recession than the quantitative forecasts did.

For the US, Stekler and Symington (2016) convert the minutes of the Federal Open Market Committee (FOMC) into a numeric indicator that reflects optimism and pessimism in the FOMC's outlook for the US economy. That is, they translate the qualitative forecasts made by the FOMC into a quantitative index. Then, they focus on how the FOMC evaluated the Great Recession when it was happening. They find that the FOMC had information that could have allowed its prediction in advance, but did not foresee the possibility that the recession would materialize. Ericsson (2016) complements the previous paper and shows that Stekler and Symington's sentiment index is very close to the GDP growth rate forecasts released in the Greenbook, the Federal Reserve Board (Fed)'s forecasts publication. Since the latter are published 5 years after they are presented to the FOMC, whereas the minutes of the FOMC meetings are published 3 weeks after the meeting itself, the index proves to be an extremely powerful tool to know the FOMC's forecasts in advance. Similarly, Sharpe et al. (2017) construct a sentiment indicator to measure the tone (optimistic/pessimistic) of the narratives describing the Fed's forecasts (Greenbook). The resulting text-based indicator closely mimics the point forecast for unemployment and inflation. In addition, they find that this index has predictive power for GDP growth and unemployment up to four quarters in advance.[1]

Finally, Catalfamo (2018) uses sentiment analysis to assess the ability of the Bank of France's staff and of the French media (one French economic newspaper) to nowcast the state of the US economy correctly at the time of the Great Recession. She focuses on the Bank of France Annual Reports and speeches, as well as on press articles on the US economy released in that period. Results suggest that both the French central bank and the media understood well the state of the US economy during the Great Recession.

To carry out our analysis, we rely on the Bank of Spain's Quarterly Economic Bulletin on the Spanish economy, published online since the first quarter of 1999. We follow the aforementioned literature and create a dictionary of positive and negative terms in Spanish that are typically used in the economic language to describe the economy, e.g., positive terms like *grow, strong, increase* and negative ones like *decrease, reduction.* We assign weights to the resulting counts of positive and negative terms. Then, for each report we sum up the weighted counts of terms in the dictionary and divide the resulting number by the length of the report. Then, we compare the resulting text-based index with the GDP series ("flash estimate") and the GDP growth projections produced each quarter by the Bank of Spain, which until 2006 were recorded internally and from 2007 started to be published.

We find a significant dynamic relationship between these series: the narrative text-based indicator follows the Spanish cycle and increases or decreases when the quantitative projections do. Note, the flash estimate of the GDP series is published after the publication of the report, so that at the time of writing each report the value of the GDP for the current quarter is still unknown. By contrast, the GDP projections produced by the Bank in the current quarter are known. In addition, the comparison shows that the reports are informative not only at the short-term forecast horizon, but, even more so, at the one-to-two year forecast horizon. The textual indicator shows the highest correlation with next quarter's projections. This is consistent with the hypothesis that the narrative of the reports anticipates a discussion on potential risks that will actually materialize, to some extent, and will be included in GDP numerical projections of the next quarter. This means that a "sophisticated" reader could infer GDP growth projections based on the text of the reports, somewhat beyond what is told in just the numbers, condesing as well the (forward-looking) risks assessment.

The rest of the paper is organized as follows. Section 2 describes the construction of our text-based index, while in Section 3 we present the data on the quantitative GDP projections

[1]Sinclair et al. (2010) studied the Fed's forecasts released in the Greenbook and concluded that the Fed knows the state of the economy for the current quarter but cannot predict it one-quarter ahead. Romer and Romer (2008) studied the set of information based upon which the FOMC sets US monetary policy. In particular, they compared the FOMC forecasts (forecasts produced by the Board of Governors staff before each FOMC meeting) with the Fed staff's forecasts (Greenbook) and conclude that the latter contain more information than the FOMC forecasts. They claim that the fact that the Board of Governors staff produces its own forecasts is at the very least inefficient, and that optimal monetary policy should rely on the Fed staff forecasts.

of the Bank of Spain. In Section 4, we compare our index with the bank's GDP projections. Finally, Section 5 sums up our findings and draws some conclusions.

2 Building the index

We consider the Quarterly Economic Bulletin on the Spanish economy published online (pdf format) since the first quarter of 1999. Specifically, we consider the overview section. With this information we construct a sentiment indicator from Q1 1999.

First, we create a dictionary of positive and negative terms in Spanish. To come up with this list of words, we read a sample of the reports under analysis and identify the terms that are most commonly used to describe the macroeconomic situation (considering adjectives, adverbs, verbs, and nouns).[2] We select 47 terms, some of which are roots, i.e., we have removed word endings.[3]

Second, we manually score the terms that express positive and negative sentiment with a value ranging between $[-1.5, 1]$. This scoring is based our own judgment and takes into account the nuances of the economic language of the Bank of Spain's reports. We cluster the selected words in our dictionary according to whether they are typically used to describe the following situations: strong growth ($value = 1$), normal growth (0.75), modest growth (0.5), slow growth (0.25), slowdown (-0.25), stagnation (-0.5), downturn (-0.75), recession (-1), strong recession (-1.5).[4] Our dictionary of positive and negative words is reported in Table B.1 in Section B of the Appendix.

Third, we count how many times each word in the dictionary appears in each text and weight each result with its scoring.

Fourth, in order to control for incorrect signs, we neutralize the terms in our dictionary around the words *unemployment*, *employment*, or *deficit*. That is, when one of these words appears in the text, whatever term appears within up to 9 words before or 9 words after is not considered in our counting. This is to avoid interpreting as positive the tone of words that are scored with a positive sign when they are actually associated with negative phenomena, and viceversa (e.g. the term *increasing* has a positive score in our dictionary, but when in the text it is associated with *unemployment*, its meaning is negative.).

