

The Role of ECB Speeches in Nowcasting German GDP

*Necmettin Alpaya Koçak**

Abstract:

The literature shows that the nowcasting models generally use structured data such as real, financial and survey indicators. Recent research has focused on finding the way how to use the unstructured data in the nowcasting models. The search items such as sentiments or emotions were gathered from internet platforms and used as unstructured data. In this study, it is analysed how the ECB presidents' speeches are included in the nowcasting model and to what degree they affect the quarterly gross domestic product (GDP) of Germany. First, ECB presidents' speeches are analysed to obtain the emotion indicators with assistance of the newly harmonised complex dictionary. These emotion indicators are next added to the unbalanced and mixed frequency data and the nowcasting model estimation for GDP is performed with these data using the expectation-maximisation algorithm in the dynamic factor model representation. Moreover, the news analysis is performed to show how the revisions in the real-time data, including emotion indicators, affect the nowcasts for the current and next quarter GDPs. Finally, a forecast scenario is performed to demonstrate the effects of emotion indicators in the nowcasting model of GDP which shows a slowdown for the last two years. In conclusion, it is suggested that ECB presidents' speeches may increase the performance of nowcasting models for the German GDP.

Key words: Emotion analysis; ECB speeches; Nowcasting.

JEL classification: C33; E52; E58.

1 Introduction

Gross domestic product is the single most important economic indicator which shows the general economic situation of the given country. Especially, the underlying dynamics of the quarter-on-quarter percentage growth rate of the quarterly gross domestic product (hereafter GDP) have always been the subject of research since it is very important to know its possible future value in order to assess or steer of the given economy in the short-term. In recent years, it has been more desirable to estimate the values of the current quarter using the nowcasting techniques rather than forecasting GDP for the future quarters since it has a relatively delayed publication schedule ($t + 45$ days) compared

* Necmettin Alpaya Koçak; Hacettepe University, Faculty of Economics and Administrative Sciences, Department of Economics, Beytepe, 06800, Ankara, Turkey, <alpaya kocak@hacettepe.edu.tr>, ORCID: 0000-0002-4232-9985.

to the reference quarter (t). Therefore, there is extensive literature on nowcasting GDP with chief purpose to find the best method and variable combination to nowcast the current period GDP.

The literature shows that the variable set generally consists of the real sector indicators and/or financial sector indicators and/or indicators obtained from consumer-producer expectation surveys. The common feature of these studies is their use of structured data. On the other hand, the use of unstructured data in nowcasting models is a new subject. Varian and Choi (2009), McLaren and Shanbhogue (2011), Fondeur and Karamé (2013), Francesco and Marcucci (2017) and Bortoli and Combes (2015) and Baker et al. (2016) used unstructured data in nowcasting models. They tried to improve the nowcasting/forecasting models by including the unstructured data available on the internet. Specifically, word-search characteristics over search platforms are often treated as unstructured data. One of the visionary works, which differs from the literature in terms of the data source, was presented by Combes et al. (2018). Combes et al. (2018) analysed the newspapers and constructed sentiment indicators for inclusion in the nowcasting model established for France, and they investigated the effect of sentiment indicators on the nowcasting model. However, there are very few studies that try to analyse nowcasting models by including sentiments or emotions. It has been also becoming popular in the literature to analyse emotions and associate them with economic indicators. The study by Kaminski and Gloor (2014) is another visionary paper on this topic in terms of analysing emotions rather than sentiments. They examined the micro-blog data to obtain several emotion indicators and analysed the effect of them on crypto-currencies. Studies by Bollen et al. (2010), Coviello (2014), Si et al. (2013) and Zhang et al. (2011) may also be given as examples of associating economic nowcasting with emotion indicators obtained from micro-blog platforms.

It seems that a point has been missed in the literature. It cannot be denied that the speeches of policymakers are more effective than social media platforms, both in terms of the correct information they contain and the ability to influence economic agents. There are several studies in the literature on examining the relationship between communication reports of monetary authorities and economic indicators (for instance, Lucca and Trebbi (2009), Hansen and McMahon (2015) and Eskici and Koçak (2018)) however it seems that there is obvious absence of the literature regarding the use of emotions of policy makers' speeches in nowcasting models.

This paper aims to improve accuracy of nowcasting model for German GDP by including ECB president's speeches to the model. In particular, we performed the emotion analysis of ECB presidents' speeches and analyse how they affect the German GDP in a multivariate nowcasting model. Next, we analyse how ECB

presidents' speeches make contribution to nowcast GDP and examine their effects on the news analysis as well the real-time forecast analysis.

At a first glance, it seems that the analysis of the impact of the ECB Presidents' comments may be focused on GDP growth in the euro area and the comments of the ECB presidents may not concern individual euro area countries, but the euro area as a whole. However, this study aims to discuss the effects of ECB Presidents on the specific country rather than euro area, and Germany is particularly selected for this purpose. This choice may be seen as acceptable considering the key role and economic size of Germany in the euro area (Eurostat, 2020). Although the main purpose of monetary policy is to ensure price stability and take measures maintain it, in this study, the effects of the speeches of ECB Presidents on GDP, which is a real indicator, is to be investigated rather than its effects on inflation.

The paper is organised as follows: Section 2 provides several important details on ECB presidents' speeches. Section 3 investigates the emotion analysis of ECB speech data to obtain emotion indicators using an extended dictionary. Section 4 is the part of the paper that presents the result of the nowcasting model. The rest of the paper examines the effect of the news on the model and on the real-time forecasts for the last two years (Section 5). Final remarks are provided in the conclusion.

2 ECB Speeches: A Closer Look

2.1 Data

ECB (2019) provides speech data with metadata on the content of all speeches made by ECB to assist researchers in the field of central bank communication. The data is currently updated every two months and presented in comma-separated-value format. All related information can be found in ECB (2019). The data, downloaded as of 25 October 2019, details 2,330 speeches with 5 different characteristics. These characteristics are the columns for the date, speakers, title, subtitle, and contents. The date column extends from 1997-02-07 to 2019-10-18 in daily format. In the speaker's column, there are 23 different speakers with their names and surnames. In line with the aim of the analysis, the presentations and speeches of other officials from the ECB are excluded since the impact of the ECB on the markets is based on the presidents' speeches and it is assumed that all the officials consistently follow the official line of the ECB policy. The speeches of Alexandre Lamfalussy and Willem F. Duisenberg, who were the Presidents of the European Monetary Institute (EMI), are covered in the analysis since the EMI was converted to the ECB. Thus, this study analyses the speeches of four presidents (Alexandre Lamfalussy [1997–1998],

Willem F. Duisenberg [1998–2003], Jean-Claude Trichet [2003–2011] and Mario Draghi [2011–2019]). The subtitle column is excluded because it is unnecessary information for the analysis. The contents column contains the text of the speech in the text format. Consequently, this study is focused on 654 speeches made by four presidents between February-1997 and October-2019.

