LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

ESSAYS ON THE EUROPEAN CENTRAL BANK'S COMMUNICATION

Armando Marozzi

A thesis submitted to the Department of International Relations of the London School of Economics and Political Science for the degree of Doctor of Philosophy, London

September 2022

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I declare that my thesis consists of 30,853 words.

Abstract

This thesis studies the communication of the European Central Bank (ECB) from different angles. In particular, the first essay analyses the effects of the ECB's fiscal communication on the macroeconomy of the euro area. The findings are threefold: first, a fiscally hawkish communication significantly lowers output and inflation; second, industrial production and price levels would have been higher during the Great Recession and the sovereign-debt crisis had the ECB held constant its fiscal communication; third, government bond purchases are inversely related to the ECB's "fiscal stance", that is, the ECB becomes more fiscally hawkish after a monetary policy easing. The second essay proposes to nowcast the ECB's press conferences exploiting the flow of conventional and textual data that become available between two consecutive press conferences. 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. The third essay investigates the effectiveness of monetary policy conditional on the level of agreement in the ECB Governing Council. The results illustrate significant asymmetry: while expansionary monetary policy propagates as documented in the literature in the scenario of high agreement, the effects of monetary policy accommodation vanish when disagreement is high.

Contents

Introduction

Since the Great Recession, central bank communication has become a key tool for central banks in their attempts to shape markets' expectations. As a result, the literature has extensively documented the effectiveness of central bank communication to convey current and future monetary policy decisions. This thesis contributes to this literature by examining different aspects of communication by the European Central Bank (ECB). In particular, I focus on three topics: central bank communication on fiscal policy, forecasting monetary policy press conferences and investigating the effects of central bankers' disagreements on the transmission of monetary policy.

In the first paper, I study how the communicated degree of fiscal policy accommodation (hereafter DFPA) of the ECB affects the macroeconomy of the euro area and how it interacts with government bond surprises. To do so, I proceed in three steps. First, to capture the communicated "fiscal stance" of the ECB, I subset the ECB press conferences to extract fiscal sentences that are normative in nature. I then apply Wordscore (Laver et al., 2003), a supervised machine learning model, to extract fiscal policy dimensions ("hawkish"/"dovish"). Second, to study the macroeconomic effects of the ECB's fiscal communication, I use Bayesian Local Projections (BLP) (Miranda-Agrippino and Ricco, 2021) recursively identified with the DFPA variable ordered last. I also run historical decomposition to estimate the impact of DFPA shocks on the development of inflation and output. Third, to investigate how the DFPA variable responds to government bond surprises, I employ a Factor-Augmented VAR (FAVAR) identified with different bond-buying shocks derived from the *Euro Area Monetary Policy Event-Study Database* (Altavilla et al., 2019).

The main results are the following. First, I document the existence of a quantifiable communicated ECB fiscal stance. Second, communicating a fiscally hawkish stance significantly lowers output and inflation. In particular, the fiscal stance of the euro area tends to react to the fiscal recommendations of the ECB by cutting expenditures and raising taxes. This response however varies across blocks and conditional on the level of interest rates. On the one hand, southern countries appear to be more responsive than northern ones to the ECB's fiscal reprimands. On the other hand, both southern and northern blocks are more responsive when their fiscal capacity is constrained by an environment of higher rates. Third, historical decomposition provides empirical evidence that output and inflation would have been higher during the Great Recession as well as the sovereign-debt crisis, had the ECB held its communicated "fiscal stance" constant. Fourth, I find a statistically significant inverse relationship between government bond purchases and the "fiscal stance" of the ECB: the monetary authority becomes more fiscally conservative given a bond-buying easing.

It shall be however noticed that these results may suffer from endogeneity as the DFPA variable may likely be endougenous with fiscal variables that in turn depend on the business cycle. In further releases, the results would benefit from the inclusion in the dataset of a measure of discretionary fiscal polcies that is deemed orthogonal to business cycle dynamics.

In the second paper, I provide an econometric framework which enables interested parties to track systematically 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.

A wide range of conventional and non-conventional 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.

This paper attempts to exploit the entire flow of information that occurs in the times-pan between two consecutive ECB press conferences. This implies the construction of a dataset that contains not only conventional data such as financial and macroeconomic releases but also news broadcasted through the media. To create a coherent econometric framework that embeds both conventional and textual data, I proceed with the following steps. 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 and finance-related information. Third, I model the total dataset containing around 140 variables as a Dynamic Factor Model (DFM) with flow and stock variables. 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 over the period 2015-2020.

In the third paper, I investigate whether monetary policy is more powerful in a regime of agreement or disagreement among central bankers. In particular, I focus on disagreements within the Governing Council (GC) of the ECB. The hypothesis of the paper is that while in a regime of agreement monetary policy propagates as documented by the literature, in a regime of disagreement the impact of a monetary policy easing might vanish due to second-round effects through financial markets. In fact, the market may interpret disagreement as a fragile commitment that may lead the monetary authority to unwind monetary policy support more quickly than actually warranted.

To test this hypothesis, I proceed in two steps: first, since the ECB does not publish voting records, I use a dictionary approach based on the ECB's press conferences to measure the level of agreement/disagreement within the GC over monetary policy decisions. I call this proxy "consensus index" (CI); second, I use smooth transition-local projection model (STLPM) employed in Auerbach and Gorodnichenko (2012) and Ramey and Zubairy (2014) to analyze fiscal policy and Tenreyro and Thwaites (2016) to study the effects of a monetary policy easing conditional on the agreement within the GC.

The results support the hypothesis that monetary policy easing significantly differs depending on whether members of the GC agree or disagree over a monetary policy decision. While, in the consensus regime, a 1% monetary easing leads to results in line with the literature, the same shock in a disagreement regime has no statistically significant effect on macroeconomic as well as financial variables. As in the first paper, endogeneity may be an issue in this specification as well. In fact, the consensus index which measures divergence/convergence about monetary policy decisions might reflect different views on the evolution of key macroeconomic variables. Further robustness checks are therefore warranted to corroborate the results.

Overall, the thesis provides methodological innovations as well as novel empirical results. Methodologically, the first paper provides, to the best of my knowledge for the first time, a method to quantify the "fiscal stance" of a central bank over the polarized policy dimension "hawkish"/"dovish". Similarly, it is also novel the endeavor to identify DFPA shocks. The second paper illustrates a mixed-frequency econometric framework for describing and predicting central banks' reaction function. The third paper proposes a new index to measure the agreement within the GC of the ECB.

Empirically, the first paper assesses that a DFPA shock might be one of the factors that bring inflation and output down. It also documents an inverse relationship between government bond purchases and the fiscal stance of the ECB. The second paper sheds light on the role of news broadcasted through the media as a fundamental channel of expectation formation. It captures, in fact, the prominent contribution of textual variables in explaining forecast revisions from 2015 on, that is, a period of renewed importance of communication as a policy tool. The third paper reports that monetary policy easing has no effect in a regime of heightened disagreement among central bankers.

The papers refer to different streams of literature. In particular, the first paper adds to the literature on central bank communication. On the one hand, many papers relied on natural language processing (NLP) techniques to measure central banks' monetary policy stance and its predictive power (Hansen and McMahon, 2016; Hansen et al., 2017; Picault and Renault, 2017; Bennani et al., 2020; Baranowski et al., 2021; Gorodnichenko et al., 2021; Marozzi, 2021). On the other hand, several scholars investigated the macroeconomic and financial effects of central banks' announcements (Brand et al., 2010; Altavilla et al., 2014; Hansen and McMahon, 2016; Jarociński and Karadi, 2020; Swanson, 2021). While the literature on monetary policy communication is abundant, the one on central bank communication on fiscal policy is much scarcer (see for example Allard et al., 2013 and Ferrara, 2020). The focus in this literature is primarily on measuring the intensity of central banks' fiscal communication, that is, how much central banks talk about fiscal policy. However, no attempt has been made to quantify the direction (e.g. "hawkish" vs "dovish") of that communication.

The second paper contributes to two 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 variables[1]. Thorsrud (2020) and Cimadomo et al. (2020) come also close to this paper: the former employs news text to derive a daily business cycle index and predict quarterly 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.

[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).

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

The third paper concerns the literature on central banks' disagreement. The paper is most closely related to Falck et al. (2018) that investigate how disagreement about inflation expectations interacts with the efficacy of monetary policy. In times of high disagreement, they estimate that a 100 basis points (bps) increase in the U.S. policy rate leads to a significant short-term increase in inflation and in inflation expectations of up to 1 percentage point, whereas in times of low disagreement they find a significant decline of close to 1 percentage point. Moreover, Seelajaroen et al. (2020) studies how monetary policy transparency of the Bank of England has information content in reducing disagreement about interest rate forecasts. The authors find that disagreement among the Monetary Policy Committee in policy rate decisions is associated with lower disagreement among professional forecasters on interest rate outlook, whereas neither announcement of changes in policy rates nor publication of inflation reports affects forecast disagreement. Similarly, Jitmaneeroj et al. (2019) provide evidence that greater transparency surrounding monetary policy reduces uncertainty of interest rates and inflation, primarily by reducing uncertainty that is common to agents rather than disagreement between agents.

This thesis is organised as follows. Section 1 presents the first paper; Section 2 illustrates the second paper and Section 3 outlines the third paper. The last section concludes the thesis.

1. Beware of Fiscal Signalling: The Effect of the ECB's Fiscal Communication in the Euro Area

1.1. Introduction

Since the outbreak of the global financial crisis in 2007, the academic literature has extensively studied the impact of central bank communication on the general public and financial market. However, while the established literature on central bank communication focused on the communication of monetary policy messages, the more specific issue of central bank communication on fiscal policy has so far received scarce attention.

In this paper, I study how the communicated degree of fiscal policy accommodation (hereafter DFPA) of the European Central Bank (ECB) affects the macroeconomy of the euro area and how it interacts with government bond surprises.

I proceed in three steps. First, to capture the communicated "fiscal stance" of the ECB, I subset the ECB press conferences to extract fiscal sentences that are normative in nature. I then apply Wordscore (Laver et al., 2003), a supervised machine learning model, to extract fiscal policy dimensions ("hawkish"/"dovish"). Second, to study the macroeconomic effects of the ECB's fiscal communication, I use Bayesian Local Projections (BLP) (Miranda-Agrippino and Ricco, 2021) recursively identified with the DFPA variable ordered last. I also run historical decomposition to estimate the impact of DFPA shocks on the development of inflation and output. Third, to investigate how the DFPA variable responds to government bond surprises, I employ a Factor-Augmented VAR (FAVAR) identified with three bond-buying shocks − Sovereign-Debt Crisis (SDC), Quantitative Easing (QE) and Large-Scale Government Bond Purchases (LSBP) − derived from the *Euro Area Monetary Policy Event-Study Database* (Altavilla et al., 2019). Throughout the paper, whenever I study the impulse response of fiscal variables, I group them into three categories: aggregate, North and South. This is to reflect the heterogeneity of fiscal positions in the euro area.

The main results are the following. First, I document the existence of a quantifiable communicated ECB fiscal stance. Second, communicating a fiscally hawkish stance may result in lowering output and inflation. In particular, the fiscal stance of the euro area tends to react to the fiscal recommendations of the ECB by cutting expenditures and raising taxes. This response however varies across blocks and conditional on the level of interest rates. On the one hand, the South appears to be more responsive than the North to the ECB's fiscal reprimands. On the other hand, both southern and northern blocks are more responsive when their fiscal capacity is constrained by an environment of higher rates. Third, historical decomposition provides empirical evidence that output and inflation would have been higher during the Great Recession as well as the sovereign-debt crisis, had the ECB held its communicated "fiscal stance" constant. Fourth, I find a statistically significant inverse relationship between government bond purchases and the "fiscal stance" of the ECB: the monetary authority becomes more fiscally conservative given a bond-buying easing.

This paper provides both methodological and empirical contributions. Methodologically, it is, to the best of my knowledge, the first attempt to quantify the "fiscal stance" of a central bank over the polarized policy dimension "hawkish"/"dovish". Similarly, it is also novel the endeavor to identify DFPA shocks. Empirically, the paper documents the effects of fiscal policy communication in the macroeconomy of the euro area and unearths a new pattern between government bond purchases and DFPA.

These contributions add to the literature on central bank communication. One stream of papers focused on monetary policy communication. On the one hand, many papers relied on natural language processing (NLP) techniques to measure central banks' monetary policy stance and its predictive power (Hansen and McMahon, 2016; Hansen et al., 2017; Picault and Renault, 2017; Bennani et al., 2020; Baranowski et al., 2021; Gorodnichenko et al., 2021; Marozzi, 2021). On the other hand, several scholars investigated the macroeconomic and financial effects of central banks' announcements (Brand et al., 2010; Altavilla et al., 2014; Hansen and McMahon, 2016; Jarociński and Karadi, 2020; Swanson, 2021). While the literature on monetary policy communication is abundant, the one on central bank communication on fiscal policy is much scarcer (see for example Allard et al., 2013 and Ferrara, 2020). The focus in this literature is primarily on measuring the intensity of central banks' fiscal communication, that is, how much central banks talk about fiscal policy. However, no attempt has been made to quantify the direction (e.g. "hawkish" vs "dovish") of that communication.

The paper proceeds as follows: Section 1.2 illustrates the quantification of DFPA shocks as well as the identification strategy. Section 1.3 presents the macroeconomic effects of DFPA shocks in the euro area. Section 1.4 documents the response of the DFPA variable to bond-buying surprises. Section 1.5 concludes.

1.2. Econometric Methodology

This section explains the econometric models used in the paper. In particular, Section 1.2.1 documents the methodology adopted to quantify the "fiscal stance" of the ECB, Section 1.2.2 outlines the identification strategy and Section 1.2.3 explains the historical decomposition.

1.2.1 The ECB's Communicated Degree of Fiscal Policy Accommodation (DFPA)

In this section, I first justify why it is consistent with economic theory to expect that an independent monetary authority has a "fiscal stance" and then I detail the procedure employed to quantify the DFPA variable.

1.2.1.1 Monetary Authorities and Fiscal Communication

Why should central banks in a monetary dominance regime have a "fiscal stance"?[1] Following the seminal paper of Sargent and Wallace (1981) and the subsequent work on the fiscal theory of price level (Sims, 1994; Woodford, 1995; Woodford, 2001; Cochrane, 2001), the answer is that monetary and fiscal actions are linked through the intertemporal government's budget constraint. This interface between monetary and fiscal authorities can be formally seen in the following equation:

$$b_{t-1} = R^{-1} \sum_{i=0}^{\infty} R^{-i} s^f_{t+i} + R^{-1} \sum_{i=0}^{\infty} R^{-i} s_{t+i} \tag{1.1}$$

where b_{t-1} denotes real debt, $R = 1+r$ is the gross real interest rate, $s^f_t = t_t - g_t$ is the primary fiscal surplus (i.e., tax revenues minus expenditures excluding interest payments and seigniorage revenue) and $s_t = Rh$, where h is the change in the monetary base, represents real seigniorage revenue. Equation 1.1 states that the current real liabilities of the government must be financed, in present value terms, by either a fiscal primary surplus or seigniorage. This means that in a regime of monetary dominance, in order for Equation 1.1 to hold, the fiscal authority must adjust to the central bank's actions. For example, if seigniorage decreases - either because the monetary authority sells in an open market operation interest-bearing debt by withdrawing monetary base (non-interest-bearing debt) or because it increases the nominal interest rate -, the fiscal authority must run higher primary fiscal surpluses (either via higher taxes or lower expenditures). It follows that a monetary authority, in a regime of monetary dominance, needs the backing of a fiscal authority to find an equilibrium price level[2].

