

Marozzi, Armando

Working Paper

The ECB's tracker: Nowcasting the press conferences of the ECB

ECB Working Paper, No. 2609

Provided in Cooperation with:

European Central Bank (ECB)

Suggested Citation: Marozzi, Armando (2021) : The ECB's tracker: Nowcasting the press conferences of the ECB, ECB Working Paper, No. 2609, ISBN 978-92-899-4862-3, European Central Bank (ECB), Frankfurt a. M.,
<https://doi.org/10.2866/214406>

This Version is available at:

<https://hdl.handle.net/10419/246186>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



EUROPEAN CENTRAL BANK
EUROSYSTEM

Working Paper Series

Armando Marozzi The ECB's tracker: nowcasting
the press conferences of the ECB

No 2609 / October 2021

Abstract

This paper proposes an econometric framework for nowcasting the monetary policy stance and decisions of the European Central Bank (ECB) exploiting the flow of conventional and textual data that become available between two consecutive press conferences. Decompositions of the updated nowcasts into variables' marginal contribution are also provided to shed light on the main drivers of the ECB's reaction function at every point in time. In out-of-sample nowcasting experiments, the model provides an accurate tracking of the ECB monetary policy stance and decisions. The inclusion of textual variables contributes significantly to the gradual improvement of the model performance.

JEL classification: E37, E47, E52

Keywords: Natural Language Processing, Dynamic Factor Model, Monetary Policy, Forecasting

Non-technical summary

When the European Central Bank (ECB) was founded in 1999, the Governing Council used to take policy decisions twice a month with a press conference taking place only on the first meeting of the month. Afterwards, from November 2001 there was only one policy meeting per month coming along with the press conference. From January 2015, the frequency of monetary policy meetings changed again to occur every six weeks. Since its inception, thus, the “blind spot” between two consecutive ECB press conferences has increased. At the same time, however, a wide range of data are published between two consecutive Governing Council meetings at daily or monthly frequencies and they contain valuable information on the expected path of monetary policy. Therefore, although the contemporaneous monetary policy stance and decisions of the incoming press conference are not available, they can be estimated exploiting higher-frequency variables that are released in a more timely manner.

This paper provides an econometric framework which enables researchers to track the real time evolution of the monetary policy stance and decisions of the ECB on the basis of the increasing amount of information that becomes available between two consecutive press conferences. First, ECB monetary policy stance is derived by applying an ECB field-specific dictionary to the introductory statements of the press conferences. Second, the flow of information includes conventional and textual data with a sampling frequency which is equal or higher than monthly. While conventional data encompass a wide range of macro and financial variables, textual data refer to a dataset with around 300,000 documents (newspaper articles, magazines, etc.) that explicitly mention the ECB in their content. Based on these data sources, the model outputs real-time forecasts of the ECB monetary policy stance at every new release of data as well as the expected probability that the ECB, conditional on the incoming data, will actually announce a monetary policy decision. Additionally, the model allows to decompose the model-based news from each data release and assess its contribution to the forecast revision. This enables researchers to understand what drives ECB monetary policy at every point in time.

The model is estimated recursively from January 2005 to December 2020 and updated for every week between two press conferences. The empirical results are the following. First, I develop a Dynamic Factor Model (DFM) with mixed-frequency conventional and textual variables to estimate the contemporaneous monetary policy stance of the ECB. Second, the model

provides an accurate tracking of the ECB monetary policy stance and decisions at historical ECB announcements. Third, the model proves to be useful in forecasting the Euro Overnight Index Average (EONIA) rates from January 2008 to December 2009. Fourth, the model provides higher forecast accuracy than competing models. Last, the inclusion of textual variables in the dataset contributes significantly to the improvement of the forecasting performance.

1. Introduction

Central banks usually announce changes to their monetary policy stance in press conferences. The interval between two consecutive press conferences, however, can be significantly long. Even if macroeconomic and/or financial conditions change abruptly, central bank watchers need to wait the nearest press conference to see confirmed or denied their expectations on central banks' decisions. Market participants and policy-makers can thus only rely on non-systematic strategies to interpret the future path of monetary policy.

This paper provides an econometric framework which enables interested parties to track systematically the real time evolution of the monetary policy stance and decisions of a central bank on the basis of the increasing amount of information that becomes available between two consecutive press conferences.

I focus on the European Central Bank (ECB). Since its inception, the “blind spot” between two consecutive ECB press conferences has increased. In 1999, the Governing Council used to take policy decisions twice a month with a press conference taking place only on the first meeting of the month. Afterwards, from November 2001 there was only one policy meeting per month coming along with the press conference. From January 2015, the frequency of monetary policy meetings changed again to occur every six weeks. At the same time, however, a wide range of conventional and textual data are published between two consecutive Governing Council meetings at daily or monthly frequencies that contain valuable information on the expected path of monetary policy. Therefore, although the contemporaneous monetary policy stance and decisions of the incoming press conference are not available, they can be estimated exploiting higher-frequency variables that are released in a more timely manner.

First, I construct ECB field-specific dictionaries and apply them to subsets of the introductory statements of the press conferences to derive the indexes of ECB monetary policy stance, economic and inflation outlook. Second, I structure a textual dataset with around 300,000 documents into daily time series with macroeconomic information. Third, I model the total dataset containing around 140 variables as a Dynamic Factor Model (DFM) with flow and stock variables estimated at daily frequency. The DFM is augmented with an auxiliary equation that takes the specification of a multinomial logit with three possible monetary policy outcomes: ease, constant and hike. Last, I set up a “pseudo” Taylor rule to assess the performance of

the DFM. Overall, the model produces three pieces of information: the nowcast of the ECB monetary policy stance, the forecast of the conditional probability that the ECB will actually take a monetary policy decision at time $t + 1$ and the block of variables that drives the revision of each nowcast at every point in time.

The empirical results are the following. First, I develop a DFM with mixed-frequency conventional and textual variables to estimate the contemporaneous monetary policy stance of the ECB. Second, the model provides an accurate tracking of the ECB monetary policy stance and decisions at historical ECB announcements. Third, the model proves to be useful in forecasting the Euro Overnight Index Average (EONIA) rates from January 2008 to December 2009. Fourth, the model provides higher forecast accuracy than competing models. Last, the inclusion of textual variables in the dataset contributes significantly to the improvement of the forecasting performance.

This paper refers to three strands of literature. The first one is on nowcasting. Since the release of [Giannone et al. \(2008\)](#)'s seminal paper, the nowcasting literature has significantly developed methodologically and empirically (see [Bok et al., 2018](#) for a survey). In particular, this paper is directly related to [Bańbura and Modugno \(2014\)](#) as it draws on their Expectation Maximization (EM) algorithm to estimate the DFM and to [Bańbura et al. \(2013\)](#) for laying down the strategy to model mixed frequency flow and stock variables¹. [Thorsrud \(2020\)](#) and [Cimadomo et al. \(2020\)](#) come also close to this paper: the former employs textual variables for nowcasting Norwegian GDP, while the latter, among other things, proposes a mixed-frequency VAR to forecast the Fed Funds rate given the latest news on US economic conditions.

The second strand of literature studies forecasting interest rate decisions. Following [Taylor \(1993\)](#), it has become common to characterize central bank policy as an interest rate rule (the so-called Taylor-rule) that responds to inflation and the output gap or other combinations of macro variables (for a survey, see [Wieland and Wolters, 2013](#)). More recently, with the ascent of text-mining techniques, macroeconomists augmented the stylized Taylor rule with textual variables capturing central bank communication. In particular, many papers attempted to study whether ECB communication helps predict future monetary policy ([Sturm and de Haan, 2011](#); [Picault](#)

¹ Although the main focus of the paper is on state-space models, it is worth acknowledging that a significant part of the literature on nowcasting also includes MIXed Data Sampling (MIDAS) regressions (see [Ghysels et al., 2004](#); [Ghysels et al., 2007](#); [Clements and Beatriz, 2008](#); [Kuzin et al., 2009](#); [Marcellino and Schumacher, 2010](#)).

and Renault, 2017; Bennani and Neuenkirch, 2017; Bennani et al., 2020; Baranowski et al., 2021).

The communication of the Federal Open Market Committee (FOMC) also attracted significant attention. For instance, Lucca and Trebbi (2009) quantify the statements released by the FOMC and find that FOMC communication is a more important determinant of Treasury rates than contemporaneous policy rate decisions. Similarly, Hansen and McMahon (2016), measuring the information released by the FOMC, find that shocks to forward guidance are more important than the FOMC communication of current economic conditions in terms of their effects on market and real variables. Hansen et al. (2017), taking advantage of the release of tape-recorded FOMC minutes (October 1993) and of text mining tools such as latent Dirichlet allocation (LDA), show that the decision to increase transparency results in a trade-off between discipline and conformity. Shapiro and Wilson (2019), estimating the FOMC's loss function from FOMC transcripts, minutes and members' speeches, find that the FOMC had an implicit inflation target of approximately 1.5% over 2000-2013.

The last stream of studies deal with the application of text-analysis techniques beyond central bank communication (see Gentzkow et al., 2019 for a survey). To quote a few recent papers, Ke et al. (2019) introduce a supervised learning framework to extract sentiment information from news articles to predict asset returns. Bybee et al. (2020) propose instead an approach to measuring the state of the US economy structuring 800,000 Wall Street Journal articles over the period 1984-2017. Babii et al. (2021) advance a new structured machine learning regression approach for high-dimensional mixed-frequency time series data. In the practical application, besides conventional macro and financial data, they exploit Bybee et al. (2020)'s textual dataset, to nowcast US GDP and document superior nowcasting performance with respect to the state-of-the-art state space model approach implemented by the Federal Reserve Bank of New York.

To the best of my knowledge, however, no study exploits the flow of information that becomes available between press conferences to continuously and systematically update the nowcasts of the monetary policy stance and decisions of a central bank. Therefore, the main contribution of this paper is the set-up of a unified econometric framework that is novel for describing central banks' reaction function.

The rest of the paper is organized as follows: Section 2 describes the datasets, the text-mining techniques and the estimated news topics. Section 3 details the models. Section 4 illustrates the results and Section 5 concludes the paper.

2. Data and Textual Methods

[Section 2.1](#) derives three indexes from the introductory statements of the press conferences of the ECB: a monetary policy stance index, an economic outlook index and an inflation outlook index; [Section 2.2](#) details the procedural steps to structure and quantify textual data; finally, [Section 2.3](#) describes the total dataset used in the model.

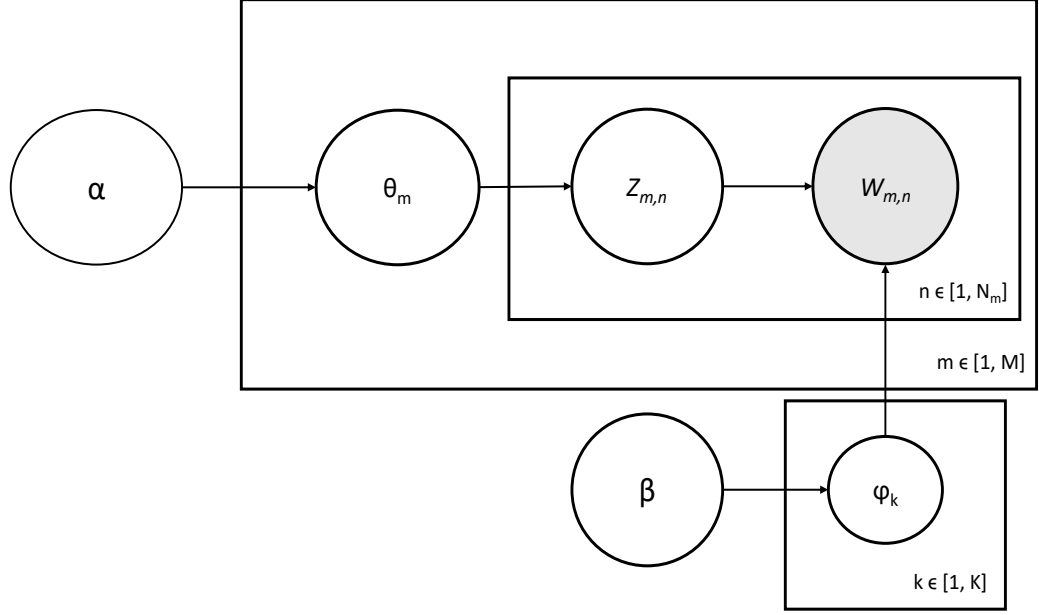
2.1. Quantifying the Press Conferences of the ECB

The ECB's press conferences have a pivotal role in transmitting monetary policy information and are therefore best suited to extracting macro and policy signals. In particular, the aim is to derive indexes of monetary policy stance, economic and inflation outlook. While the first one is the variable to nowcast, the other two convey significant information for the path of monetary policy and, for this reason, will be part of the DFM.

To quantify these indexes, I proceed in three steps: first, I gather press conferences from January 2002 to December 2020 by webscraping the ECB's website. I then remove the Q&A part from every press conference and subset the texts into paragraphs. To further reduce the noise in the data, I apply some common pre-processing steps such as removing stopwords, punctuation, numbers and deleting general expressions with no economic content (e.g. greetings, welcome statements and the like).

Second, I apply Latent Dirichlet Allocation (LDA) model ([Blei et al., 2003](#)) to the corpus of press conferences to identify and cluster together paragraphs that belong to the same topic. This is possible since the fundamental idea of LDA is that documents are represented as a distribution of latent topics, where each topic is characterized by a distribution over words. [Figure 1](#) provides a graphical illustration of the model.

Figure 1: Graphical Representation of LDA



Note: The figure illustrates the Bayesian network of LDA using plate notation. The circle associated with $w_{m,n}$ is gray colored since it indicates that these are the only observable variables in the model.

Allowing bold font indicate the vector version of the variables, it is possible to write the model formally. Let $\mathbf{w} = (w_1, w_2, \dots, w_N)$ be a document formed by a sequence of words where w_n indicates the n^{th} word and let $D = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_M\}$ denote the entire corpus composed by M documents with $N = \sum_{m=1}^M N_m$ being the total number of words in all documents and $N_m = \sum_{n=1}^N w_n$ the total number of words for the m^{th} document. Let also \mathbf{z} be a set of K latent topics. A corpus D has a distribution of topics for each document given by $\boldsymbol{\theta}_m$ and, in turn, each topic has a distribution of words denoted by $\boldsymbol{\varphi}_k$, with both $\boldsymbol{\theta}_m$ and $\boldsymbol{\varphi}_k$ assumed to have conjugate Dirichlet distributions with hyper parameters α and β . Each document \mathbf{w} in the corpus D is an iterated choice of topics $z_{m,n}$ and words $w_{m,n}$ drawn from the multinomial distribution using $\boldsymbol{\theta}_m$ and $\boldsymbol{\varphi}_k$. This can be formally expressed, for the m^{th} document, with the joint distribution of all known and latent variables given the hyper parameters as follows:

$$P(\mathbf{w}_m, \mathbf{z}_m, \boldsymbol{\theta}_m, \boldsymbol{\Phi}; \alpha, \beta) = \underbrace{\prod_{n=1}^{N_m} \underbrace{P(w_{m,n} | \boldsymbol{\varphi}_k) P(z_{m,n} | \boldsymbol{\theta}_m)}_{\text{word level}}}_{\text{document level (1 document)}} \cdot P(\boldsymbol{\theta}_m; \alpha) \cdot \underbrace{P(\boldsymbol{\Phi}; \beta)}_{\text{topic level}} \quad (1)$$

where $\Phi = \{\varphi_k\}_{k=1}^K$ is a $K \times N$ matrix. As equation (1) shows, there are three levels in the LDA model: a document level where every document is a mixture of latent topics, a topic level where every document has a probability to belong to a topic and a word level where every word has a probability to belong to a topic. The final step is to determine K , so far, assumed to be fixed. Perplexity cross-validation measure across Markov chain Monte Carlo (MCMC) iterations (Heinrich, 2009) and minimization of the average cosine distance of topics (Cao et al., 2009) suggest that 8 topics provide the best decomposition of the press conferences. The model is then estimated integrating θ_m and φ_k out of equation (1) and using Gibbs sampling simulations as in Griffiths and Steyvers (2004). Further technical details are described in Appendix A. The estimated words' probabilities (φ_k) allow me to identify the three topics of interest (monetary policy stance, economic outlook and inflation outlook)¹. Then, once the topics are identified, I use the probability of each document to belong to one of these three topics (θ_m) to combine together paragraphs with the same content in datasets that I denote PC^{MP} for introductory statements with monetary policy information, PC^{IO} for introductory statements with inflation content, and PC^{EO} for introductory statements with growth outlook indications².

The third step implies the quantification of these datasets. Drawing on Picault and Renault (2017)'s methodology, I create ECB-field specific dictionaries for each topic c . I first subset each dataset PC^c into combinations of words (hereafter n -gram) and manually classify each n -gram into the topic c (monetary policy, inflation and economic outlook) with tone κ (hawkish, neutral, dovish for monetary policy and positive, neutral, negative for inflation and economic outlook)³. I then compute the probability that every n -gram belongs to each of the corresponding category. Next, each n -gram is classified as positive, 1, or negative, -1, on the basis of which category has the highest probability and kept only if the probability is greater than 50%. In parallel, I create an ECB-specific dictionary for valence shifters (i.e., negators, amplifiers, de-amplifiers and adversative conjunctions). Once the polarized dictionary is constructed, I apply Rinker (2019)'s methodology to measure the tone of a document. The algorithm breaks each press conference into sentences and, in turn, each sentence into an ordered bag of words. The word w in each

¹ Appendix A provides insights on topics' identification.

² Since every document m has a probability θ_m to belong to any of the three topics, the m^{th} document is assigned to topic k only if the topic with the highest probability (θ_m^{MAX}) exceeds 35%.