2.2 Pre-processing findings

Pre-processing is an analysis technique which aims to prepare the text data for text analysis. The first stage of pre-processing is generally tokenisation. Tokenisation is the process of breaking down a text document into words (Welbers et al., 2017). Descriptive statistics on ECB's speech data are given in Table 1 obtained by tokenisation. In Table 1, it is remarkable that almost half of the speeches were made by Jean-Claude Trichet, although his governance term length is the same as Mario Draghi. Besides, it is seen that Mario Draghi's is the lowest mean among others. It is understood that Duisenberg is the one who uses the highest number of words on average (4,040 words) in his speeches.

Tab. 1 Number of Tokens in Speeches by Presidents

Statistics	Speakers			
	Alexandre Lamfalussy	Willem F. Duisenberg	Jean-Claude Trichet	Mario Draghi
Start Year	1997	1997	2003	2011
End Year	1997	2003	2011	2019
Number of Speeches	7	152	315	180
Minimum Token	1856	224	211	320
Mean Token	3183	4040	3525	11254
Maximum Token	4145	21697	11254	14809

Source: Authorial computation.

The second stage of pre-processing is the determination of the words (stop-words) to be excluded from the analysis. This is a recursive process (Loughran and McDonald, 2016). In this study, the stop-word lists are used which are already available in Benoit et al. (2018), Rinker (2018), Rinker (2020), Benoit et al. (2019), Silge and Robinson (2016) and Feinerer et al. (2008) studies in addition to a long user-defined stop-words list. After the data is adjusted from the stop-words, digitising the text data is performed by creating a weighted document-term matrix as tf-idf (term frequency– inverse document frequency). One measure of how important a word maybe is its tf (term frequency), which shows how

frequently a word occurs in a document. Another approach is to look at the term's idf (inverse document frequency), which decreases the weight for commonly used words and increases the weight for words that are not used very often in a collection of documents. In this study, it is intended to measure how important a word is to a document in a collection of documents by combining tf and idf (Salton and Buckley, 1988).

2.3 n-Gram analysis

The analysis of word groups in addition to a single word provides more detailed information about the content of the text. For this reason, this study examines the group of two or three words (bigram and trigram, respectively) as well as a single word (unigram). An n-gram is a sequence of n adjacent elements from a string of tokens (Jurafsky and Martin, 2008). A bigram is an n-gram for $n = 2$ and a trigram is an n-gram for $n = 3$. For simplicity, the relationship of Bayesian conditional probability is given in Eq. (1) only for a bigram which provides the conditional probability of a token given the preceding token.

$$P(W_n|W_{n-1}) = \frac{P(W_{n-1}, W_n)}{P(W_{n-1})} \quad (1)$$

That is, the probability $P()$ of a token W_n given the preceding token W_{n-1} is equal to the probability of their bigram, or the co-occurrence of the two tokens $P(W_{n-1}, W_n)$ divided by the probability of the preceding token. Table 2 presents together the unigrams, bigrams and trigrams that each president frequently uses over the years. When n-grams are given together, it is easy to understand which economic conditions prevail in the relevant year.

According to Table 2, the words “stability” and “foreign exchange” were the most frequently used by both Lamfalussy and Duisenberg during the 1997 Asian financial crisis. Also, Duisenberg underlined the phrases “stability”, “harmonised price index”, and “structural reforms” during and after the 1998 Russian financial crisis. Duisenberg emphasised “real GDP growth trend”, “cross border payments/retail”, “structural reforms” between the years 2000 and 2003. During the period from 2003 to 2005, Trichet used the terms “stability” and “structural reforms” as before 2000, but an important point that draws attention is the emphasis on “labour productivity growth” and “cross border” in the years before the 2009 global economic crisis (2005–2008). Then, the global economic crisis and the measures taken are frequently mentioned and Trichet used “systemic risk”, “macroprudential supervision” and “unit labour cost” during the years between 2009 and 2011.

Tab. 2 Most used unigrams, bigrams and trigrams by year and president

Year	President	Unigram-Bigram-Trigram
1997	Alexandre Lamfalussy	stability-foreign exchange-marginal lending facility
1997	Willem F. Duisenberg	stability-foreign exchange-wide banking
1998	Willem F. Duisenberg	stability-stability oriented-consumer prices HICP
1999	Willem F. Duisenberg	stability-structural reforms-marginal lending facility
2000	Willem F. Duisenberg	growth-cross border-cross border retail
2001	Willem F. Duisenberg	growth-cash changeover-cross border payments
2002	Willem F. Duisenberg	growth-cash changeover-trend potential growth
2003	Willem F. Duisenberg	growth-structural reforms-real gdp growth
2003	Jean-Claude Trichet	stability-structural reforms-potential output growth
2004	Jean-Claude Trichet	stability-structural reforms-sound fiscal policies
2005	Jean-Claude Trichet	integration-cross border-labour productivity growth
2006	Jean-Claude Trichet	growth-productivity growth-labour productivity growth
2007	Jean-Claude Trichet	growth-productivity growth-labour productivity growth
2008	Jean-Claude Trichet	stability-cross border-labour productivity growth
2009	Jean-Claude Trichet	crisis-systemic risk-macro prudential supervision
2010	Jean-Claude Trichet	crisis-global governance-macro prudential oversight
2011	Jean-Claude Trichet	crisis-unit labour-unit labour costs
2011	Mario Draghi	fiscal-fiscal compact-substantial downside risks
2012	Mario Draghi	stability-real economy-single supervisory mechanism
2013	Mario Draghi	union-banking union-single resolution mechanism
2014	Mario Draghi	union-structural reforms-single supervisory mechanism
2015	Mario Draghi	reforms-structural reforms-expanded asset purchase
2016	Mario Draghi	growth-structural reforms-systemic risk board
2017	Mario Draghi	growth-financing conditions-capital markets union
2018	Mario Draghi	growth-risk sharing-private risk sharing
2019	Mario Draghi	growth-risk sharing-sovereign debt crisis

Source: Authorial computation.

While Draghi, who took office in 2011, emphasised “fiscal” issues in the first year of his duty, he made speeches on “structural reforms”, “union”, “single supervisory / resolution mechanism” in the following years. Draghi used “growth”, “financing conditions”, “private risk sharing” and “sovereign debt crisis” most frequently in the period from 2016 to 2019.

3 Emotion Analysis

3.1 An extended emotion dictionary

Emotion analysis is a typical text mining analysis and aims to categorise words in the text data regarding several pre-defined emotions. It is a different approach to the sentiment analysis which categorise words into symmetric measures such as “positive”, “neutral”, and “negative”. Emotions are defined in a dictionary in the content analysis. In this study, the first step is to define a dictionary that includes a list of terms with different types of emotion connotations from an economic viewpoint. Several studies present emotion lexicons. The main financial-purpose lexicons considered to harmonise with this study are as follows:

- bing dictionary (Hu and Liu, 2004),
- National Research Council of Canada (NRC) dictionary (Mohammad and Turney, 2013),
- Lexicoder Sentiment Dictionary (Young and Soroka, 2012),
- Harvard-IV dictionary as used in the General Inquirer software (Harvard University, 2000),
- Henry’s Financial dictionary (Henry, 2008),
- Loughran-McDonald Financial dictionary (Loughran and McDonald, 2011),
- AFINN dictionary (Nielsen, 2011)

All these lexicons are based on unigrams. These lexicons contain many English words and the words are assigned to emotions. The Loughran and McDonald (2011) and Mohammad and Turney (2013) lexicons categorise words into emotions, the remaining lexicons are for sentiments. The AFINN lexicon only assigns words with a score that runs between -5 and 5, with negative scores indicating negative sentiment and positive scores indicating positive sentiment. Due to the scale problem, the AFINN dictionary is excluded from the scope of the analysis. After harmonisation of all available lexicons, there are 22,298 words (15,349 of them unique) in the new dictionary and the number of total emotions is 8. These emotions can be listed as negative, trust, fear, surprise, positive, uncertainty, constraining, anticipation. The distribution of words to emotions is given in Table 3. The words are mostly flagged as negative-positive separation, then fear, trust, etc. according to the distribution. Of course, it is a fact that a word

represents more than one emotion. As a result, it appears that a word roughly represents two ($1.5 \approx 2$) different emotions on average.