[1] I follow Leeper (1991) in defining regimes of monetary and fiscal dominance. A regime of monetary dominance is assumed to occur when fiscal policy adjusts to ensure that the government's intertemporal budget is always in balance while monetary policy is free to set the nominal money stock or the nominal interest rate. Conversely, a fiscal dominance regime is one in which the fiscal authority sets its expenditure and taxes without regard to any requirement of intertemporal budget balance and monetary policy must adjust to balance the government's budget.

[2] This is so because, in equilibrium, the real quantity of money equals the real demand for money. Hence, if fiscal policy affects the real demand for money, the equilibrium price level will also depend on fiscal factors. Fiscal theories of price level instead maintain that there may be multiple price levels consistent with a given nominal quantity of money and equality between money supply and demand.

I considered so far the representative case of a single central bank and a single fiscal authority. However, the focus of this paper is on the euro area where a single central bank has to deal with many fiscal authorities. In such a peculiar architectural set-up, the task of the monetary authority is even more challenging. In fact, to deliver on price stability, the ECB needs the backing of several fiscal authorities with heterogeneous cyclical and structural positions (see Debrun et al., 2021). The absence of a centralised fiscal capacity therefore leaves the ECB exposed to fiscal policies set by possibly non-cooperative and self-oriented member states. This poses a threat to monetary dominance since the ECB (being the only institution in the euro area able to cushion an adverse macroeconomic shock) may be forced to depart from its primary objective (i.e. maintaining price stability) in order to avoid a sovereign default. This would violate the Treaty on the Functioning of the European Union (TFEU) and institute a regime of fiscal dominance.

One may argue that the ECB is shielded by the European Commission (EC) that is the institution entrusted with supervisory authority and executive power over fiscal matters in the euro area. While this is descriptively true, the Commission is ultimately a political body whose budgetary recommendations need to pass through the ECOFIN Council where politics reigns as it is made up of the finance ministers from all member states. Therefore, given the non-binding nature of fiscal surveillance procedures in the euro area[3], the ECB remains overexposed to member states' fiscal discipline.

Given the inherent nexus between monetary and fiscal policy, this paper then hypothesizes that an independent central bank might retain a "fiscal stance", that is, a preference towards the optimal course of fiscal policy *from a monetary standpoint* that enables the monetary authority to achieve its inflation target. Having a "fiscal stance", thus, does not entail breaching monetary dominance or overstepping the remits of a central bank, it rather follows from the "unpleasant arithmetic" of independent monetary institutions. In other words, following Equation 1.1, a monetary authority under monetary dominance is

Fiscal policy may then determine which of these is the equilibrium price level.

[3] For an account and explanation of the times the EC failed to sanction countries non-compliant with the Stability and Growth Pact (SGP) see, for instance, de Haan et al. (2003) and Sacher (2021).

expected to have a fiscal stance, not despite, but *because of* its independent status[4]. This is supposed to be even more pronounced in an institutional framework like the European Monetary Union (EMU) where the independence of the monetary authority is much more vulnerable[5].

After making the case that an independent monetary authority might retain a fiscal stance, I also claim that the central bank has an incentive in communicating its preferred path of fiscal policy *from a monetary standpoint*. The reason is that when the central bank communicates, financial markets price in an expectation that may then constrain the fiscal authority. As before, this incentive is stronger in a setting like the EMU where the ECB is threatened by the potential fiscal misconduct of many fiscal authorities. Under this regard, the fiscal communication of the monetary authority can be interpreted as an attempt to lock in government(s), thereby increasing the odds to achieve price stability while preserving monetary dominance.

The last remark is on the meaning of "fiscal stance" in the euro area (see Ferdinandusse et al., 2017; Cimadomo et al., 2017). Given the absence of a centralised fiscal capacity, the euro area fiscal stance can be regarded as the aggregation of the national fiscal stances measured by the structural primary balance. The aggregated euro area fiscal stance may then hide different positions at the domestic level with, for instance, fiscal tightening in some member states and loosening in others. This is to say that the fiscal stance of the ECB might then reflect the cross-sectional variation among euro area governments. As a result, I will take this heterogeneity into account throughout the empirical applications of this paper.

[4] The argument laid down in this paper could be a variation of Mervin King (1995)'s quote at Jackson Hole while being Chief Economist of the Bank of England: "Central banks are often accused of being obsessed with inflation. This is untrue. If they are obsessed with anything, it is with fiscal policy". This paper contends that since central banks are obsessed with inflation, then they must be obsessed with fiscal policy.

[5] The most prominent evidence in support of the claim that the ECB retains a "fiscal stance" consists in the leaked letters that the ECB sent to the Spanish (José Luis Rodríguez Zapatero) and Italian (Silvio Berlusconi) Prime Ministers on August 5 2011. The letters required Italy and Spain to follow a detailed fiscal plan. Despite this type of communication being highly confidential, it proves that the monetary authority had a very clear understanding of the desired course of fiscal policy from a monetary perspective.

1.2.1.2 Quantifying DFPA

In this paragraph, I provide evidence in support of the claim that the ECB regularly conveys its preferred fiscal stance in official communication.

I focus on the ECB's press conferences given their pivotal role in transmitting policy signals to the market[6]. In particular, I distinguish two types of fiscal communication. On the one hand, the ECB may comment *descriptively* on fiscal policy, regarding it as a factor that affects macroeconomic and price developments. On the other hand, the ECB may release *normative* statements on fiscal policy, that is, statements that call on government(s) to take some defined actions. These statements can either refer to single governments or to the euro area as a whole. In line with the previous paragraph and to fully capture the cross-sectional variation among member states, I refer to this normative dimension.

Accordingly, I manually subsetted every press conference from January 2002 to December 2019 to extract only fiscal sentences that are normative in nature. I then validated the final dataset by Latent Dirichlet Allocation (LDA) (Blei et al., 2003). Figure 1.1 shows the evolution of the share of normative fiscal sentences in the ECB's press conferences.

[6] There are five types of documents available: Governing Council (GC) decisions, press conferences (PC), Monetary policy accounts, speeches and interviews. This paper focuses on the PCs led by the President and the Vice-President of the ECB. GC decisions were excluded since numbers are more frequent than text. Yet, this choice does not affect the reliability of the study since PCs simply expand in words (and in much more detail) what is contained in the GC decisions. Monetary policy accounts are also excluded from the sample because they are available only from 2015. Finally, speeches and interviews were excluded due to the high number of topics discussed.

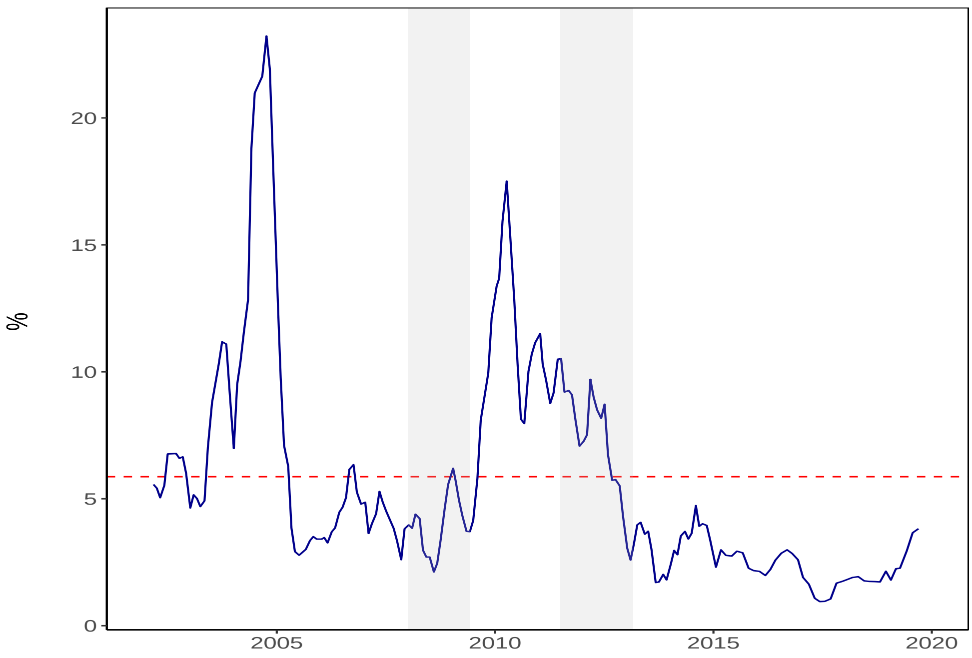


Figure 1.1: Share of Normative Fiscal Sentences in the ECB's Press Conferences. The figure shows the share of normative fiscal sentences in the ECB's press conferences. The dashed red line indicates the average from 2002 to 2019, which is around 6%.

Figure 1.1 shows two results: first, from 2002 to 2019, the ECB has *never* missed the opportunity to express its normative views over the course of fiscal policy in the euro area. On average, around 6% of the ECB's communication in the press conferences is devoted to outlining fiscal recommendations. Second, the percentage of statements dedicated to the DFPA varies over time. Notably, over the time frame 2002-2004, when France and Germany were violating the fiscal rules in the euro area, the ECB spent a considerable amount of time (up to 23%) recommending fiscal actions as well as praising the importance of fiscal discipline. Moreover, throughout the sovereign-debt crisis, the ECB significantly increased the proportion of normative statements concerning fiscal policy.

However, while Figure 1.1 documents how much the ECB talks normatively about fiscal policy, it is not informative about the actual "fiscal stance" of the ECB. Therefore, after pre-processing the normative dataset, I apply Wordscore (Laver et al., 2003), a supervised machine learning model, to extract fiscal policy dimensions ("hawkish"/"dovish").

Formally, the model estimates the scores of L out-of-sample documents, the virgin texts, given R reference texts with known positions or scores on a fiscal policy dimension.

Start with a set of R reference texts, represented by an $R \times J$ document-feature matrix F_{rj}, where r indexes the r^{th} document and j indicates the j^{th} word. Wordscores for the reference texts are generated as:

$$\underset{(1\times J)}{S} = \underset{(1\times R)}{a} \cdot \underset{(R\times J)}{P} \tag{1.2}$$

where S is the $1 \times J$ vector containing the single score for every word in the reference texts, a_{rd} is the assumed *a priori* position of reference text r on dimension d and $P_{rj} = \frac{F_{rj}}{\sum_{r=1}^{R} F_{rj}}$ indicates a $R \times J$ matrix of conditional probabilities where F_{rj} is the relative frequency, as a proportion of the total number of words in the text, of each different word used in reference text r. Each element in P_{rj} tells us the probability that we are reading reference text r, given that we are reading word j.

Given S, the aim is now to obtain a single score for any new text (virgin text). This is obtained as follows:

$$\underset{(1\times V)}{V} = \underset{(1\times J)}{S} \cdot \underset{(J\times V)}{W} \tag{1.3}$$

where V is the $1 \times V$ vector of document score (the mean dimension score of all of the scored words that it contains, weighted by the frequency of the scored words), S is the $1 \times J$ vector of wordscores derived above and $W_{jv} = \frac{P_{vj}}{\sum_{v=1}^{V} P_{vj}}$ is a $J \times V$ matrix of relative frequency of each word in the virgin text word as a proportion of the total number of words in the virgin text.

In the empirical specification, the *a priori* policy dimension is as simple as -1 for the fiscally hawkish reference texts and 1 for the fiscally dovish ones[7]. Moreover, while virgin texts are the ECB press conferences subsetted as above, reference texts contain normative fiscal sentences qualitatively classified as either fiscally hawkish or fiscally dovish from ECB press conferences, speeches, monetary hearings and interviews (from 2002 to 2019)[8].

[7] Fiscally hawkish is defined a document whose content encourages governments to keep government budgets under control (sizeable reduction of government deficit and stabilization of government debt) by advocating for cuts in government spending or higher taxes or a combination of both. Fiscally dovish is defined a document whose content encourages governments to increase government deficit (and hence possibly debt) to finance sizeable government spending to sustain aggregate demand.

[8] Details on how I built reference texts are in Appendix 1.7.

For example, assume that the hawkish reference text H, with known policy position -1, is given below[9]:

> At the current juncture, [governments] are especially encouraged to press ahead with structural reform, and governments are particularly urged to avoid a pro-cyclical loosening of fiscal policies. This is particularly important in those countries where there is a risk of overheating. We urge governments, in this high growth period, not to exert pro-cyclical impulses through tax reductions which would endanger the inflation performance of the euro area economy. So, by implication, that means that we hope these extra revenues, as they come in, be they from privatisation or from cyclically higher tax revenues, will be used to the maximum extent possible to strengthen the process of consolidation of the public finances, in other words, to reduce the deficit and to reduce the burden of debt. In terms of fiscal policy, the overall pace of consolidation is disappointing against the background of the favourable outlook for growth. Consolidation targets continue to be at risk, notably in a number of countries with an excessive deficit.

and assume that the dovish reference text D, with known policy position 1, is the following:

> In view of the weakened economic outlook, [...] governments with fiscal space should be ready to act in an effective and timely manner. All governments should intensify their efforts to achieve a more growth-friendly composition of public finances. It would be very welcome to have other policies join the monetary policy in order to support the reduction of slack and to arrive at that growth potential and why not exceed it? So that would be very welcome, both in fiscal terms and in structural reform terms as well.

[9] What follows is just an example and while sentences are extracted from actual reference texts, they have not been pre-processed to favour readability.

I compute scores as follows. For instance, "governments" appear 3 times every 163 words in H and 2 times per 88 words in D. It is straightforward to calculate the probability that one is reading H or D while reading "governments" $-$ 0.6 and 0.4, respectively. Now, since the reference text H has a known policy position of -1 and the reference text D has a policy position of 1, the score of "governments" is then $0.6(-1)+(0.4)(1) = -0.2$. This procedure holds for every word in the reference texts. The out-of-sample score of the j^{th} press conference is then obtained by multiplying the score of each word in the reference text by the relative frequency of the corresponding words in the j^{th} press conference. Averaging over all words' scores results in the out-of-sample estimate for the j^{th} press conference.

Table 1.1 shows the most representative and extreme words with their relative score in the reference texts:

Table 1.1: Estimated Wordscores in the Training Model

	Hawkish		Dovish
budget balance	−0.30	aggregate demand	0.99
rebuild fiscal buffer	−0.66	fiscal space	0.98
fiscal consolidation	−0.54	government expenditure	0.47
get their house in order	−0.34	fiscal expansion	0.69
avoid any procyclical bias	−0.34	flexibility	0.63
expenditure restraint	−0.86	automatic stabiliser	0.66
budget retrenchment	−0.81	public investment	0.99
no case for fiscal activism	−0.59	public sector	0.69
ricardian effect	−0.34	fiscal union	0.41
deficit reduction	−0.86	growth friendly	0.65
debt sustainability	−0.81	ease social hardship	0.69
fiscal stability	−0.77	infrastructure	0.41

Note: The table shows the estimated wordscores of the reference texts. A negative score means that the word is a fiscally hawkish word, while a positive score indicates a fiscally dovish one.