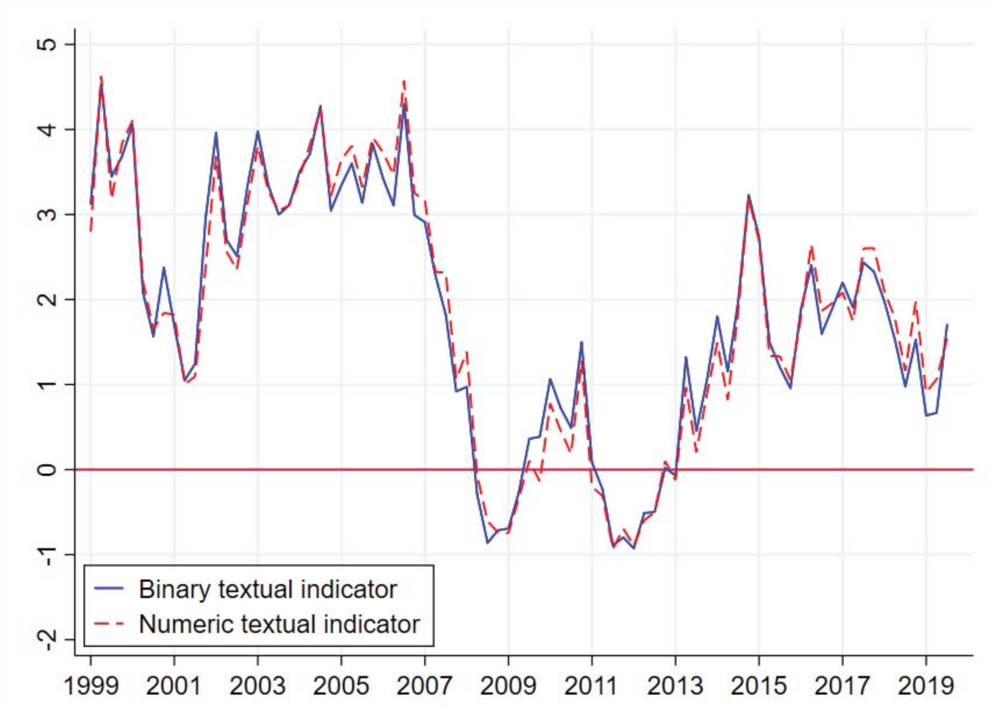
Fifth, we sum all weighted occurrences in each text and divide the resulting number by the total length of the text.

[2] All words we consider appear in at least 8% of the documents of the corpus. In addition, on average, the selected terms appear in more than 50% of the documents, i.e. we focus on terms that are very commonly used to describe the economic situation. This is in line with - but considerably more demanding than - what is done in the literature: for instance, Loughran and McDonald (2011) consider terms that appear in more than 5% of the documents.

[3] One root can identify several terms simultaneously. For instance, the root *crec* refers to the noun *crecimiento* (growth), the adjective *creciente/es* (growing), and the verb *crecer* (to grow).

[4] Our scoring is asymmetric. We attribute a value of -1.5 to terms describing *strong recessions* because these are used very carefully in the reports. For all other categories, the scoring ranges between $[-1, 1]$. A similar approach has been used in Catalfamo (2018).

Figure 1: Text-based sentiment indicators



Note: The dashed red line represents the numeric text-based indicator, in which the terms in the dictionary are weighted by values ranging between $[-1.5, 1]$. The solid blue line represents the binary text-based indicator, in which positive terms are weighted by $+1$ and -1, respectively. Both indicators are rescaled so that they have the same mean as the quarterly composite forecast for the GDP growth rate at the one-year horizon (this series is presented in Section 3).

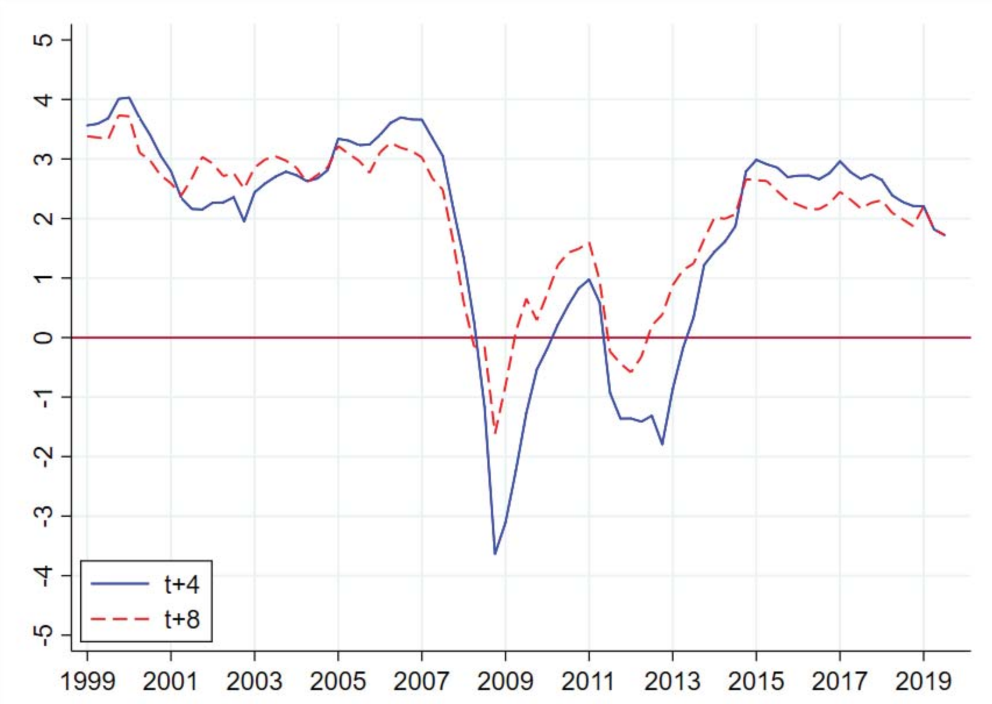
Finally, to ensure that our results are not driven by our scoring values, we construct an alternative sentiment index in which the words that express a positive sentiment are scored with $+1$, while those that are associated with a negative tone are scored with -1. The resulting indexes are depicted in Figure 1 below. This comparison reveals that they are practically identical when using the difference between the number of positive terms and the number of negative terms (i.e. the second weighting scheme based on $+1$ and -1).[5]

3 The data

To evaluate the quality of our text-based sentiment indicators, we use both the observed GDP growth rate (flash series) as well as the Bank of Spain's yearly predictions, which are produced each quarter. The former is published by the Spanish National Institute of Statistics (INE) approximately 30 days after the end of the quarter, i.e. the first estimate of the GDP for the current quarter is not known when the report is written and published. By contrast, the projections have been released to the public in each quarter only starting

[5]There exists other techniques to construct the dictionary, spanning from term frequency to bootstrapping. Our approach is closer to the former, since we consider words that appear very frequently in the corpus and assign them a polarity based on their meaning.

Figure 2: Composite forecast indicators



Note: $t + 4$ indicates the quarterly composite forecast at the one-year-ahead horizon. $t + 8$ indicates the quarterly composite forecast at two-years-ahead horizon. The construction of both indicators is explained in Section 3 of the main text. The y-axis measures the GDP growth rate (in percentage points).

from 2014. From 2007 to 2013, the numeric point forecasts were published only once per year. Before that they were not published at all. Nevertheless, the Bank keeps records of all results from the quarterly forecast exercises ever produced.