Tab. 3 Properties of dictionary

Emotions	Number of tokens
Negative	9,439
Positive	5,171
Fear	1,476
Trust	1,231
Anticipation	839
Surprise	534
Uncertainty	297
Constraining	184

Source: Authorial computation.

3.2 Findings of emotions analysis

There is a variety of methods that exist for evaluating the emotion in text data. In this study, emotions are extracted from ECB presidents' speech text data using the methods explained in Rinker (2019) through the dictionary created in Section 3.1. The analysis is done at the sentence level in the data. Emotion weight is calculated at the sentence level. This approach was originally invented in Plutchik (1962) and Plutchik (2001) studies and the scores are between 0 (no emotion in the sentence) and 1 (all vocabulary used in the sentence represents emotion). It should be noted that emotion words prefixed with 'un-' are treated as a negation. For example, "unhappy" would be treated as "not happy".

The daily analysis results are aggregated to the monthly frequency since it is aimed to use the output of this analysis in the nowcasting model in Section 4. Various approaches can be used in this aggregation process. For instance, daily speech texts made in the relevant month can be evaluated as a single text if a speech is made once a month. However, in this approach, there is a disadvantage that the effects of the speech at the end of any month are only reflected to that month. For this reason, the speeches are analysed on daily basis, and their emotion effects are assumed to last until the next speech day in this study. Afterwards, the aggregation is performed by taking monthly averages of each emotion.

Table 4 presents the results of emotion analysis by presidents. The figures in Table 4 express “the proportion of emotion words in a sentence of ten words”. To give an example, Willem F. Duisenberg uses a total of 2.66 emotion words on average in a representative 10-word sentence. According to the total emotion indicator in Table 4, Alexandre Lamfalussy performs the most emotion speeches. Draghi is the second place in this ranking. The most positive and negative speeches are made by Mario Draghi. While the most uncertain and surprising speeches are made by Alexandre Lamfalussy, most angry and fearful speeches are made by Alexandre Lamfalussy and Mario Draghi, respectively.

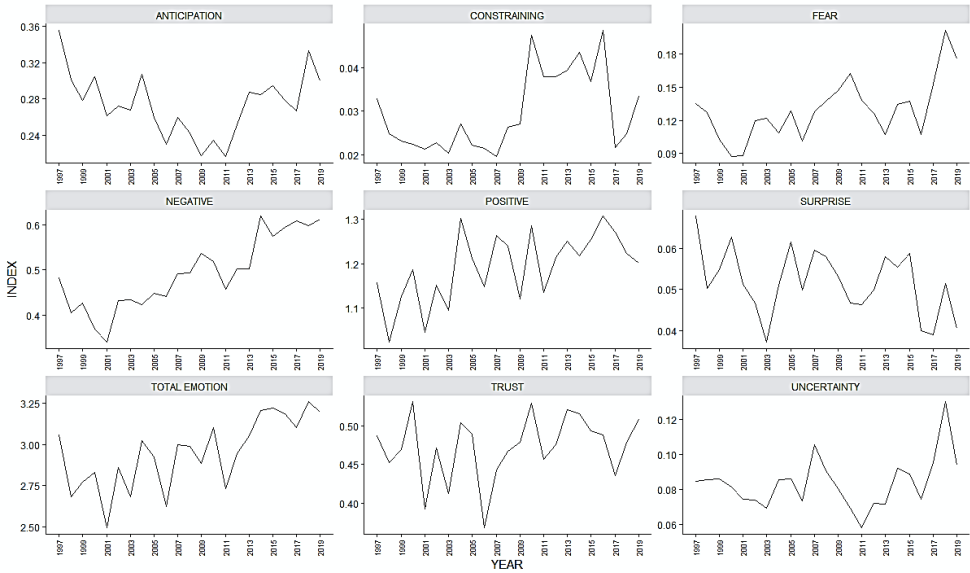
Tab. 4 Emotions by president

Emotions	Speakers			
	Alexandre Lamfalussy	Willem F. Duisenberg	Jean-Claude Trichet	Mario Draghi
Anticipation	0,33	0,27	0,25	0,27
Constraining	0,04	0,02	0,02	0,04
Fear	0,13	0,11	0,13	0,14
Negative	0,50	0,41	0,49	0,58
Positive	1,15	1,06	1,19	1,19
Surprise	0,07	0,05	0,06	0,05
Trust	0,48	0,44	0,45	0,46
Uncertainty	0,10	0,09	0,09	0,09
Total	2,80	2,45	2,68	2,81

Source: Authorial computation.

The emotion indicators are presented annually for convenience notation in Figure 1, although they are calculated at monthly frequency. Fear and Uncertainty indicators show an upward trend after the year 2011. Following a decrease, the anticipation indicator shows an upward trend (‘U’ shape) until the year 2011. The trust indicator is said to be relatively horizontal. On the other hand, the positive indicator increases with volatility, while the negative indicator increases in a stable pattern. It is seen that the total emotion indicator shows an increasing trend, but it has shown sharp decreases in several years (2001, 2006, 2011). It is important to note that in 2001, there was a significant drop in the US manufacturing sector, very low inflation in 2006, and a rapid decrease in unit labour costs in 2011 across the EU.

Fig. 1 Emotions by year



Source: Authorial computation.

4 Nowcasting German GDP

4.1 Data

This study uses data that contains several indicators at monthly and quarterly frequencies between January 1991 and December 2019. The data consist entirely of the first estimation of the indicators. Sub-sequent revisions of the indicators are ignored. Therefore, it allows for the real-time analysis. The indicators can be summarised in three groups such as real, survey, and emotion obtained in Section 3. Detailed information on the indicators is given in Appendix 1.

The group of real indicators contains five variables. The first one is the quarterly GDP (flash estimates in constant prices) which is the target variable of the nowcasting model. Flash estimates of quarterly GDP are usually published with a delay of 45 days compared to the reference period. The second variable represents historical estimates of quarterly GDP which is finalised approximately two years later. In addition to these two variables, industrial production, factory orders and retail sales are also included in real indicators. These three variables are published with a minimum delay of 30 days compared to the reference periods. There are six variables in the survey indicators group. The Purchasing Managers IndexTM (PMITM) related to the manufacturing and service sectors published

separately and at the monthly frequency by IHS-MARKIT company. The Business Climate Index, the Business Expectations Index and the Export Expectations Index, which are published at monthly frequency by IFO-CES Institution, are released 7 days prior to the end of the reference period. The last variable of the survey group is the Economic Sentiment Index published monthly by ZEW Institution, which is published 14 days prior to the end of the reference period.