Since the choice of reference texts involved discretionary calls, I built an external validation exercise in the attempt to minimize the degree of subjectivity in choosing the reference texts. The validation exercise consisted of an expert survey asking respondents to score 15 press conferences randomly assigned. The scoring is as simple as: -1, 0, 1, where -1 stands for fiscally hawkish, 1 for fiscally dovish and 0 for fiscally neutral. I gathered around 200 responses and each press conference received, on average, 15 scores. For further details on the external validation exercise I refer to Appendix 1.7. Figure 1.2 plots the results of the supervised model and the expert survey:

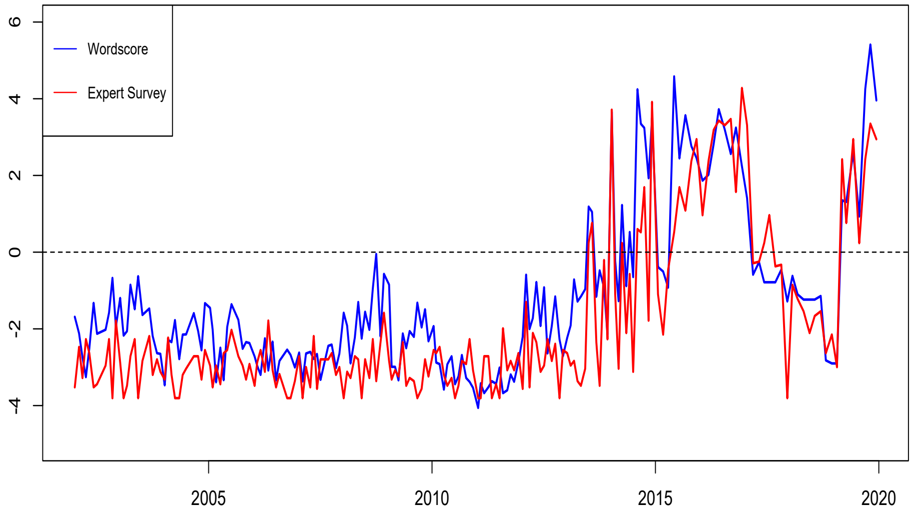


Figure 1.2: Estimated Degree of Fiscal Policy Accommodation. The figure shows the estimated DFPA for Wordscore (blue line) and the mean of the responses in the expert survey (red line). Above 0 indicates a *dovish* score, while below 0 a *hawkish* one. The results of the expert survey were standardized to make comparability easier.

The two series are significantly correlated (86%). More importantly, the signs of the series differ only in 4 cases (2% of the total observations) that, upon closer inspection, appear to be among the most contentious press conferences (see Appendix 1.7). Overall, the "fiscal stance" appears to maintain a robust hawkish bias until the end of 2014, with a mild and short-lived rebound at the peak of the Great Recession, before bouncing back in a dovish territory over the timeframe 2015-2017. Afterwards, DFPA plummets in 2018, it reaches a low in mid-2019 and then skyrockets to its all-time high through the end of 2019.

1.2.2 Identification Strategy

What are the macroeconomic effects, if any, of the ECB's normative fiscal communication? The answer to this question requires the identification of a DFPA shock which I define as a surprise change in the communicated fiscal stance of the ECB. In particular, a negative DFPA shock corresponds to an unexpected communication on contractionary fiscal policy that the ECB encourages European countries to implement.

The identification strategy is based on Bayesian Local Projections (BLP) (Miranda-Agrippino and Ricco, 2021) recursively identified with the DFPA variable ordered last[10].

Table 1.2 shows the ten variables that enter the model. Since there is no European "fiscal stance", I capture the heterogeneity of fiscal positions among member states by running BLP for three different blocks: aggregate, South and North as shown in Table 1.2[11].

Table 1.2: Variables Employed in Bayesian Local Projections

Aggregate	**South**	**North**	**Transf.**	**Symbol**
EA Shadow Rate	EA Shadow Rate	EA Shadow Rate	Level	i_t
Industrial Production	–	–	Log	y_t
HICP	–	–	Log	π_t
Government Revenue	–	–	Percent	gr_t
Government Expenditure	–	–	Percent	ge_t
Gov. Budget Balance	–	–	Percent	bb_t
Debt-to-GDP	–	–	Percent	$debt_t$
CISS	CISS	CISS	Level	$ciss_t$
EURO STOXX 50	EURO STOXX 50	EURO STOXX 50	Log	$stoxx_t$
DFPA	DFPA	DFPA	Level	$dfpa_t$

Note: The table shows the ten variables employed in the BLP. – indicates the variables that have been changed compared to the "Aggregate".The other variables remain constant across different specifications. Following Hachula et al. (2020), fiscal variables were seasonally adjusted and linearly interpolated to get monthly frequency.

The identification of a DFPA shock is challenging since the ECB is expected to make most statements about fiscal policy in response to what has happened or what is expected to happen in the economy. Therefore, it is key to minimize the probability that a DFPA

[10] I refer to Appendix 1.6 for further details.

[11] Countries in the Southern block are: Italy, Spain, Portugal, Malta, Greece and Ireland. Countries in the Northern block are: Germany, Netherlands, Finland, Slovakia, Slovenia, France, Belgium, Austria, Lithuania, Latvia and Estonia. Data are aggregated using weights based on real output in 2019. Results are robust to moving or excluding France and Ireland from their respective categories.

shock captures an ECB statement about what the market has already priced in. For this reason, I include in the model two fast-moving market variables sensitive to expectations about future fiscal policy: the Composite Index of Systemic Stress (CISS) (Kremer et al., 2012) and the index of European stocks (EURO STOXX 50)[12]. Moreover, as the time span of the analysis includes zero lower bound (ZLB) observations, shadow rates for the euro area (Wu and Xia, 2016) are preferred to the Euro OverNight Index Average (EONIA) rates to trace monetary policy stance.

Formally, the identification strategy goes as follows. Assume y_{t+1} to be the (10×1) vector of macroeconomic variables outlined and ordered as in Table 1.2. Local projections (LP) are a set of linear regressions estimated independently at each horizon and take the following form:

$$y_{t+1} = B^{(h)} y_t + \varepsilon_{t+h}^{(h)}, \qquad \varepsilon_{t+h}^{(h)} \sim N(0, \textstyle\sum_{\varepsilon}^{(h)}) \qquad \forall \ h = 1, \ldots, H \tag{1.4}$$

where $B^{(h)}$ is the n-dimensional matrix of coefficients and $\varepsilon_{t+h}^{(h)}$ is a (10×1) vector of reduced-form innovations or one-step-ahead forecast errors serially correlated and heteroscedastic (Jordà, 2005). The horizon-h impulse response functions from LP are given by:

$$IRF_h = B^{(h)} A_0 \tag{1.5}$$

where A_0 identifies the mapping between the structural shocks u_t and the reduced-form one-step-ahead forecast errors, i.e. $\varepsilon_t = A_0 u_t$.

The challenge is now to identify A_0 in order to retrieve a structural DFPA shock (u_t^{dfpa})[13]. I propose to improve on the recursive identification strategy which is standard

[12] The inclusion of fast-moving market variables also addresses the concern that a DFPA shock may instead be reflecting the restrictive domestic fiscal stance driven by market pressure or even some announcements made by the European Commission over fiscal policy. In fact, by construction, a DFPA is exogenous to what has already priced in by the market. Results are consistent using different measures of government bond spreads vis-à-vis Germany instead of the CISS across the three blocks. Similarly, results are robust to using domestic indexes (or a weighted average) of stock markets in South and North specifications, rather than keeping EURO STOXX 50 constant.

[13] The remaining shocks are identified from their respective equations, but left uninterpreted.

for the identification of monetary policy shocks (Christiano et al., 1999; Christiano et al., 2005). With the DFPA variable ordered last, a lower triangular Cholesky decomposition of A_0 would be interpreted as follows: macroeconomic variables do not simultaneously react to DFPA, while a simultaneous reaction from the macroeconomic environment to DFPA is accounted for. In other words, the ECB's DFPA reaction function responds to any variable in Table 1.2 placed before $dfpa_t$, whereas a DFPA surprise propagates through those variables with a lag (one month). While it is reasonable to assume that macroeconomic variables respond with a lag to DFPA shocks, such an identification strategy rules out a potentially important transmission channel between DFPA, bond markets and stock prices (Bjørnland and Leitemo, 2009). I therefore impose the alternative identifying restriction that a DFPA shock can have an immediate impact on bond markets (CISS) and stock prices (EURO STOXX 50). This is obtained by imposing only seven zero restrictions on the relevant coefficients in the tenth column in the A_0 matrix:

$$\begin{bmatrix} \varepsilon^{i} \\ \varepsilon^{y} \\ \varepsilon^{\pi} \\ \varepsilon^{gr} \\ \varepsilon^{ge} \\ \varepsilon^{bb} \\ \varepsilon^{debt} \\ \varepsilon^{ciss} \\ \varepsilon^{stoxx} \\ \varepsilon^{dfpa} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & 0 & 0 & 0 & 0 & 0 \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} & 0 & 0 & 0 & 0 \\ a_{71} & a_{72} & a_{73} & a_{74} & a_{75} & a_{76} & a_{77} & 0 & 0 & 0 \\ a_{81} & a_{82} & a_{83} & a_{84} & a_{85} & a_{86} & a_{87} & a_{88} & a_{89} & a_{8,10} \\ a_{91} & a_{92} & a_{93} & a_{94} & a_{95} & a_{96} & a_{97} & a_{98} & a_{99} & a_{9,10} \\ a_{10,1} & a_{10,2} & a_{10,3} & a_{10,4} & a_{10,5} & a_{10,6} & a_{10,7} & a_{10,8} & a_{10,9} & a_{10,10} \end{bmatrix} \begin{bmatrix} u_t^{i} \\ u_t^{y} \\ u_t^{\pi} \\ u_t^{gr} \\ u_t^{ge} \\ u_t^{bb} \\ u_t^{debt} \\ u_t^{ciss} \\ u_t^{stoxx} \\ u_t^{dfpa} \end{bmatrix} \tag{1.6}$$

Equation 1.6, however, calls for two additional restrictions to be imposed. Drawing on Bjørnland and Leitemo, 2009), I then assume that DFPA shocks have no long-run effects on stock prices and bond markets in the euro area. Denoting $B^{(h)}A_0 = C(L)$, this is taken care by setting the infinite number of relevant lag coefficients, $\sum_{j=0}^{\infty} C_{8,10,j}$ and

$\sum_{j=0}^{\infty} C_{9,10,j}$, equal to zero[14].

Despite the attempt of avoiding endogeneity by introducing high-frequency and forward-looking market variables, the model may still suffer from endogeneity. In fact, the fiscal variables that enter the model (e.g. government spending, tax revenues, budget deficit and debt-to-GDP) are *de facto* endogenous to the business cycle. The model would therefore benefit from the introduction of a measure of discretionary fiscal policies (e.g. change in primary structural balance) that is orthogonal to the business cycle. It follows that this unresolved endogeneity issue leads to the problem that causality statements about how the ECB fiscal pronouncements affect macroeconomic variables must be taken with a degree of caution.

To sum up, recursive Cholesky restrictions identify the non-zero parameters above the DFPA equation, whereas the remaining parameters are uniquely identified from long-run restrictions. That is, I assume that DFPA shocks have no immediate impact on macroeconomic variables, while having a contemporaneous effect on CISS and EURO STOXX 50. Moreover, the "fiscal stance" of the ECB is responsive to all macroeconomic and financial variables listed in Table 1.2 before $dfpa_t$. That said, endogeneity issues might still persist due to the absence of a variable that controls for discretionary fiscal policies. Therefore, the "causal" interpretation of the results must be taken with caution.

1.2.3 Historical Decomposition

Next, I propose to study how much DFPA shocks affected the historically observed developments of output and inflation. To do so, I build on the model above and use historical decomposition (HD) as counterfactual similarly to Kilian and Lee (2014). HD can in fact be computed from covariance stationary VAR models in three steps: first, computing the structural MA coefficient matrices $\Theta_0, \ldots, \Theta_{T-1}$; second, identifying the structural shocks $w_t = A_0 u_t, t = 1, \ldots, T$ and, finally, matching up each shock j with the appropriate impulse weight to form $T \times 1$ vectors of fitted values for variable k, denoted $\hat{y}_t^{(j)} = \sum_{s=0}^{t-1} \Theta_s w_{t-s}$.

[14] It follows that the stock price shock and the CISS shock are differentiated from the DFPA shock by allowing them to have a long-run impact on stock prices and bond markets.

Therefore, constructing the counterfactual is as simple as:

$$y_{kt} - \hat{y}_{kt}^{(j)}, \tag{1.7}$$

where y_{kt} denotes the k^{th} actual time series variable and $\hat{y}_{kt}^{(j)}$ is the cumulative contribution of shock j to the evolution of variable k up to date t. The counterfactual series indicates how the variable of interest would have evolved had one replaced all realizations of shock j by zero while preserving the remaining structural shocks in the model. In the current specification, the goal is to investigate the development of output and inflation after having turned off DFPA surprises.

1.3. Macroeconomic Effects of DFPA Surprises

Section 1.3.1 illustrates the effects of a negative DFPA shock in the euro area, while Section 1.3.2 documents what would have been the evolution of inflation and output had the ECB maintained a neutral fiscal communication.

1.3.1 Impulse Responses to a Hawkish DFPA Shock

The model is estimated using monthly data from 2002:01 to 2019:12. Impulse responses are produced for the aggregate, North and South blocks. The optimal number of lag is set to ten for the three specifications following Akaike Information Criterion (AIC). Using ten lags, the model satisfies the stability condition (no eigenvalues, i.e., inverse roots, of the autoregressive characteristics polynomial lie outside the unit circle) and basic diagnostic tests (i.e., there is no evidence of autocorrelation or heteroscedasticity in the model residuals). All impulse responses are for a standard deviation decrease in DFPA, that is, an exogenous communicated hawkish fiscal stance[15]. Endogeneity concerns outlined above need to be taken into account when interpreting the following results.

Figure 1.3 displays the results for the euro area, while Figure 1.4 shows the results for

[15] Prior selection is VAR-based (VAR with 10 lags) on all data points available. I refer to Appendix 1.6 for further details.

the Northern (solid blue line) and Southern block (solid red line). All impulse responses are for a standard deviation decrease in DFPA, that is, a communicated hawkish fiscal stance. 90% posterior confidence bands are provided.

Focusing on the aggregate response, there are three results to highlight. First, a hawkish DFPA shock significantly lowers industrial production and inflation. While HICP remains negative throughout the forecast horizon, industrial production hits a low after a year and a half from the shock and then recovers from the twentieth month onward. Second, the European stock market welcomes on impact the recommendation of a fiscal consolidation before gradually reversing the gains a year after the shock. Instead, CISS displays a neat hump-shaped response, that is, systemic risk is tilted upward after the verbal intervention of the ECB as markets might be discounting fiscal issues. Third, fiscal variables suggest that, overall, the euro area adjusts its fiscal position to follow the recommendations of the ECB. In fact, on the one hand, the hump-shaped response of government revenues indicates that governments in the euro area raised taxes and, on the other hand, the decline in government expenditure for around 15 months denotes that member states cut government spending. As a result, government budget balance improves and debt-to-GDP drops[16].

[16] The response of shadow rates is also noteworthy: a hawkish fiscal communication pushes rates gradually up to 0.2%. Further research is however warranted before drawing conclusions.