³ For instance, the n^{th} n -gram of PC^{IO} is assigned to one of the three categories for inflation: inflation outlook positive, inflation outlook neutral and inflation outlook negative. The same procedure holds for all the datasets.

sentence τ is then compared to the dictionaries of polarized words just described. These polarized words form a polar cluster $\gamma_{w,\tau}$, that is, a subset of a sentence ($\gamma_{w,\tau} \subset \tau$) where every polarized word (w_γ^p) in the cluster is preceded and succeeded by valence shifters that weight the impact of the reference word by a factor η set by the researcher. Amplifiers w_a (de-amplifiers w_d) increase (decrease) the polarity by η in such a way that $w_a = \sum[\eta \cdot (w_\gamma^{neg} \cdot w_\gamma^a)]$ where $w_\gamma^{neg} = (-1)^{2+\sum w_\gamma^n}$ and w_γ^n stands for the n^{th} negator in the j^{th} cluster⁴. Amplifiers become de-amplifiers w_d if there is an odd number of negators w_γ^n in the cluster. This is so because w_γ^{neg} is positive for an even number of negators and negative otherwise; such a logic is based on the rule that two negatives equal a positive, three negatives a negative, and so on. As a result, negations can also change the sign of the polarized word. On the other hand, an adversative conjunction w_{advcon} *before* the polarized word up-weights the cluster by $1 + [\eta \cdot (w_{advcon})]$, whereas an adversative conjunction *after* the polarized word down-weights the cluster by $1 + [(w_{advcon} - 1) \cdot \eta]$. This resembles the belief that an adversative conjunction augments the weight of the next clause while reducing the weight attributed to the prior clause. Overall, the score for each sentence s is computed following the equation:

$$\psi_\tau = \frac{\gamma_{w,\tau}^n}{\sqrt{\sum_{n=1}^N w_n}} \quad (2)$$

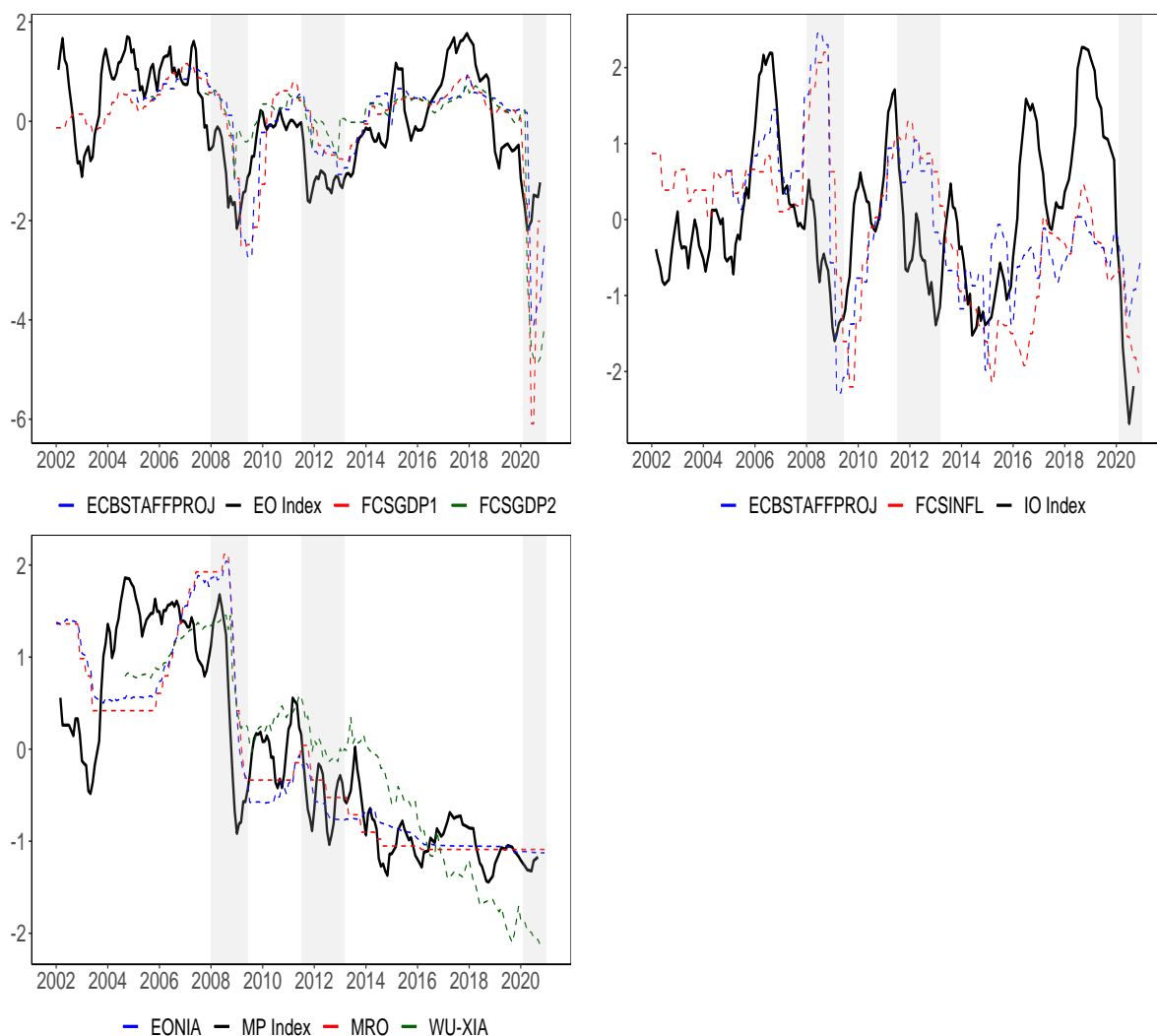
where $\gamma_{w,\tau}^n = \sum[(1 + w_a + w_d) \cdot w_\gamma^p \cdot w_\gamma^{neg}]$ is the sum of single polar clusters and $\sqrt{\sum_{n=1}^N w_n}$ is the square root of the total number of words in a sentence. To obtain the mean of all sentences within a press conference I simply calculate the average sentiment score $PC^d = \frac{1}{n} \sum \psi_\tau$.

This method departs from the frequently used [Loughran and McDonald \(2011\)](#) (LM) and the [Apel and Blix Grimaldi \(2012\)](#) (AG) dictionaries in at least two dimensions: on the one hand, it is crafted on the ECB communication on a specific topic, that is, it avoids misinterpreting words' tone because of a different context; on the other hand, it considers a mixture of single words as well as sequence of n words, thereby preventing from attributing the wrong sign to specific n -grams (e.g. "lower unemployment" vs "unemployment"). Besides, unlike [Picault and Renault \(2017\)](#), I calculate the average tone of a press conference controlling for valence shifters. Such a method provides a much more robust inference since it is able to handle linguistic nuances

⁴ The summation symbol in $w_a = \sum[\eta \cdot (w_\gamma^{neg} \cdot w_\gamma^a)]$ indicates the sum of amplifiers in the j^{th} polar cluster.

that approximate the natural language. Figure 2 displays the evolution of the economic outlook, inflation outlook and monetary policy stance indexes against relevant macroeconomic data.

Figure 2: Indexes Based on the Introductory Statements of the Press Conferences



Note: The figure plots the time series of economic outlook, inflation outlook and monetary policy stance against key macro variables. The variables are standardized to facilitate the comparison. Since the indexes are somewhat jagged, I show the moving average.

In particular, the economic outlook index is strongly correlated with the ECB's GDP projections as well as the Bloomberg weighted-average of private (FCSGDP1) and official (FCSGDP2) GDP forecasts. The inflation outlook index, instead, underperformed the ECB (ECBSTAFFPROJ) and private (FCSINFL) inflation forecasts in 2008 and 2013, while it outperformed them from 2016 to 2019. Lastly, the monetary policy stance index closely resembles the dynamics of EONIA rates, the marginal refinancing operations (MRO) and the Wu-Xia ECB shadow rate.

2.2. Words as Data: Augmenting the Nowcasting Exercise with Textual Data

In addition to conventional macro and financial data, I augment the model with information extracted from textual data at daily frequency. The dataset is taken from an internal ECB's database and contains around 300,000 documents from 19 September 2004 to 31 December 2020⁵. This dataset includes newspapers' articles, online websites, magazines, TV news, etc. that explicitly mention the ECB in their content (see [Appendix A](#) for details).

To make this dataset applicable for time series analysis, I follow a similar procedure to the one employed in [Section 2.1](#). After pre-processing the dataset, I decompose the textual corpus into news topics using LDA estimated with $K = 80$ ⁶. Out of 80 topics, I identify and label 60 of them which I reduce to 40 once filtering for monetary policy relevance⁷. The remaining 40 topics are then clustered into 7 meta-topics on the basis of the similarity among topics. These 7 meta-topics are: "Financial Crisis", "Eurexit", "European Banks", "Inflation Outlook", "Economic Outlook", "Monetary Policy" and "Fiscal Policy"⁸. [Figure 3](#) shows the evolution of the topic probability of the first three topics. I refer to [Appendix A](#) for further details.

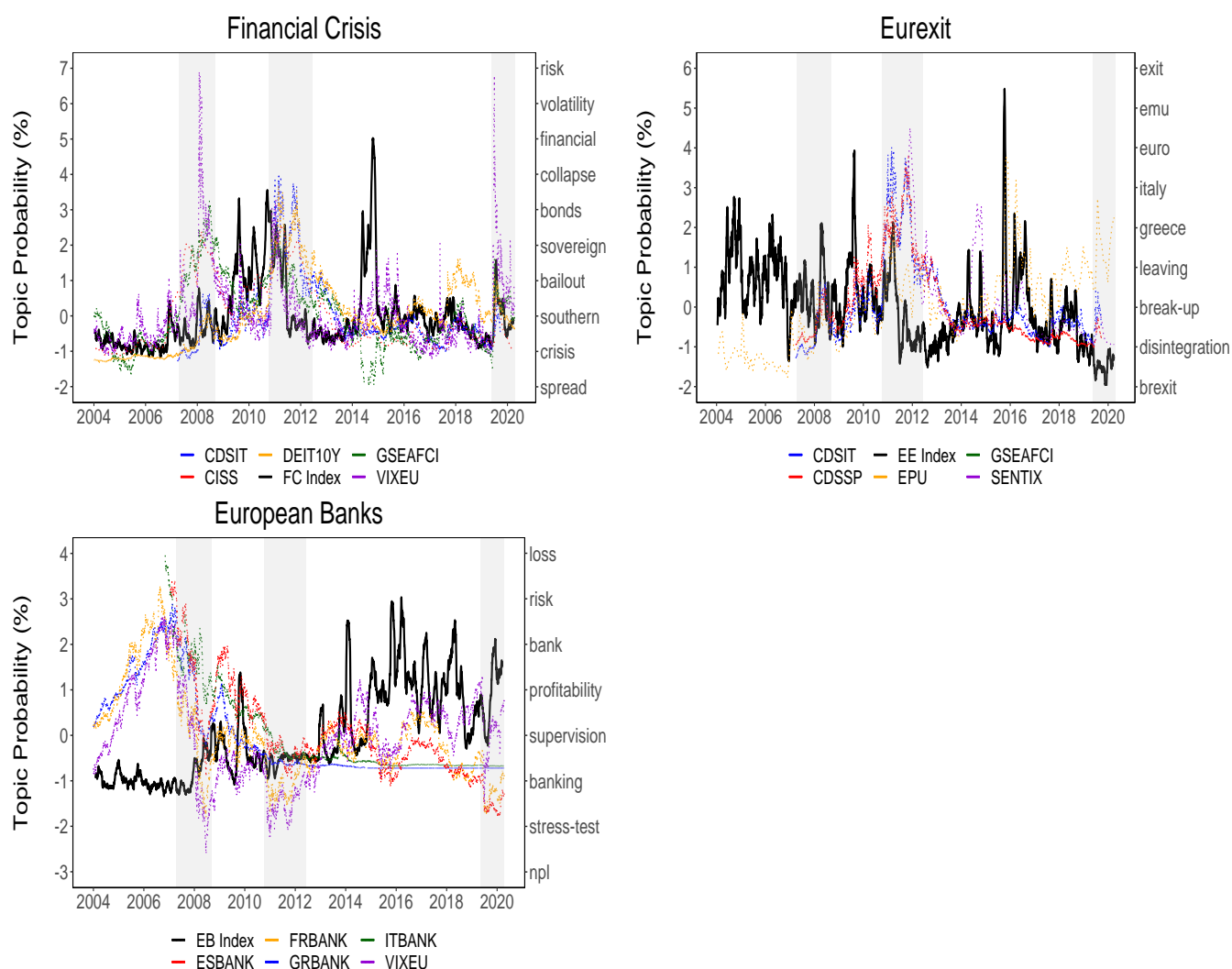
⁵ "Internal" does not mean "private". In fact, the textual information is public. ECB staff only translated every non-English article into English and every non-textual information (Radio interviews and podcasts) into text.

⁶ Similar to the previous section, I estimate the optimal number of topics using perplexity cross-validation measure across Markov chain Monte Carlo (MCMC) iterations ([Heinrich, 2009](#)) and minimization of the average cosine distance of topics ([Cao et al., 2009](#)).

⁷ I identify and label only 60 topics out of 80 since the remaining topics were not clearly identifiable.

⁸ It should be noted that while the indexes for inflation and economic outlook in [Section 2.1](#) represent the ECB internal assessment of those macro variables, the ones derived here reflect the tone of a much broader "audience".

Figure 3: Topic Probabilities for Daily Indexes



Note: The figure plots the evolution of the topic probabilities derived from the DTM for “Financial Crisis”, “Eurexit” and “European banks”. The variables are standardized to facilitate the comparison. The second y-axis shows the words with the highest probability to belong to a certain topic. Since the indexes are somewhat jagged, I display the moving average.

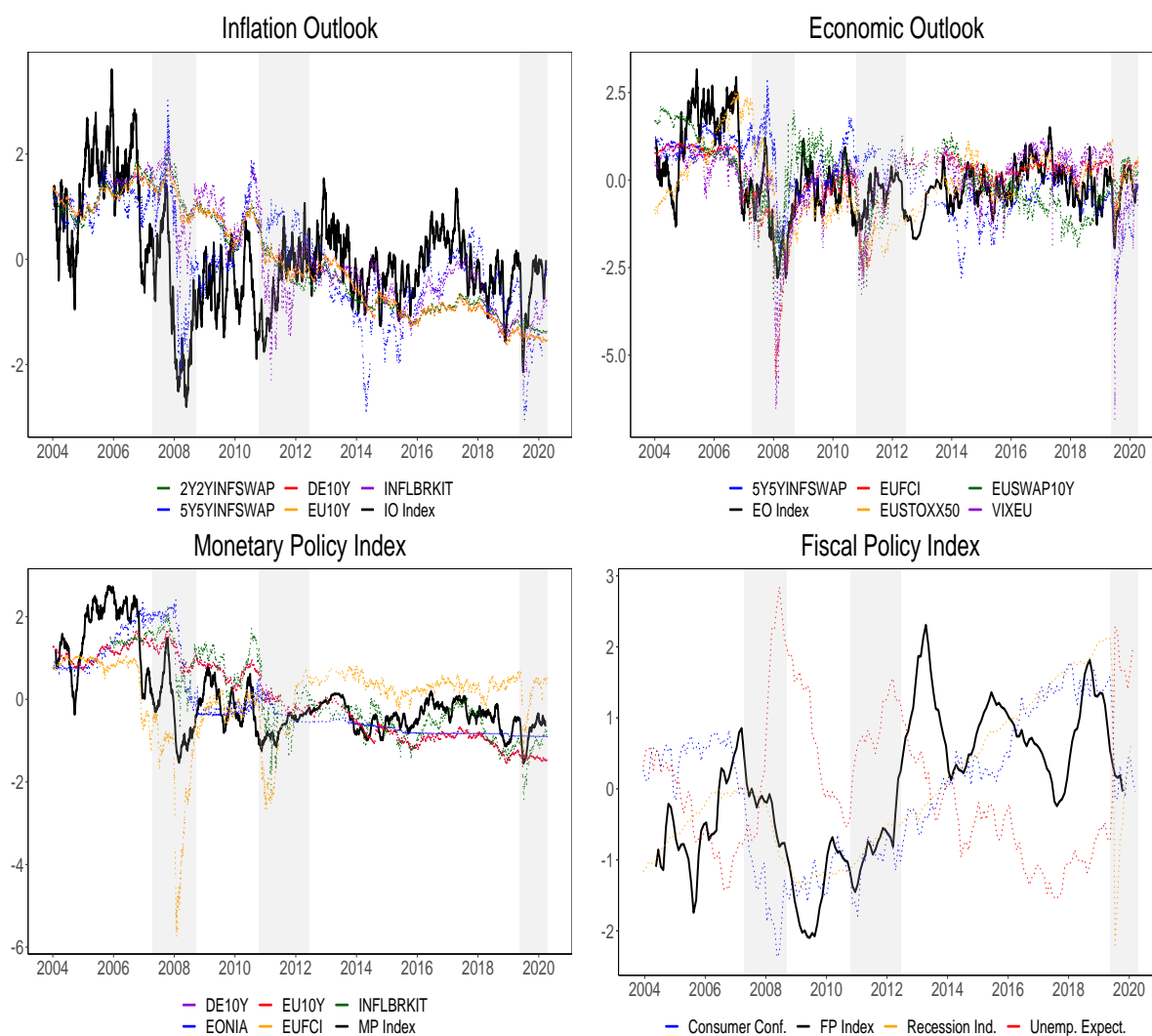
As the figure reports, the “Financial Crisis” and “Eurexit” series to a large extent correlate with a wide range of indicators from EU volatility Index (VIXEU) to the composite indicator of systemic stress (CISS) (Kremer et al., 2012), while the “European banks” index, especially from 2008 on, follows a similar evolution to a variety of bank-related indicators⁹.

⁹ In the figure’s legend, “FRBANK” refers to FR CAC 40, “ITBANK” to MIB, “ESBANK” to IBEX 30 and “GRBANK” to FTSE/ATHEX Banks Index. Additionally, “GSEAFCI” indicates the Goldman Sachs index of financial conditions in the euro area, “EPU” stands for the Economic Policy Uncertainty Index for the euro area

While for the topics in [Figure 3](#) the creation of field-specific dictionaries is not necessary as it would not add much information to the evolution of the topic probabilities, the remaining four topics require a quantification step to be made. To do so, I proceed in five steps that I describe in more detail in [Appendix A](#): first, as in [Section 2.1](#), I exploit LDA’s topic probabilities to create content-specific datasets. Second, I apply word embedding ([Stoltz and Taylor, 2019](#)) to identify patterns of the co-occurrence of words within a “context window” centred around, respectively, “inflation”, “economy”, “monetary”, and “fiscal”, where these key words simply correspond to the topics of the field-specific dictionaries. Third, I compare the vector of words just obtained with the field-specific dictionaries in [Section 2.1](#). If the word in the vector does not appear in the existing dictionary, I manually classify it to be either positive or negative and add it to the relative field-specific dictionary. For the fiscal policy index I use the same procedure but the comparison occurs with a dictionary of fiscal words developed in [Marozzi \(2021\)](#). Fourth, I apply equation (2) to score and aggregate each sentence for each content-specific dataset. Fifth, I filter out the bias that can come from the country of origin of the document by regressing the sentiment score on countries’ dummies. [Figure 4](#) presents the results.

and “SENTIX” is the Sentix Investor Confidence.

Figure 4: Quantified Indexes Derived from the Textual Dataset



Note: The figure plots the quantified indexes of inflation outlook, economic outlook, monetary policy index and fiscal policy index derived from the textual dataset. The variables are standardized to facilitate the comparison. Since the indexes are somewhat jagged, I display the moving average.

Notably, while the index for inflation outlook closely follows 5Y5Y inflation swap rates, the dynamics of the economic outlook index somewhat correlates with VIX EU and the European stock market index (EUSTOXX50). Moreover, the index for monetary policy resembles the evolution of a bunch of risk-free rates in the euro area. Instead, the index for fiscal policy appears to have its closest fit with the European Commission’s Consumer Confidence Indicator¹⁰.

¹⁰In the figure’s legend, “INFLBRKIT” indicates the Italian 10Y inflation break-even rate, “EUFCI” stands for the OIS-LIBOR spread and “EUSWAP10Y” is Euro 10Y swaps.

2.3. Total Dataset

The total dataset consists of 140 macroeconomic indicators for the euro area at daily, monthly and irregular frequency, with the latter being denoted as the frequency of the introductory statements of the press conferences. Given the mixed frequency nature of the variables, I follow [Mariano and Murasawa \(2003\)](#) in writing the dataset at the highest frequency, i.e. daily, and therefore assuming that lower frequency variables are missing periodically¹¹. Data are collected in December 2020 with the sample starting in January 2002 and transformed to induce stationarity. Since real-time vintages for every series are not available, the dataset is to be considered a “pseudo” real-time dataset.