The group of emotion indicators consists of the emotion variables. Eight monthly emotion variables and their sum are considered in the group. A graphical presentation of the variables is given in Appendix 2 to give an idea of their movement over time. Also, the number of speeches made by ECB presidents is added as a monthly variable to the group of the emotion indicators. It can be assumed that emotion group, which consists of thirteen variables, is published simultaneously without delay.

Main restriction of the analysis is to include only the ECB's comments among the factors influencing emotions and thus also the GDP forecast. However, each central bank also uses other tools in communication with the public, such as announcing a change in interest rates, introducing other monetary policy tools, publishing minutes or an inflation report with an up-to-date macroeconomic forecast.

4.2 Methodology

In this study, the dynamic factor model (DFM) specification is used to build a nowcasting model (Stock and Watson, 2005). Mixed frequency data is used in DFM by following the method explained in the study by Camacho and Perez-Quiros (2010). The DFM used in this study also allows using unbalanced data, where not all data cover the same time interval following the method of the study by Bańbura and Modugno (2014). They suggest a solution for the problems of estimating missing observations in the system estimation in the case of mixed frequency data.

In DFM, a large number of time series are represented by a smaller number of factors, both by carrying the time series features and by reducing their dimensions. A DFM can be represented as given in Eq. (2) – (4):

$$x_{i,t} = \Lambda_i f_t + \epsilon_{i,t} \quad i = 1, \dots, n \quad (2)$$

$$f_{i,t} = \varphi(L)f_{i,t-1} + \eta_{i,t} \quad \eta_t \sim i.i.d. N(0, R) \quad (3)$$

$$\epsilon_{i,t} = \alpha_i \epsilon_{i,t-1} + v_{i,t} \quad v_t \sim i.i.d. N(0, \sigma_i^2) \quad (4)$$

Let x represent the standardised variables in terms of mean and variance, Λ is the $n \times r$ dimension matrix and represents the effects of x 's on invisible factors

(f), which represents the loadings. These factors are assumed to be followed by a VAR process in p lag. ϵ_t shows idiosyncratic residuals and they are assumed to be an AR(1) process. The loadings of quarterly observations on monthly factors are determined by using restrictions in Mariano and Murasawa (2003). Besides, the year-on-year growth rates are calculated for the monthly factor loadings following the approach of Giannone et al. (2013). The transformations of variables in the loadings of DFM are given in last column of Table 8.

Expectation-maximisation (EM) algorithm is used to estimate DFM in this study. Although this algorithm is derived from the study by Shumway and Stoffer (1982), Marta and Michele (2010) first used it in the DFM approach. The lag length is determined as three for the factors according to the Akaike information criterion. As proposed by Doz et al. (2012), the restriction of the effects of variables on the factors is applied in a subjective approach in the form of real/surveys/emotion separation explained in Section 4.1.

4.3 Findings

Three separate DFM models are estimated to understand the effect of the emotional indicators group to the nowcasting model. The “Benchmark” Model is the base model where the emotion indicators group is not included, that is, only real and survey groups of indicators are used. The second model is the model (Speech Model V-1) in which the emotion indicators group is included as a single factor to the nowcasting model. The third model is the model (Speech Model V-2) in which the emotion indicators group is included as three factors to the nowcasting model. All three DFMs cover the full time-span which is from January 1991 to December 2019. The Speech Model V-1 shows better performance than the Speech Model V-2 according to the RMSE measure. Then, the effect of the emotion indicators group on the nowcasting model can be seen by comparing the Benchmark Model and the Speech Model V-1. Table 5 presents RMSE and gain/loss values for the three DFMs.

Tab. 5 Properties of nowcasting models for the GDP

Models	Number of Factors	RMSE	Gain in RMSE (respect to Benchmark)
Benchmark Model	2	0,0097	-
Speech Model V-1	3	0,0052	46%
Speech Model V-2	5	0,0109	-12%

Source: Authorial computation.

The RMSE value is 0.0097 in the Benchmark Model and 0.0052 in the Speech Model V-1. Thus, 46% gain is achieved by the Speech Model V-1 compared

to the Benchmark Model. However, there is a loss of 12% in the RMSE if the Speech Model V-2 is preferred to the Benchmark Model. As a result, it can be claimed that the inclusion of the emotion indicators group as a single factor to the Benchmark Model creates a noticeable gain in RMSE and in the nowcasting performance for GDP. Table 6 contains the nowcasted figures for GDP produced by the Benchmark Model and the Speech Model V-1. It can be claimed that the Speech Model V-1 is superior to the Benchmark Model.

Tab. 6 Nowcasting performances for the GDP growth by DFM models

Period	GDP	Benchmark Model	Speech Model V.1
2017-Q1	0.6	0.74	0.70
2017-Q2	0.6	0.65	0.63
2017-Q3	0.8	0.78	0.79
2017-Q4	0.6	0.62	0.62
2018-Q1	0.3	0.34	0.32
2018-Q2	0.5	0.49	0.49
2018-Q3	-0.2	-0.09	-0.11
2018-Q4	0.0	-0.01	0.00
2019-Q1	0.4	0.27	0.32
2019-Q2	-0.1	-0.10	-0.10
2019-Q3	0.1	0.03	0.04

Source: Authorial computation.

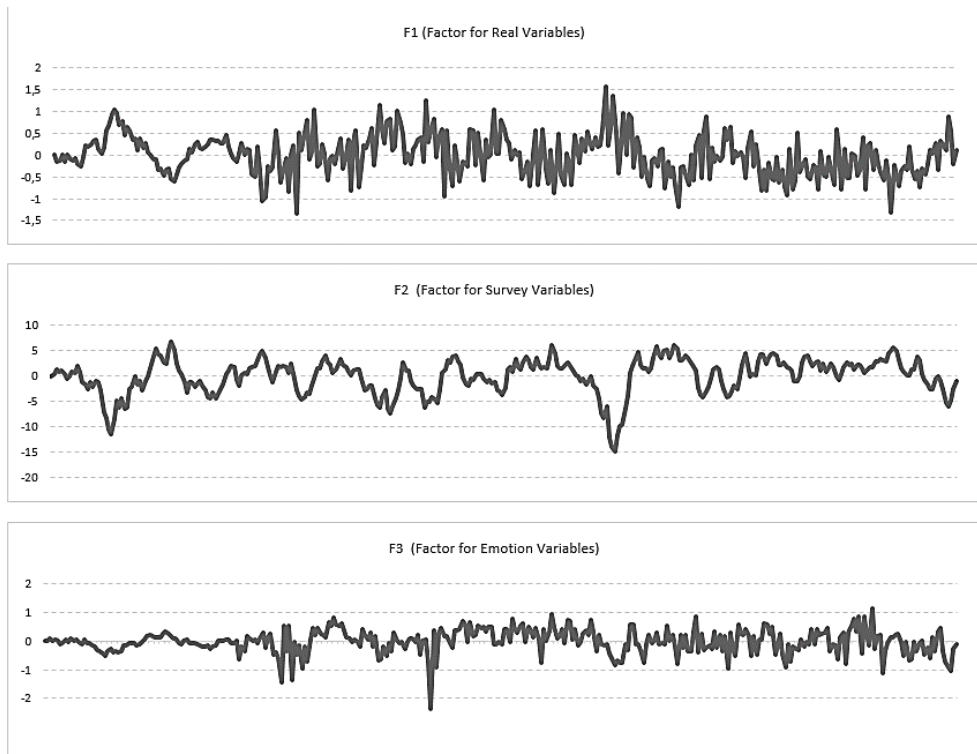
In Table 6, the nowcasted figures of the Benchmark Model and Speech Model V-1 are compared in terms of their proximity to flash estimates of GDP for 2017 and beyond. From this point, the Benchmark Model shows that the importance of indicators of real and survey groups cannot be denied. Then, it can be said that the contribution of the emotion indicators group to the real and survey groups is not very high, as expected, but it affects positively in terms of proximity to flash estimates.