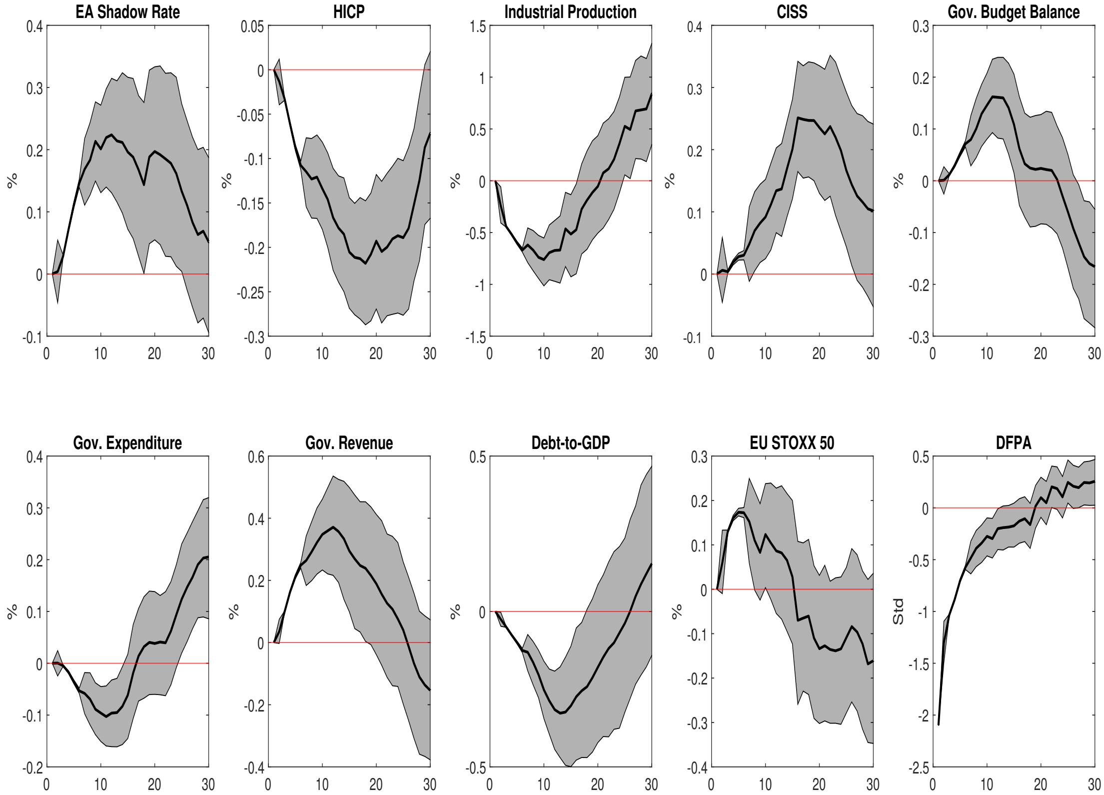


Figure 1.3: Impulse Responses to a DFPA Shock for the euro area as a whole. Shaded areas are 90% posterior coverage bands.

While the results for the macro and financial variables are consistent across blocks[17], fiscal variables display some variation. Notably, Figure 1.4 shows that Southern and Northern governments, rather than cutting, increase government spending by 0.2% after one year from the shock. However, while Southern countries compensate higher expenditure with higher taxes, Northern countries maintain a mildly expansionary fiscal stance regardless of the recommendations of the ECB. Such a divergent response may be explained by structural initial positions: while Northern countries can afford to ignore the ECB due to their sounder fiscal positions, peripheral countries cannot dare to do so.

[17] One interesting exception is the response of the European stock market that declines much faster in the Northern case and it is weakly positive only for a few months. This might indicate that while the equity market welcomes the possibility of a fiscal consolidation in the South, it considers harmful a restrictive fiscal stance in the North.

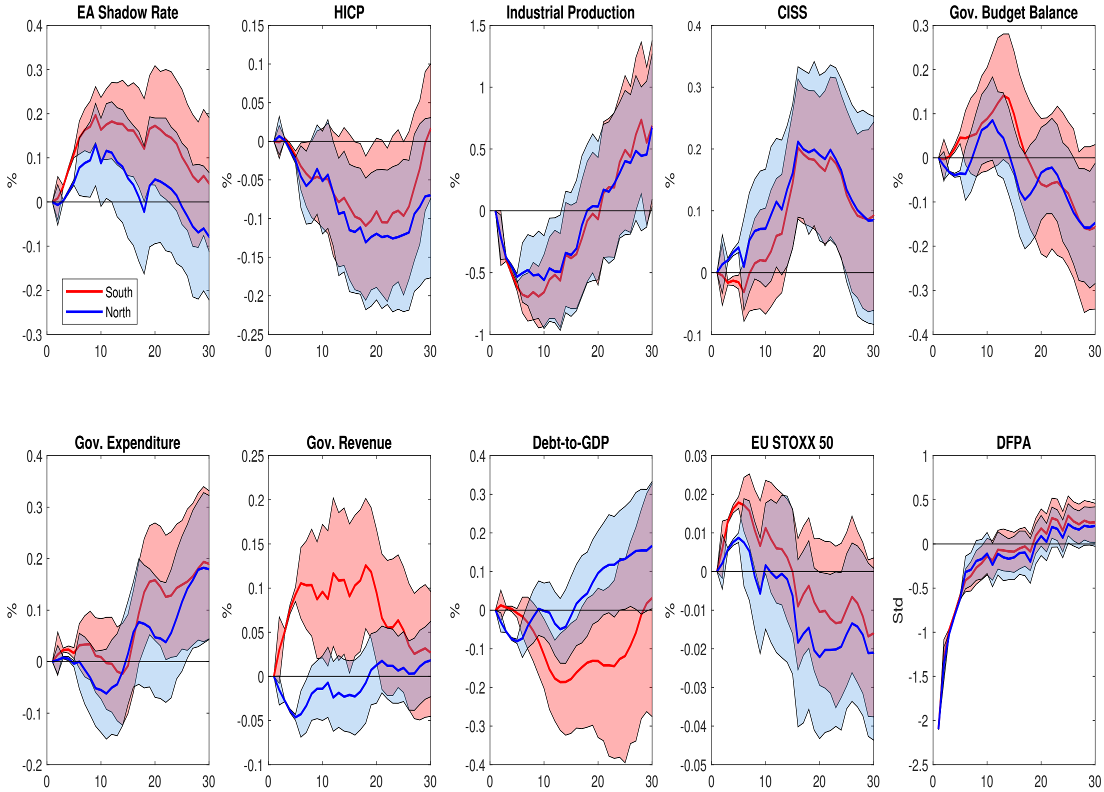


Figure 1.4: Impulse Responses to a DFPA shock for the North (solid blue line) and South (solid red line). Shaded areas are 90% posterior coverage bands.

The heterogeneity among European countries is not however the only confounding factor. The estimation time span (2002-2019) contains a very diverse period when it comes to monetary policy. To control for such a regime change, I employed euro are shadow rates instead of EONIA ones. Yet, it is of interest to further investigate how a negative DFPA shock propagates conditional on the level of interest rates. Exploiting the flexibility of local projections, I introduce non-linearities in Equation 1.4 as in Tenreyro and Thwaites (2016). Figure 1.5 illustrates the impulse responses for the euro area, South and North in a regime of conventional monetary policy (solid red line) and ZLB (solid blue line)[18].

[18] ZLB is identified for $i_t \leq 0.5$, while a regime of conventional policy is set by $i_t > 0.5$. See Appendix 1.6.

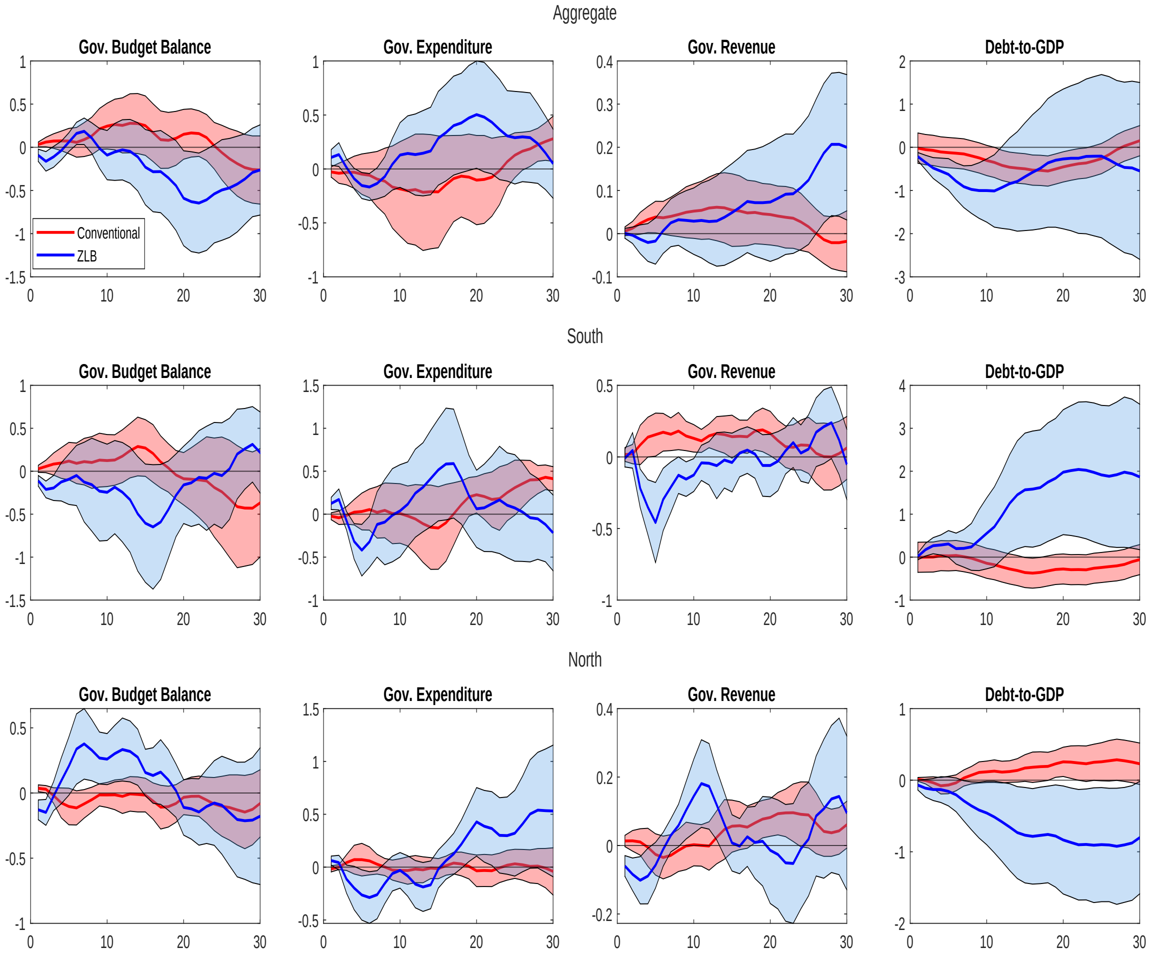


Figure 1.5: Non-linear Impulse Responses to a DFPA Shock for aggregate, North and South. The solid blue lines indicate the impulse responses to a DFPA shock in a regime of conventional monetary policy, while the solid red lines denote the impulse responses to a DFPA shocks at the ZLB. Every impulse response indicates the posterior median with its 90% confidence intervals.

Figure 1.5 shows the heterogeneity between the regimes under investigation. To begin with the aggregate block, the regime of conventional monetary policy displays a familiar pattern: assuming endogeneity away, a DFPA shock might be one of the factors that drive countries to reduce expenditure and raise taxes, thereby improving the budget balance. Instead, when rates are at the ZLB government spending increases more than the tax increase, leading to a deterioration of the budget balance. In the southern block, when rates are high, member states hardly change government expenditure for the first twenty months while they increase taxes by around 0.25%. At the ZLB, southern countries look through the ECB's fiscal reprimands and fail to enact any fiscal consolidation. Similarly, in the North, when rates are high the adjustment comes from the tax side, while at the ZLB higher taxes are compensated by higher government spending. Once again, it is worth recalling that endogeneity issues might still bias the results and overstate the nature of the "causal" nexus.

To sum up, a hawkish DFPA surprise might have contractionary effects on inflation and output. On balance, relaxing endogeneity issues, the fiscal stance of the euro area may react to the fiscal recommendations of the ECB by reducing expenditures and raising taxes. However, this response varies across blocks and conditional on the level of interest rates. On the one hand, southern countries appear to be more responsive than northern ones to the ECB's fiscal reprimands. On the other hand, both North and South block are more responsive when their fiscal capacity is constrained by an environment of higher rates. Instead, at the ZLB, countries are more prone to look through the fiscal recommendations of the ECB.

1.3.2 Cumulative Effect of DFPA Shocks

Based on the the structural shocks identified above, Figure 1.6 provides the historical counterfactual of output and inflation. In the following interpretation, I relax the endogeneity issues underlined above.

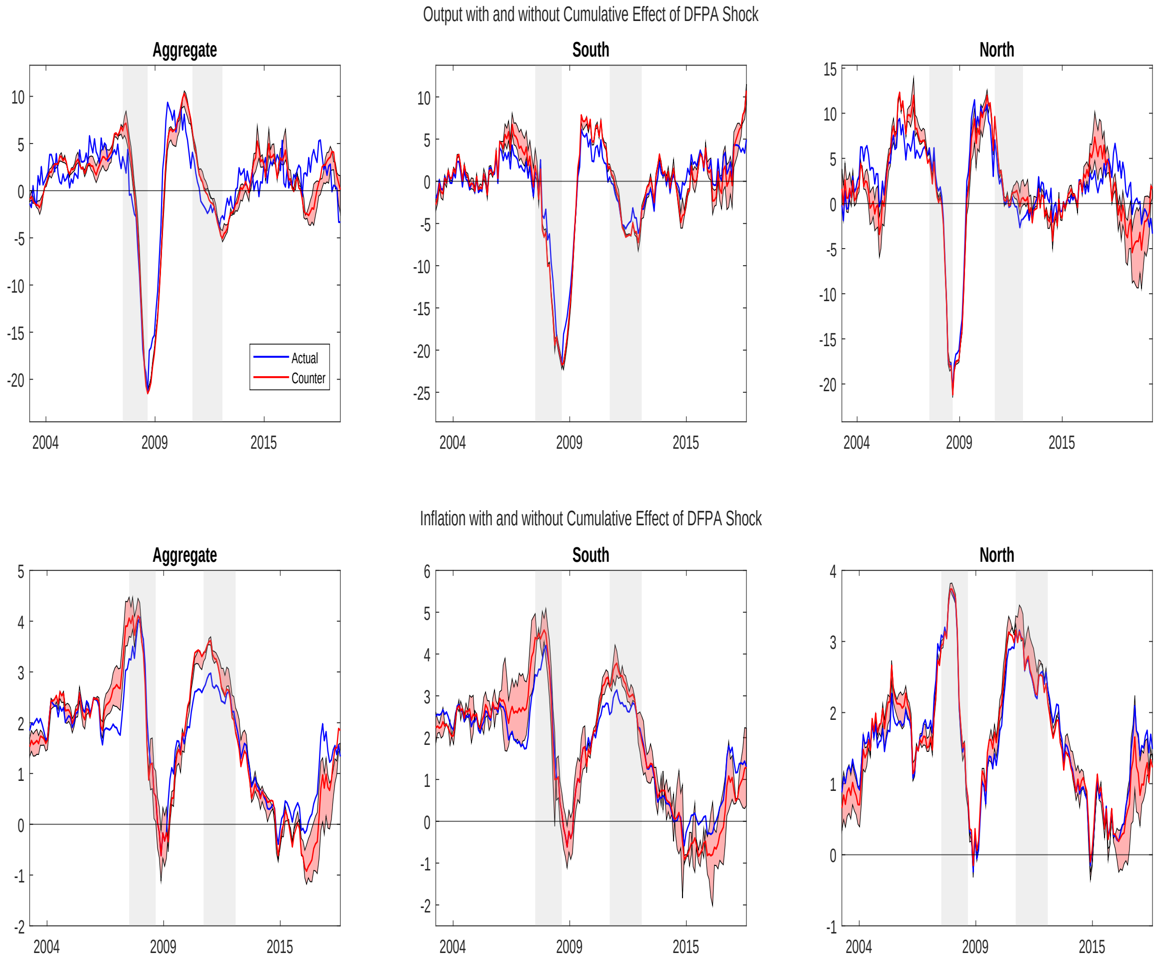


Figure 1.6: Historical Counterfactual. The figure shows output (proxied by industrial production) and inflation from January 2002 to December 2019 with and without the cumulative effect of the DFPA shock. I also display 90% confidence bands for the counterfactual series. Results are for year-on-year growth rate at monthly frequency.