We use this confidential information for this exercise. In particular, we consider the quarterly predictions for the yearly GDP growth rate for the current year (nowcast), for the next year (one-year-ahead forecast), and for 2 years ahead; we combine them so as to obtain a proxy of a fixed-horizon forecast: a quarterly composite forecast at one-year- and two-years-ahead horizons ($t + 4$ and $t + 8$, respectively). In doing this, we follow the procedure in Dovern et al. (2012) and approximate fixed horizon forecasts as a weighted average of fixed event forecasts. In line with this literature, we use the following weighting scheme, for each quarter: in quarter 1 of the year, to construct the $t + 4$ composite forecast, we assign 0.75 to the nowcast and 0.25 to the one-year-ahead forecast; in quarter 2, we assign 0.5 to each prediction; in quarter 3, we assign 0.25 to the nowcast and 0.75 to the forecast; in quarter 4, we assign 1 to the nowcast and zero to the one-year-ahead forecast. To construct the $t + 8$ composite forecast, we apply the same weighting scheme to the two-years-ahead horizon forecasts. The resulting forecast series are shown in Figure 2. For consistency, we construct the observed GDP series applying the same weighting scheme,[6] so that all series are comparable.

[6]We use the yearly flash estimates of the GDP growth rate.

Figure 3 reports both text-based indicators (solid blue and black line, respectively) against the first estimate of the GDP growth rate (dashed red line). This evidence suggests that the narrative reflected in the text of the economic bulletins by the Bank of Spain follows very closely the observed GDP growth rate, which is not known at the time when the report is written.

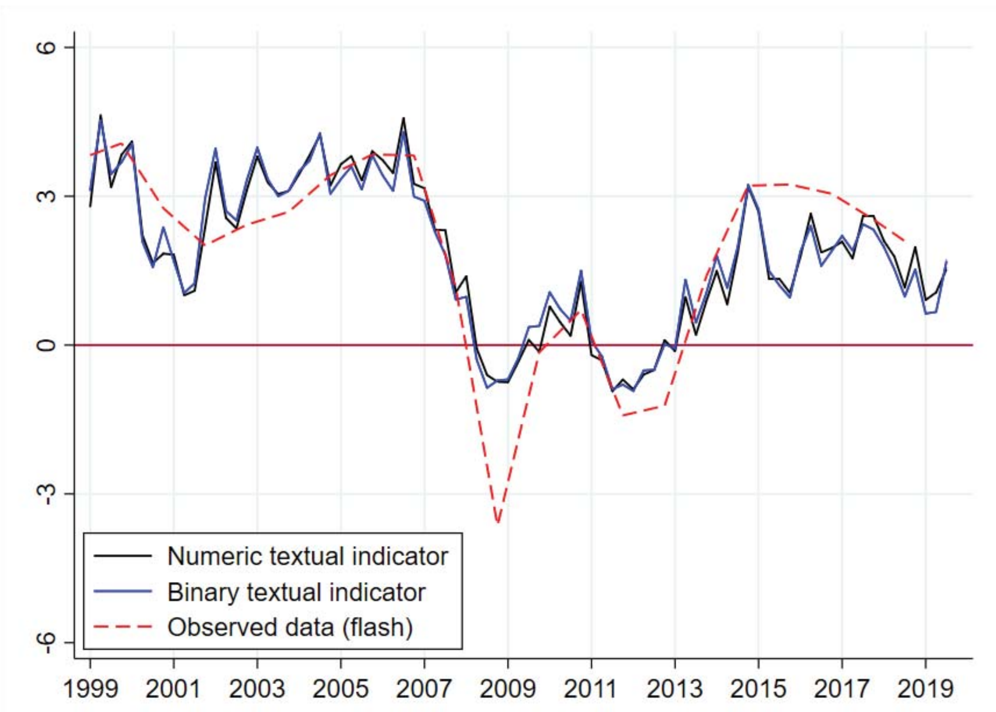
4 Empirical analysis

In this section, we compare our sentiment indicators with the quantitative projections that the Bank of Spain produces on a quarterly basis.

As a first piece of evidence, we compute the correlation between the sentiment indicators and the $t+4$ and $t+8$ composite forecast series. We consider their contemporaneous relation but also how lagged and future values of each forecast series correlate with the sentiment indicators. The correlation values are shown in Figures 4 and 5 for the $t+4$ and $t+8$ composite forecast series, respectively.

In each figure, the solid black line represents the correlation of the forecast series with itself. The correlation is 1 when the variable is correlated with itself contemporaneously (this corresponds to the third category on the x-axis). The categories on the x-axis to the right and left of the contemporaneous one show the correlation of the composite forecast series with its own lagged and future values, respectively. Similarly, the black dotted line refers to

Figure 3: Text-based indicators vs. observed GDP (flash series)



Note: The graph shows the binary and numeric text-based indicators (solid blue and black line, respectively) against the first estimate of the GDP growth rate (dashed red line). The y-axis measures the GDP growth rate (in percentage points).

the flash series of the GDP growth rate (i.e. the first publication of the estimated annual value of the GDP growth rate by the Spanish National Statistical Office), and in particular, it depicts the correlation of the flash series with its own contemporaneous terms, lags, and leads. Finally, the solid blue line represents the correlation of the binary sentiment indicator with contemporaneous terms, lags, and leads of the composite forecast series (Figures 4 and 5 consider the $t+4$ and $t+8$ composite forecasts, respectively), while the solid red line represents the same statistics for the numeric sentiment indicator. The exact values of these pairwise correlation coefficients are reported in Section A of the Appendix (Tables A.1 and A.2 for the $t+4$ and $t+8$ composite forecasts, respectively).

In Figure 4, both sentiment indicators correlate the most with the $t+4$ composite forecast when they are leading by one quarter: this suggests that the text embeds information that is related to the next quarter's forecast for the yearly GDP growth rate (e.g., the 1999q1 report discusses elements that will be used to obtain the quantitative projection, produced in 1999Q2, for the GDP growth rate for 1999q2–2000q2). Similarly, Figure 5 suggests that the sentiment indicators are mostly correlated with the contemporaneous $t+8$ composite predictions (e.g., the 1999q1 discusses information that is used to obtain the annual prediction, produced in 1999q1, for the GDP growth rate for 2000q1–2001q1).[7] Hence, the

Figure 4: Correlations between sentiment indicators and the $t+4$ composite forecast