In the Speech Model V-1, the first factor, which represents the group of real indicators, is relatively more volatile and can be said to capture fluctuations in GDP. The second factor, which represents the group of survey indicators, is relatively more stable and can be said to capture the main trend in GDP. The third factor, which represents the emotion indicators group, is the factor with the highest volatility and represents the portion of GDP that cannot be captured

by the first and second factor. The three factors are respectively represented in Figure 2.

The variance decomposition (VD) indicates the amount of information each factor contributes to GDP in the Speech Model V-1 in Figure 3. It determines how much the forecast error variance of GDP can be explained by exogenous shocks to the three factors. The comparison between the Benchmark Model and the Speech Model V-1 by VD analysis is useful to understand the effect of the emotion indicators group on the explained variance of GDP. Although an effect as large as the real and survey indicators is not expected, an acceptable forecast error variance explanation rate is desired from the factor of the emotion indicators group. Indeed, it is seen that the emotion indicators group explains small but important (approximately 5%) part of the forecast error variance of the GDP.

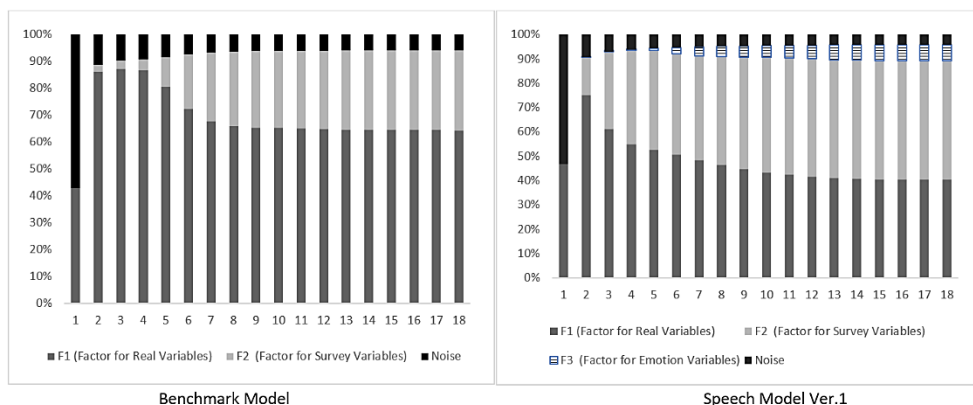
Fig. 2 Factors from Speech Model V1



Source: Authorial computation.

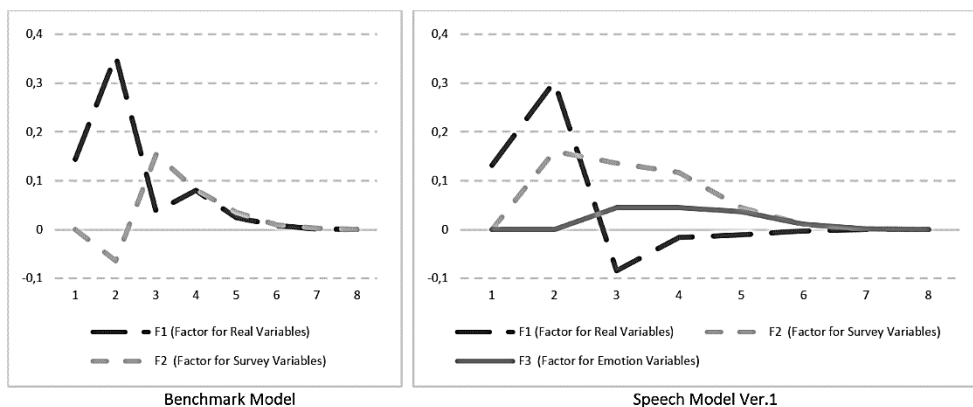
Impulse-response functions (IRF) are analysed in this study to describe the evolution of factors in reaction to a shock in GDP. It measures the changes in the future responses of all factors in the DFM when GDP is shocked by an impulse in one unit of standard deviation. It is understood from r.h.s. of Figure 4 that the responses of the first and second factors occur in a positive direction to one standard deviation impulse in GDP as expected in the Speech Model V-1. Also, the third factor shows a positive but small response of approximately 0.05 units. It should be noted that responses of the first and second factors in the Benchmark Model are more complicated than the Speech Model V-1.

Fig. 3 Variance decomposition of DFM Models



Source: Authorial computation.

Fig. 4 Impulse-response function of DFM Models



Source: Authorial computation.

5 News and Scenario Analysis

5.1 News Analysis

The news analysis is important to show revisions in the forecasts in terms of revisions/upcoming information in new data releases. In this section, it is aimed to measure the effect of emotion indicators group to the news analysis. The approach presented by Basselier et al. (2017) is followed in this study. Using the Speech Model V-1 model structure, the model is refreshed in terms of loadings by adding new data for each month starting from January 2019 until December 2019. It can be tracked the effects of factors (aggregated by indicators) on the nowcasted value for GDP by months. The results are given in Table 7 by quarterly aggregation for convenience notation.

Table 7 shows the effects of new data points added to the data by months during the year 2019 on the nowcasted values for GDP for the years 2019 and 2020. Seemingly complex Table 7 is explained with a short example. The value of 0.066, in the intersection of the second row and the third column of the table, is the forecast produced for Q2-2019 using information until February 2019 (i.e. January-2019 period and before). After updating the data with February-April 2019 upcoming data, the new forecast value for Q2-2019 is revised upwards in the amount of 0.288 (it is the value in the “sum of news” in the third row). The effect of real indicators is 0.111, the effect of survey indicators is 0.121, and the effect of emotion indicators is 0.05 in this revision. The total effect is 0.288.

As a result, following the update of data, the new forecast value for Q2-2019 becomes 0.354 (in the seventh row). The realised value for Q2-2019, which is given as extra information to evaluate the forecasting performance, is 0.1 (True Values row in the table).

In Table 7, it is understood that the forecast performance is low as the variable selection is made for nowcasting, not forecasting. To sum up, the speeches made throughout the year mostly affect the GDP forecast for the Q4-2019 period considering the contributions of emotion indicators to the forecasts. This result is as expected because the expectations of market agents guided by the presidents’ speeches turn into a real effect that occurs with a certain delay (approximately 3 quarters).