To provide more detail on the structure of the dataset, the variables have been aggregated into eleven blocks with similar economic content and release day ([Table 1](#))¹². For a complete breakdown of the dataset I refer to [Appendix B](#). Starting with the daily frequency, I group the variables into four blocks: *rates and spreads*, *financial*, *forecasts*, and *textual*. While *rates and spreads* contains government bond rates and spreads, risk-free rates, overnight index swaps (OIS), mortgage and break-even rates, the *financial* block includes stock indexes, banking and credit data and credit default swaps (CDS) for the main European countries. Moreover, the *forecasts* block comprises private and official forecasts for GDP, inflation and unemployment, whereas the *textual* block is composed of the variables derived in [Section 2.2](#)¹³. As for the monthly frequency, while *output* spans industrial production, unemployment and exchange rates variables, *price* contains various inflation indexes; *surveys* denotes Purchasing Managers Indexes (PMIs), EuroCOIN indicator and the European Commission’s surveys; the *mixed* group indicates variables with the same release period ranging from sovereign CISS to indexes of financial stress and economic sentiment; ECB’s loans to household, financial and non-financial institutions, ECB’s holdings of securities, M3 and a range of ECB key rates are subsumed into the *monetary* group; the *US* block contains US variables ranging from PMIs to CPI. Finally, the last block for irregular data simply includes the textual variables extracted from press conferences in

¹¹ In so doing, I also avoid applying any transformation to the variables at irregular frequency, among which there is the variable to nowcast.

¹² Because the timing and order of data releases vary only slightly from month to month, I assume that the pattern of data availability is unchanged throughout the evaluation sample.

¹³ To have forecasts in a unique release block, when forecasts were not available at a daily frequency, I filled missing observations with previous values.

Table 1: Total Dataset by Blocks

Block	Timing	Delay	Frequency	Number
Financial	Daily	No delay	Daily	16
Forecasts	Daily	No delay	Daily	6
Rates and Spreads	Daily	No delay	Daily	41
Textual Newspapers	Daily	No delay	Daily	7
Prices	Mid-month	One month	Monthly	9
Output	Mid-month	One month	Monthly	10
Surveys	End of month	No delay	Monthly	14
Mixed	End of month	One Month	Monthly	8
Monetary	End of month	One month	Monthly	19
US	Mixed	Mixed	Monthly	7
Textual PC	Press conferences	No Delay	Irregular	3

Note: The first column reports the block in which the released variable are included. The second column indicates the official dates of the publication. The third one reports the lag with which the data are released. The frequency of the data is reported in the fourth one, while in the last column is displayed the number of variables per group. Data have been collected from Haver and Bloomberg.

3. Methodology

Section 3.1 details the features of the Dynamic Factor Model (DFM) used to track in real time the developments of the ECB monetary policy stance and decisions and Section 3.2 explains the benchmark model employed to compare the performance of the DFM.

3.1. The Dynamic Factor Model

I build on Modugno (2013), Bańbura et al. (2013) and Bańbura and Modugno (2014) to develop a mixed-frequency Dynamic Factor Model with flow and stock variables. Let $Y_t^{k,n}$ be the collection of variables $n \in \{f, s\}$ in Section 2.3 where f denotes a flow variable and s a stock variable at frequency $k \in \{i, m, d\}$ where i stands for irregular, m for monthly and d for daily. Let $F_t^{k,n}$ also denote the corresponding unobservable factor for $Y_t^{k,n}$. The measurement equation can be written as follows:

$$\begin{bmatrix} Y_t^{i,f} \\ Y_t^{m,f} \\ Y_t^{m,s} \\ Y_t^d \end{bmatrix} = \begin{bmatrix} \tilde{\Lambda}^{i,f} & 0 & 0 & 0 \\ 0 & \tilde{\Lambda}^{m,f} & 0 & 0 \\ 0 & 0 & \Lambda^{m,s} & 0 \\ 0 & 0 & 0 & \Lambda^d \end{bmatrix} \begin{bmatrix} \tilde{F}_t^{i,f} \\ \tilde{F}_t^{m,f} \\ F_t^{m,s} \\ F_t^d \end{bmatrix} + \begin{bmatrix} E_t^{i,f} \\ E_t^{m,f} \\ E_t^{m,s} \\ E_t^d \end{bmatrix}, \quad (3)$$

where $\Lambda^{k,\cdot}$ are the factor loadings for each frequency k and $\tilde{\Lambda}^{k,f} = (\Lambda^{k,f}, 0)$ is needed to match $\tilde{F}_t^{k,f'} = (F_t^{k,f'}, \bar{F}_t^{k,f'})$ that contains an additional auxiliary aggregator $\bar{F}_t^{k,f}$ for the flow variables; $E_{i,t}$ are idiosyncratic errors such that $E_{i,t} = i.i.d.N(0, \Sigma_E)$ where Σ_E is diagonal. I leave technical details on aggregators in [Appendix C](#). The transition equation with time-varying coefficients can then be written as:

$$\begin{bmatrix} I_{2r} & 0 & 0 & \mathcal{W}_t^{i,f} \\ 0 & I_{2r} & 0 & \mathcal{W}_t^{m,f} \\ 0 & 0 & I_r & \mathcal{W}_t^{m,s} \\ 0 & 0 & 0 & I_r \end{bmatrix} \begin{bmatrix} \tilde{F}_t^{i,f} \\ \tilde{F}_t^{m,f} \\ F_t^{m,s} \\ F_t^d \end{bmatrix} = \begin{bmatrix} \mathcal{I}_t^{i,f} & 0 & 0 & 0 \\ 0 & \mathcal{I}_t^{m,f} & 0 & 0 \\ 0 & 0 & \mathcal{I}_t^{m,s} & 0 \\ 0 & 0 & 0 & \Xi \end{bmatrix} \begin{bmatrix} \tilde{F}_{t-1}^{i,f} \\ \tilde{F}_{t-1}^{m,f} \\ F_{t-1}^{m,s} \\ F_{t-1}^d \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ U_t \end{bmatrix}, \quad (4)$$

where $\mathcal{W}_t^{k,n}$ contain aggregation weights for every frequency higher than the daily one, $\mathcal{I}_t^{k,n}$ are selection matrices of time varying coefficients that take the value of zero the day after each release of data with frequency k and one elsewhere, Ξ is a matrix of autoregressive coefficients for the daily factors that are assumed to follow a VAR(p) process and $U_{i,t} \sim i.i.d.N(0, \Sigma_U)$ where Σ_U is a diagonal matrix.

The model is then estimated using the methodology developed by [Bańbura and Modugno \(2014\)](#) that adapts from [Dempster et al. \(1977\)](#) the Expectation Maximization (EM) algorithm to handle arbitrary patterns of missing data. The algorithm consists of three steps: first, initial parameters are estimated via principal components on a sample of data where missing observations are handled using splines. Afterwards, the algorithm iterates until convergence to a local maximum between the expectation step, where the missing data in the likelihood are filled in by the Kalman filter and smoother, and the maximization step where this likelihood with complete data is optimized (further details in [Appendix C](#)).

Equations (3) and (4) output the nowcasts of monetary policy stance. However, they are not

informative on the expected probability that the ECB, conditional on the incoming data, will actually take a monetary policy decision. To add this important piece of information, I augment the model with a bridge equation whose specification is:

$$P(y_{t+h} = j | \hat{X}_{t+h}) = \Phi(\alpha + \beta' \hat{X}_{t+h}). \quad (5)$$

Equation (5) is a multinomial logit model where y_{t+h} is a categorical variable that takes $j = 3$ values: y_{t+h} equals 1 if the ECB hikes at time $t + h$, 0 if there is no actual change and -1 if the ECB eases at time $t + h$ ¹; α and β are, respectively, a constant and a vector of parameters; $\Phi(\cdot)$ denotes the cumulative distribution function of the logistic distribution; and \hat{X}_{t+1} is a vector containing a set of predictors for which the DFM provided the forecasts at time $t + 1$ ². This bridge equation thus yields updated nowcasts of the conditional probability attached to each outcome whenever new information becomes available.

This model finally provides a framework to structure the flow of data releases in real time to trace out which variables drive the evolution of the nowcast. Drawing on Bańbura and Modugno (2014), the decomposition of the source of nowcast revision can be written formally. Let Ω_t and Ω_{t-1} denote two consecutive data vintages where $\Omega_{t-1} \subset \Omega_t$. Let I_{new} be the observations for the j^{th} variable (y_t^j) in Ω_t , missing in Ω_{t-1} , that the new release made available³. The interest is in inspecting how the release of I_{new} affected the updated nowcast of the ECB monetary policy stance at time t (y_t^{ECB}). Considering two consecutive nowcast updates, $\mathbb{E}(y_t^{ECB} | \Omega_t)$ and $\mathbb{E}(y_t^{ECB} | \Omega_{t-1})$, the expression for the nowcast revision can be written as⁴:

$$\underbrace{\mathbb{E}(y_t^{ECB} | \Omega_t)}_{\text{new forecast}} = \underbrace{\mathbb{E}(y_t^{ECB} | \Omega_{t-1})}_{\text{old forecast}} + \mathbb{E}(y_t^{ECB} | \underbrace{I_{new}}_{\text{news}}). \quad (6)$$

Note that in equation (6), $\mathbb{E}(y_t^{ECB} | I_{new}) \neq 0$ only if $I_{new} = y_t^j - \mathbb{E}(y_t^j | \Omega_{t-1}) \neq 0$, that is, the

¹ Appendix C provides details on how I turned a continuous variable into a categorical one.

² The set of regressors are: core inflation, industrial production, M3, EONIA, 10Y BUND-BTP spread, trade-weighted exchange rate and forecasts of inflation and GDP.

³ To be precise, Ω_{t-1} and Ω_t can differ for two reasons: the first reason is that new observations for the j^{th} variable become available and the second one is that some of the past data might have been revised. I abstract from the latter case.

⁴ Following Bańbura et al. (2013), I exploit the properties of conditional expectation as an orthogonal projection operator to derive the equation. $I_{new} \perp \Omega_{t-1}$ since I_{new} is the “unexpected” (with regard to the model) part of the release. Therefore, the forecast revision is the projection on the space orthogonal to the old forecast. As orthogonality holds, the revision is also uncorrelated to the old forecast.

nowcast is updated only if the new observations in y_t^j are different from the forecasts made at time $t - 1$ ⁵. Therefore, it is not the new release that leads to the revision of the forecast but the unexpected component, the *news*, from the latest release. Abstracting from the problem of parameter uncertainty, the expression for the revision can be developed further as follows⁶:

$$\mathbb{E}(y_t^{ECB}|I_{new}) = \mathbb{E}(y_t^{ECB}I'_{new}) \mathbb{E}(I_{new}I'_{new})^{-1}I_{new} , \quad (7)$$

where, for the state space model written as in equations (3) and (4):

$$\begin{aligned} \mathbb{E}(y_t^{ECB}I'_{new}) &= \Lambda^{k,\cdot} \mathbb{E}[(F_t^{k,n} - F_{t|\Omega_{t-1}}^{k,n})(F_t^{k,n} - F_{t|\Omega_{t-1}}^{k,n})'] \Lambda_j^{k,\cdot'} \\ \mathbb{E}(I_{new}I'_{new}) &= \Lambda_j^{k,\cdot} \mathbb{E}[(F_{tj}^{k,n} - F_{tj|\Omega_{t-1}}^{k,n})(F_{tj}^{k,n} - F_{tj|\Omega_{t-1}}^{k,n})'] \Lambda_j^{k,\cdot'} + \sum_j j , \end{aligned} \quad (8)$$

where $\Lambda^{k,\cdot}$ and $\Lambda_j^{k,\cdot}$ are the factor loadings of the observation equation with the rows corresponding, respectively, to the ECB monetary policy stance variable and the j^{th} variable and \sum_j is a diagonal matrix filled by j^{th} 's data. The expectations are estimated using the Kalman filter and smoother. As a result, it is possible to find a vector of weights $A = (a_1, \dots, a_J)$ such that:

$$\underbrace{\mathbb{E}(y_t^{ECB}|\Omega_t)}_{\text{new forecast}} - \underbrace{\mathbb{E}(y_t^{ECB}|\Omega_{t-1})}_{\text{old forecast}} = A \underbrace{(y_t^j - y_t^j|\Omega_{t-1})}_{\text{news}} . \quad (9)$$

Equation (9) states that the nowcast revision can be decomposed as a weighted average of the news in the most recent release. This relationship allows me to compute the weighted marginal contribution of each variable to the updated nowcast. More precisely, since in the dataset used for the empirical application there is simultaneous release of several variables, the results of the decomposition have been aggregated into groups of series that share approximately the same release day and economic characteristics as outlined in Table 1.

⁵ For example, let us assume that the difference between Ω_t and Ω_{t-1} is the release of some PMI data for the period t . The news is $I_{new}^{PMI} = y_t^{PMI} - \mathbb{E}(y_t^{PMI}|\Omega_{t-1})$ where y_t^{PMI} is a vector containing the last figures released

⁶ The equality follows from the definition of conditional expectations in \mathbb{L}^2 space, that is, the conditional expectation of y given x is equal to the linear regression of y on x , assuming Gaussianity.

3.2. Benchmark Model

In the empirical section, I will compare the performance of the DFM with that of a forecast-based policy rule with contemporaneous and forward-looking measures of inflation and output gap (Jansen and De Haan, 2009) where the dependent variable, rather than being the nominal interest rate, is the monetary policy stance of the ECB (ECB_t) derived in Section 2.1. The choice of a forecast-based policy rule compared to a simple outcome-based one is led by the fact that the ECB is found to set interest rates in a forward-looking manner (see Gerlach, 2007; Gorter et al., 2008)⁷. I then estimate the model in the following specification:

$$ECB_t = \alpha + \phi_\pi(\pi_t - \pi^*) + \phi_\gamma(\gamma_t - \gamma_t^*) + \phi_{\pi,h}\mathbb{E}_t\pi_{t+h} + \phi_{\gamma,h}\mathbb{E}_t\gamma_{t+h} + \epsilon_t, \quad (10)$$

where α is just a constant, $\pi_t - \pi^*$ indicates the difference between the growth rate of core inflation and the ECB inflation target, $\gamma_t - \gamma_t^*$ denotes the difference between the growth rate of output and the output gap, π_{t+h} and γ_{t+h} are respectively the inflation and output growth ECB staff forecasts four quarters ahead ($h = 4$)⁸; ϕ_π , ϕ_γ , $\phi_{\pi,h}$, $\phi_{\gamma,h}$ indicate the ECB's response parameters to be estimated and ϵ_t is a monetary policy shock capturing deviations from the systematic policy response to output and inflation.

4. Empirical Results

The DFM is parametrized with one factor per frequency and variable type and one lag¹. It is then estimated recursively on the time frame ranging from 1 January 2005 to 31 December 2020. The bridge equation is also estimated as part of the entire econometric framework. The performance of this model is thus assessed from different perspectives². First, Figure 5 reports

⁷ Since the literature abounds with different Taylor rule specifications, the choice to focus on one single “benchmark model” may be excessively discretionary. For this reason, in Appendix C I compare the performance of the DFM with alternative benchmark models.

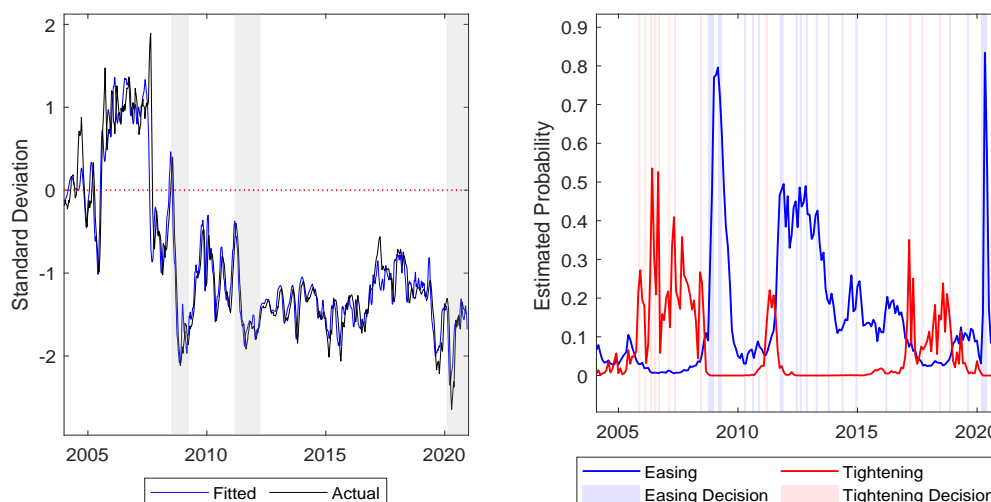
⁸ Constant four-quarter ahead forecasts are derived as follows: December, $y_{t+4} = y_{t+4}$; March, $y_{t+4} = 0.75y_{t+3} + 0.25y_{t+7}$; June, $y_{t+4} = 0.5y_{t+2} + 0.5y_{t+6}$; September, $y_{t+4} = 0.25y_{t+1} + 0.75y_{t+5}$. The series have then been interpolated to fit the regression model. I used industrial production as a measure of output and derived the output gap using the Hodrick-Prescott (HP) filter.

¹ This choice is motivated by simplicity and by the fact that results based on two factors and/or more lags are qualitatively similar to those based on one factor.

² Since the DFM is built to nowcast a variable that I developed, in Appendix D I test the “external” validity of the model nowcasting variables related to the ECB's monetary policy stance but not developed by myself. The

the in-sample estimates of the actual vs fitted values of ECB monetary policy stance (first column) as well as the in-sample fit of the bridge equation (second column).

Figure 5: In-Sample Results



Note: The left figure shows the fitted values (blues) against the actual values (black) of the ECB monetary policy stance, while the figure on the right displays the estimated probability of an easing (blue) and a tightening (red) against actual monetary policy decisions to tighten (shaded red) and ease (shaded blue).

As the figure shows, the fitted values track closely the actual index of monetary policy stance. Besides, the estimated probability of tightening and easing significantly increase in correspondence of introductory statements where such decisions were taken. More precisely, the bridge equation correctly predicts 86% of tightening decisions and 90% of easing decisions³.