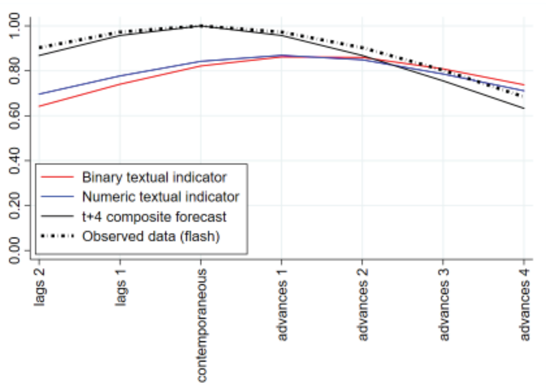
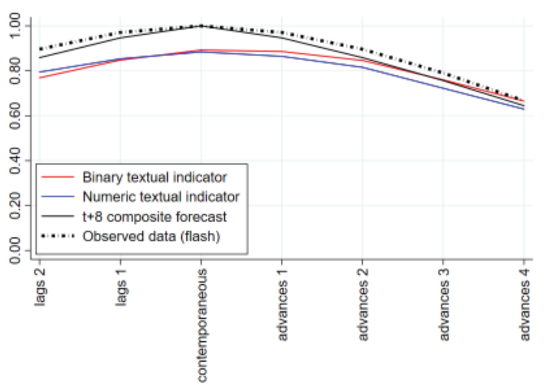


Figure 5: Correlations between sentiment indicators and the $t+8$ composite forecast



Note: Figure 4 (5) shows the correlation between the sentiment indicators and the numeric $t+4$ ($t+8$) composite forecast series of the Bank of Spain. The x-axis orders the correlation of each indicator with different versions of the composite forecast series: its lagged values, its contemporaneous values, and its future values. The solid blue (red) line represents the correlation of the binary (numeric) sentiment indicator with the contemporaneous terms, lags, and leads of the forecast series (Figures 4 and 5 consider the $t+4$ and $t+8$ forecast series, respectively). In both Figures, the solid black line represents the correlation of the forecast series with its own contemporaneous terms, lags, and leads, while the dotted black line depicts the correlation of the flash estimate of the GDP growth rate with its own contemporaneous terms, lags, and leads.

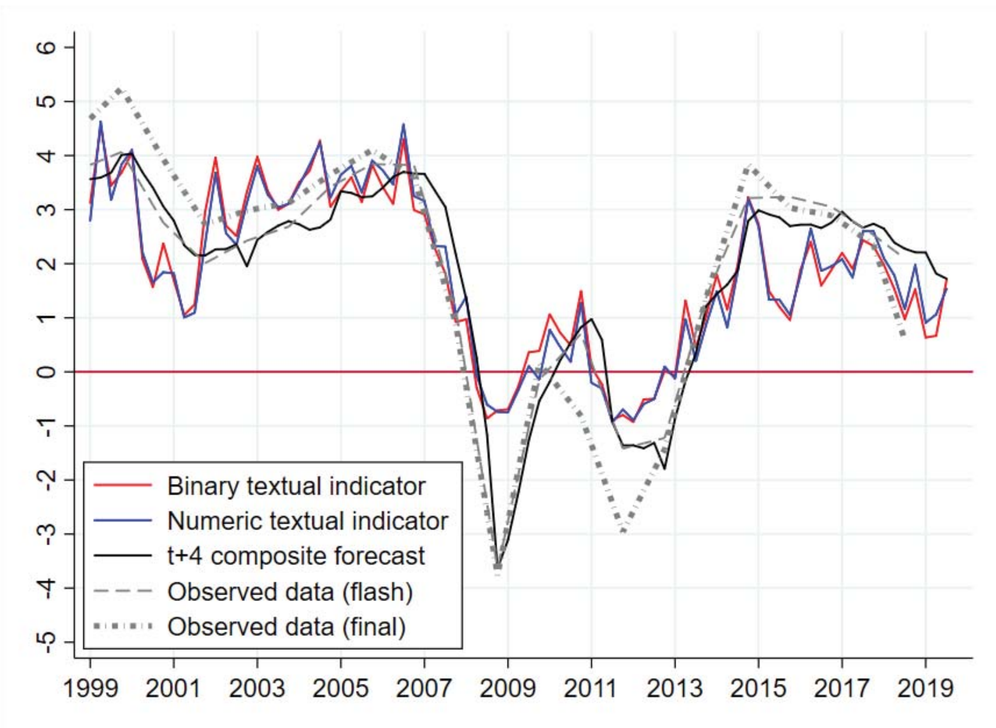
[7] As a robustness exercise, we compute the correlation between the sentiment indicators and the Bank of Spain's quantitative projections by means of simple regressions, which allow us to obtain the significance of the estimates. The correlations between the sentiment indicators and the quantitative projections are always significant at the 1% level, and their magnitude confirms the aforementioned discussion. Results are reported in Section A of the Appendix (see Table A.3 and A.4 for the $t+4$ and $t+8$ composite forecast, respectively).

reports are informative not only at the short-term forecast horizon but even more so at the one-to-two-year forecast horizon. In other words, the narrative of the reports also tends to discuss about possible future risks that will be included in the next projections, although this evidence is not statistically significant.[8] However, in a next exercise we do find strong evidence that the narrative contains information about the future expected evolution of the GDP growth rate that is not included in the quantitative forecast. To do that, we regress the sentiment indicators on the GDP growth rate controlling for the quantitative forecast and find a significant relation between the tone of the text and the GDP growth rate (see column 3 and 4 of Table 1).

Overall, this comparison suggests that the qualitative narrative contained in the reports contains similar information to that conveyed by quantitative forecasts. A "sophisticated" reader could infer GDP growth projections based on the text of the reports.

As a second exercise, we compare the sentiment indicators with the observed data (both the flash and the final estimation released by the Spanish National Statistical Office) and the composite forecast series of the Bank of Spain (see Figure 6 and 7 for $t+4$ and $t+8$,