Tab. 7 Nowcasting performances for the GDP growth by DFM models

Periods	Measurements	Impact II-2019	Impact III-2019	Impact IV-2019	Impact I-2020	Impact II-2020	Impact III-2020
M2-2019:M4- 2019	Old Forecasts	0.066	0.045	0.190	-	-	-
	Sum of news	0.288	0.201	0.193	-	-	-
	F1 -Real Variables	0.111	0.034	0.010	-	-	-
	F2- Survey Variables	0.121	0.121	0.070	-	-	-
	F3 - Emotion Variables	0.056	0.046	0.113	-	-	-
	New Forecasts	0.354	0.246	0.383	-	-	-
	True Values	0.100	0.100	-	-	-	-
M5-2019:M7- 2019	Old Forecasts	-	0.155	0.157	0.258	-	-
	Sum of news	-	0.536	0.768	0.559	-	-
	F1 -Real Variables	-	0.022	0.028	0.004	-	-
	F2- Survey Variables	-	0.493	0.603	0.454	-	-
	F3 - Emotion Variables	-	0.021	0.137	0.101	-	-
	New Forecasts	-	0.691	0.925	0.817	-	-
	True Values	-	0.100	-	-	-	-
M8-2019:M10- 2019	Old Forecasts	-	-	0.282	0.097	0.033	-
	Sum of news	-	-	0.343	0.329	0.302	-
	F1 -Real Variables	-	-	0.088	0.036	0.010	-
	F2- Survey Variables	-	-	0.215	0.282	0.223	-
	F3 - Emotion Variables	-	-	0.040	0.011	0.069	-
	New Forecasts	-	-	0.625	0.426	0.335	-
	True Values	-	-	-	-	-	-
M11-2019:M12- 2019	Old Forecasts	-	-	0.064	0.136	0.199	0.333
	Sum of news	-	-	0.257	0.167	0.170	0.086
	F1 -Real Variables	-	-	0.209	0.043	0.001	0.001
	F2- Survey Variables	-	-	0.032	0.123	0.104	0.076
	F3 - Emotion Variables	-	-	0.016	0.001	0.065	0.009
	New Forecasts	-	-	0.321	0.303	0.369	0.419
	True Values	-	-	-	-	-	-

Source: Authorial computation.

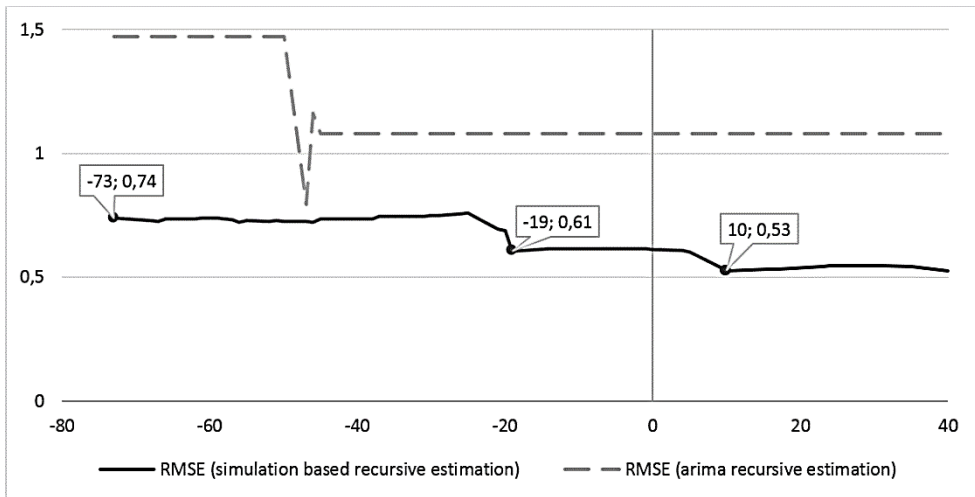
* Periods represents the reference periods of new data added to the data set.

5.2 Scenario Analysis

The German economy has been recently questioned whether it will come out in the case of a slowdown. A scenario analysis may be useful to understand how nowcasting model is reliable to detect GDP at the earliest phase. Therefore, it is examined here how long the GDP nowcasts converge to the true values in the past. A scenario analysis is performed for the last two years using the Speech Model V-1 structure.

Figure 5 presents the relationship between the nowcasting period and RMSE of the Speech Model V-1. It gives the distribution of the RMSE of the nowcasted value to the nowcasting date (the straight thick line in Figure 5) obtained by the Speech Model V-1. $(-73; 0.74)$ represent that the average RMSE value for the nowcasted value of GDP is 0.74 at the day $t - 73$ (73 day before t reference).

Fig. 5 Average RMSE for quarterly GDP growth by forecast horizon

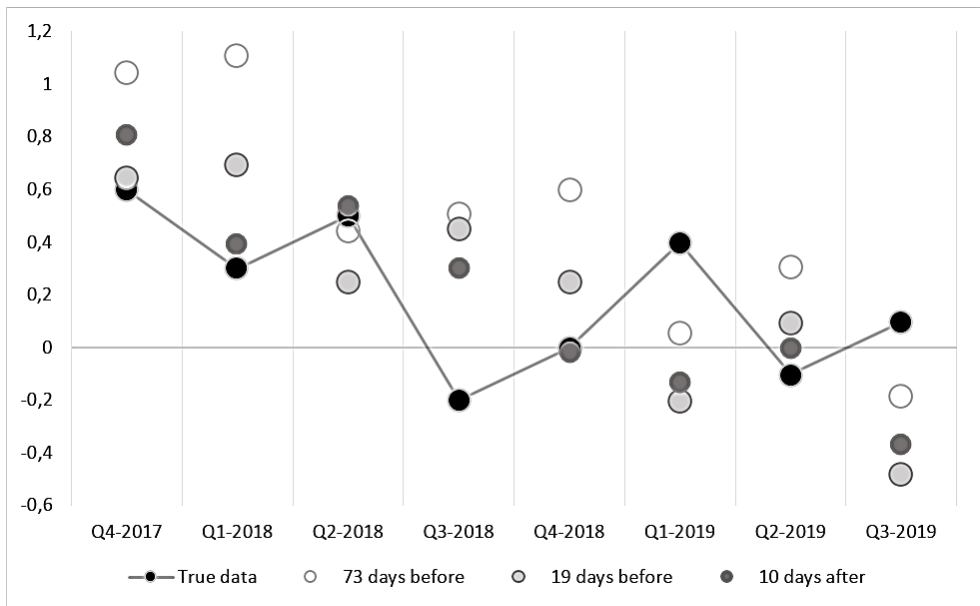


Source: Authorial computation.

The average RMSE value is 0.61, and 0.53 for the day $t + 10$ if the model is run at the day $t - 19$. GDP is generally known to be published at $t + 45$. It is understood that nowcasting model run at $t + 10$ brings the highest performance for the nowcasting of GDP. In this part, the results of the estimated ARIMA model (Gomez and Maravall, 2001), which are also reported by estimation day, are given for comparison with the Speech Model V-1, represented by a dashed line in Figure 5. RMSE values obtained from the ARIMA model do not fall below the Speech Model V-1's in any prediction lagged or leading period.