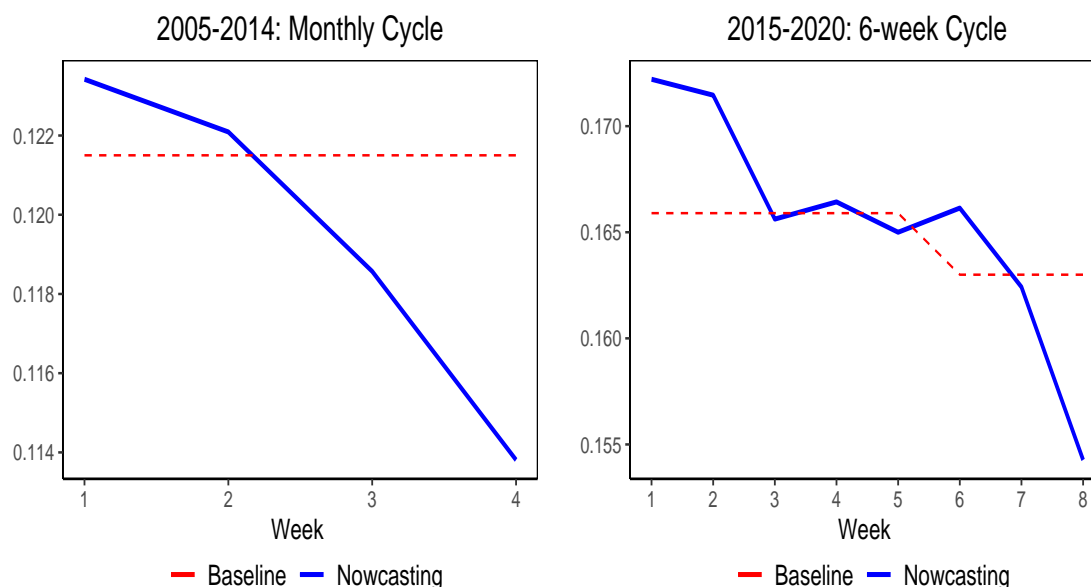
Second, I evaluate the average precision of the nowcasts by running a recursive out-of-sample exercise that produces nowcasts for the nearest press conference. The model is updated every week over the evaluation sample 2005–2020 from the day after the press conference to the day before the press conference and replicates at each point of the forecast evaluation sample the “pseudo” real-time data availability to that point in time. The model is re-estimated each time in order to take into account parameter uncertainty. In particular, I focus on point forecast evaluation using Root Mean Squared Forecast Error (RMSFE) statistics. The accuracy of the DFM-based predictions are then compared to the benchmark model described in [Section 3.2](#). Since the frequency of the ECB press conferences is approximately monthly until 2014 before

³ last exercise carried out in [Section 4](#) provides further evidence on the robustness of the DFM.

³ To understand how these percentages were calculated see [Appendix C](#).

moving to a six-week cycle, I calculate the RMSFE for each subperiod. [Figure 6](#) reports the result.

Figure 6: Root Mean Squared Forecast Error (RMSFE)



Note: The figure shows the RMSFE of the nowcasting model compared to the baseline model outline in [Section 3.2](#). The first column displays the RMSFE for the period 2005-2014, while the second one exhibits the RMSFE from 2015 to 2020.

In either periods, the RMSFE of the DFM displays a downward-sloping behaviour, that is, the more the data become available, the more the accuracy of the prediction improves. In contrast, the RMSFE of the baseline model outlined in [Section 3.2](#) remains constant in the monthly cycle and reduces only in the fifth week of the six-week cycle in correspondence of the release of real activity data. Although the baseline model outperforms the DFM at the beginning of the forecasting period, it tends to lag behind as the weeks progress. These conclusions are also supported by the formal Diebold-Mariano test statistic ([Diebold and Mariano, 1995](#)) whose results are presented in [Table 2](#). The DFM, therefore, maintains an informational advantage compared to the benchmark model due to the ability to process mixed-frequency data in a unified framework.

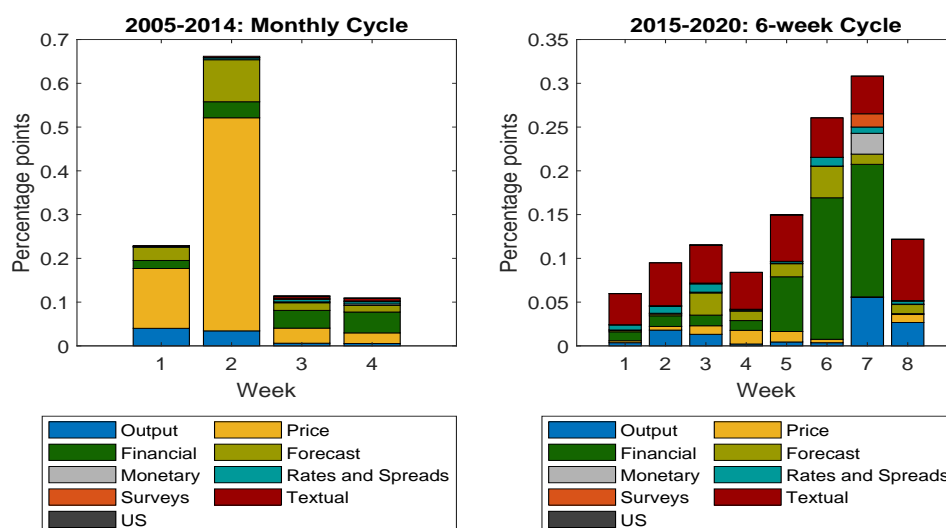
Table 2: Forecast Evaluation

Horizon	2005-2014		Horizon	2015-2020	
	DFM	BM		DFM	BM
1	0.123	0.121	1	0.172	0.165
2	0.122	0.121	4	0.166	0.165
3	0.118**	0.121	6	0.166**	0.163
4	0.113***	0.121	8	0.154***	0.163

Note: The table shows the RMSE for the DFM and the benchmark model (BM) with the relative result of the Diebold-Mariano test statistic (Diebold and Mariano, 1995). The test-statistics are based on out-of-sample forecast errors for the period 2005-2014 and 2015-2020. The column “Horizon” indicates the number of weeks between two consecutive press conferences. *, **, and *** denote, respectively, the 10%, 5%, and 1% significance level.

To complement this analysis, I show the most important variables driving the nowcast of the ECB monetary policy stance. Figure 7 documents the average absolute weekly impact of each block on the updated nowcast computed in real time. In other words, Figure 7 sheds light on the reaction function of the ECB.

Figure 7: Average Absolute Contribution



Note: The figure reports the average (absolute) weekly impact of each series, grouped by category, computed in real time over the evaluation sample 2005–2020. The first column shows the monthly-cycle (2005-2014), while the second one the six-week cycle (2015-2020).

From this figure, three features can be highlighted. First, while in the period 2005-2014 the main contributors to changes in nowcasts are, in order of importance, price, forecast, financial,

output and US data, in the period 2015-2020, the most prominent blocks are financial, textual, forecast, output, price and rates and spread data. Second, the monthly-cycle period displays a bell shape indicating that the most useful information for the nowcast arrives in the second week of the nowcasting period, when most of the hard data first become available. Conversely, the six-week cycle approximates a bimodal distribution with the third and seventh week being respectively the minor and major mode. In absolute terms, the most informational weeks are the sixth and seventh week where the second release of hard data become available. Instead, news move the nowcast relatively less at the beginning, where signals are likely to be weak, and at the end of the period, where the room for improvement has significantly reduced. Third, the contribution of textual variables appear to have been large in the period 2015-2020. To formally test the contribution of textual data to the nowcast performance, I run the DFM excluding textual variables (DFM_{notext}) and then compare it with the DFM with all the variables included (DFM_{all}) drawing on the Diebold-Mariano test statistic. Results are reported in [Table 3](#). While in the period 2005-2014 the null hypothesis of equal accuracy between the models fails to be rejected, from 2015 to 2020 the inclusion of textual variables improves the forecast accuracy in a statistically significant manner at any horizon under consideration. These results therefore bring new evidence to the usefulness of text-data in forecasting purposes.

Table 3: Do Textual Variables Matter?

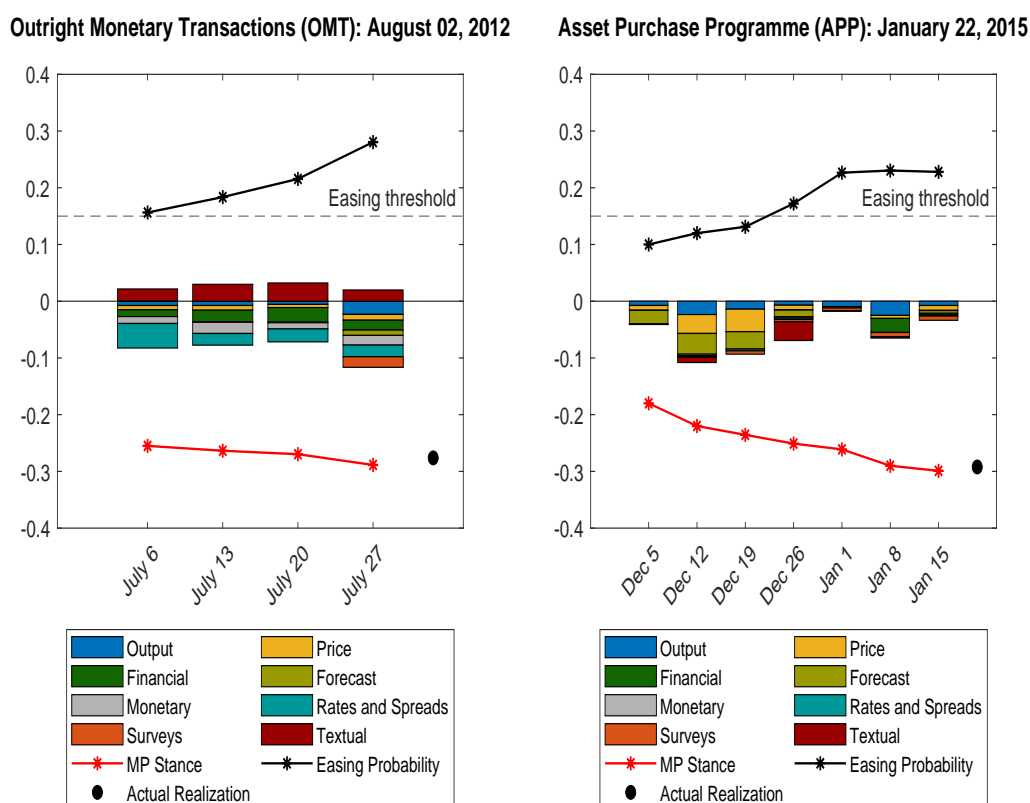
Horizon	2005-2014		Horizon	2015-2020	
	DFM_{all}	DFM_{notext}		DFM_{all}	DFM_{notext}
1	0.123	0.124	1	0.172***	0.181
2	0.122	0.123	4	0.166***	0.175
3	0.118	0.118	6	0.166**	0.176
4	0.113	0.115	8	0.154**	0.166

Note: The table shows the RMSE for the DFM with all the variables (DFM_{all}) and the DFM excluding textual variables (DFM_{notext}) with the relative result of the Diebold-Mariano test statistic ([Diebold and Mariano, 1995](#)). The test-statistics are based on out-of-sample forecast errors for the period 2005-2014 and 2015-2020. The column “Horizon” indicates the number of weeks between two consecutive press conferences. *, **, and *** denote, respectively, the 10%, 5%, and 1% significance level.

To further test the model performance, I replicate a real-time nowcast for two historical “easing” announcements - the Outright Monetary Transactions (OMT) on August 2, 2012 and the Asset Purchase Programme (APP) on January 22, 2015 and two historical “tightening”

ones - the 25 basis points (bps) interest rise on July 3, 2008 and April 7, 2011. [Figure 8](#) and [Figure 9](#) show the evolution of the nowcast for the events of interest: the red line indicates the point forecasts of the monetary policy stance, while the black line represents the nowcast of the expected probability of an easing decision.

Figure 8: Nowcasting Historical Episodes - Easing



Note: The figure shows the evolution of the nowcast for the announcements of the OMTs and the APP. While the black line represents the nowcast of the expected probability of an easing decision, the red line indicates the nowcast of the monetary policy stance. News contribution have been rescaled for better visualization.

Focusing on the announcement of the OMTs, it is noteworthy recalling that such an announcement was preceded by then-President Mario Draghi’s “whatever it takes” at the Global Investment Conference in London on July 26 2012. Setting 15% as a threshold for a monetary policy easing (see [Appendix C](#)), an easing decision (black line) is robustly predicted four weeks ahead of the actual Governing Council meeting and, even more importantly, three weeks ahead of President Draghi’s “whatever it takes”. Moreover, the nowcast of the ECB’s monetary policy stance (red line) starts with a significant dovish realization before gradually turning more negative approximating the actual value of the press conference (black circle). This nowcast

is mainly driven by the significant downward surprises from financial, monetary and rates and spread data, indicating that the evolution of the nowcast is due to a tightening of financial and monetary conditions as well as a widening of sovereign rate differentials⁴. Turning to the APP, while the monetary policy stance (red line) becomes more accurate as the weeks progress, the model anticipates a monetary easing (black line) four weeks in advance of the press conference. For this event, the nowcasts are mostly driven by lower-than-expected forecast, price and output data⁵.

Moving to tightening decisions, at the peak of the financial crisis, on July 3, 2008, the Governing Council of the ECB decided to increase the key interest rates by 25 bps motivating the decision as follows: “This decision was taken [...] to counteract the increasing upside risks to price stability over the medium term. HICP inflation rates have continued to rise significantly since the autumn of last year”⁶. Coherently, the first column of [Figure 9](#) shows that the downward pressure stemming from spread and rates, output and survey is mainly counteracted by higher-than-expected price-related variables and to a lesser extent by the forecast block. Additionally, considering a tightening threshold of 8% (see [Appendix C](#)), the model predicts a hike four weeks in advance of the actual press conference. Turning to the 25 bps rate increase on April 7, 2011 “warranted in the light of upside risks to price stability that we have identified in our economic analysis”⁷, the second column of [Figure 9](#) displays a mildly hawkish monetary policy stance driven by the price and forecast blocks. Moreover, although the conditional probability of a tightening decision declines after the third week, it remains always above the tightening threshold throughout the nowcasting period.

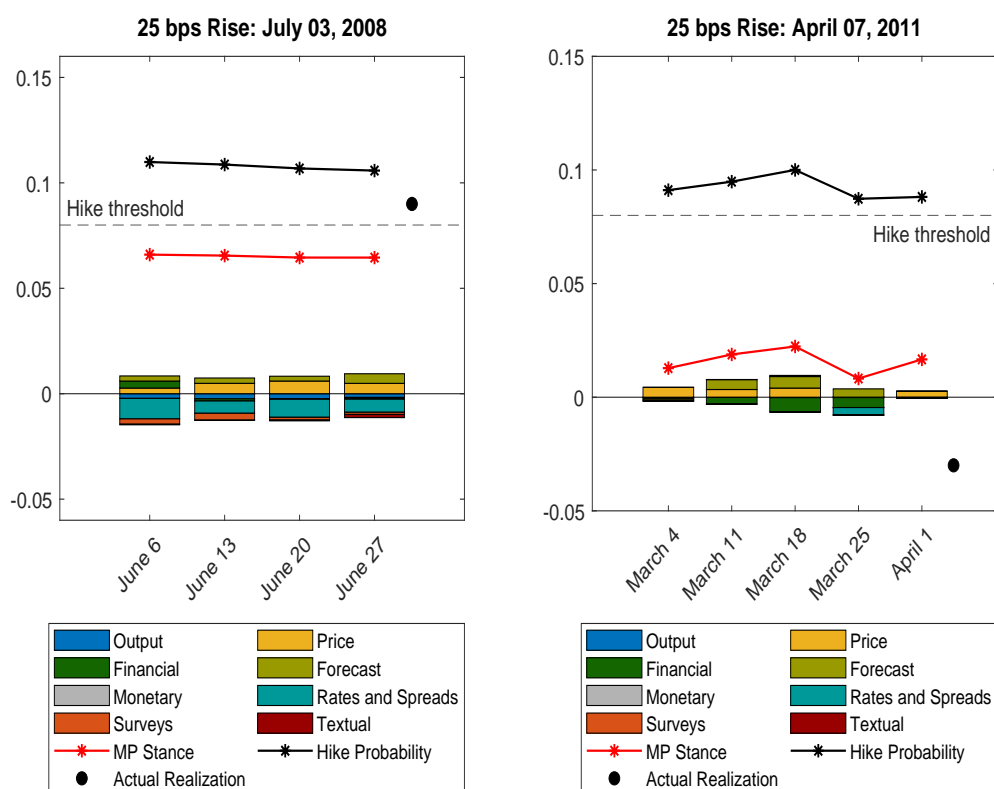
⁴ The ECB Governing Council justified the OMTs stating that “the severe malfunctioning in the price formation process in the bond markets of euro area countries [...] need to be addressed in a fundamental manner”. *Introductory statement to the press conference (with Q&A)*, Mario Draghi, President of the ECB, Vítor Constâncio, Vice-President of the ECB, Frankfurt am Main, 2 August 2012.

⁵ The ECB decided to launch the APP as a result of “inflation dynamics [...] [being] weaker than expected [...] [and as a result of] a further fall in market-based measures of inflation expectations over all horizons and [of] the fact that most indicators of actual or expected inflation stand at, or close to, their historical lows”. *Introductory statement to the press conference (with Q&A)*, Mario Draghi, President of the ECB, Frankfurt am Main, 22 January 2015.

⁶ *Introductory statement to the press conference (with Q&A)*, Jean-Claude Trichet, President of the ECB, Vítor Constâncio, Vice-President of the ECB, Frankfurt am Main, 3 July 2008.

⁷ *Introductory statement to the press conference (with Q&A)*, Jean-Claude Trichet, President of the ECB, Vítor Constâncio, Vice-President of the ECB, Frankfurt am Main, 7 April 2011.

Figure 9: Nowcasting Historical Episodes - Tightening

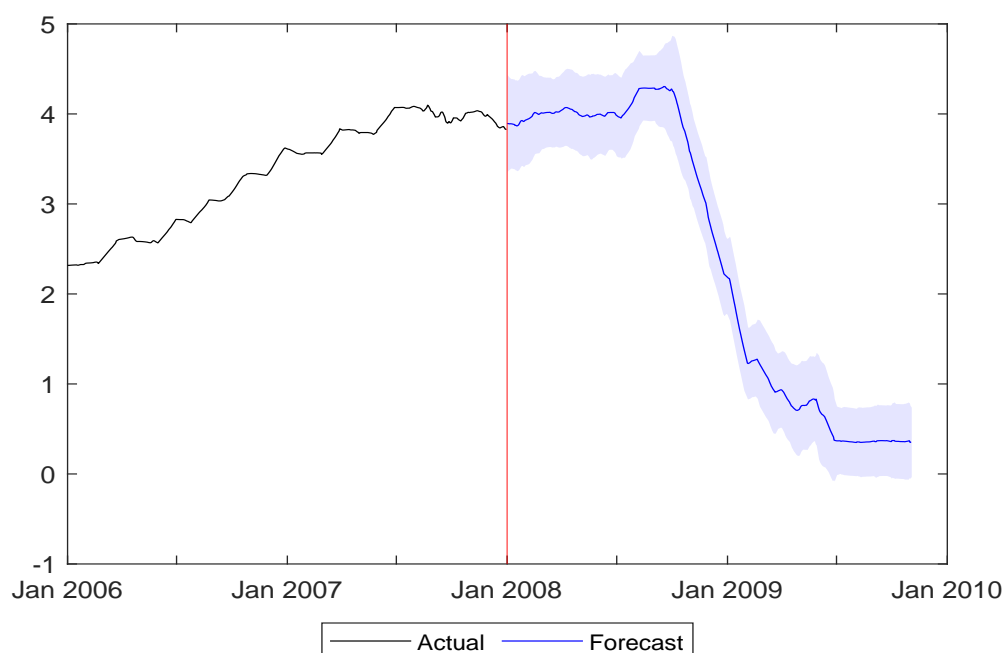


Note: The figure shows the evolution of the nowcast for the announcements of the 25bps rate increase on July 3, 2008 and on April 7, 2011. While the black line represents the nowcast of the expected probability of an easing decision, the red line indicates the nowcast of the monetary policy stance. News contribution have been rescaled for better visualization.