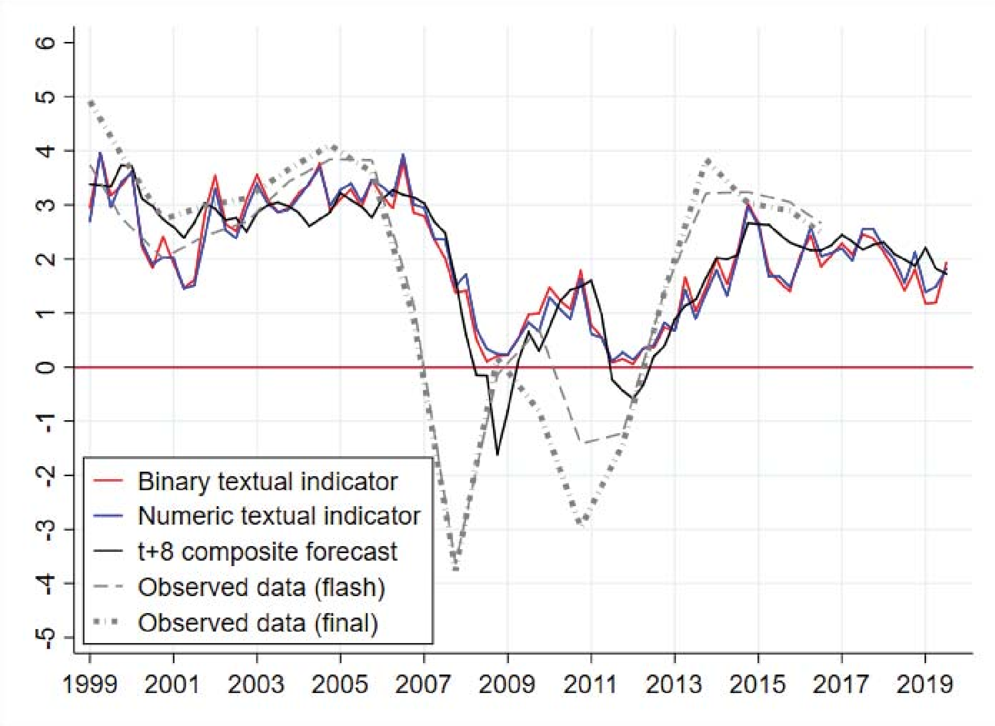
Figure 6: Our text-based indicators vs the $t+4$ composite forecasts



Note: The solid red (blue) line represents the binary (numeric) textual indicator. The solid black line represents the $t+4$ composite forecast series produced by the Bank of Spain. The dashed (dotted) gray line depicts the flash (final) estimate of the GDP growth rate data released by the Spanish National Statistical Office.

[8]For both horizons (i.e. $t+4$ and $t+8$), the correlations between the sentiment indicators and the contemporaneous projections are not statistically different from the correlations between the sentiment indicators and the next quarter's projections.

Figure 7: Our text-based indicators vs. the $t+8$ composite forecasts



Notes: The solid red (blue) line represents the binary (numeric) textual indicator. The solid black line represents the $t+4$ composite forecast series produced by the Bank of Spain. The dashed (dotted) gray line depicts the flash (final) estimate of the GDP growth rate data released by the Spanish National Statistical Office.

respectively). We find a strong dynamic relationship between our text-based indicators and the composite forecasts. That is, the sentiment indicators follow the Spanish cycle and increase or decrease when the quantitative projections do. The text is not as good at reflecting the composite forecasts during downturns. This is in line with the hypothesis that the Bank staff is very prudent during recessions and uses moderate language. There are two periods in which it seems that the text of the report shows some marked deviation from actual cyclical developments. The first is during the international Dot–com crisis of 2000-2001, which hit the US and some of the countries in the Euro Area, but not Spain. The reports covering this period were very cautious, despite the fact that Spain was not affected by the crisis, given the overall economic situation. A second example refers to a period from 2015 onwards, in which Spain started recovering from the financial crisis, at a faster pace than other European countries. Despite the composite forecasts signaling a period of positive growth, the Bank's reports maintained a moderated tone, suggesting that the Bank's staff was held back by doubts about the strength of the Spanish recovery, considering the state of the economy in the rest of the Euro Area.

In a third exercise, we ask whether the sentiment indicators have predictive power. We estimate a simple AR(1) model for the GDP growth rate adding one sentiment indicator at a time to the right-hand side of the equation. We measure the GDP growth rate using the flash series.[9] The coefficients associated with the sentiment indicators are positive and significant at the 1% level (see column 1 and 2 of Table 1). This confirms that the information contained in the text of the quarterly reports is useful to nowcast the GDP growth rate. In addition, in column 3 and 4 of Table 1 we regress the GDP growth rate on a constant, the $t + 4$ composite projection, and one sentiment at a time. Both sentiment indicators are positive and significant at the 1% level when controlling for the projection, which suggests that the narrative embeds additional information which is not contained in the quantitative forecast.

Table 1: Testing the predictive power of sentiment indicators

	GDP growth rate		GDP growth rate		$t + 4$ forecast error	
	(1)	(2)	(3)	(4)	(5)	(6)
Num. text indicator	0.185***		0.198***		0.065*	
	(0.058)		(0.064)		(0.035)	
Bin. text indicator		0.230***		0.234***		0.081**
		(0.054)		(0.061)		(0.035)
Lagged GDP growth	x	x				
$t + 4$ composite forecast			x	x		
Constant	x	x	x	x	x	x
Observations	78	78	79	79	79	79
R-squared	0.952	0.957	0.944	0.947	0.043	0.063

Note: The GDP growth rate is the flash release. The t+4 composite forecast error is the difference between the flash release of the GDP growth rate and the t+4 composite forecast error. Column 1 and 2 report the results from regressing the GDP growth rate on a constant, its lag and one sentiment indicator at the time (the numeric and the binary textual indicator, respectively). Column 3 and 4 result from regressing the GDP growth rate on a constant, the t+4 composite forecast, and one sentiment indicator at the time. Column 5 and 6 result from regressing the t+4 composite forecast errors to a constant and one sentiment indicator at the time. $***$, $**$, $*$ significant at the 1%, 5%, and 10% level.

Finally, we check whether the sentiment indicators explain the $t + 4$ composite forecast error, defined as the difference between the observed values (the flash data for the GDP growth rate) and the predicted ones (the t+4 composite forecast). Results are reported in column 2 of Table 1. The estimated relationships suggest that the qualitative information embedded in the quarterly reports is positively associated with the forecast error (significant at the 5% and 10% levels for the binary and numeric text-based indicators, respectively). This means that when the text of the report is particularly positive, the forecast error is also likely to be positive (that is, the observed value is more likely to be higher than the forecast). On average, the dependent variable in this sample is positive, i.e., the prediction tends to underestimate the GDP growth rate. The results from the estimation exercise suggest that this conservative bias is stronger for the quantitative figures of the forecast than for the text that describes them. In other words, the narrative compensates for the conservative bias of the quantitative forecast, so that the entire package of the report (narrative plus projections) convey a more precise picture of the expected GDP growth rate.

[9]Results are similar if we use the GDP growth rate revised series.

5 Conclusions

In this paper we construct a text-based sentiment indicator that reflects the narratives of the Bank of Spain's Quarterly Economic Bulletin on the Spanish economy and compare it with both the observed GDP growth rate and the quantitative forecasts produced by the Bank of Spain. In a nutshell, we first calculate the frequency of positive and negative terms in the text and then define the sentiment index as the difference (with some weights) between the two frequencies. That is, a text has a positive (negative) sentiment when the frequency of positive terms is higher (or lower) than that of the negative terms. Then, we compare the sentiment indicator with the Bank of Spain's quarterly composite forecasts of the GDP growth rate at one- and two-year horizons, which is produced at the same time when the report is written, and the observed GDP growth rate series (first release), which is published after the publication of the report and hence it is not known at the time of writing.