The counterfactual exercise shows how output and inflation would have evolved had one been able to replace all realizations of the DFPA shock by zero while preserving the other shocks in the model. If the counterfactual exceeds the actual series, this means that the DFPA shock lowered the actual series and vice versa. The first row, focusing on output, highlights that industrial production, in the absence of DFPA surprises, would have been significantly higher at the start of the Great Recession as well as during the period of the sovereign-debt crisis. In contrast, throughout 2017-2018, DFPA shocks appear to have stimulated output in the euro area. This patter varies across the fiscal blocks under investigation: the South would have mainly benefited from a different fiscal stance right before the sovereign-debt crisis. Instead, the North, while not being affected during the crises, could have aimed for higher output growth from 2005 to 2006 − when the ECB was highly concerned with the fiscal discipline in the euro area. In 2017, instead, an expansionary fiscal communication helped Northern industrial production edge higher.

Turning to inflation (second row), in the absence of a DFPA shock, HICP would have been higher during the period preceding the Great Recession and throughout the sovereign-debt crisis. The contractionary effects of DFPA surprises appear to be much more prominent for the South than the North where the difference between the actual and counterfactual series is never statistically significant.

To sum up, had the ECB held a neutral fiscal stance, output and inflation would have been higher during the two major crises. This finding is consistent in the South. Instead, in the North, while inflation is never affected by DFPA shocks, output could have been higher in the run-up to the Great Recession and lower in 2017.

1.4. Government Bond Surprises and DFPA

Figure 1.5 suggests that accommodative monetary policy results in weakened fiscal discipline. Reduced fiscal discipline might, in turn, expose governments to debt servicing and refinancing issues once interest rates rise because of monetary policy normalisation. Such a scenario could potentially pose a threat to monetary dominance as the ECB may face a trade-off between price stability and debt stabilisation.

Building on this result, I propose to study how the DFPA variable responds to bond-buying shocks. I first identify three unconventional monetary policy shocks from the *Euro Area Monetary Policy Event-Study Database* (EA-MPD) by Altavilla et al. (2019): sovereign-debt crisis (SDC) shocks, Quantitative Easing (QE) shocks and Large-Scale Bond Purchases (LSBP)[19]. I then set up a monthly Factor-Augmented VAR (Bernanke et al., 2005) identified with these instruments. The sample consists of monthly observa-

[19] I follow Wright (2019)'s comments on Altavilla et al. (2019) to improve on the extraction of the monetary policy factors. Details on the derivation of monetary policy shocks can be found in Appendix 1.8.

tions for the period 2002:1-2019:12[20]. Based on the AIC, the lag length in the FAVAR is set to four. All impulse responses are for a 0.25% decrease in the policy indicator and I show results for a forecast horizon of 30 months. I report 90% confidence intervals computed using a recursive wild bootstrap using 10,000 replications.

Figure 1.7 displays the response of the DFPA variable to an expansionary SDC, QE and LSBP shock. The results show that the DFPA plunges between 1.2% and 1.6% on impact and remains significantly negative for twelve months after the shock. The ECB, therefore, calls on governments to constrain fiscal policies after an unconventional monetary easing. Notably, as showed in Appendix 1.9, the significantly hawkish response is specific to a government-purchase shock. Neither a conventional monetary policy shock nor a forward guidance surprise is, in fact, able to trigger such a reaction.

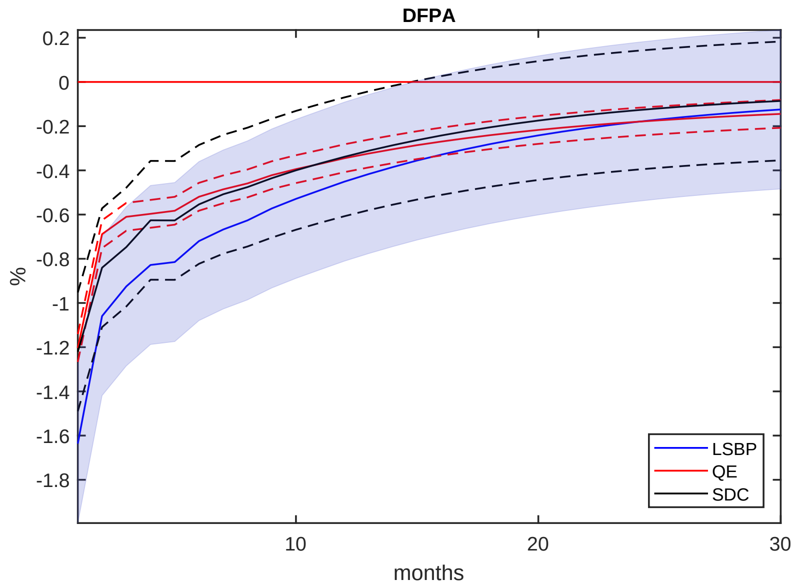


Figure 1.7: Impulse Response to an Expansionary SDC, QE and LSBP Shock. The figure shows the response to a 0.25% decrease in the policy indicator due to SDC (black line), QE (red line) and LSBP (blue line) shocks. 90% bands are displayed.

To further inspect the rationale underpinning this result, I report the behavior of some

[20] Since the instruments are available for a shorter period, I first use the full sample to estimate the lag coefficients and obtain the reduced form residuals. Then, I use the factors and residuals for the corresponding period to identify unconventional monetary policy surprises as in Gertler and Karadi (2015).

key fiscal variables given the shocks under investigation. Figure 1.8 displays the results.

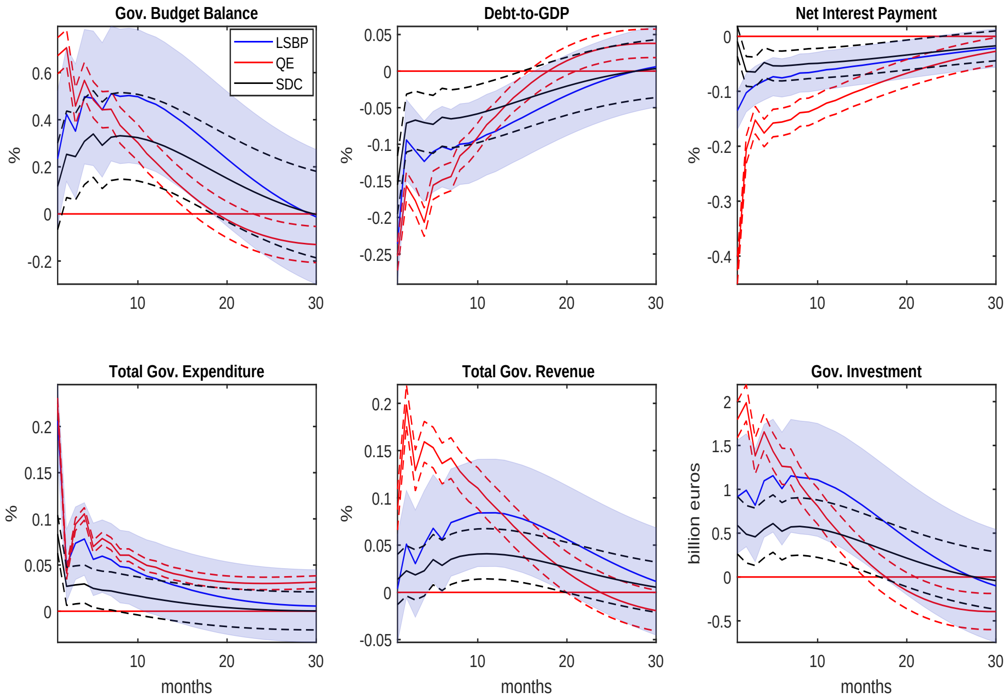


Figure 1.8: Fiscal Variables. Fiscal responses at the euro area level. The figure shows the response to a 0.25% decrease in the policy indicator due to a SDC (black line), QE (red line) and LSBP (blue line) shock. 90% bands are displayed.

The figure shows that government bond purchases have a sizeable windfall for EA countries, that is, government budget improves, debt-to-GDP and net interest payments significantly decline and government revenue increases. Besides, governments internalize the positive fiscal effects of monetary measures and increase government spending and investment, thereby proving, as in Hachula et al. (2020), that the response of government spending behaves countercyclically[21].

The results can thus shed some light on the interface between government bond purchases and DFPA. On the one hand, a protracted period of low rates reduce the fiscal discipline of member states. On the other hand, government bond purchases might drive the fiscal stance of the ECB into a hawkish territory. Taken together these results might suggest that the ECB leans toward a more prudent fiscal approach since it is concerned

[21] I refer to Appendix 1.9 for comprehensive results on the distinction between South and North blocks.

about the solvency of member states once it has become their main creditor. In other words, the more the ECB purchases government bonds, the more it becomes a stakeholder in the sustainability of fiscal positions in the euro area. This leads the fiscal preferences of the ECB towards recommending fiscal discipline.

1.5. Conclusion

This paper investigates the macroeconomic effects of the ECB fiscal policy communication as well as its interaction with government bond purchases. In particular, I run BLP recursively identified with the DFPA variable ordered last to assess the macroeconomic impact of DFPA surprises in the euro area. I also carry out historical decomposition to study whether output and inflation would have been higher had the ECB held constant its "fiscal stance". Lastly, using a FAVAR identified with external instruments, I study how government bond purchases shocks propagate through the degree of fiscal policy accommodation of the ECB. Although the econometric identification of the models attempts to take care of endogeneity, endogeneity may still persist due to the lack of a measure of discretionary fiscal policy in the models proposed. Hence, the "interpretation" of the results must be mindful.

The paper provides novel results. First, I document the existence of a quantifiable ECB "fiscal stance". Second, a fiscally hawkish DFPA shock may be one of the factors contributing to a significant downward effect on output and inflation. Moreover, the counterfactual simulation shows that output and inflation would have been higher had the ECB held its "fiscal stance" constant. These contractionary effects are more prominent in the South than in the North and are mainly related to the Great Recession and the sovereign-debt crisis. Third, the fiscal recommendations of the ECB tend to be more effective in a regime of conventional monetary policy than at the ZLB. Fourth, there appears to be a statistically significant inverse relationship between purchases of government bonds and DFPA: the ECB may require fiscal consolidation to come with the purchase of government bonds. This finding can be explained by the fact that the ECB, having become the main creditor of the governments in the euro area, is more exposed to solvency concerns.

1.6. APPENDIX A: Bayesian Local Projections

I first give more detail about Bayesian Local Projections (BLP) and then present sensitivity analysis.

The Model

In this section I outline the gist of the Bayesian approach to local projections following Miranda-Agrippino and Ricco (2021). For a more comprehensive treatment of the method I refer to Miranda-Agrippino and Ricco (2015).

The reason to prefer BLP over VAR or local projections (Jordà, 2005) is that BLP improves on both frameworks by retaining the efficiency of VAR models while being robust to misspecifications. Let's try to understand how and why it is so. VARs recover impulse responses by iterating up to the relevant horizon the coefficients of a system of one-step ahead reduced-form equations:

$$y_{t+1} = By_t + \varepsilon_{t+h}, \qquad \varepsilon_{t+h} \sim N(0, \textstyle\sum_{\varepsilon}) \tag{1.8}$$

On the other hand, LPs estimate the IRFs directly from a set of linear regressions of the form:

$$y_{t+1} = B^{(h)}y_t + \varepsilon^{(h)}_{t+h}, \qquad \varepsilon^{(h)}_{t+h} \sim N(0, \textstyle\sum^{(h)}_{\varepsilon}) \qquad \forall\ h = 1, \ldots, H \tag{1.9}$$

Assuming the VAR to be the true description of the data generating process (DGP), the coefficients and residuals of an iterated VAR can then be mapped into those of LP:

$$B^{(h)} \leftrightarrow B^{(VAR,h)} = B^h \tag{1.10}$$

$$\varepsilon^{(h)}_{t+h} \leftrightarrow \varepsilon^{(VAR,h)}_{t+h} = \textstyle\sum_{j=1}^{h} B^{h-j}\varepsilon_{t+h} \tag{1.11}$$

The map in Equation 1.10 and Equation 1.11 is exploited to inform the priors for the BLP. For the coefficients of Equation 1.9 at each horizon h, and ignoring for the moment the structure of the projection residuals, it is possible to specify standard conjugate

Normal-inverse Wishart informative priors of the form:

$$\textstyle\sum_{\varepsilon}^{h} |\lambda^{h} \sim IW(\psi_0^{(h)}, d_0^{(h)}) \tag{1.12}$$

$$\beta^{(h)} | \textstyle\sum_{\varepsilon}^{(h)}, \lambda^{(h)} \sim (\beta_0^{(h)}, \textstyle\sum_{\varepsilon}^{(h)} \otimes \Omega_0^{(h)}(\lambda^{(h)})) \tag{1.13}$$

where $\beta^{(h)}$ is the vector containing all the local projection coefficients at horizon h. $\lambda^{(h)}$ is the hyperparameter that regulates the variance of the coefficients in $\beta^{(h)}$, and thus effectively determines the overall tightness of the priors. The prior mean $\beta_0^{(h)}$ is informed by the iterated coefficients of a similarly specified VAR estimated over a pre-sample. For the model in Equation 1.8 this writes

$$\beta_0^{(h)} = vec(\beta_{T_0}^{h}) \tag{1.14}$$

where $B_{T_0}^{h}$ is the h^{th} power of the autoregressive coefficients estimated over a pre-sample T_0, that is then discarded. As it is standard in Bayesian econometrics modelling, $d_0^{(h)}$ is fixed to equal to the number of variables minus 2, such that the prior mean of $\sum_{\varepsilon}^{(h)}$ exists and the remaining hyperparameters in $\psi_0^{(h)}$ and $\Omega_0^{(h)}$ using sample information (see e.g. Kadiyala and Karlsson, 1997; Banbura et al., 2010).