The final exercise I propose is to nowcast, out-of-sample, the evolution of the EONIA rate from January 2008 to December 2009 exploiting the forecasts of the daily factor of ECB monetary policy stance, output and inflation-related variables⁸. This period is particularly important since the European overnight swap rates for the first time approximated the (then-perceived) zero lower bound as a result of the intensification of the 2007-2009 global financial crisis. Figure 10 graphs what would have been the forecast for the EONIA rates conditional on the information available at the time of the analysis.

⁸ EONIA is assumed to be unknown during each month even if its data are released daily. Output variables include industrial production, public and private GDP forecast, while inflation variables refer to core inflation, public and private HICP forecast.

Figure 10: Forecasting EONIA



Note: The figure shows the actual values of EONIA rates (black line) and the out-of-sample nowcast with its 90% confidence intervals (blue line).

As the figure shows, the model proves to be valuable also in tracking the evolution of the EONIA rates in a period of high volatility and uncertainty. In fact, since the end of 2008, the forecasts of the EONIA sharply decline reflecting the validity of the forecasting strategy.

5. Conclusion

This paper proposes an econometric framework to track in real time the monetary policy stance and decisions of the ECB exploiting conventional and textual data at a sampling frequency higher or equal than monthly that become available between two consecutive press conferences. This mixed-frequency dataset is modeled as a dynamic factor model estimated at daily frequency and augmented with a bridge equation that takes the specification of a multinomial logit. Such a framework outputs three pieces of information: the nowcast of the ECB monetary policy stance, the forecast of the conditional probability that the ECB will actually take a monetary policy decision and the block of variables that drives the revision of each nowcast at every point in time.

The empirical results are the following. First, I develop a DFM with mixed-frequency conventional and textual variables to estimate the contemporaneous monetary policy stance of the ECB. Second, the model provides an accurate tracking of the ECB monetary policy stance and decisions at historical ECB announcements. Third, the model proves to be useful in forecasting the EONIA rates from January 2008 to December 2009. Fourth, the model provides higher forecast accuracy than competing models. Last, the inclusion of textual variables in the dataset contributes significantly to the improvement of the forecasting performance.

APPENDIX

A. Text Mining: Data and Techniques

Section A.1 provides details on the estimation of the LDA model; Section A.2 compares the indexes derived from the press conferences with alternative dictionaries; Section A.3 assesses the selection and identification of the topics; Section A.4 describes the textual dataset and Section A.5 sheds additional light on the quantification of the time series extracted from the textual dataset.

A.1. Estimating the Latent Dirichlet Allocation (LDA) Model

This paper implements the LDA model developed by Blei et al. (2003) and follows the estimation algorithm described by Griffiths and Steyvers (2004). For a full explanation of the mechanics of the model I refer to Heinrich (2009); here I will only deploy the essential features.

Formally, recall that $D = \{\mathbf{w}_1, \dots, \mathbf{w}_M\}$ denotes the entire corpus composed by M documents where $N = \sum_{m=1}^M N_m$ is the total number of words, $N_m = \sum_{n=1}^N w_n$ is the total number of words for the m^{th} document, \mathbf{z} is the set of K latent topics and V indicates the size of the vocabulary. A corpus D has a distribution of topics given by $\boldsymbol{\theta}_m$ and, in turn, each topic has a distribution of words denoted by $\boldsymbol{\varphi}_k$, with both $\boldsymbol{\theta}_m$ and $\boldsymbol{\varphi}_k$ assumed to have conjugate Dirichlet distributions with hyper parameters α and β . Note that bold-font variables denote vectors. Each document \mathbf{w} in the corpus D is an iterated choice of topics $z_{m,n}$ and words $w_{m,n}$ drawn from the multinomial distribution using $\boldsymbol{\theta}_m$ and $\boldsymbol{\varphi}_k$. Moreover, let t be a term in V and denote $P(t|z = k)$, the mixture component, one for each topic, by $\boldsymbol{\Phi} = \{\boldsymbol{\varphi}_k\}_{k=1}^K$ ($M \times V$). Finally, let $P(z|D = m)$ define the topic mixture proportion for the m^{th} document, with one proportion for each document $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_m\}_{m=1}^M$ ($M \times K$). The aim of the algorithm is then to approximate the distribution:

$$P(\mathbf{z}, \mathbf{w}; \alpha, \beta) = \frac{P(\mathbf{w}, \mathbf{z}; \alpha, \beta)}{P(\mathbf{w}; \alpha, \beta)} \quad (\text{A.1})$$

using Gibbs simulations, where α and β are the (hyper) parameters controlling the prior conjugate Dirichlet distributions for $\boldsymbol{\varphi}_k$ and $\boldsymbol{\theta}_m$, respectively.

With these definitions, the probability of the model can be written as follows:

$$P(\mathbf{z}, \mathbf{w}, \Theta, \Phi; \alpha, \beta) = \prod_{k=1}^K P(\varphi_k; \beta) \prod_{m=1}^M P(\theta_m; \alpha) \prod_{n=1}^N P(z_{m,n} | \theta_m) P(w_{m,n} | \varphi_k) \quad (\text{A.2})$$

Integrating out the parameters φ and θ :

$$\begin{aligned} P(\mathbf{z}, \mathbf{w}; \alpha, \beta) &= \int_{\Theta} \int_{\Phi} P(\mathbf{z}, \mathbf{w}, \Theta, \Phi; \alpha, \beta) d\Phi d\Theta \\ &= \int_{\Phi} \prod_{k=1}^K P(\varphi_k; \beta) \prod_{m=1}^M \prod_{n=1}^N P(w_{m,n} | \varphi_k) d\Phi \int_{\Theta} \prod_{m=1}^M P(\theta_m; \alpha) \prod_{n=1}^N P(z_{m,n} | \theta_m) d\Theta \quad (\text{A.3}) \end{aligned}$$

In equation (A.3), the first integral does not include θ nor the second integral contains φ . As a result, φ and θ can be solved separately. Thus, drawing on the properties of the conjugate Dirichlet distribution it can be shown that:

$$\int_{\Phi} \prod_{k=1}^K P(\varphi_k; \beta) \prod_{m=1}^M \prod_{n=1}^N P(w_{m,n} | \varphi_k) d\Phi = \frac{\Gamma(\sum_{k=1}^K \alpha_k)}{\prod_{k=1}^K \Gamma(\alpha_k)} \frac{\prod_{k=1}^K \Gamma(n_m^{(k)} + \alpha_k)}{\Gamma(\sum_{k=1}^K n_m^{(k)} + \alpha_k)} \quad (\text{A.4})$$

and

$$\int_{\Theta} \prod_{m=1}^M P(\theta_m; \alpha) \prod_{n=1}^N P(z_{m,n} | \theta_m) d\Theta = \prod_{k=1}^K \frac{\Gamma(\sum_{t=1}^V \beta_t)}{\prod_{t=1}^V \Gamma(\beta_t)} \frac{\prod_{m=1}^M \Gamma(n_k^{(t)} + \beta_t)}{\Gamma(\sum_{t=1}^V n_k^{(t)} + \beta_t)} \quad (\text{A.5})$$

where here $n_m^{(k)}$ denotes the number of word tokens in the m^{th} document assigned to the k^{th} topic, and $n_k^{(t)}$ is the number of times the t^{th} term in the vocabulary has been assigned to the k^{th} topic.

Since in equation (A.1) $P(\mathbf{w}; \alpha, \beta)$ is invariant to \mathbf{z} , the conditional distribution $P(\mathbf{z} | \mathbf{w}; \alpha, \beta)$ can be derived from $P(\mathbf{w}, \mathbf{z}; \alpha, \beta)$ directly using Gibbs simulation and the conditional probability:

$$P(z_{(m,n)} | \mathbf{z}_{-(m,n)}, \mathbf{w}; \alpha, \beta) = \frac{P(z_{(m,n)}, \mathbf{z}_{-(m,n)}, \mathbf{w}; \alpha, \beta)}{P(\mathbf{z}_{-(m,n)}, \mathbf{w}; \alpha, \beta)} \quad (\text{A.6})$$

where $z_{(m,n)}$ denotes the hidden variable of the n^{th} word token in the m^{th} document, and $z_{-(m,n)}$ denotes all z but $z_{(m,n)}$. Denoting the index of a word token by $\iota = (m, n)$, and using the expressions in (A.4) and (A.5), cancellation of terms (and some extra manipulations exploiting the properties of the gamma function) yields:

$$P(z_\iota = k | \mathbf{z}_{(\iota)}, \mathbf{w}; \alpha, \beta) \propto (n_{m,\iota}^{(k)} + \alpha_k) \frac{n_{k,\iota} + \beta_t}{\sum_{t=1}^V n_{k,\iota}^{(t)} + \beta_t} \quad (\text{A.7})$$

where the counts $n_{\cdot,\iota}^{(\cdot)}$ indicate that token ι is excluded from the corresponding document or topic. Therefore, sampling topic indexes using equation (A.7) for each word in a document and across documents until convergence allows us to approximate the posterior distribution given by (A.1).

The model can be estimated as described once the researcher sets three parameters: the number of topics K and the two parameter vectors of the Dirichlet priors α and β . While I already discussed in the paper how to optimally estimate K , α and β are defined as a function of the number of topics and unique words (Griffiths and Steyvers, 2004) as $\alpha = \frac{50}{K}$ and $\beta = \frac{200}{N}$.

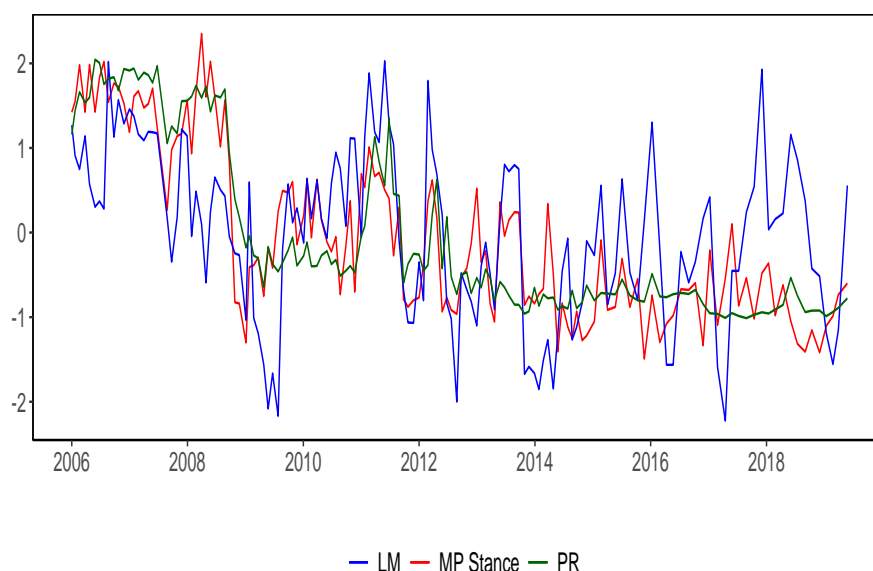
A few final remarks concern the out-of-sample estimation exercise¹. To avoid potential look-ahead biases when the full-sample-based news topic is used, the news corpus is considered up to end of the last day of 2010, and K new news topics are estimated using the LDA model. Then, using the estimated topic distributions from the “trained” corpus, the period 2010–2020 is classified, and topic time series for the whole sample period are constructed as described above. The out-of-sample classification is implemented following Heinrich (2009), Hansen et al. (2018) and Thorsrud (2020), where a procedure for querying documents outside the set on which the LDA is estimated is outlined. In a nutshell, the strategy consists in using the same Gibbs simulations just described, but with the difference that the sampler is run with the estimated word distributions (from the training sample) held constant.

A.2. Alternative Tone’s Measures

Figure A.1 compares the monetary policy stance index derived in Section 2.1 with alternative measures available in the literature.

¹ Updating the topic model recursively is infeasible, as one estimation round takes approximately 10 days using my infrastructure. Thus, at the end of the evaluation period, I am essentially using topic distributions based on 10-year old estimates. Another caveat with re-estimating the LDA recursively is the lack of topics identifiability.

Figure A.1: Alternative Tone's Measures



Note: The figure shows the evolution of the monetary policy stance (red) index against alternatives specifications available in the literature.

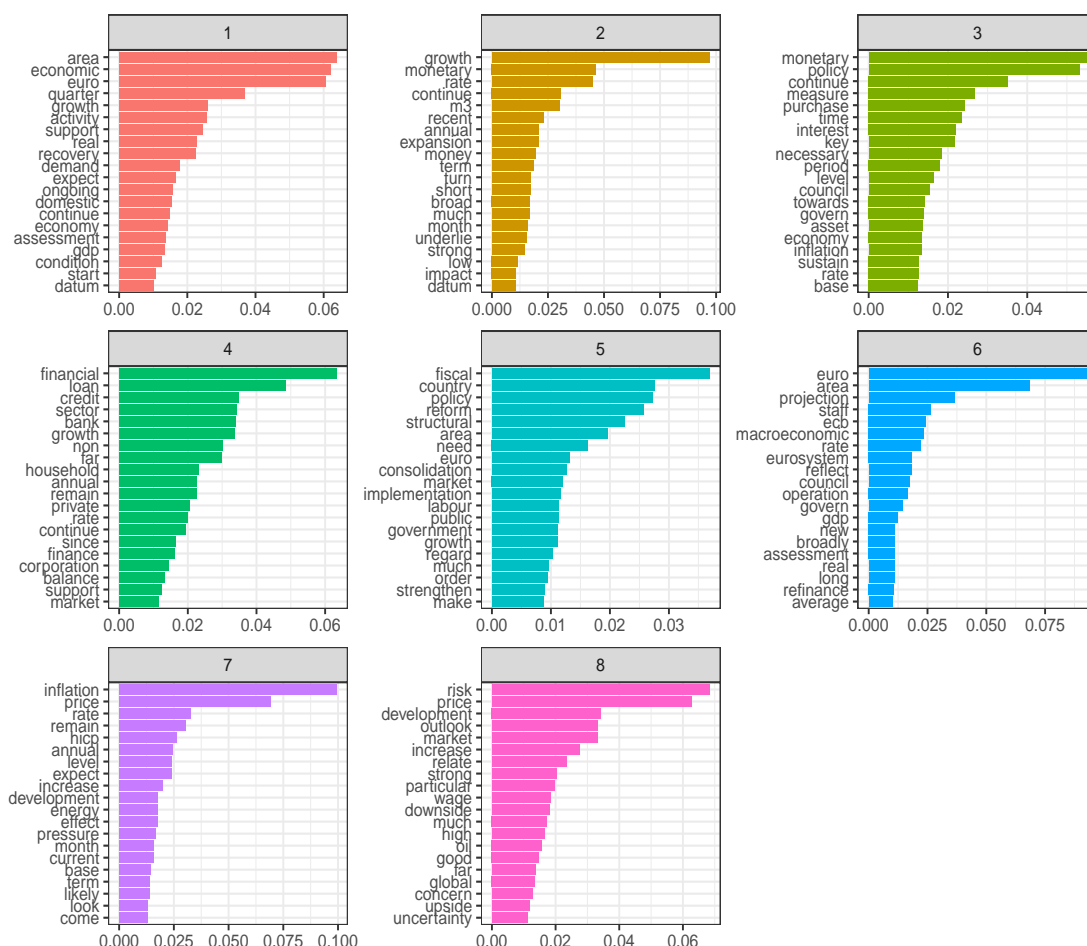
The figure shows a strong correlation (0.90) with the more sophisticated index developed by [Picault and Renault \(2017\)](#), while it departs from the index based on [Loughran and McDonald \(2011\)](#)'s dictionary with a correlation of 0.42.

A.3. Assessing Topics' Identification

The identification of topics in the LDA model requires a dose of subjective judgment. Therefore, to prove the plausibility of the choices made in the paper I show a few additional results. To this end, [Figure A.2](#) illustrates the probability that each word belongs to a specific topic²:

² Besides quantitative and graphical evidence, I checked qualitatively the plausibility of the identification strategy. In fact, once the probability of a topic to belong to a paragraph is retrieved, it is straightforward to investigate whether the content of the paragraph matches the labeling. Upon closer qualitative inspection, the labeling proves to be plausible.

Figure A.2: Words per Topic



Note: The figure displays the terms that are most common within each topic.

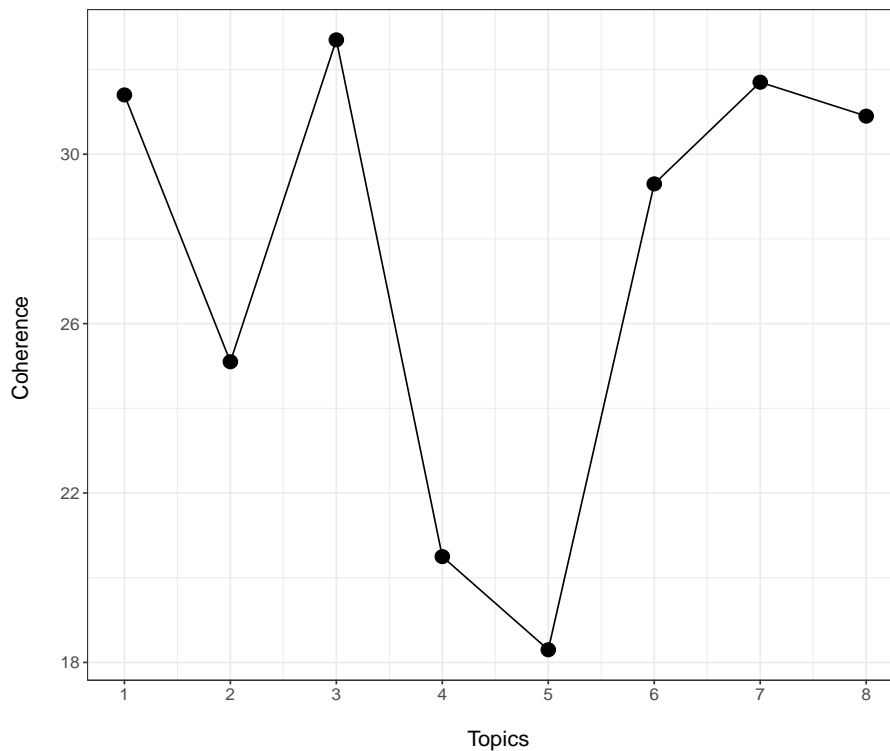
On the basis of the outcome displayed in [Figure A.2](#), I label the topics as follows³:

³ The results are robust to different pre-processing choices. Topics are also roughly stable for a number of latent topics ranging from 6 to 9.

List of Topics	Identified Topics
Topic 1	Economic Outlook
Topic 2	Monetary Analysis
Topic 3	Monetary Policy
Topic 4	Financial Conditions
Topic 5	Fiscal and Structural Reforms
Topic 6	GDP Forecasts
Topic 7	Inflation Outlook I
Topic 8	Inflation Outlook II

To further validate the identification exercise, I measure the coherence of each topic. Topic coherence provides a measure of the degree of semantic similarity between high scoring words in the topic. Such a measurement therefore help distinguish between topics that are semantically interpretable and topics that are artifacts of statistical inference. The results are shown in Figure A.3.