We find that the narrative reflected in the text of the reports by the Bank of Spain follows very closely the evolution of the observed GDP growth rate, and hence the Spanish business cycle. All in all, this evidence suggests that the text of the report is consistent with the underlying story told by the institution's GDP growth projections. This means that a "sophisticated" reader could infer GDP growth projections based on the text of the reports. In addition, not only the narrative is consistent with the quantitative projections, but it also complements them by adding new information (e.g. for instance, a discussion related to the international economic context, when the latter differs from the Spanish economic conjuncture). Furthermore, while the quantitative bias tends to underestimate the GDP growth rate especially during upturns, the narrative tends to outweigh this conservative bias. Thus, the combination of quantitative forecast and narrative providess a more precise picture of the expected GDP growth rate.

References

Catalfamo, E. (2018). French Nowcasts of the US Economy during the Great Recession: A Textual Analysis. Working Papers 2018–001, The George Washington University, Department of Economics, Research Program on Forecasting.

Clements, M. P. and J. J. Reade (2016, November). Forecasting and Forecast Narratives : The Bank of England Inflation Reports November 2016. Working Paper ICM–2016–10, ICMA Centre, University of Reading.

Dovern, J., U. Fritsche, and J. Slacalek (2012). Disagreement Among Forecasters in G7 Countries. *The Review of Economics and Statistics 94*(4), 1081–1096.

Ericsson, N. R. (2016). Eliciting GDP forecasts from the FOMC's minutes around the financial crisis. *International Journal of Forecasting 32*(2), 571–583.

Jones, J. T., T. Sinclair, and H. Stekler (2019). A textual analysis of the bank of england growth forecasts. Working Papers 2018–005, The George Washington University, Department of Economics, Research Program on Forecasting.

Loughran, T. and B. McDonald (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10–Ks. *The Journal of Finance 66*(1), 35–65.

Romer, C. D. and D. H. Romer (2008, May). The FOMC versus the Staff: Where Can Monetary Policymakers Add Value? *American Economic Review 98*(2), 230–35.

Sharpe, S. A., N. R. Sinha, and C. A. Hollrah (2017). What's the story? a new perspective on the value of economic forecasts. Finance and Economics Discussion Series 2017–107, Board of Governors of the Federal Reserve System (U.S.).

Sinclair, T. M., F. Joutz, and H. Stekler (2010). Can the fed predict the state of the economy? *Economics Letters 108*(1), 28–32.

Stekler, H. and H. Symington (2016). Evaluating qualitative forecasts: The FOMC minutes, 2006–2010. *International Journal of Forecasting 32*(2), 559–570.

Appendix

A Further Results

Table A.1: Pairwise correlations between the text-based indicators and the $t+4$ composite forecast

	Numeric textual indicator	Binary textual indicator	$t+4$ composite forecast	observed data (flash)
Composite forecast lagged by 2 quarters	0.70	0.64	0.87	0.90
Composite forecast lagged by 1 quarter	0.78	0.74	0.96	0.97
Composite forecast contemporaneous	0.84	0.82	1.00	1.00
Composite forecast 1 quarter in advance	0.87	0.86	0.96	0.97
Composite forecast 2 quarters in advance	0.85	0.86	0.87	0.90
Composite forecast 3 quarters in advance	0.79	0.81	0.76	0.80
Composite forecast 4 quarters in advance	0.71	0.74	0.63	0.68

Note: The table shows the pairwise correlations between the sentiment indicators and the $t+4$ composite forecast series of the Bank of Spain that are plotted in Figure 4 of the main text. For the first three columns, each row of the table refers to the correlation of each sentiment indicator with different versions of the $t+4$ composite forecast series: its lagged values, its contemporaneous values, and its future values. The first (second) column refers to the correlation of the numeric (binary) sentiment indicator with contemporaneous terms, lags, and leads of the forecast series. The third column represents the correlation of the $t+4$ composite forecast series with its own contemporaneous terms, lags and leads. The fourth column depicts the correlation of the flash series of the GDP growth rate with its own contemporaneous terms, lags and leads.

Table A.2: Pairwise correlations between the sentiment indicators and the $t + 8$ composite forecast

	Numeric textual indicator	Binary textual indicator	$t+8$ composite forecast	Observed data (flash)
Composite forecast lagged by 2 quarters	0.79	0.77	0.86	0.90
Composite forecast lagged by 1 quarter	0.85	0.85	0.95	0.97
Composite forecast contemporaneous	0.88	0.89	1.00	1.00
Composite forecast 1 quarter in advance	0.86	0.89	0.95	0.97
Composite forecast 2 in advance	0.82	0.85	0.86	0.90
Composite forecast 3 in advance	0.72	0.76	0.76	0.79
Composite forecast 4 in advance	0.63	0.67	0.64	0.67

Note: The table shows the pairwise correlations between the sentiment indicators and the $t + 8$ composite forecast series of the Bank of Spain that are plotted in Figure 5 of the main text. For the first three columns, each row of the table refers to the correlation of each sentiment indicator with different versions of the $t+8$ composite forecast series: its lagged values, its contemporaneous values, and its future values. The first (second) column refers to the correlation of the numeric (binary) sentiment indicator with contemporaneous terms, lags, and leads of the forecast series. The third column represents the correlation of the $t + 8$ composite forecast series with its own contemporaneous terms, lags and leads. The fourth column depicts the correlation of the flash series of the GDP growth rate with its own contemporaneous terms, lags, and leads.

Table A.3: Correlations between the sentiment indicators and $t + 4$ composite forecast

	Numeric textual indicator			Binary textual indicator		
	Coef.	Std.Err.		Coef.	Std.Err.	
Forecast lagged 2 Q	0.65	0.06	***	0.61	0.06	***
Forecast lagged 1 Q	0.57	0.07	***	0.51	0.07	***
Forecast contemporaneous	0.71	0.05	***	0.68	0.05	***
Forecast 1 Q in advance	0.73	0.05	***	0.71	0.05	***
Forecast 2 Q in advance	0.72	0.05	***	0.71	0.05	***
Forecast 3 Q in advance	0.67	0.06	***	0.67	0.06	***
Forecast 4 Q in advance	0.61	0.07	***	0.62	0.06	***

Note: The table shows the correlations between the sentiment indicators and the $t + 4$ composite forecast series of the Bank of Spain. These correlations are obtained by separated regressions, each time regressing the sentiment indicator on a constant and a specific version of the $t + 4$ composite forecast series: its contemporaneous values, its lagged values, or its future values. For each regression we report the estimated coefficient of the relation between the sentiment indicator and the composite forecast and its standard errors; $***$, $**$, and $*$ refer to 10%, 5%, and 1% confidence levels, respectively. The first (second) row reports the correlation between the sentiment indicator and the forecast series lagged by one (two) quarter. The third column represents the correlation between the sentiment indicator and the contemporaneous terms of the composite forecast. The forth (fifth, sixth, seventh) row represents the correlation between the sentiment indicator and the forecast series shifted forward by one (two, three, four) quarter.