During the period from Q2-2017 to Q3-2019, the comparison of the nowcasted values and true values of the GDP is presented in Figure 6 obtained by applying the Speech Model V-1 on the days for $t - 73$, $t - 19$ and $t + 10$. It is considered that the nowcasted and true values show a similar trend until the last year and it can be suggested that the run for nowcasting model appears more successful at the day $t + 10$. However, it can be said that the noisy figures of the true data for the year 2019 prevent reaching a general conclusion for all the scenario periods.

Fig. 6 Average RMSE for quarterly GDP growth by forecast horizon



6 Conclusions

In this study, emotion analysis is performed for the speeches of ECB presidents from 1997 to the end of 2019, and the indicators obtained from the emotion analysis are used as an additional input to the nowcasting model of the quarterly German GDP. Besides, important intermediate results are obtained in the text analysis of the speeches. The single and multi-term (unigram and n-grams) expressions used by presidents provide useful information to define the economic situation over time. Unlike the literature, the performance of the analysis is increased by applying to multiple dictionaries instead of using a single dictionary in emotion analysis. As a result of the emotion analysis, monthly indicators are obtained for the nowcasting model.

The emotion indicators are included the data contains the real and survey indicators which have unbalanced time span and mixed frequency. The specification of the nowcasting model is chosen as a dynamic factor model (DFM) and the expectation-maximisation (EM) approach is adopted as the estimation method. Three DFMs are predicted (each with 3 lags) and DFMs are classified as the Benchmark Model (2 factors without emotion indicators) and the Speech Model V-1 (3 factors with emotion indicators), and the Speech Model V-2 (5 factors with emotion indicators). The Speech Model V-1 shows better performance than other models according to RMSE. It is understood that the emotion indicators made significant contributions to the variance decomposition and impulse-response analysis for the GDP. Of course, emotion indicators cannot be expected to be entirely successful in explaining GDP, but it can be claimed that they may provide statistical contribution to a nowcasting model using the real and survey indicators. In the news analysis, it is determined that the effects of emotion indicators on forecasts are limited and increase towards the end of the year, and it may be concluded that the model containing contains emotion indicators can nowcast the quarterly German GDP with the lowest error at the day $t + 10$.

References

- Baker, S.R., Bloom, N., Davis, S.J., 2016. Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics* 131, 1593–1636.
- Bañbura, M., Modugno, M., 2014. Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics* 29, 133–160.
- Basselier, R., Antonio Liedo, D. de, Geert Langenus, 2017. Nowcasting real economic activity in the euro area : Assessing the impact of qualitative surveys (Working Paper Research No. 331). National Bank of Belgium.
- Benoit, K., Muhr, D., & Watanabe, K. 2019. stopwords: Multilingual stopword lists (Version 1.0). Available from: <CRAN.R-project.org/package=stopwords>. [11 September 2020].
- Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., Matsuo, A., 2018. Quanteda: An r package for the quantitative analysis of textual data. *Journal of Open Source Software* 3, 774.
- Bollen, J., Mao, H., Zeng, X., 2010. Twitter mood predicts the stock market. *CoRR* abs/1010.3003. Available from: <arxiv.org/pdf/1010.3003&>. [11 September 2020].

Bortoli, C., Combes, S., 2015. Contribution from google trends for forecasting the short-term economic outlook in france: Limited avenues. Institut National de la Statistique et des Études Économiques. Available from: <insee.fr/en/statistiques/1408911>. [11 September 2020].

Camacho, M., Perez-Quiros, G., 2010. Introducing the euro-sting: Short-term indicator of euro area growth. *Journal of Applied Econometrics* 25, 663–694. Available from: <jstor.org/stable/40607051>. [11 September 2020].

Combes, S., Renault, T., Bortoli, C., 2018. Nowcasting gdp growth by reading newspapers.

Coviello, Y.A.K., Lorenzo AND Sohn, 2014. Detecting emotional contagion in massive social networks. *PLOS ONE* 9, 1–6. DOI: 10.1371/journal.pone.0090315.

Doz, C., Giannone, D., Reichlin, L., 2012. A quasi–maximum likelihood approach for large, approximate dynamic factor models. *Review of economics and statistics* 94, 1014–1024. Available from: <jstor.org/stable/23355337>. [11 September 2020].

ECB, 2019. Speeches dataset. Available from: <ecb.europa.eu/press/key/html/downloads.en.html>. [11 September 2020].

Eskici, H.B., Koçak, N.A., 2018. A text mining application on monthly price developments reports. *Central Bank Review* 18, 51–60.

Eurostat, 2020. Germany: Trade and economy. Available from: <europa.eu/european-union/about-eu/countries/member-countries/germany_en>. [11 September 2020].

Feinerer, I., Hornik, K., Meyer, D., 2008. Text mining infrastructure in R. *Journal of Statistical Software* 25, 1–54.

Fondeur, Y., Karamé, F., 2013. Can google data help predict French youth unemployment? *Economic Modelling* 30, 117–125.

Francesco, D’Amuri, Marcucci, J., 2017. The predictive power of google searches in forecasting us unemployment. *International Journal of Forecasting* 33, 801–816.

Giannone, D., Agrippino, S.M., Modugno, M., 2013. Nowcasting china real GDP. Mimeo New York, NY.

Gomez, V., Maravall, A., 2001. Automatic modeling methods for univariate series. A course in time series analysis. Available from: <onlinelibrary.wiley.com/doi/pdf/10.1002/9781118032978#page=186>. [11 September 2020].

Hansen, S., McMahon, M., 2015. Shocking Language: Understanding the macroeconomic effects of central bank communication (Discussion Papers No. 1537). Centre for Macroeconomics (CFM).

Henry, E., 2008. Are investors influenced by how earnings press releases are written? *The Journal of Business Communication* (1973) 45, 363–407.

Hu, M., Liu, B., 2004. Mining and summarizing customer reviews, in: *Proceedings of the Tenth Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, KDD '04*. ACM, New York, NY, USA, pp. 168–177.

Jurafsky, D., Martin, J.H., 2008. *Speech and language processing: An introduction to speech recognition, computational linguistics and natural language processing*. Upper Saddle River, NJ: Prentice Hall.

Kaminski, J., Gloor, P.A., 2014. Nowcasting the bitcoin market with twitter signals. Available from: <[arxiv.org/pdf/1406.7577](https://arxiv.org/pdf/1406.7577v1.pdf)>. [11 September 2020].

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

Loughran, T., McDonald, B., 2016. Textual analysis in accounting and finance: A survey. *Journal of Accounting Research* 54, 1187–1230.

Lucca, D.O., Trebbi, F., 2009. *Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements* (NBER Working Papers No. 15367). National Bureau of Economic Research, Inc.

Mariano, R.S., Murasawa, Y., 2003. A new coincident index of business cycles based on monthly and quarterly series. *Journal of applied Econometrics* 18, 427–443.

Marta, B., Michele, M., 2010. Maximum likelihood estimation of factor models on data sets with arbitrary pattern of missing data. European central bank. Working Paper Series.

McLaren, N., Shanbhogue, R., 2011. Using internet search data as economic indicators. *Bank of England Quarterly Bulletin* 51, 134–140.

Mohammad, S.M., Turney, P.D., 2013. CROWDSOURCING a word–emotion association lexicon. *Computational Intelligence* 29, 436–465.