Figure A.3: Coherence within Topics

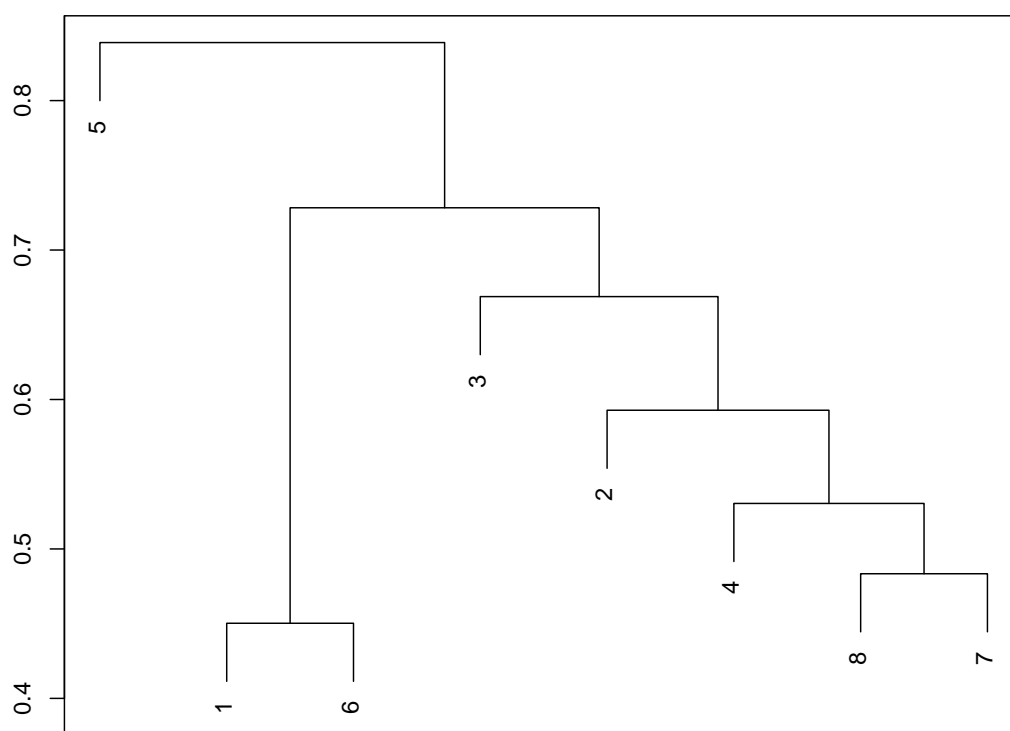


Note: The figure shows the coherence score for each topic.

Figure A.3 shows that the topics with the highest coherence score are: 1, 3, 6, 7 and 8. This result is important since these topics are those quantified using the ECB field-specific dictionaries outlined above. In fact, Topic 1 and 2 form the topic “Economic Outlook”, Topic 7 and 8 “Inflation Outlook” and Topic 3 “Monetary Policy”⁴.

Finally, Figure A.4 shows a cluster dendrogram that uses Hellinger distance (distance between two probability vectors) to decide whether the topics are closely related.

Figure A.4: Cluster Dendrogram



Note: The figure illustrates the dendrogram for the eight topics estimated from the model

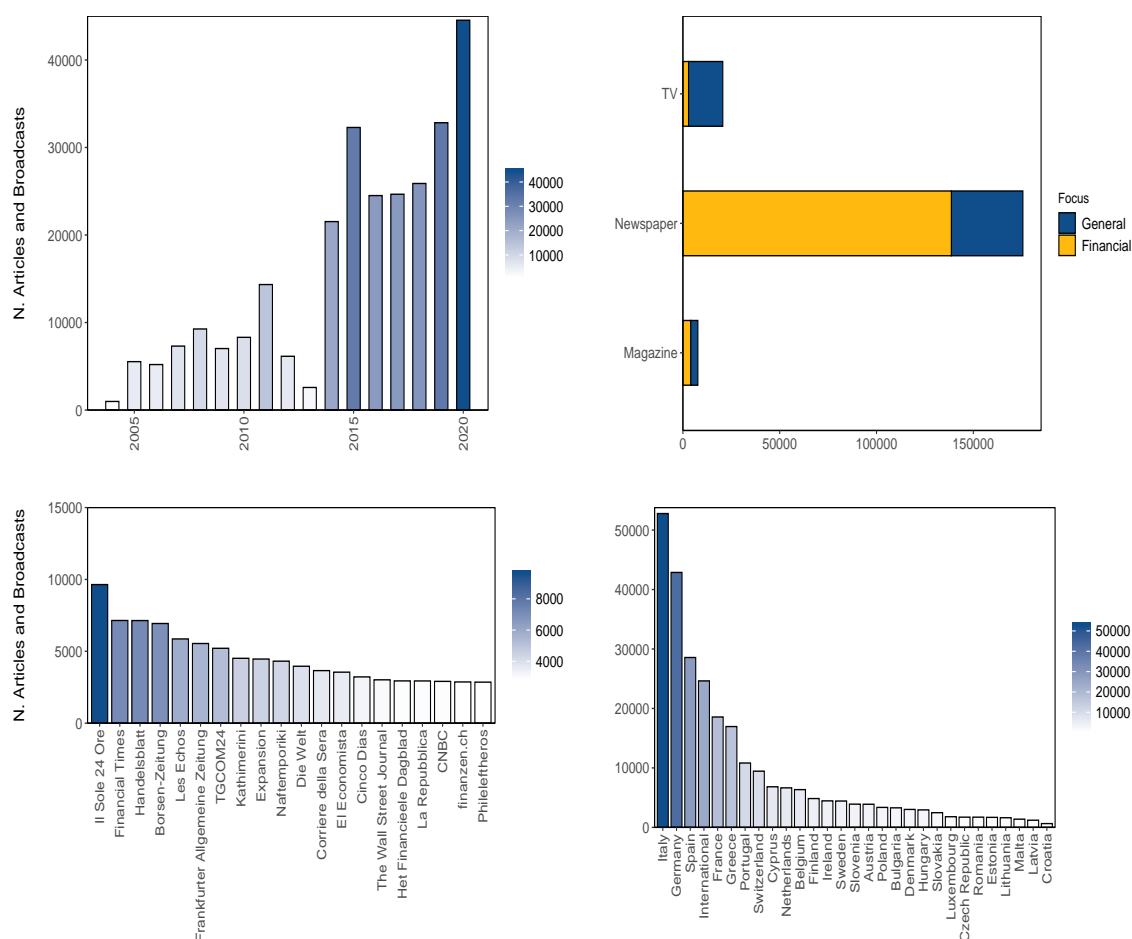
The dendrogram, in addition to structuring the relations among the topics in a economically meaningful way, supports the choice to consider Topic 7 and 8 as a single topic, likewise Topic 1 and 6.

⁴ It is not surprising that Topic 4 and Topic 5 have lower score. In fact, the words listed in “Financial Conditions” span from market-related words to households and banks. Similarly, “Fiscal and Structural Reforms” intersect both fiscal and labour related issues.

A.4. The Textual Dataset

In Figure A.5, I provide some descriptive statistics of the textual dataset collected in December 2020 from an ECB internal database. The database consists of English detailed summaries of any communication event (mainly but not only newspapers) that mentioned the ECB in their content.

Figure A.5: Descriptive Statistics of Textual Data



Note: The figure shows descriptive statistics for the textual dataset. Starting with the first row, the first column presents the number of documents per year, while the second column the type of document and the field of specialization. In the second row, the first column details the number of documents by outlet, while the second column the number of documents by country of origin.

It includes around 300,000 documents whose breakdown by year is not constant as shown by the figure in the the first column of the first row. In fact, from 2004 to 2013, the average number of documents per year stands around 5,000, while from 2014 on, the average is tilted

significantly upwards to approximately 33,000 documents per year. The figure in the second column of the first row shows the composition of the documents where the large majority of them is composed of specialized newspaper articles. As for the second row, the first column breaks the number of articles into outlets' names and the second one into the country of origin of the outlet. To control for any bias, in the analysis I filtered out the time series regressing them on countries' dummies.

A.5. Quantifying Textual Data

As mentioned in the paper, I applied LDA to the textual dataset with $K = 80$. Out of eighty topics, only sixty were clearly identifiable and I list them in the first column of the table below. These identified topics were further scaled down to forty following the criterion of monetary policy relevance (second column). Finally, these topics were manually grouped into seven meta-topics (third column) on the basis of the conceptual similarity between them. Therefore, the table below provides a summary of how this exercise of dimensionality reduction was carried out.

Identified Topics	Relevant Topics	Meta-Topics
Equity developments		
French politics		
European Commission President		
Eastern EU politics		
EU crisis	X	Financial crisis
Bank of England policy		
Core countries growth	X	Economic outlook
Banking crisis	X	European banks
Macro uncertainty	X	Economic outlook
Exchange rates	X	Economic outlook
Lending conditions	X	Monetary Policy
Stress tests	X	European banks
Eurexit	X	Eurexit
NPLs	X	European banks
Italian Politics		
Greek bonds	X	Financial crisis
Greek bailout	X	Financial crisis
Banknotes		
Rating Agencies	X	Financial crisis
Mario Draghi		
ECB Board Members		
Italian Banks	X	European banks

Greek Banks	X	European banks
Spanish Banks	X	European banks
German Banks	X	European banks
Digital Currency		
Lagarde		
FGCJ Ruling	X	Monetary Policy
IMF	X	Financial crisis
Banking supervision	X	European banks
Fixed-Income Developments		
Sovereign Bonds	X	Financial crisis
Financial stress	X	Financial crisis
Spreads	X	Financial crisis
Sovereign-debt crisis	X	Financial crisis
Fiscal rules	X	Fiscal policy
Fiscal stance	X	Fiscal policy
German politics		
Inflation dynamics	X	Inflation outlook
Deflation	X	Inflation outlook
TLTROs	X	Monetary policy
Negative rates	X	Monetary policy
Forward guidance	X	Monetary policy
Asset purchases	X	Monetary policy
Fed policies		
US politics		

Trade war	X	Economic outlook
US economy		
NGEU	X	Fiscal policy
Market expectations		
European growth	X	Economic outlook
Core inflation	X	Inflation outlook
Brexit	X	Eurexit
Bank profitability	X	European Banks
Liquidity support	X	Monetary policy
Bank strategies	X	European banks
Monetary statistics	X	Monetary Policy
Climate change		
Cryptocurrency		
Bank of Japan policies		

These seven topics were treated differently. In fact, while “financial crisis”, “eurexit” and “European banks” were left as topic probabilities, “monetary policy”, “economic outlook”, “inflation outlook” and “fiscal policy” needed a further quantification step. In quantifying these topics I followed the same steps as [Section 2.1](#). However, in [Section 2.1](#) the field-specific dictionaries were based on the ECB press conferences that have much narrower number of words compared to the ECB’s dataset. To inspect such a large dataset and look for polarized words relevant to the already existing field-specific dictionaries, I relied on word embedding ([Stoltz and Taylor, 2019](#)) which is a tool for identifying similarities between words that occur within a “context window”. The algorithm is based on the Word Mover’s Distance (WMD) that attempts to find the “closest neighbour” for each word so that the “cost” of moving all the words in one collection (sentences, subsection, etc.) to the positions of all the words in another collection is minimized. Thus, collections sharing many semantically similar words should have smaller distances than collections with very different words. Formally, the WMD algorithm finds the values of a matrix \mathbf{T} that minimize “moving” one document, D , to the other, D' :

$$WMD_{ij} = \min_{\mathbf{T} \geq 0} \sum_{i,j=1}^n T_{ij} c(i, j) \quad (\text{A.8})$$

where $c(i, j)$ indicates the cosine distance between the i^{th} and j^{th} word in the n -dimensional embedding space and under the constraints such that:

$$\sum_{i=1}^n T_{ij} = d_i, \forall i \in \{1, \dots, n\} \quad (\text{A.9})$$

$$\sum_{j=1}^n T_{ij} = d'_j, \forall i \in \{1, \dots, n\} \quad (\text{A.10})$$

where equation [\(A.9\)](#) says that the sum of row word i in \mathbf{T} is equal to the relative frequency of i in document D (after any word removal), d_i , and equation [\(A.10\)](#) that the sum of column word j in \mathbf{T} is equal to the relative frequency of j in D' (again after word removal), d'_j . An example can clarify: if the word “fiscal” had a relative frequency of 0.25 in a document after any word removal, then the sum of the row and column with the word “fiscal” must each sum to 0.25. The intuition is to weight each ij cosine distance by “how much” of the relative frequency of i in D will move to j in D' .

To measure “how close” certain words are to a focal concept, however, an additional step,

called Relaxed Word Mover’s Distance (RWMD), is required. With RWMD, the flow matrix weighting for each i, j pair is solved twice: once with just the constraint from equation (A.9) removed, and then once with just the constraint from equation (A.10) removed. The RWMD for each i, j pair is then calculated twice following equation (A.8), and the final reported RWMD score for the i, j pair is:

$$RWMD_{ij} = \max\left(\min_{\mathbf{T} \geq 0} \sum_{i,j=1}^n T_{ij}c(i, j), \min_{\mathbf{T} \geq 0} \sum_{i,j=1}^n T'_{ij}c(i, j)\right) \quad (\text{A.11})$$

In the paper, for example, I applied equation (A.11) to find the closest words to the focal concepts of “inflation”, “economy”, “monetary”, and “fiscal”, where the focal concepts simply correspond to the topics of the field-specific dictionaries. What this equation provides is a list of words that are closely related to each focal point. Once these lists are derived, the last steps consist in finding which words are not present in the dictionaries derived in Section 2.1. These words are then manually classified as “positive” or “negative” using discretion with regard to the topic of interest.

B. Total Dataset

The table below lists all the variables used in the model with its respective transformation, publication lag, block and frequency. In the transformation column: 1 corresponds to the variable being left in level, 2 to first difference, 3 to log-difference and 4 to percentage change; in the frequency column, D stands for daily, M for monthly and I for irregular; finally, in the column that indicates the publication lag, the numbers denote months.

Series	Transformation	Publication Lag	Frequency
Real Activity			
Industrial Production Total	3	1	M
Industrial Production (construction)	3	1	M
Industrial Production: manufacturing	3	1	M
Unemployment (labour force)	4	1	M
Harmonised Index of Consumer Prices (HICP)	3	1	M
HICP (%)	2	1	M
HICP: services	3	1	M
HICP: excl. energy and food	3	1	M
HPPI: Industry excluding construction	3	1	M
Oil Prices (Brent)	3	1	M
Consumer Price Index (CPI)	3	1	M
CONSUMER PRICES OF PETROLEUM PRODUCTS (PUMP PRICES)	3	1	M
HWVI COMMODITY PRICES FOR THE EURO AREA: Total Excluding Energy	3	1	M
Surveys			
EUROCOIN	1	0	M
Composite Output EA (PMI)	1	0	M
Manufacturing (PMI)	1	0	M
Services (PMI)	1	0	M
Construction (PMI)	1	0	M
Capacity Utilization (PMI)	1	0	M
Euro Area Big 2 (PMI)	1	0	M
Employment expectations	1	0	M
Unemployment expectations	1	0	M
Italy CLIFS	1	0	M
Spain CLIFS	1	0	M
Portugal CLIFS	1	0	M
Business Climate Indicator	1	0	M

Series	Transformation	Publication Lag	Frequency
Economic sentiment indicator	1	0	M
Financial			
DJ U.S. Total Stock Market Index	3	0	D
SP 500 Index	3	0	D
NASDAQ Composite	3	0	D
STOXX Euro	3	0	D
EUSTOXX Banks	3	0	D
VIXEU	3	0	D
CDS Italy	3	0	D
CDS Spain	3	0	D
CDS Portugal	3	0	D
CDS France	3	0	D
CDS Germany	3	0	D
BARCLAYSMBSFIXRATE	2	0	D
FTSEMIB	2	0	D
IBEX 30	2	0	D
CAC 40	2	0	D
DAX 30	2	0	D
Rates and Spreads			
10Y GDP-Weighted Nominal Yields	2	0	D
5Y5YINFSWAP rates	2	0	D
2Y2YINFSWAP rates	2	0	D
Goldman Sachs Financial Condition Index	2	0	D
EURINFLSWAP0COUP	2	0	D
EONIA	2	0	D
EURIBOR 3M	2	0	D
LIBOR-OIS	2	0	D
OISEONIA	2	0	D

Series	Transformation	Publication Lag	Frequency
Repo funds DE	2	0	D
Repo funds IT	2	0	D
Repo funds FR	2	0	D
EUFCI	2	0	D
EUDESPREAD	2	0	D
EUSWAP10Y	2	0	D
IT10Y	2	0	D
DE10Y	2	0	D
1W SWAP	2	0	D
1Y1Y EONIA Forward	2	0	D
EU10Y	2	0	D
FR10Y	2	0	D
DEIT10Y	2	0	D
1Y1YSPRD	2	0	D
3MEOIS	2	0	D
3MGLBR	2	0	D
3MGLOIS	2	0	D
DE2Y10Y	2	0	D
DEIT2Y	2	0	D
DEIT5Y	2	0	D
DEGR10Y	2	0	D
DESP10Y	2	0	D
DEPT10Y	2	0	D
EUREONIA	2	0	D
EUROIS3M	2	0	D
FR2Y10Y	2	0	D
GBPOIS1M	2	0	D
DE2Y10Y	2	0	D

Series	Transformation	Publication Lag	Frequency
ES2Y10Y	2	0	D
FR2Y10Y	2	0	D
PT2Y10Y	2	0	D
IT2Y10Y	2	0	D
Forecasts			
Bloomberg Weighted Average Private GDP Forecast	4	0	D
Bloomberg Weighted Average Official GDP Forecast	4	0	D
Bloomberg Weighted Average Private IP Forecast	4	0	D
Bloomberg Weighted Average Official IP Forecast	4	0	D
Bloomberg Weighted Average Private CPI Forecast	4	0	D
Bloomberg Weighted Average Private Unemployment Forecast	4	0	D
Textual Newspaper			
ECB Monetary Policy Stance (Newspaper)	1	0	D
Economic Outlook Index (Newspaper)	1	0	D
Inflation Outlook Index (Newspaper)	1	0	D
Eurexit Index	1	0	D
Financial Crisis Index	1	0	D
European Banks Index	1	0	D
Fiscal Policy	1	0	M
Textual PC			
Monetary Policy Stance Index	1	0	I
Economic Outlook Index	1	0	I
Inflation Outlook Index	1	0	I
Monetary			
M3	2	1	M
MFI loans to euro area residents excluding MFIs and general government	2	1	M
Securities other than shares issued in euro by non-MFI corporations	2	1	M
Total assets	2	1	M

Series	Transformation	Publication Lag	Frequency
Lending to euro area credit institutions in euro	2	1	M
Main refinancing operation	2	1	M
Longer-term refinancing operations	2	1	M
Securities of euro area residents in euro	2	1	M
Deposit facility rate	2	1	M
Main refinancing rate	2	1	M
Marginal lending facility rate	2	1	M
Loans to insurance corporations and pension funds	2	1	M
Loans to other financial intermediaries (including investment funds)	2	1	M
Loans to nonfinancial corporations	2	1	M
Households	2	1	M
Lending for house purchase	2	1	M
Interest Rates New Deposits from households (overnight)	2	1	M
New deposits: Deposits from nonfinancial corporations	2	1	M
New deposits: repos	2	1	M
Mixed			
ZEW FINANCIAL MARKET SURVEY: current macro	4	0	M
ZEW FINANCIAL MARKET SURVEY: expected macro	4	0	M
Sentix overall economic index	4	0	M
Sentix break-up economic index	4	0	M
Euro Area Composite Financial Conditions Index	2	0	M
Inflation factor	3	1	M
Uncertainty factor	3	1	M
Sovereign CISS (weighted)	2	1	M
Exchange Rates			
Japanese yen	3	0	M
UK pound	3	0	M
US dollar	3	0	M

Series	Transformation	Publication Lag	Frequency
Harmonized indicators against 38 trading partners	3	0	M
Nominal effective exchange rate of the euro: EER-19	3	0	M
Real (CPI) effective exchange rate of the euro: EER-19	3	0	M
US			
JP Morgan Global PMI	3	0	M
PMI North America	3	0	M
Inflation Expectations (1Y), Michigan University	3	1	M
Inflation Expectations (5Y), Michigan University	3	1	M
Federal funds (effective)	2	0	D
Treasury yields 1Y	2	0	D
Treasury yields 10Y	2	0	D

C. State Space Representation and Estimation

Section C.1 explains the temporal aggregation implemented in the Dynamic Factor Model, Section C.2 provides details of the Expectation Maximization (EM) algorithm, Section C.3 sheds light on the bridge equation and Section C.4 considers alternative benchmark models to compare the DFM with.