Table A.4: Correlations between the sentiment indicators and $t+8$ composite forecast

	Numeric textual indicator			Binary textual indicator		
	Coef.	Std.Err.		Coef.	Std.Err.	
Forecast lagged by 2 Q	0.75	0.05	***	0.74	0.05	***
Forecast lagged by 1 Q	0.68	0.06	***	0.66	0.06	***
Forecast contemporaneous	0.77	0.05	***	0.78	0.05	***
Forecast 1 Q in advance	0.77	0.05	***	0.79	0.05	***
Forecast 2 Q in advance	0.73	0.06	***	0.75	0.05	***
Forecast 3 Q in advance	0.65	0.07	***	0.68	0.07	***
Forecast 4 Q in advance	0.57	0.08	***	0.61	0.08	***

Note: The table shows the correlations between the sentiment indicators and the $t+4$ composite forecast series of the Bank of Spain. These correlations are obtained by separated regressions, each time regressing the sentiment indicator on a constant and a specific version of the $t+4$ composite forecast series: its contemporaneous values, its lagged values, or its future values. For each regression we report the estimated coefficient of the relation between the sentiment indicator and the composite forecast and its standard errors; $***$, $**$, and $*$ refer to 10%, 5%, and 1% confidence levels, respectively. The first (second) row reports the correlation between the sentiment indicator and the forecast series lagged by one (two) quarter. The third column represents the correlation between the sentiment indicator and the contemporaneous terms of the composite forecast. The fourth (fifth, sixth, seventh) row represents the correlation between the sentiment indicator and the forecast series shifted forward by one (two, three, four) quarter.

B Data

Table B.1: Dictionary for sentiment analysis

Positive sentiment			Negative sentiment		
Word	Counts[a]	Value[a]	Word	Counts	Value
dinami	372	1	desequilibrio	82	-0,25
elevad	425	1	fragil	8	-0,25
expansi	361	1	menor ritmo	9	-0,25
fortaleza	130	1	debil	225	-0,5
fuerte	127	1	descen	442	-0,5
robust	13	1	deterior	143	-0,5
aceler	609	0,75	estanc	22	-0,5
crec	1615	0,75	impacto negativo	18	-0,5
increment	357	0,75	inestabilidad	53	-0,5
notable	180	0,75	inferior	164	-0,5
ascendente	22	0,5	ralentiz	92	-0,5
considerable	57	0,5	redu	507	-0,5
estable	148	0,5	retroce	156	-0,5
favorable	286	0,5	perdida de impulso	9	-0,5
impulso	97	0,5	perdida de dinamismo	22	-0,5
mejor	417	0,5	caida	145	-0,75
superior	176	0,5	disminu	185	-0,75
moderad	247	0,25	empeoramiento	43	-0,75
modest	78	0,25	turbulencias	18	-0,75
recuper	487	0,25	desfavorable	28	-0,75
			decrec	16	-0,75
			desaceler	448	-0,75
			contraccion	67	-1
			destru	29	-1
			grav	46	-1
			crisis	106	-1,5
			recesion	13	-1,5

Notes: Authors' own elaboration. This table describes the economic dictionary used for sentiment analysis based on the Quarterly Economic Bulletins published by the Bank of Spain.

[a]: *Counts* refers to the number of times each word appears in the entire corpus of text.

[a]: *value* is the weight attributed to each word, depending on its economic meaning and its strength.

BANCO DE ESPAÑA PUBLICATIONS

WORKING PAPERS

1940 MYROSLAV PIDKUYKO: Heterogeneous spillovers of housing credit policy.

1941 LAURA ÁLVAREZ ROMÁN and MIGUEL GARCÍA-POSADA GÓMEZ: Modelling regional housing prices in Spain.

1942 STÉPHANE DÉES and ALESSANDRO GALESI: The Global Financial Cycle and US monetary policy in an interconnected world.

1943 ANDRÉS EROSA and BEATRIZ GONZÁLEZ: Taxation and the life cycle of firms.

1944 MARIO ALLOZA, JESÚS GONZALO and CARLOS SANZ: Dynamic effects of persistent shocks.

1945 PABLO DE ANDRÉS, RICARDO GIMENO and RUTH MATEOS DE CABO: The gender gap in bank credit access.

1946 IRMA ALONSO and LUIS MOLINA: The SHERLOC: an EWS-based index of vulnerability for emerging economies.

1947 GERGELY GANICS, BARBARA ROSSI and TATEVIK SEKHPOSYAN: From Fixed-event to Fixed-horizon Density Forecasts: Obtaining Measures of Multi-horizon Uncertainty from Survey Density Forecasts.

1948 GERGELY GANICS and FLORENS ODENDAHL: Bayesian VAR Forecasts, Survey Information and Structural Change in the Euro Area.

2001 JAVIER ANDRÉS, PABLO BURRIEL and WENYI SHEN: Debt sustainability and fiscal space in a heterogeneous Monetary Union: normal times vs the zero lower bound.

2002 JUAN S. MORA-SANGUINETTI and RICARDO PÉREZ-VALLS: ¿Cómo afecta la complejidad de la regulación a la demografía empresarial? Evidencia para España.

2003 ALEJANDRO BUESA, FRANCISCO JAVIER POBLACIÓN GARCÍA and JAVIER TARANCÓN: Measuring the procyclicality of impairment accounting regimes: a comparison between IFRS 9 and US GAAP.

2004 HENRIQUE S. BASSO and JUAN F. JIMENO: From secular stagnation to robocalypse? Implications of demographic and technological changes.

2005 LEONARDO GAMBACORTA, SERGIO MAYORDOMO and JOSÉ MARÍA SERENA: Dollar borrowing, firm-characteristics, and FX-hedged funding opportunities.

2006 IRMA ALONSO ÁLVAREZ, VIRGINIA DI NINO and FABRIZIO VENDITTI: Strategic interactions and price dynamics in the global oil market.

2007 JORGE E. GALÁN: The benefits are at the tail: uncovering the impact of macroprudential policy on growth-at-risk.

2008 SVEN BLANK, MATHIAS HOFFMANN and MORITZ A. ROTH: Foreign direct investment and the equity home bias puzzle.

2009 AYMAN EL DAHRAWY SÁNCHEZ-ALBORNOZ and JACOPO TIMINI: Trade agreements and Latin American trade (creation and diversion) and welfare.