The posterior distribution for the BLP coefficients can then be obtained by combining the priors in Equation 1.12 and Equation 1.13 with the likelihood of the data conditional on the parameters, where the autocorrelation of the projection residuals is not taken into account. This modelling choice has three implications. First, the priors are conjugate, hence the posterior distribution is of the same Normal inverse-Wishart family as the prior probability distribution. Second, the Kronecker structure of the standard macroeconomic priors is preserved. However, there is a third implication that outweighs the first two: the shape of the true likelihood is asymptotically Gaussian and centred at the Maximum Likelihood Estimator (MLE), but has a larger variance than the misspecified posterior distribution. This implies that if one were to draw inference about $\beta^{(h)}$ – i.e. the horizon-h responses – from the misspecified posterior distribution, one would be underestimating the variance albeit correctly capturing the mean of the distribution of the regression coefficients.

Müller (2013) shows that posterior beliefs constructed from a misspecified likelihood such as the one discussed here are "unreasonable", in the sense that they lead to inadmissible decisions about the pseudo-true values, and proposes a superior mode of inference – i.e. of asymptotically uniformly lower risk –, based on artificial "sandwich" posteriors. Hence, similarly to the frequentist practice, inference about $\beta^{(h)}$ is conducted by replacing the original posterior with an artificial Gaussian posterior centred at the MLE but with a HAC-corrected covariance matrix:

$$\textstyle\sum_{\varepsilon,HAC}^{(h)} |\lambda^{(h)}, y \sim IW(\psi_{HAC}^{(h)}, d) \tag{1.15}$$

$$\beta^{(h)}| \textstyle\sum_{\varepsilon,HAC}^{(h)}, \lambda^{(h)} \sim N(\tilde{\beta}^{(h)}, \textstyle\sum_{\varepsilon,HAC}^{(h)} \otimes \omega^{(h)}) \tag{1.16}$$

This allows to remain agnostic about the source of model misspecification as in Jordà (2005). It is important to observe that BLP IRFs have been engineered to span the space between VARs and local projections. To see this, note that given the prior in Equation 1.12 and Equation 1.13, the posterior mean of BLP responses takes the form

$$\beta_{BLP}^{(h)} \propto (X'X + (\Omega_0^{(h)}(\lambda^{(h)}))^{-1})^{-1}((X'X)B_{LP} + (\Omega_0^{(h)}(\lambda^{(h)}))^{-1}B_{VAR}^{h}) \tag{1.17}$$

where $X \equiv (x_{h+2}, \ldots, x_T)'$, and $x_t \equiv (1, y_{t-h}, \ldots, y_{t-(h+1)})$, while the posterior variance of BLP coefficients is equal to

$$Var(B_{BLP}^{(h)}) = \textstyle\sum_{\varepsilon,HAC}^{(h)} \otimes (X'X + (\Omega_0^{(h)}(\lambda^{(h)}))^{-1})^{-1} \tag{1.18}$$

At each horizon h, the relative weight of VAR and LP responses is set by $\Omega_0^{(h)}(\lambda^{(h)})$ and is a function of the overall level of informativeness of the prior $\lambda^{(h)}$. When $\lambda^{(h)} \to 0$, BLP IRFs collapse into VAR IRFs (estimated over T_0). Conversely, if $\lambda^{(h)} \to \infty$ BLP IRFs coincide with those implied by standard LP.

The formulation of BLP therefore allows to address the bias-variance trade-off by estimating IRFs that are an optimal combination of LP and VAR-based IRFs at each horizon. In fact, extending the argument in Giannone et al. (2019), $\lambda^{(h)}$ is treated as an additional model parameter for which a Gamma prior probability distribution is specified and estimated at each horizon as the maximiser of the posterior likelihood in the spirit of hierarchical modelling. This allows to effectively balance bias and estimation variance at all horizons, and therefore solve the trade-off in a fully data driven way. Since the marginal likelihood is available in closed-form, $\lambda^{(h)}$ can be estimated at each horizon as the maximiser of the posterior likelihood (see Giannone et al., 2019, for details). This step requires a numerical optimisation. Conditional on a value of $\lambda^{(h)}$, the BLP coefficients $[\beta^{(h)}, \sum_{\varepsilon,HAC}^{(h)}]$ can then be drawn from their posterior, which is Normal-inverse-Wishart with HAC-corrected posterior scale.

Another advantage of LP is their flexibility to be adapted to non-linear specifications. Therefore, to study the transmission of DFPA shocks conditional on the level of interest rates, I build on Auerbach and Gorodnichenko (2012), Ramey and Zubairy (2014) and Tenreyro and Thwaites (2016) and augment Equation 1.9 with the logistic function $F(z_t)$ as follows:

$$y_{t+1} = F(z_t)(B^{(h,conv)}y_t) + (1 - F(z_t))(B^{(h,ZLB)}) + \varepsilon_{t+h}^{(h)} \tag{1.19}$$

where $F(z_t)$ is a smooth increasing function of the indicator z_t (shadow rates) defined as follows:

$$F(z_t) = \frac{exp(\theta\frac{z_t - c}{\sigma_z}))}{1+exp(\theta\frac{(z_t - c)}{\sigma_z})} \tag{1.20}$$

where c is a parameter that controls what proportion of the sample spent in either state and σ_z is the standard deviation of the state variable z. The parameter θ determines how violently the interest rate switches from conventional to ZLB when z_t changes. In the empirical specification, I set $\theta = 3$ to give an intermediate degree of intensity to the regime switching and $c = 0.45$[22]. The results are robust to these choices.

[22]Note that 45% is the proportion of ZLB observations in the sample.

Robustness Checks

I carried out three main sensitivity tests. To save space, I only display results for the euro area. Results are however consistent across blocks and are available upon request.

First, I checked the sensitivity of the model to different lag length (Figure 1.9).

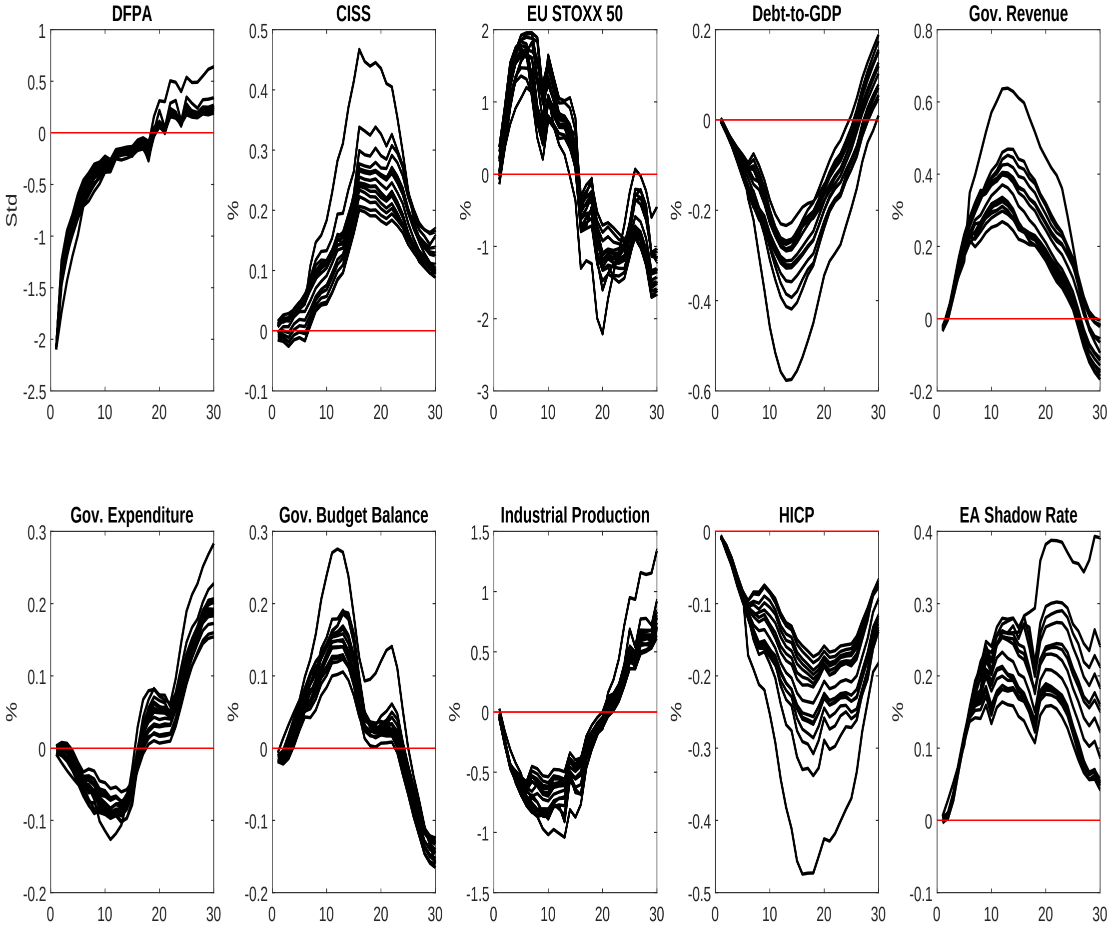


Figure 1.9:
Lag Length. The figure shows the mean impulse response for lag length ranging from 1 to 15.

Second, since I interpolated fiscal variables to monthly frequency, I run BLP at quarterly frequency (Figure 1.10).

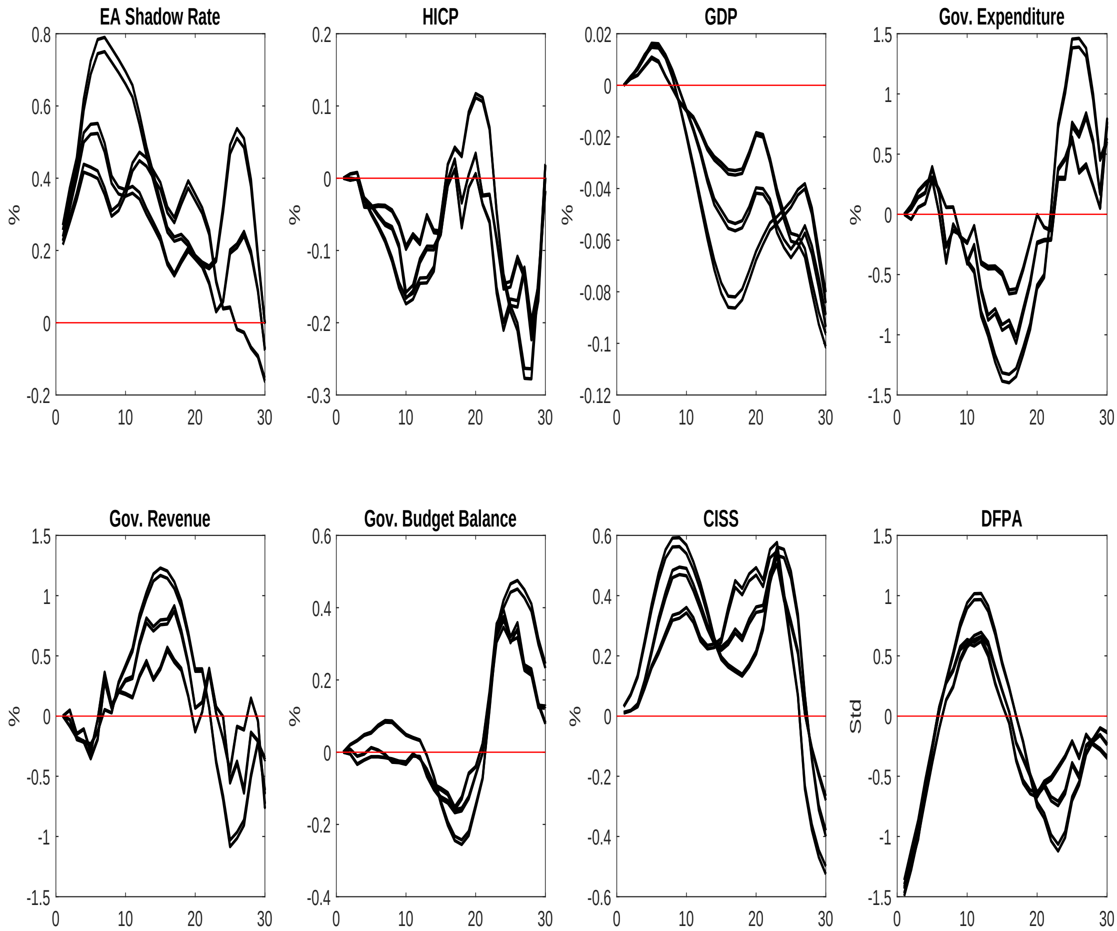


Figure 1.10: Quarterly Frequency. The figure displays mean impulse responses from 1 to 4 lags. BLP are run at quarterly frequency.

Third, I run BLP using different priors (Figure 1.11).

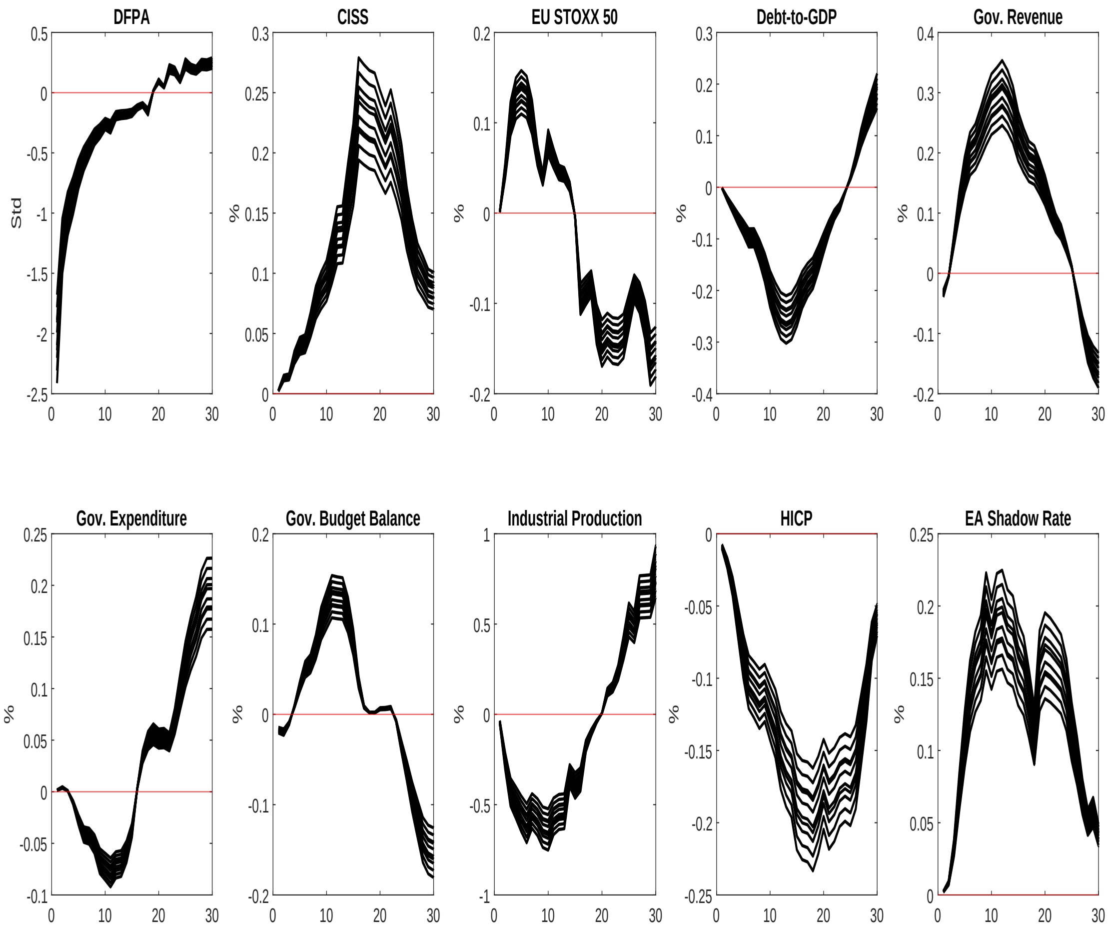


Figure 1.11: Different Priors. The figure shows mean impulse responses to different priors in BLP.