C.1. Temporal Aggregation

Letting bold-font variables indicate the vector version of the variables, I write the model in equations (3) and (4) compactly as follows:

$$\mathbf{Y}_t = \mathbf{A}_t \mathbf{F}_t + \mathbf{e}_t \quad (\text{C.12})$$

$$\mathbf{F}_t = \mathbf{A}_t \mathbf{F}_{t_1} + \mathbf{u}_t \quad (\text{C.13})$$

with $\mathbf{F}_t = [\mathbf{F}_t^{k,i} \mathbf{F}_t^{k,m} F_t^d]'$. The last element in the vector \mathbf{F}_t is the scalar F_t^d and can be interpreted as the latent common daily monetary policy stance index. F_t^d is the only scalar in \mathbf{F}_t since the other elements are vectors containing aggregator variables used to handle the mixed-frequency property of the model. In mixed-frequency models, lower frequency variables are usually treated as the highest frequency series, i.e. daily in this case, with missing observations (see Foroni and Marcellino (2013) for an overview). For a generic variable Y_t^k , time aggregation from higher to lower frequency is restricted as follows:

$$\begin{aligned} y_t^k &= \log(v_{1,t}^k) - \log(v_{1,t-k}^k) \\ &\approx \log\left(\sum_{i=0}^{k-1} v_{1,t-i}^k\right) - \log\left(\sum_{i=k}^{2k-1} v_{1,t-i}^k\right) \\ &\approx \sum_{i=0}^{k-1} \log(v_{1,t-i}^k) - \sum_{i=0}^{2k-1} \log(v_{1,t-i}^k) \\ &= \sum_{i=0}^{2k-2} \omega_i^k y_{1,t-i}^k, \quad t = k, 2k, \dots, \end{aligned} \quad (\text{C.14})$$

where y_t^k is the observed low-frequency growth rate, v_t^k its level, and $\omega_i^k = i + 1$ for $i = 1, \dots, k - 1$; $\omega_i^k = 2k - i - 1$ for $i = k, \dots, 2k - 2$; and $\omega_i^k = 0$ otherwise. It follows from equation (C.14) that imposing a common factor structure to y_t^k gives:

$$y_t^k = \sum_{i=0}^{2k-2} \omega_i^k y_{1,t-i} = \sum_{i=0}^{2k-2} \omega_i^k (\Lambda F_{t-i}^d + e_{t-i}) \quad (\text{C.15})$$

However, equation (C.15) makes the inference rather challenging since it significantly increases the number of state variables in the model. To sort this issue and limit the size of the state vector, I employ the double cumulator approach in Bańbura et al. (2013) where the temporal aggregator variables are recursively updated such that at the end of each respective period we have:

$$F_t^k = \sum_{i=0}^{2k-2} \omega_i^k F_{t-i} \quad t = k, 2k, \dots \quad (\text{C.16})$$

As shown in Bańbura et al. (2013), this expression can be computed recursively with the help of two (additional) state variables. In particular, by introducing the auxiliary variable \bar{F}_t^k and denoting $R(\cdot, k)$ the positive remainder of the division by k , \tilde{F}_t^k can be obtained recursively as follows:

$$\tilde{F}_t^{k,f} = \begin{pmatrix} \bar{F}_t^{k,f} \\ F_t^{k,f} \end{pmatrix} = \begin{cases} \begin{pmatrix} \bar{F}_t^{k,f} + \omega_{k-1}^{k,f} F_t \\ 0 \end{pmatrix}, & t = 1 + k + 1, 2k + 1, \dots \\ \begin{pmatrix} F_{t-1}^{k,f} + \omega_{R(k-t,k)}^{k,f} F_t \\ 0 \end{pmatrix}, & \text{otherwise.} \end{cases} \quad (\text{C.17})$$

For the stock variables only single aggregator variable is necessary:

$$F_t^{k,s} = \begin{cases} \omega_{k-1}^{k,s} F_t, & t = 1 + k + 1, 2k + 1 \\ F_{t-1}^{k,s} + \omega_{k-1}^{k,s} F_t, & \text{otherwise.} \end{cases} \quad (\text{C.18})$$

This is implemented via the transition equation (4) with the following weight vector \mathcal{W}_t^k and selection matrix \mathcal{I}_t^k :

$$\mathcal{W}_t^{k,f} = \begin{cases} \begin{pmatrix} -\omega_{k-1}^{k,f} \\ 0 \end{pmatrix}, & t = 1 + k + 1, 2k + 1, \dots \\ \begin{pmatrix} -\omega_{R(k-t,k)}^{k,f} \\ -\omega_{R(k-t,k)+k}^{k,f} \end{pmatrix}, & \text{otherwise.} \end{cases} \quad (\text{C.19})$$

$$\mathcal{I}_t^{k,f} = \begin{cases} \begin{pmatrix} 0 & I_r \\ 0 & 0 \end{pmatrix}, & t = 1 + k + 1, 2k + 1, \dots \\ I_{2r}, & \text{otherwise,} \end{cases} \quad (\text{C.20})$$

and for the stock variables simply: $\mathcal{W}_t^{k,s} = -\omega_{R(k-t,k)}^{k,s}$ and $\mathcal{I}_t^{k,s} = \begin{cases} 0, & t = 1, k + 1, \dots \\ I_r, & \text{otherwise.} \end{cases}$

In general, equations (C.17)-(C.20) can handle temporal aggregation from higher to lower frequencies for a range of k values. In the model specification outlined in Section 3, both $k = k_i, k_m, k_d$ are considered, where the k refers to the (average) number of days in the irregular, monthly and daily frequencies, respectively.

To handle different number of days per month or irregular, I follow Bańbura et al. (2013) in approximating the flow variables as follows:

$$z_t^k = \frac{k}{k_t} \sum_{i=0}^{k_t-1} z_{t-i}, t = k_1, k_1 + k_{k_1+1}, \dots, \quad (\text{C.21})$$

where k_t is the number of business days in the period (month or irregular) that contains day t and k is the average number of business days per period over the sample. As a result, $\gamma_t^k = z_t^k - z_{t-k}^k$ becomes:

$$\gamma_t^k = k \left(\sum_{i=0}^{k_t-1} \frac{i+1}{k_t} \gamma_{t-i} + \sum_{i=k_t}^{k_t+k_t-k_t-2} \frac{k_t+k_t-k-i-1}{k_t-k} \gamma_{t-i} \right), t = k_1, k_1 + k_{k_1+1}, \dots, \quad (\text{C.22})$$

Hence, this results in time-varying weights and the formulas above should be updated with: $\omega_{t,i}^{k,f} = k \frac{i+1}{k_t}$ for $i = 0, 1, \dots, k_t-1$, $\omega_{t,i}^{k,f} = k \frac{k_t+k_t-k-i-1}{k_t-k}$ for $i = k_t, k_{t+1}, \dots, k_t + k_{k_t+1} - 2$ and

$\omega_{t,i}^{k,f} = 0$ otherwise.

C.2. The Expectation Maximization Algorithm

To describe the Expectation Maximization (EM) algorithm for the daily model in equations (3) and (4), I follow once again Bańbura and Modugno (2014). Assuming one lag in the factor VAR, I can group the parameters as follows: $\theta = (\mu, \Lambda, \Xi, \Sigma_E, \Sigma_U)$, where the only restriction is that Σ_E is diagonal. Let $T_v = \max_n T_v(n)$ also denote the time index of the most recent observation in Ω_v . It is then possible to write the log-likelihood in terms of $Y_t^{k,n} = (Y_1^{k,n}, Y_2^{k,n}, \dots, Y_{T_v}^{k,n})$ and $F_t^{k,n} = (F_1^{k,n}, F_2^{k,n}, \dots, F_{T_v}^{k,n})$ as $l(Y_t^{k,n}, F_t^{k,n}; \theta)$. After that an initial estimate of the parameters $\theta(0)$ is computed on a sample of data where missing observations are handled using splines, the EM algorithm proceeds in two steps as follows:

$$\begin{aligned} E - step : L(\theta, \theta(j)) &= E_\theta[l(Y_t^{k,n}, F_t^{k,n}; \theta) | \Omega_v], \\ M - step : \theta_{j+1} &= \underset{\theta}{\operatorname{argmax}} L(\theta, \theta(j)). \end{aligned} \quad (\text{C.23})$$

The new parameter estimates in the M-step can be obtained in two steps, first $\Lambda(j+1)$ and $\Xi(j+1)$ are given by:

$$\begin{aligned} \operatorname{vec}(\Lambda^{k,\cdot}(j+1)) &= \left(\sum_{t=1}^{T_v} E_{\theta(j)}[Y_t^{k,n} F_t^{k,n'} | \Omega_v] \otimes \mathcal{S}_t \right) \left(\mathcal{S}_t \sum_{t=1}^{T_v} E_{\theta(j)}[Y_t^{k,n} F_t^{k,n'} | \Omega_v] \right)^{-1}, \\ \Phi(j+1) &= \left(\sum_{t=1}^{T_v} E_{\theta(j)}[F_t^{k,n} F_{t-1}^{k,n'} | \Omega_v] \right) \left(\sum_{t=1}^{T_v} E_{\theta(j)}[F_{t-1}^{k,n} F_{t-1}^{k,n'} | \Omega_v] \right)^{-1}. \end{aligned} \quad (\text{C.24})$$

where \mathcal{S}_t is a selection matrix, that is, a diagonal matrix that takes the value of one for non-missing observations in $Y_t^{k,n}$ and zero otherwise. \mathcal{S}_t allows the estimation to work with arbitrary pattern of missing data. $\Lambda^{k,\cdot}$ are estimated blockwise, by frequency and by stock or flow type, using the corresponding block of $Y_t^{k,n}$ and aggregator variable.

Second, given the new estimates of $\Lambda^{k,\cdot}$ and Ξ , the covariance matrices can be obtained as follows:

$$\begin{aligned}
\sum_E(j+1) = & \text{diag}\left(\sum_{t=1}^{Tv} (\mathcal{S}_t Y_t^{k,n} Y_t^{k,n'} \mathcal{S}_t' - \mathcal{S}_t Y_t^{k,n} E_{\theta(j)}[F_t^{k,n'} | \Omega_v] \Lambda^{k,\cdot} (j+1)' \mathcal{S}_t \right. \\
& - \mathcal{S}_t \Lambda^{k,\cdot} (j+1) E_{\theta(j)}[F_t^{k,n} | \Omega_v] Y_t^{k,n'} \mathcal{S}_t + \mathcal{S}_t \Lambda^{k,\cdot} (j+1) E_{\theta(j)}[F_t^{k,n} F_t^{k,n'} | \Omega_v] \Lambda^{k,\cdot} (j+1)' \mathcal{S}_t \\
& \left. + (I_N - \mathcal{S}_t) \sum_{E(j)} (I_N - \mathcal{S}_t) \right),
\end{aligned} \tag{C.25}$$

and

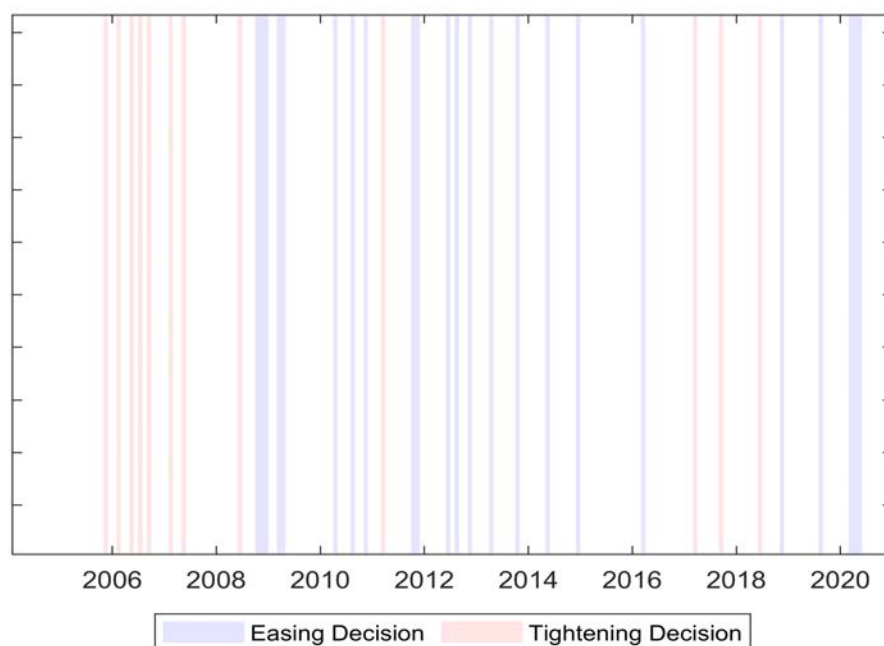
$$\sum_U(j+1) = \frac{1}{T} \left(\sum_{t=1}^{Tv} E_{\theta(j)}[F_t^{k,n} F_t^{k,n'} | \Omega_v] - \Xi(j+1) \sum_{t=1}^{Tv} E_{\theta(j)}[F_t^{k,n} F_t^{k,n'} | \Omega_v] \right) \tag{C.26}$$

where the expectations $E_{\theta(j)}(\cdot)$ are obtained via the Kalman filter and smoother, and the estimates for Ξ and \sum_U follow from taking the elements from the conditional covariances of the state vector corresponding to $F_t^{k,n}$.

C.3. The Bridge Equation

The bridge equation is introduced to equip the model with the possibility to forecast the probability that the ECB will take a monetary policy decision given some incoming information. It owes its definition to its function, that is, it “bridges” the forecast variables from the DFM with the dependent variable. In this case, the specification of the bridge equation follows a multinomial logit. To create the categorical variable of monetary policy decisions, I then followed some basic principles. First, a tightening decision to be classified as such must contain one (or a combination) of the following announcements: an increase in interest rates, a reduction in asset purchases and a hawkish revision of the forward guidance; second, an easing decision to be classified as such must instead include one (or a combination) of the following announcements: an interest rate cut, the launch of an asset purchase programme or a type of long-term refinancing operations (LTROs), an increase in asset purchases, a dovish revision of the forward guidance and an extension of the collaterals eligible for repos; last, monetary policy remains constant if no decision in the first or second group is announced. As a result, from January 2002 to December 2020 tightening decisions total 13, while easing one amount to 33. [Figure A.5](#) displays the tightening monetary policy decisions (shaded red) and the easing ones (shaded blue).

Figure C.1: Out-of-Sample Application



Note: The figure shows actual tightening (shaded red) versus easing decisions (shaded blue).

The performance of the bridge equation is assessed by two statistics: the mean absolute forecast squared error (MAFE) and the pseudo R^2 statistics. The MAFE is simply given by the mean absolute difference between the model-implied probabilities and the actual outcome of the categorical monetary policy variable. Therefore, once the probabilities have been estimated is straightforward to compute:

$$MAE = \frac{1}{T} \sum_{t=1}^T (|y_t - P(y_t = j)|) \quad (C.27)$$

On the other hand, pseudo R^2 is a measure of fit defined as the weighted sum of fractions of correctly identified easing, tightening and constant decisions. Formally, it can be written:

$$R^2 = \frac{n_{tt}}{n_t} \cdot \frac{n_t}{n} + \frac{n_{ee}}{n_e} \cdot \frac{n_e}{n} + \frac{n_{cc}}{n_c} \cdot \frac{n_c}{n} = \frac{n_{tt} + n_{ee} + n_{cc}}{n} \quad (C.28)$$

where n is the number of press conferences for which forecasts have been computed, n_t , n_e and n_c are, respectively, the number of tightening, easing and constant decisions and n_{tt} , n_{ee} and n_{cc} are the number of correctly identified tightening, easing and constant decisions. It is important to note that the pseudo R^2 relies on a conversion of the model-implied recession probabilities into a categorical variable. That is, given an estimated probability of recession

$\hat{P}(y_t = j)$ for some time t , a mapping $D : [0, 1] \rightarrow \{0, 1\}$, $\hat{y}_t = D[\hat{P}(y_t = j)]$, needs to be implemented to convert probabilities into alleged monetary policy decisions. Usually, such a decision rule depends on a threshold c^* such that $\hat{y}_t = j$, if $\hat{P}(y_t = j) > c^*$. Hence, the choice of the threshold c^* is key. Setting $c^* = 0.5$ is often too conservative a criterion, especially when the overall fraction of tightening and easing decisions in the sample is relatively small (see [Green, 2012](#) on discrete choice models). A more natural choice is to set c^* equal to the frequency of easing and tightening decisions measured over a long period. For this reason, I set $c^* = 0.15$ for easing decisions and $c^* = 0.08$ for tightening ones. While in the out-of-sample exercises, this threshold could be adjusted period by period, I treat it as constant so that results are independent of such time variation.

One concern with this choice of threshold may be that the probability of a tightening (p_T) and the probability of an easing (p_E) can be true at the same time. However, such an outcome is found to occur twice in roughly 900 forecasts (0.2%). Similarly, the same situation can happen for the probability of no decisions (p_N) and the probability of an easing and for the probability of no decisions and the probability of a tightening. Also in this case, however, the probability for that to happen is not even one in ten. In fact, $p_N \vee p_T \approx 6\%$, while $p_N \vee p_T \approx 8\%$.

C.4. Alternative Benchmark Models

The literature on forecasting interest rates abounds with different Taylor rule specifications. In the main text, however, I only considered one single specification as a benchmark model. As a complement, in this paragraph, I compare the performance of the DFM with alternative “pseudo” Taylor-rule models¹. To do so, I follow [Blattner and Margaritov \(2010\)](#) who apply the “thick-modelling” approach suggested by [Granger and Jeon \(2004\)](#) to the euro area. The strategy consists in estimating all plausible specifications of a model and then pooling the parameter estimates according to some efficiency criteria.

Let me consider a very general family of monetary policy rules that nests a variety of specifications:

¹ “Pseudo” because my dependent variable is not “interest rate” but the ECB monetary policy stance series derived above.

$$ECB_t = \rho ECB_{t-1} + (1 - \rho)(r^* + \pi^* + \beta_\pi(\mathbb{E}_t[\pi_{t+h}] - \pi^*)) + \beta_y[y_{t+k} - y_{t+k-1}] \quad (\text{C.29})$$

where ECB_t is the ECB monetary policy stance, r^* the real equilibrium interest rate, π^* the inflation target of the monetary authority, π_t a measure of price developments and $(y_t - y^*)$ a measure of economic slack².