2010 ALFREDO GARCÍA-HIERNAUX, MARÍA T. GONZÁLEZ-PÉREZ and DAVID E. GUERRERO: Eurozone prices: a tale of convergence and divergence.

2011 ÁNGEL IVÁN MORENO BERNAL and CARLOS GONZÁLEZ PEDRAZ: Sentiment analysis of the Spanish Financial Stability Report. (There is a Spanish version of this edition with the same number).

2012 MARIAM CAMARERO, MARÍA DOLORES GADEA-RIVAS, ANA GÓMEZ-LOSCOS and CECILIO TAMARIT: External imbalances and recoveries.

2013 JESÚS FERNÁNDEZ-VILLAVERDE, SAMUEL HURTADO and GALO NUÑO: Financial frictions and the wealth distribution.

2014 RODRIGO BARBONE GONZALEZ, DMITRY KHAMETSHIN, JOSÉ-LUIS PEYDRÓ and ANDREA POLO: Hedger of last resort: evidence from Brazilian FX interventions, local credit, and global financial cycles.

2015 DANILO LEIVA-LEON, GABRIEL PEREZ-QUIROS and EYNO ROTS: Real-time weakness of the global economy: a first assessment of the coronavirus crisis.

2016 JAVIER ANDRÉS, ÓSCAR ARCE, JESÚS FERNÁNDEZ-VILLAVERDE and SAMUEL HURTADO: Deciphering the macroeconomic effects of internal devaluations in a monetary union.

2017 FERNANDO LÓPEZ-VICENTE, JACOPO TIMINI and NICOLA CORTINOVIS: Do trade agreements with labor provisions matter for emerging and developing economies' exports?

2018 EDDIE GERBA and DANILO LEIVA-LEON: Macro-financial interactions in a changing world.

2019 JAIME MARTÍNEZ-MARTÍN and ELENA RUSTICELLI: Keeping track of global trade in real time.

2020 VICTORIA IVASHINA, LUC LAEVEN and ENRIQUE MORAL-BENITO: Loan types and the bank lending channel.

2021 SERGIO MAYORDOMO, NICOLA PAVANINI and EMANUELE TARANTINO: The impact of alternative forms of bank consolidation on credit supply and financial stability.

2022 ALEX ARMAND, PEDRO CARNEIRO, FEDERICO TAGLIATI and YIMING XIA: Can subsidized employment tackle long-term unemployment? Experimental evidence from North Macedonia.

2023 JACOPO TIMINI and FRANCESCA VIANI: A highway across the Atlantic? Trade and welfare effects of the EU-Mercosur agreement.

2024 CORINNA GHIRELLI, JAVIER J. PÉREZ and ALBERTO URTASUN: Economic policy uncertainty in Latin America: measurement using Spanish newspapers and economic spillovers.

2025 MAR DELGADO-TÉLLEZ, ESTHER GORDO, IVÁN KATARYNIUK and JAVIER J. PÉREZ: The decline in public investment: "social dominance" or too-rigid fiscal rules?

2026 ELVIRA PRADES-ILLANES and PATROCINIO TELLO-CASAS: Spanish regions in Global Value Chains: How important? How different?

2027 PABLO AGUILAR, CORINNA GHIRELLI, MATÍAS PACCE and ALBERTO URTASUN: Can news help measure economic sentiment? An application in COVID-19 times.

2028 EDUARDO GUTIÉRREZ, ENRIQUE MORAL-BENITO, DANIEL OTO-PERALÍAS and ROBERTO RAMOS: The spatial distribution of population in Spain: an anomaly in European perspective.

2029 PABLO BURRIEL, CRISTINA CHECHERITA-WESTPHAL, PASCAL JACQUINOT, MATTHIAS SCHÖN and NIKOLAI STÄHLER: Economic consequences of high public debt: evidence from three large scale DSGE models.

2030 BEATRIZ GONZÁLEZ: Macroeconomics, Firm Dynamics and IPOs.

2031 BRINDUSA ANGHEL, NÚRIA RODRÍGUEZ-PLANAS and ANNA SANZ-DE-GALDEANO: Gender Equality and the Math Gender Gap.

2032 ANDRÉS ALONSO and JOSÉ MANUEL CARBÓ: Machine learning in credit risk: measuring the dilemma between prediction and supervisory cost.

2033 PILAR GARCÍA-PEREA, AITOR LACUESTA and PAU ROLDAN-BLANCO: Raising Markups to Survive: Small Spanish Firms during the Great Recession.

2034 MÁXIMO CAMACHO, MATÍAS PACCE and GABRIEL PÉREZ-QUIRÓS: Spillover Effects in International Business Cycles.

2035 ÁNGEL IVÁN MORENO and TERESA CAMINERO: Application of text mining to the analysis of climate-related disclosures.

2036 EFFROSYNI ADAMOPOULOU and ERNESTO VILLANUEVA: Wage determination and the bite of collective contracts in Italy and Spain: evidence from the metal working industry.

2037 MIKEL BEDAYO, GABRIEL JIMÉNEZ, JOSÉ-LUIS PEYDRÓ and RAQUEL VEGAS: Screening and Loan Origination Time: Lending Standards, Loan Defaults and Bank Failures.

2038 BRINDUSA ANGHEL, PILAR CUADRADO and FEDERICO TAGLIATI: Why cognitive test scores of Spanish adults are so low? The role of schooling and socioeconomic background

2039 CHRISTOPH ALBERT, ANDREA CAGGESE and BEATRIZ GONZÁLEZ: The Short- and Long-run Employment Impact of COVID-19 through the Effects of Real and Financial Shocks on New Firms.

2040 GABRIEL JIMÉNEZ, DAVID MARTÍNEZ-MIERA and JOSÉ-LUIS PEYDRÓ: Who Truly Bears (Bank) Taxes? Evidence from Only Shifting Statutory Incidence.

2041 FELIX HOLUB, LAURA HOSPIDO and ULRICH J. WAGNER: Urban air pollution and sick leaves: evidence from social security data.

2042 NÉLIDA DÍAZ SOBRINO, CORINNA GHIRELLI, SAMUEL HURTADO, JAVIER J. PÉREZ and ALBERTO URTASUN: The narrative about the economy as a shadow forecast: an analysis using Banco de España quarterly reports.

BANCO DE ESPAÑA
Eurosistema

Unidad de Servicios Generales I
Alcalá, 48 - 28014 Madrid
E-mail: publicaciones@bde.es
www.bde.es