Overall, the model proves to be robust to lag length, frequency and prior specification.

1.7. APPENDIX B: Quantifying DFPA

In what follows, I detail the pre-processing decisions to clean the dataset, explain the classification model used to check the dataset I applied Wordscore to, give further detail the choice of reference texts, shed light on the external validation exercise and run some robustness checks.

Pre-processing ECB Press Conferences

The aim of this paragraph is to provide guidance on the pre-processing steps taken to cleanse the corpus of ECB press conferences. Beyond common pre-processing steps such as removing stopwords, punctuation, lemmatization, etc., I proceeded in three stages: first, despite including Q&A in the corpus, I removed journalists' questions (but kept president's and vice-president's answers) since they may inflate or deflate the score of the document. Second, I included in the corpus only fiscal sentences that are normative in nature and refer both to the euro area as a whole and to specific countries. The intuition for such a choice is that since the euro area does not have a common "fiscal stance" the time-series is constructed to reflect the cross-sectional variation among member states. Third, I check for negative statements. For instance, a statement like "countries should not consolidate" is classified as a dovish statement in the training set rather than a hawkish one.

Latent Dirichlet Allocation (LDA)

The selection of the dataset containing sentences/paragraphs on normative fiscal policy was computed manually. The following quote is an example of a normative paragraph:

> In order to support confidence, sustainable growth and employment, the Governing Council calls upon governments to restore sound fiscal positions. Commitments under the Stability and Growth Pact need to be fully honoured and weaknesses in competitiveness forcefully addressed. National policy-makers

need to fully meet their responsibilities to ensure fiscal sustainability.

To avoid any bias in the selection procedure, I apply Latent Dirichlet Allocation (LDA) (Lafferty and Blei, 2006) to retrieve the share of fiscal policy topics discussed in the ECB's press conferences. This is possible since the fundamental idea of LDA is that documents (press conferences) are represented as a distribution of latent topics, where each topic is characterized by a distribution over words. In other words, LDA assigns to every document the probability to belong to a topic. I run LDA with 10 topics as suggested by 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)[23].

Figure 1.12 shows the share of (general) fiscal policy topics (green) in the ECB's press conferences versus normative fiscal policy (yellow).

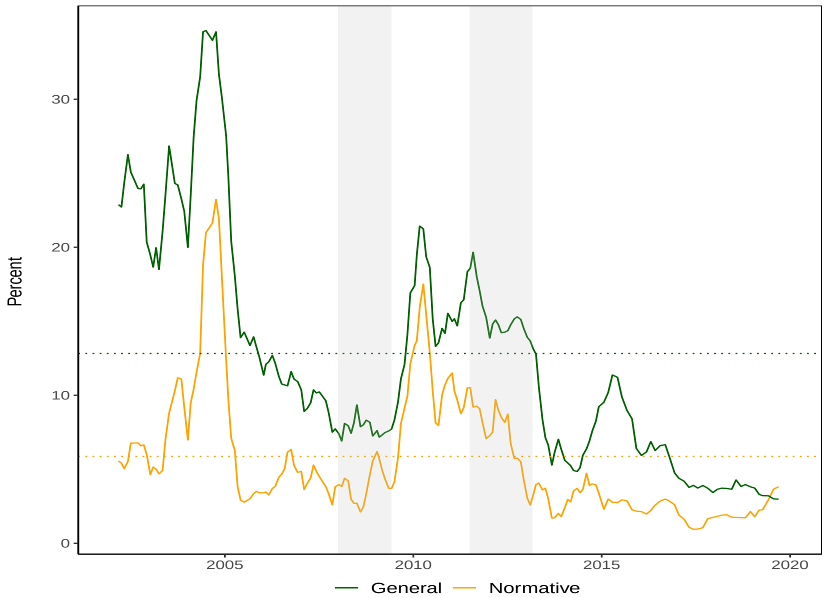


Figure 1.12: Share of Normative Fiscal Sentences in the ECB's Press Conferences. The figure shows the share of normative fiscal sentences (yellow line) and the share of fiscal sentences (green) in the ECB's press conferences. The dashed lines indicates the average from 2002 to 2019, which is around 6% for normative fiscal policy and around 13% for fiscal policy in general.

Figure 1.12 shows the high correlation between the two series (0.67). Qualitatively

[23]The "fiscal policy" topic contains the following words: "fiscal", "measures", "consolidation", "reforms", "budget", "bond", "structural", "deficit", "risks", "ecb".

comparing the text corpus derived manually with the one obtained via LDA indicates that there is a strong degree of overlapping since the former is always contained in the latter. However, the automated text also includes sentences that are fiscal but not normative in nature. Both datasets are available upon request.

Reference Texts

In choosing the reference texts I follow the guidelines outlined in Laver et al. (2003). First, reference texts should use the same lexicon, in the same context, as the virgin texts being analyzed. Second, policy positions of the reference texts should "span" the policy dimensions under study. Last, reference texts should contain as many different words as possible. While the first and third recommendations are easily met as any document used as reference texts comes from the official communication of the ECB, the second one proved to be more challenging as it requited qualitative judgment.

I proceeded in the following way: first, I collected ECB Press Conferences, Speeches, Monetary Hearings and Interviews from 2002 to 2019 and extracted only the paragraphs that referred to fiscal policy using LDA, obtaining a corpus of around 5000 documents that I, in turn, subsetted into sentences. Second, I inspected the entire dataset qualitatively classifying each sentence according to the three categories defined in the paper: fiscally hawkish, fiscally neutral and fiscally dovish. Finally, I only kept those sentences that clearly belonged to either "fiscally hawkish" or "fiscally dovish" category and, among these, I chose the ones I considered to be more neatly representative of the two policy dimensions. This strategy led me to have a total of 6133 words for the dovish reference text and 36026 for the hawkish one – such an unbalanced dataset is consistent with the results of the expert survey where the number of fiscally hawkish documents is much bigger than the number of dovish ones, respectively, 157 (83%) and 33 (17%)[24]. To make the reference texts homogeneous to the virgin text, I followed the same pre-processing steps: I removed

[24] I ran Wordscore way before the results of the expert survey. Therefore, the initial inspection was qualitative.

stopwords, currency symbols, punctuation, numbers, separators and features that, across all documents, do not occur at least ten times; I also created bigrams and trigrams and lemmatized the corpus to better score economic words (i.e. "fiscal consolidation" instead of "fiscal" and "consolidation"). When the corpus of reference texts was transformed into a dfm, it resulted in 2,207 features (33.7% sparse); instead the dfm of the virgin text contained 2,266 features (96.0% sparse). The words that were present in the virgin text and absent in the reference text were carefully checked and deemed insignificant.

To deal with the unbalanced nature of the corpus of reference texts, I set the smoothing parameter equal to 0. In fact, if I smoothed by 1 as in Lafferty and Blei (2006), I would attach a high weight to the terms associated with the smaller training class (the fiscally dovish class). This would result in an inappropriate scoring since although dovish terms may occur more often in the hawkish texts, they have higher relative frequency in the shorter dovish text. To show the plausibility of my choice, one can consider Figure 1.13 that shows the predicted document position of reference texts:

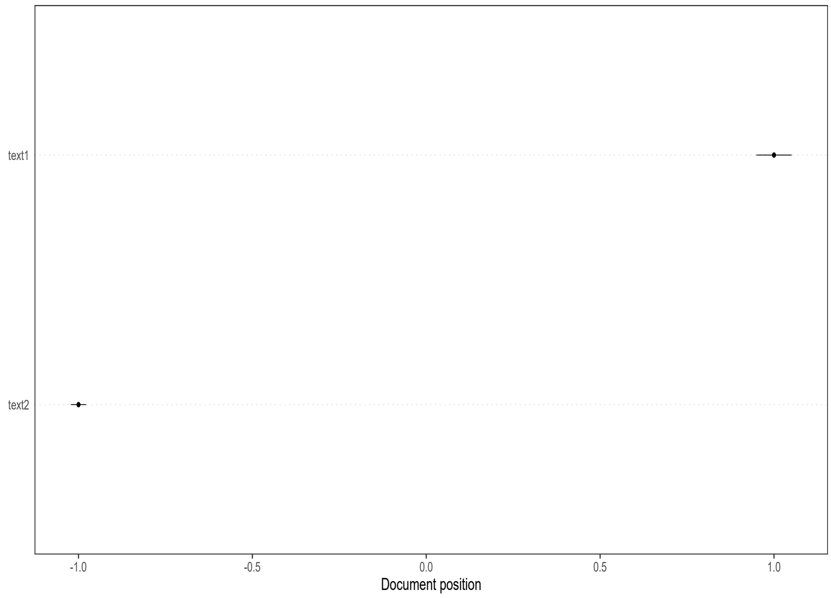


Figure 1.13: Predicted document position using reference texts. The figure shows the predicted document position of reference texts. The horizontal line is the standard error.

The figure shows that the the composition of the reference text, despite being unbalanced, is appropriate since the predicted position coincides with the *a priori* definition. However, this does not lead to the conclusion that I identified the "true" reference texts.

As I said before, choosing reference texts involves a great deal of qualitative and subjective decisions and the reference texts I used are in no way the only possible ones. Researchers can experiment many other combinations as long as the result is validated by any reliable source.

External Validation

As any machine learning algorithm, the selection and composition of an appropriate training set is of crucial importance. The problem that occurs in this case is that the choice of reference texts requires qualitative and subjective decisions that are not validated by any source. In fact, being the first time that such a methodology is applied to central banks' documents, there are no official references that guide the selection. For this reason, I set up an expert survey to validate the choices I made.

The external validation exercise consists of an expert survey that I carried out in *Qualtrics* between May 8 and July 5 2020. Every expert is randomly assigned 15 documents (average length 300 words) containing normative fiscal policy paragraphs in the Press Conferences. Since necessary information was provided on the context of the survey and the content of the study, randomization was to prevent the respondents from directly recognizing the time period. Moreover, Presidents' and officials' names were removed for similar reasons and dates kept only when deemed necessary to the correct execution of the task. The sample of experts was composed of macro and political economists whose expertise lies between monetary and fiscal policy in the euro area. The survey was sent to each expert by email. I collected around 200 responses whose breakdown by sector is summarized in Table 1.3[25].

[25] Anonymity was fully guaranteed. I also said in advance that nothing but the profession was going to be recorded.

Table 1.3: Respondents by Sector

Sector	*Share*
Prefer not to say	7.5%
Central Bank	4%
Academia	57.2%
Financial Sector	17.3%
Generic Economist	14%

The task of the experts was to score each document. The scoring was as simple as: -1, 0 and 1, where -1 stands for *fiscally hawkish*; 1 for *fiscally dovish* and 0 for *fiscally neutral*. The categories are defined as in the paper: fiscally hawkish is a document whose content encourages governments to keep government budgets under control (sizeable reduction of government deficit and stabilization of government debt) by advocating for cuts in government spending or higher taxes or a combination of both; fiscally dovish is a document whose content encourages governments to increase government deficit (and hence possibly debt) to finance sizeable government spending to sustain aggregate demand; fiscally neutral is a document whose content cannot be clearly identified as -1 or 1 for conflicting and/or scarce information and/or moderate policy advise. Figure 1.14 and Figure 1.15 show, respectively, the introductory part of the survey and an example of a Press Conference to be scored:

Information

This is an external validation exercise for a work-in-progress paper. Focusing on the European Central Bank (ECB), the paper studies the interface between unconventional monetary policy, institutional architecture and the degree of fiscal policy accommodation.

The dataset is made up of subsections/sentences in which the Presidents of the ECB talk normatively about fiscal policy over the period 2002(1)-2020(4). The documents are extractions from Press Conferences (Q&A included). Due to the nature of the dataset, there may be conceptual jumps between sentences. If the same document appears more than once, it simply means that the ECB kept the communication constant over a period of time. Moreover, Presidents' names were removed to minimize bias. Similarly, dates were kept only when deemed necessary to the correct execution of the task.

Every expert will be randomly assigned 15 documents (average length 300 words). It shouldn't take more than 15 minutes.

Anonymity is taken seriously and fully guaranteed. Nothing but the profession will be recorded, so that I will not be able to link the respondent to the response. In the paper, I will only use descriptive statistics of the responses.

Thank you very much for your cooperation.

Armando Marozzi
PhD Candidate
email: a.marozzi@lse.ac.uk
London School of Economics

Instructions

Your task is to score 15 documents. The scoring is as simple as: -1, 0, 1, where -1 stands for *fiscally hawkish,* 1 for *fiscally dovish* and 0 for *fiscally neutral*. These are defined in the following way:

Fiscally hawkish = a document whose content encourages governments to keep government budgets under control (sizeable reduction of government deficit and stabilization of government debt) by advocating for cuts in government spending or higher taxes or a combination of both.

Fiscally neutral = a document whose content cannot be clearly identified as -1 or 1 for conflicting and/or scarce information and/or moderate policy advise.

Fiscally dovish = a document whose content encourages governments to increase government deficit (and hence possibly debt) to finance sizeable government spending to sustain aggregate demand.

Q0. State your Profession

Figure 1.14: Introduction to Expert Survey

Q. Fiscal policies will make their best contribution to stability, growth and confidence if prevailing imbalances are tackled as part of determined and well-designed reform programmes. A rigorous implementation of the revised Stability and Growth Pact would reinforce the credibility of reform plans and boost expectations of a sound fiscal and growth situation. Against this background, it is regrettable that the pace of fiscal consolidation remains too slow. In some countries, targets for correcting excessive deficits are at risk. Moreover, due to a very generous application of the new rules of the Pact, countries which have recently breached the 3% deficit limit are being granted relatively long periods to correct the situation. The governing council therefore urges member states to step up consolidation efforts where needed and to implement the revised rules in a manner that supports these efforts and deters future slippages.