Based on [Equation C.29](#), I estimate in total 1,290 policy rules, exploiting all plausible combinations of 43 measures of economic activity and 30 indicators of inflation and inflation expectations (see below)³.

² y^* is estimated, as in the main text, with the HP filter.

³ Data are gathered from Bloomberg, Haver and Consensus Economics. When the variables have frequency lower than monthly, I use linear interpolation. Instead, when variables have frequency higher than monthly, I average the observations. Besides, given the impossibility to retrieve real-time data for every series, unlike [Blattner and Margaritov \(2010\)](#), regressions are not implemented with real-time variables.

Series
Economic Activity
Industrial Production Total
Industrial Production (construction)
Industrial Production: manufacturing
Industrial Production Confidence Unemployment (labour force)
EUROCOIN
Composite Output EA (PMI)
Manufacturing (PMI)
Services (PMI)
Construction (PMI)
Capacity Utilization (PMI)
Euro Area Big 2 (PMI)
Employment expectations
Unemployment expectations
Business Climate Indicator
Economic sentiment indicator
Bloomberg Weighted Average Private GDP Forecast
Bloomberg Weighted Average Official GDP Forecast
Bloomberg Weighted Average Private IP Forecast
Bloomberg Weighted Average Official IP Forecast
Bloomberg Weighted Average Private Unemployment Forecast
Consumer Confidence
Domestic Demand
Economic Slack (Domestic Demand - Real GDP)
Foreign annual GDP growth forecast (US, JP, UK)
Foreign industrial production growth forecast (US, JP, UK)
Next Exports
Gross Fixed Capital Formation

Series	
Gross Fixed Capital Formation (current year)	
Gross Fixed Capital Formation (next year)	
Labour Productivity	
Private Consumption Expenditure	
Private Consumption Expenditure forecast (current year)	
Private Consumption Expenditure forecast (next year)	
Real GDP	
Real GDP growth forecast (current year)	
Real GDP growth forecast (next year)	
Retail Sales Confidence	
Retail Sales Volume	
Total Employment Growth	
Unemployment Rate	
Unemployment Rate growth forecast (current year)	
Unemployment Rate growth forecast (next year)	
Inflation	
Harmonised Index of Consumer Prices (HICP)	
HICP (%)	
HICP: services	
HICP: excl. energy and food	
HICP: processed food	
HICP: unprocessed food	
HICP: consumer goods (durables)	
HICP: consumer goods (non-durables)	
HICP: consumer goods (total)	
HICP: non-energy industrial goods	
HPPI: Industry excluding construction	
Oil Prices (Brent)	

Series
Consumer Price Index (CPI)
CONSUMER PRICES OF PETROLEUM PRODUCTS (PUMP PRICES)
HWI COMMODITY PRICES FOR THE EURO AREA: Total Excluding Energy
Bloomberg Weighted Average Private CPI Forecast
Unit Labour Costs
Wage Growth Forecast (Current Year)
Wage Growth Forecast (Next Year)
Industrial Producer Prices: capital goods
Industrial Producer Prices: intermediate goods
Industrial Producer Prices: manufacturing goods
Industrial Producer Prices: total (excl. construction)
Compensation per employee
5y5y inflation swap rates
Consumer price forecast (current year)
Consumer price forecast (next year)
Break-even inflation rate (French indexed bonds)
Break-even inflation rate (German indexed bonds)
Break-even inflation rate (Italian indexed bonds)

I then apply four filters that reduce the number of specifications to a set of rules that delivers a meaningful approximation of the ECB's interest-rate setting behaviour:

1. I require the estimates of β_π and β_y to be statistically significant at least at the 10% confidence level.
2. I restrict the estimates of β_π and β_y to be strictly positive.
3. I require the equilibrium interest rate r^* to lie within -2 and 6 (see Brand et al., 2018 for a survey of the literature).
4. I discard all specifications with a standard error of the estimated coefficients of the economic growth variable larger than the threshold value of 0.5. Similarly, I ignore rules with a standard error of the smoothing coefficient larger than 0.04, a standard error of the constant larger than 0.7 and a standard error of the coefficient of the inflation weight larger than 1.5).

The criteria outlined reduce the number of models to 157 (12.17%). Table C.4 provides an overview of the least squares estimates based on the listed data. I report simple averages of the estimated coefficients, standard errors and the R^2 . That is, I assume that policy-makers assign equal weight to each of the relevant combinations of inflation and output indicators when forming their policy decision⁴.

Table C.4: Mean Estimates of Policy Rules for the Euro Area

Rules	N	% of Total	ρ	r^*	β_π	β_y	SEE	R^2
Total Rules	157		0.90 (0.03)	-0.51 (0.34)	1.20 (0.39)	0.90 (0.20)	0.2017	0.9516
Forward-looking	46	29.9%	0.91 (0.04)	-1.05 (0.38)	2.11 (0.86)	1.10 (0.19)	0.2134	0.9598
Backward-looking	51	32.4%	0.92 (0.03)	-0.36 (0.32)	0.79 (0.14)	0.92 (0.25)	0.2001	0.9511
Mixed Rules	60	38.2%	0.91 (0.04)	-0.61 (0.39)	1.52 (0.61)	0.96 (0.20)	0.2211	0.9522

⁴ Given that the whole idea behind thick modelling is to hedge against possible mispecifications, assigning equal weight to each indicator minimizes the risks associated with model uncertainty.

It is now possible to compare the performance of the DFM with the performance of alternative benchmark models. The comparison occurs with five policy rules with the smallest RMSFE as well as with the aggregate models listed in [Table C.4](#). [Table C.5](#) reports the results of the Diebold-Mariano test statistic ([Diebold and Mariano, 1995](#)).

Table C.5: Forecast Evaluation for Alternative Benchmark Models

2005-2014				
H	1	2	3	4
<i>DFM\BM₁</i>	0.75	1.22	1.90*	2.61***
<i>DFM\BM₂</i>	1.14	1.45	2.01**	2.37**
<i>DFM\BM₃</i>	1.22	1.38	1.79*	3.11***
<i>DFM\BM₄</i>	0.98	1.13	1.59	2.41**
<i>DFM\BM₅</i>	0.67	0.94	1.55	2.78***
<i>DFM\BM_{all}</i>	1.55	1.58	1.73*	2.26**
<i>DFM\BM_{for}</i>	1.41	1.49	1.65*	1.88*
<i>DFM\BM_{back}</i>	1.05	1.33	2.11**	2.90***
<i>DFM\BM_{mix}</i>	1.15	1.34	2.23**	3.08***
2015-2020				
H	1	4	6	8
<i>DFM\BM₁</i>	1.172	1.19	1.71*	3.10***
<i>DFM\BM₂</i>	1.16	1.26	1.90*	3.22***
<i>DFM\BM₃</i>	1.01	0.163	1.34	3.11***
<i>DFM\BM₄</i>	1.15	1.16	1.45	2.99***
<i>DFM\BM₅</i>	1.15	1.26	2.33**	2.65***
<i>DFM\BM_{all}</i>	1.10	1.16	2.55**	2.76***
<i>DFM\BM_{for}</i>	0.78	1.11	1.56	3.01***
<i>DFM\BM_{back}</i>	0.98	1.32	2.89***	3.44***
<i>DFM\BM_{mix}</i>	1.21	1.44	2.79***	3.21***

Note: The table shows the t-test of the Diebold-Mariano test statistic (Diebold and Mariano, 1995). The test-statistics are based on out-of-sample forecast errors for the period 2005-2014 and 2015-2020. The row “H” indicates the number of weeks between two consecutive press conferences. *, **, and *** denote, respectively, the 10%, 5%, and 1% significance level.

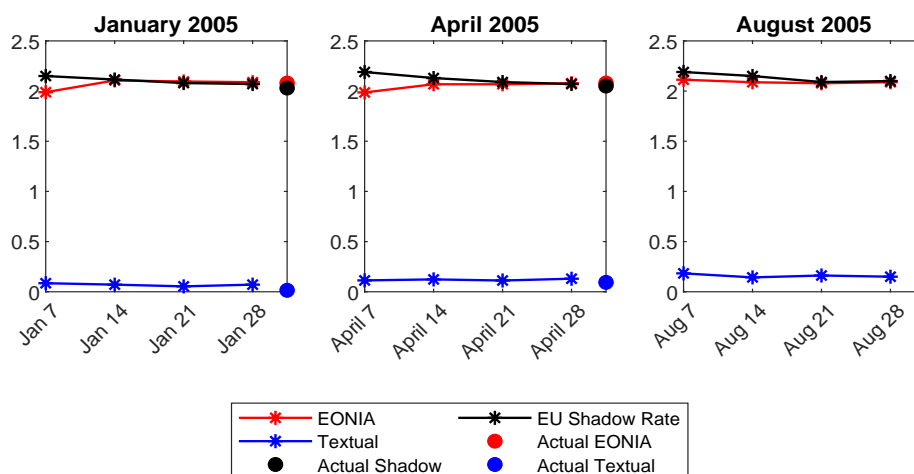
Table C.5 shows that the DFM, in the vast majority of cases, outperforms different versions of the Taylor rule at longer horizons.

D. Validating the DFM

This section attempts to validate the DFM by nowcasting the EONIA, the EU shadow rate (Wu and Xia, 2016) and a textual measure of monetary policy stance derived using Loughran and McDonald (2011)’s dictionary. Such a validation exercise aims to test whether the DFM performs well on time series related to the ECB’s monetary policy stance that are different from the one I developed in the paper.

For every series, I display the out-of-sample nowcasting of January, April and August 2005. Figure D.1 shows the results for, respectively, EONIA, EU shadow rates and the alternative textual measure of ECB monetary policy stance¹. Figure 10 should also be considered part of this exercise.

Figure D.1: Validating the DFM with Out-of-sample Nowcasts



Note: The figure shows the out-of-sample nowcasts of EONIA (red line), EU Shadow rate (black line) and an alternative textual measure of ECB monetary policy stance (blue line) for January, April and August 2005.

The validation exercise, along with the results shown in Figure 10, sheds light on the validity of the DFM beyond the application to the textual measure of ECB monetary policy stance derived in this paper. The DFM, in fact, tracks accurately the actual realizations for the variables of interest.

¹ The textual measure should not be taken as representative of the ECB monetary policy stance since it suffers from the flaws outlined in Section 2.2

References

- Apel, M. and M. Blix Grimaldi (2012). The Information Content of Central Bank Minutes. Working Paper Series 261, Sveriges Riksbank (Central Bank of Sweden).
- Babii, A., E. Ghysels, and J. Striaukas (2021). Machine learning time series regressions with an application to nowcasting. *Journal of Business & Economic Statistics* 0(0), 1–23.
- Bańbura, M., D. Giannone, M. Modugno, and L. Reichlin (2013). Now-casting and the real-time data flow. In C. G. G. Elliott and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 2, Chapter 0, pp. 195–237. Elsevier.
- Bańbura, M. and M. Modugno (2014). Maximum Likelihood Estimation of Factor Models on Datasets with Arbitrary Pattern of Missing Data. *Journal of Applied Econometrics* 21(1), 1–21.
- Baranowski, P., H. Bennani, and W. Doryń (2021). Do the ECB’s introductory statements help predict monetary policy? Evidence from a tone analysis. *European Journal of Political Economy* 66, 101964.
- Bennani, H., N. Fanta, P. Gertler, and R. Horvath (2020). Does central bank communication signal future monetary policy in a (post)-crisis era? The case of the ECB. *Journal of International Money and Finance* 104(19), 102167.
- Bennani, H. and M. Neuenkirch (2017). The (home) bias of European central bankers: new evidence based on speeches. *Applied Economics* 49(11), 1114–1131.
- Blattner, T. S. and E. Margaritov (2010). Towards a robust monetary policy rule for the euro area. Working Paper Series 1210, European Central Bank.
- Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3, 993–1022.
- Bok, B., D. Caratelli, D. Giannone, A. M. Sbordone, and A. Tambalotti (2018). Macroeconomic Nowcasting and Forecasting with Big Data. *Annual Review of Economics* 10(1), 615–643.

- Brand, C., M. Bielecki, and A. Penalver (2018). The natural rate of interest: estimates, drivers, and challenges to monetary policy JEL Classification: E52, E43. Occasional Paper Series 217, European Central Bank.
- Bybee, L., B. T. Kelly, A. Manela, and D. Xiu (2020). The Structure of Economic News. NBER Working Papers 26648, National Bureau of Economic Research, Inc.
- Cao, J., T. Xia, J. Li, Y. Zhang, and S. Tang (2009). A density-based method for adaptive LDA model selection. *Neurocomputing* 72(7-9), 1775–1781.
- Cimadomo, J., D. Giannone, M. Lenza, A. Sokol, and F. Monti (2020). Nowcasting with large Bayesian vector autoregressions. Working Paper Series 2453, European Central Bank.
- Clements, M. and A. Beatriz, Galvao (2008). Macroeconomic forecasting with mixed-frequency data. *Journal of Business & Economic Statistics* 26, 546–554.
- Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* 39(1), 1–38.
- Diebold, F. and R. Mariano (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics* 13(3), 253–63.
- Foroni, C. and M. Marcellino (2013). A survey of econometric methods for mixed-frequency data. Working Paper 2013/06, Norges Bank.
- Gentzkow, M., B. Kelly, and M. Taddy (2019). Text as data. *Journal of Economic Literature* 57(3), 535–74.
- Gerlach, S. (2007). Interest Rate Setting by the ECB, 1999-2006: Words and Deeds. *International Journal of Central Banking* 3(3), 1–46.
- Ghysels, E., P. Santa-Clara, and R. Valkanov (2004). The midas touch: Mixed data sampling regression models. Cirano working papers, CIRANO.
- Ghysels, E., A. Sinko, and R. Valkanov (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews* 26(1), 53–90.

- Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics* 55(4), 665–676.
- Gorter, J., J. Jacobs, and J. De Haan (2008). Taylor rules for the ECB using expectations data. *Scandinavian Journal of Economics* 110(3), 473–488.
- Granger, C. and Y. Jeon (2004). Thick modeling. *Economic Modelling* 21(2), 323–343.
- Green, W. H. (2012). *Econometric Analysis* (7th Edition ed.). Upper Saddle River.
- Griffiths, T. L. and M. Steyvers (2004). Finding Scientific Topics. *Pnas* 101(SUPPL.1), 5228–5235.
- Hansen, S. and M. McMahon (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics* 99, S114–S133. 38th Annual NBER International Seminar on Macroeconomics.
- Hansen, S., M. McMahon, and A. Prat (2017). Transparency and deliberation within the fomc: A computational linguistics approach. *The Quarterly Journal of Economics* 133(2), 801–870.
- Hansen, S., M. McMahon, and A. Prat (2018). Deliberation within the FOMC: a Computational Linguistics Approach. *The Quarterly Journal of Economics* 133(2), 801–870,.
- Heinrich, G. (2009). Parameter estimation for text analysis (based on Gibbs sampling). *Technical Report, Fraunhofer IGD*.
- Jansen, D.-J. and J. De Haan (2009). Has ECB communication been helpful in predicting interest rate decisions? An evaluation of the early years of the Economic and Monetary Union. *Applied Economics* 41(16), 1995–2003.
- Ke, Z. T., B. T. Kelly, and D. Xiu (2019). Predicting Returns With Text Data. NBER Working Papers 26186, National Bureau of Economic Research, Inc.
- Kremer, M., M. Lo Duca, and D. Holló (2012). CISS - a composite indicator of systemic stress in the financial system. Working Paper Series 1426, European Central Bank.
- Kuzin, V. N., M. Marcellino, and C. Schumacher (2009). MIDAS versus mixed-frequency VAR: nowcasting GDP in the euro area. Discussion Paper Series 1: Economic Studies 7, Deutsche Bundesbank.

- Loughran, T. and B. McDonald (2011). When is a liability not a liability? Textual Analysis, Dictionaries, and 10-Ks. *Journal of Finance* 66(1), 35–65.
- Lucca, D. O. and F. Trebbi (2009). Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements. NBER Working Papers 15367, National Bureau of Economic Research, Inc.
- Marcellino, M. and C. Schumacher (2010). Factor midas for nowcasting and forecasting with ragged-edge data: A model comparison for german gdp. *Oxford Bulletin of Economics and Statistics* 72(4), 518–550.
- Mariano, R. S. and Y. Murasawa (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics* 18(4), 427–443.
- Marozzi, A. (2021). The ECB and the Cost of Independence. Unearthing a New Doom-Loop in the European Monetary Union. BAFFI CAREFIN Working Papers 21152, BAFFI CAREFIN, Centre for Applied Research on International Markets Banking Finance and Regulation, Università Bocconi, Milano, Italy.
- Modugno, M. (2013). Now-casting inflation using high frequency data. *International Journal of Forecasting* 29(4), 664–675.
- Picault, M. and T. Renault (2017). Words are not all created equal: A new measure of ECB communication. *Journal of International Money and Finance* 79, 136–156.
- Rinker, T. W. (2019). *sentimentr: Calculate Text Polarity Sentiment*. Buffalo, New York. version 2.8.0.
- Shapiro, A. H. and D. J. Wilson (2019). Taking the Fed at its Word: A New Approach to Estimating Central Bank Objectives Using Text Analysis. Working Paper Series 2019-02, Federal Reserve Bank of San Francisco.
- Stoltz, D. S. and M. A. Taylor (2019). Concept Mover’s Distance: measuring concept engagement via word embeddings in texts. *Journal of Computational Social Science* 2(2), 293–313.
- Sturm, J. E. and J. de Haan (2011). Does central bank communication really lead to better forecasts of policy decisions? New evidence based on a Taylor rule model for the ECB. *Review of World Economics* 147(1), 41–58.

- Taylor, J. B. (1993). Discretion versus policy rules in practice: two critical points. A comment. *Carnegie-Rochester Confer. Series on Public Policy* 39(C), 215–220.
- Thorsrud, L. A. (2020). Words are the New Numbers: A Newsy Coincident Index of the Business Cycle. *Journal of Business and Economic Statistics* 38(2), 393–409.
- Wieland, V. and M. Wolters (2013). Forecasting and policy making. In C. G. G. Elliott and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 2, Chapter 0, pp. 239–325. Elsevier.
- Wu, J. C. and F. D. Xia (2016). Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound. *Journal of Money, Credit and Banking* 48(2-3), 253–291.

Acknowledgements

The paper was written while working at the ECB. I would like to thank for useful comments Michele Modugno, Michele Piffer and an anonymous referee of the ECB Working Paper Series.

Armando Marozzi

London School of Economics, London, United Kingdom; email: a.marozzi@lse.ac.uk

© European Central Bank, 2021

Postal address 60640 Frankfurt am Main, Germany

Telephone +49 69 1344 0

Website www.ecb.europa.eu

All rights reserved. Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the ECB or the authors.

This paper can be downloaded without charge from www.ecb.europa.eu, from the [Social Science Research Network electronic library](#) or from [RePEc: Research Papers in Economics](#). Information on all of the papers published in the ECB Working Paper Series can be found on the [ECB's website](#).

PDF

ISBN 978-92-899-4862-3

ISSN 1725-2806

doi:10.2866/214406

QB-AR-21-100-EN-N