Nielsen, F.Ä., 2011. A new anew: Evaluation of a word list for sentiment analysis in microblogs. Available from: <[arxiv.org/pdf/1103.2903](https://arxiv.org/pdf/1103.2903v1.pdf)>. [11 September 2020].

Plutchik, R., 1962. *The emotions: Facts, theories and a new model*. New york, US. Crown Publishing Group/Random House.

- Plutchik, R., 2001. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American scientist* 89, 344–350. Available from: <[jstor.org/stable/27857503](https://www.jstor.org/stable/27857503)>. [11 September 2020].
- Rinker, T.W., 2018. *lexicon: Lexicon data*. Buffalo, New York.
- Rinker, T.W., 2019. *sentimentr: Calculate text polarity sentiment*. Buffalo, New York.
- Rinker, T.W., 2020. *qdap: Quantitative discourse analysis package*. Buffalo, New York.
- Salton, G., Buckley, C., 1988. Term-weighting approaches in automatic text retrieval. *Information Processing & Management* 24, 513–523.
- Shumway, R.H., Stoffer, D.S., 1982. An approach to time series smoothing and forecasting using the EM algorithm. *Journal of time series analysis* 3, 253–264.
- Si, J., Mukherjee, A., Liu, B., Li, Q., Li, H., Deng, X., 2013. Exploiting topic based twitter sentiment for stock prediction, in: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Association for Computational Linguistics, Sofia, Bulgaria, pp. 24–29.
- Silge, J., Robinson, D., 2016. Tidytext: Text mining and analysis using tidy data principles in R. *JOSS* 1.
- Stock, J.H., Watson, M.W., 2005. Implications of dynamic factor models for var analysis. *National Bureau of Economic Research*.
- Harvard University, 2000. *Harvard iv-4 dictionary*. Available from: <wjh.harvard.edu/~inquirer/homecat.htm>. [11 September 2020].
- Varian, H., Choi, H., 2009. Predicting the present with google trends. *Economic Record* 88.
- Welbers, K., Atteveldt, W.V., Benoit, K., 2017. Text analysis in r. *Communication Methods and Measures* 11, 245–265.
- Young, L., Soroka, S., 2012. Affective news: The automated coding of sentiment in political texts. *Political Communication* 29, 205–231.
- Zhang, X., Fuehres, H., Gloor, P.A., 2011. Predicting stock market indicators through twitter “i hope it is not as bad as i fear”. *Procedia - Social and Behavioral Sciences* 26, 55–62.

Appendix 1: Data used in Nowcasting Model

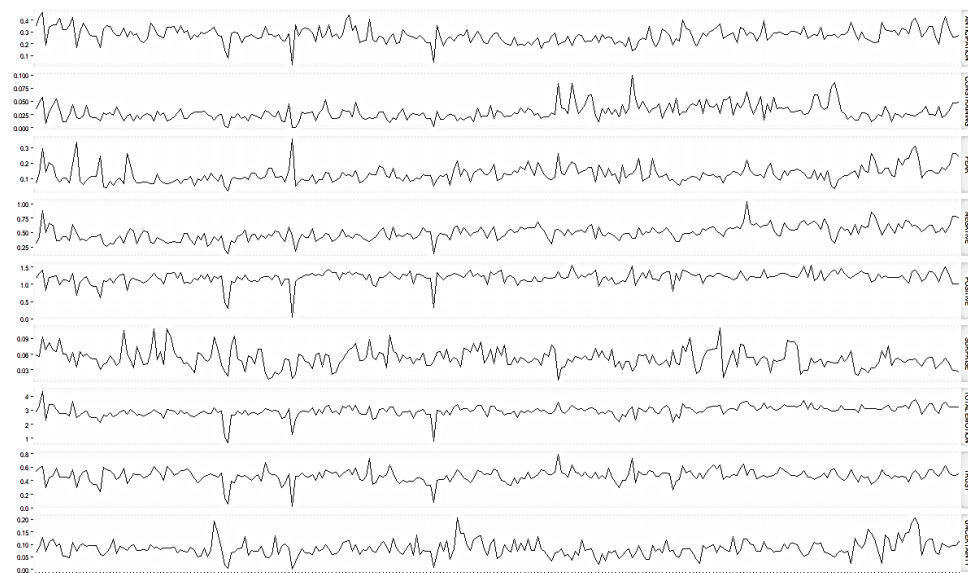
Group	Variables	Publication Delay	Frequency	Starting Date	Transformation	Transformation in DFM
Real	Real GDP (Final	2 years	Quarterly	1991-Q2	Quarter-on-quarter percentage growth	Quarterly
Real	Real GDP (Quarterly Estimates)	45 days	Quarterly	2003-Q4	Quarter-on-quarter percentage growth	Quarterly
Real	Industrial Production %MoM (first estimate)	40 days	Monthly	1996-M12	Month-on-month percentage growth	Monthly
Real	Factory Orders %MoM (total value of new	35 days	Monthly	2004-M01	Month-on-month percentage growth	Monthly
Real	Retail Sales %MoM	30 days	Monthly	2004-M02	Month-on-month percentage growth	Monthly
Survey	Markit PMI (Manufacturing-Flash)	7 days *in advance	Monthly	1996-M4	None	Quarterly
Survey	Markit PMI (Services-Flash)	7 days *in advance	Monthly	1997-M6	None	Quarterly
Survey	IFO-CES Business Climate	7 days *in advance	Monthly	1991-M01	None	Year-on-year growth
Survey	IFO-CES Business	7 days *in advance	Monthly	1991-M01	None	Quarterly
Survey	IFO-CES Export	7 days *in advance	Monthly	1991-M01	None	Quarterly
Survey	ZEW Economic Sentiment	14 days *in advance	Monthly	2004-M01	None	Monthly
Emotion	Anger	Concurrent	Monthly	1997-M02	None	Monthly
Emotion	Anticipation	Concurrent	Monthly	1997-M02	None	Monthly
Emotion	Constraining	Concurrent	Monthly	1997-M02	None	Monthly
Emotion	Fear	Concurrent	Monthly	1997-M02	None	Monthly
Emotion	Joy	Concurrent	Monthly	1997-M02	None	Monthly
Emotion	Negative	Concurrent	Monthly	1997-M02	None	Monthly
Emotion	Positive	Concurrent	Monthly	1997-M02	None	Monthly
Emotion	Sadness	Concurrent	Monthly	1997-M02	None	Monthly
Emotion	Surprise	Concurrent	Monthly	1997-M02	None	Monthly

Group	Variables	Publication Delay	Frequency	Starting Date	Transformation	Transformation in DFM
Emotion	Trust	Concurrent	Monthly	1997-M02	None	Monthly
Emotion	Uncertainty	Concurrent	Monthly	1997-M02	None	Monthly
Emotion	Total Sentiment	Concurrent	Monthly	1997-M02	None	Monthly
Emotion	Number of Speech	Concurrent	Monthly	1997-M02	None	Monthly

Source: Authorial computation.

All the indicators in the data are taken from the source as seasonally adjusted. Emotion indicators have no significant seasonality.

Appendix 2: Emotion indicators used in the nowcasting models



Source: Authorial computation.