- 1
- 0
- -1

Figure 1.15: Sample Question

The results of the expert survey were shown in the main paper. The two series are clearly correlated (86%) and also the signs of each observation correspond throughout the period. In only 4 occasions the two series show different signs. It is worth exploring the disagreement between the supervised model and the expert survey. The first controversial Press Conference is on 5 June 2014:

> In this context, the Governing Council takes note of the European Commission's recommendations on fiscal and structural policies, published on 2 June 2014, to continue the path of reducing budgetary and macroeconomic imbalances. The recommendation to the Council to abrogate the excessive deficit procedures for four euro area countries indicates continued progress in restoring sound public finances. However, euro area countries should not unravel progress made with fiscal consolidation. A full and consistent implementation of the euro area's macroeconomic surveillance framework, together with the necessary policy actions by euro area countries, will help to raise potential

> growth, increase the euro area's resilience to shocks and facilitate job creation. I would say the word, the message that the ECB has sent has always been you have to consolidate your budget, but governments have to do so in a growth-friendly way. And the growth-friendly way means lower current government expenditure, lower taxes, to the extent that it's possible within the Stability and Growth Pact higher capital goods expenditure, higher public investment and all this should be accompanied by structural reforms. If any of these pieces falls down then you don't have a growth-friendly consolidation. If governments pursue fiscal consolidation only through the increase in taxes, we have what we have today, namely that this is the area of the world where you have the highest taxes. And that doesn't seem to be conducive to growth. If governments think that they can consolidate the budget and that's enough and they don't pursue structural reforms, you can see that that's not sustainable. We shouldn't forget that one of the reasons for this crisis was the condition in which many budgets of European countries were at the beginning of the crisis and the level of debt and deficits in many countries. This was not the only cause of the crisis, but certainly it was one of the important ones. So we don't want to go back to that situation. And that's why I always insist to growth-friendly fiscal consolidation.

The document is certainly controversial to score. This difficulty is reflected in the expert survey attributing -0.57, while the supervised algorithm 0.52. In this case, both magnitudes are below one and their average would result into a plausible fiscally neutral document according to the definition I gave above.

The second conflicting result is with respect to the Press Conference on 6 November 2014 and is the one that registers the biggest deviation between the two observations:

> As regards fiscal policies, countries with remaining fiscal imbalances should not unravel the progress already made and should proceed in line with the rules of the Stability and Growth Pact. Throughout the procedural steps under the agreed framework, the Pact should remain the anchor for confidence in

> sustainable public finances. The existing flexibility within the rules should allow governments to address the budgetary costs of major structural reforms, to support demand and to achieve a more growth-friendly composition of fiscal policies. A full and consistent implementation of the euro area's existing fiscal and macroeconomic surveillance framework is key to bringing down high public debt ratios, to raising potential growth and to increasing the resilience of the euro area economy to shocks.

The expert survey scores this document -1.78; instead, wordscore results in 1.93. It very much depends on the weight that one gives to each sentence and the overall interpretation of the document. Once again, though, the difference emerges in an intrinsically controversial document. Besides, even in this scenario, averaging the two observations would result in a balanced document.

The other two controversial results have the same content and correspond to the Press Conferences that took place, respectively, on 8 June 2017 and 20 July 2017:

> Regarding fiscal policies, all countries would benefit from intensifying efforts towards achieving a more growth-friendly composition of public finances. A full, transparent and consistent implementation of the stability and growth pact and of the macroeconomic imbalances procedure over time and across countries remains essential to bolster the resilience of the euro area economy.

Being the documents the same, Wordscore gives correctly the same result: -0.78. Instead, interestingly, experts score the June Press Conference 0.24 and the July one 0.96. This largely reflects the randomization of the survey and the, by default, the mean of the scoring of different experts. That said, the document is prone to controversial results since the opening (and shorter) sentence can be considered fiscally dovish, while the second (and longer) one is more tilted toward a hawkish score - always according to the definition of the task. In this case as well, the average is very close to 0.

Overall, the documents that register different signs are the most controversial ones and some mild form of measurement error needs to be accounted for. Fortunately, these deviations are always small in magnitude and often negligible. Nonetheless, rather than trying to solve the tensions between these observations, I will run the models with both series to test whether any significant deviation occurs.

Robustness Checks

I carried out two main robustness checks as to the DFPA variable. The first check regards the frequency of the series. At its inception in 1999, the ECB Governing Council took policy decisions twice a month, whereas a press conference took place only once a month, on the first meeting of the month. After November 2001 only one meeting per month was a policy meeting, taking place on the first Thursday of the month, regularly accompanied by the press conferences, with some exceptions (see Ehrmann and Fratzscher, 2009). As of January 2015, the frequency of monetary policy meetings has moved to a six-week cycle. This is to say that the frequency of the series is not monthly; from 2002(1) to 2019(12), 26 observations are missing to obtain a monthly frequency. In the main specification, I assumed that, when a Press Conference is missing, the ECB keeps the "fiscal stance" constant, that is, it does not have the opportunity to update its considerations on fiscal policy in the euro area. In practical terms, this assumption means to fill the missing observations at time t with the lag value at time $t-1$. Alternatively, I linearly interpolated the missing observations. In either cases, the differences between the two methodologies is marginal (99% correlation in either cases). Besides, the FAVAR is not only robust to this alternative specification but also to an analysis carried out with a real-time dataset.

The second check consists in running the algorithm with different pre-processing decisions, essentially without lemmatization and ngrams (bigrams and trigrams). The result is essentially the same as the pre-processed series, 95.1% correlation. More importantly, the FAVAR is robust to different text pre-processing decisions.

1.8. APPENDIX C: Extracting Monetary Policy Shocks

In what follows, I explain how I derived the monetary policy shocks.

Extracting MP Shocks

To identify unconventional monetary policy shocks, I rely on the *Euro Area Monetary Policy Event-Study Database* (EA-MPD) by Altavilla et al. (2019) (henceforth, ABGMR). The dataset features price changes for a broad class of assets and various maturities, including OIS rates, sovereign bond yields, stock prices and exchange rates for a total of 45 series from January 1999 to December 2019[26].

I follow Wright (2019)'s comments on ABGMR to improve on the extraction of the monetary policy factors. Unlike ABGMR that only use OIS rates, I exploit the entire dataset of price changes from November 2005 to December 2019[27]. The reason for doing so is twofold: first, narrowing the identification to OIS rates would cut off a crucial source of variation since unconventional measures in the euro area work more through sovereign yields than the term structure of risk-free rates (Cour-Thimann and Winkler, 2012; Wright, 2019). Second, following ABGMR's strategy would have resulted in downplaying the essential period of the sovereign debt crisis. In fact, ABGMR capture unconventional measures only from 2015 onward[28].

Four factors are thus extracted from the entire EA-MPD dataset by using a static factor model whose baseline specification is[29]:

[26]The database is also classified into three windows: "Pre-Conference", "Conference" and "Monetary Event". The first window captures the price changes around the brief press release published at 13.45 Central European Time (CET). The second one captures market reactions around the reading of the Introductory Statement (IS) that starts at 14.30 CET and is then followed by Q&A. The last one documents the price changes for the entire monetary event spanning both the "Pre-Conference" and "Conference" window. This paper focuses on the "Monetary Event" window.

[27]Excluding the periods with missing observations resulted in my sample starting in 2005.

[28]This is due to ABGMR's QE identification strategy (see Altavilla et al., 2019).

[29]The selection of the number of factors follow the reduced rank test of Cragg and Donald (1997). I test the number of statistically significant factors over the full sample, the pre-crisis sample and the pre-QE sample. Additionally, the factors take value zero on non-announcement days.

$$X = \Lambda F + \epsilon \,, \tag{1.21}$$

where X is a vector of N observable variables, F is a vector of K latent factors where $K < N$, Λ is a $N \times K$ matrix of factor loadings and ϵ is an N-dimensional vector of uncorrelated idiosyncratic components such that $\sum_\epsilon$ is diagonal.

Although the latent factors F do not have a structural interpretation as monetary policy surprises, the second principal component appears to closely resemble a sovereign-debt crisis (henceforth, SDC) factor[30]. Table 1.4 shows the variance shares explained by this factor (almost 14%) along with the loadings of the selected variables:

Table 1.4: SDC Factor

Variance Share (%)	13.7
Loadings:	
Germany 10Y yield	0.115
France 10Y yield	0.003
Italy 10Y yield	−0.336
Spain 10Y yield	−0.330
STOXX50E Index	0.348
SX7E (Bank) Index	0.384
Euro-Dollar	0.032
EUR-Pound	0.018
EUR-Yen	0.107

Note: The table reports the variance shares and loadings for the first four principal components of all the asset price changes from November 2005 to December 2019 in the ABGMR dataset.

The SDC component raises German and French yields, lowers Italian and Spanish ones, appreciates the euro and increases stock prices. These responses can be explained by two dynamics. First, several of the unconventional measures taken during the sovereign debt crisis led to a positive reaction of German and French yields and to a negative reaction of Italian and Spanish ones. Government bonds of France and Germany were perceived as a safe haven by the market and, as unconventional measures decreased uncertainty

[30] Wright (2019) called this factor "Save the Euro".

and risk, the demand for safe-haven assets declined while the demand for riskier assets increased. Second, unconventional measures reduced the risk of a break-up of the euro area lowering redenomination risks priced in these bonds. This, other than reinforcing the inverse response of northern and southern yields, appreciated the euro and inflated stock prices. Figure 1.16 plots the time series of the SDC factor:

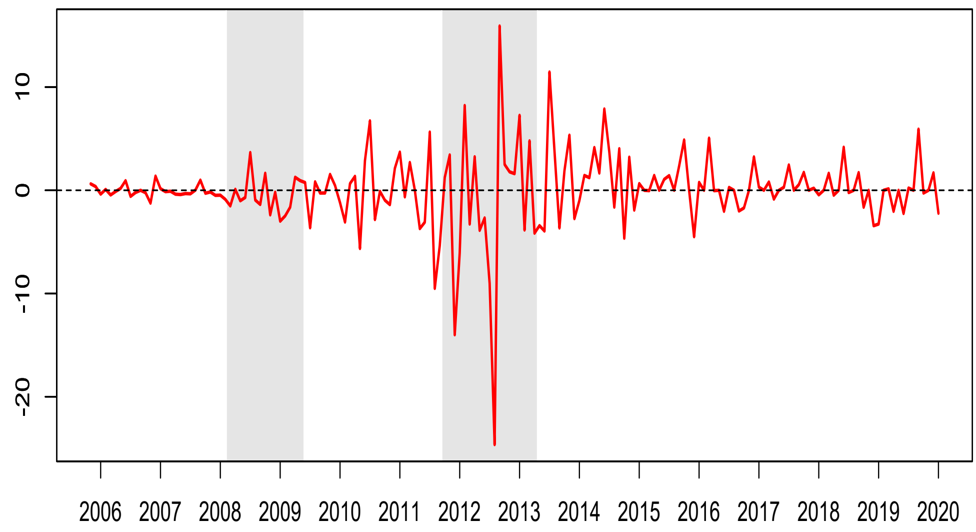


Figure 1.16: SDC Factor. This figure plots the time series of the second principal component of all the asset prices changes from November 2005 to December 2019 in the monetary event windows (from before the press release to after the press conference) in the ABGMR dataset.

The SDC factor barely exists prior to 2008 but then becomes relevant in magnitude from 2010 and even more so during the sovereign debt crisis. This component would not have appeared, had I used OIS rates alone.

As for the three remaining factors, I follow Swanson (2021)'s methodology that AGBMR applies to euro area monetary policy data[31]. The orthogonal factors are identified by imposing the following restrictions on the rotation matrix:

[31] A similar rotation strategy that does not consider QE has been used since Gürkaynak et al. (2005) and was first applied to the euro area data by Brand et al. (2010) using intraday data.

1. the first factor does not load on the one-month OIS;
2. the second factor does not load on the one-month OIS;
3. the third factor has the smallest variance in the pre-crisis period (2 January 2002 - 7 August 2008).

The third condition identifies the QE factor that, once linearly combined with the SDC one, originates the Large-Scale Bond Purchases (LSBP) factor[32]. In particular, the SDC factor goes from January 2009 to December 2013, while the QE factor goes from January 2014 to January 2020.

Finally, Figure 1.17 shows the factor loadings of the rotated factors over the seven maturities in the analysis[33]:

[32]The SDC factor is normalized to have a unit effect on 10Y German yields. Additionally, the shocks can reveal information not just about policy but also about the central bank's assessment of the economic outlook. Accordingly, to control for the presence of information frictions in the economy, I project these monetary surprises on their own lags as well as ECB staff macroeconomic projections for inflation and GDP (see Miranda-Agrippino and Ricco, 2021).

[33]The release of the *US initial jobless claims* (IJC) coincides in many occasions with the beginning of the ECB's Press Conferences as Brand et al. (2010) noted. For this reason, I control for it by employing a continuous measure of IJC surprises that is standard in the literature. The variable is obtained by subtracting the market expectation to the actual value of the release; this difference is then standardized by its standard deviation.

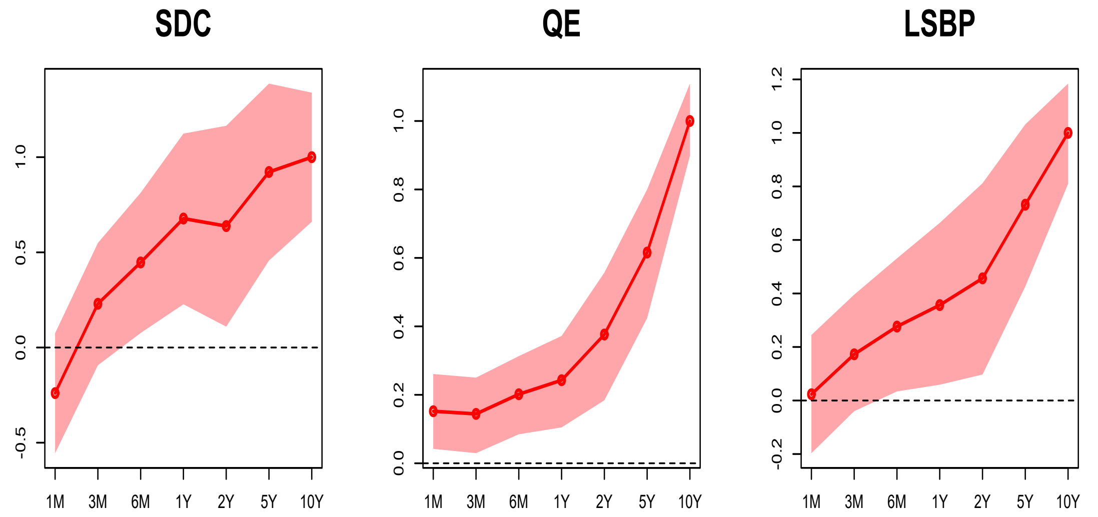


Figure 1.17: Factor Loadings. The figure shows the factor loadings. For each maturity the loadings are obtained by regressing the surprises onto the factors, controlling for the standardized surprise associated with the release of the US initial jobless claims. Every factor displayed is normalized to have unit effect on the 10-year yields. The shaded areas indicate the 90% confidence intervals.

The estimated shape in Figure 1.17 demonstrates the plausibility of the rotation strategy. In fact, similarly to the SDC and QE factor, the impact of the LSBP factor is stronger the longer the maturity. This is exactly what identifies large-scale government bond purchases.

Formally, the rotation strategy goes as follows (see Swanson, 2021). Given the $T \times n$ matrix X of n asset price responses to the T ECB announcements in my sample, I first demean and scale each column of X to have zero mean and unit variance. I then extract the first three principal components of the standardized matrix to estimate the three latent factors that explain a maximal fraction of the variance of the (standardized) data. Let F denote the $T \times 3$ matrix of these first three principal components, and Λ the $3 \times n$ matrix of loadings of the data X on F. It is straightforward to show that a 3×3 orthogonal matrix U is uniquely identified by three parameters. Thus, three identifying restrictions are required to uniquely identify the rotation U that maps the principal components F into three factors $\tilde{F} = FU$ that have a structural interpretation as 1) the surprise change in the short term rate (Jump), 2) the surprise change in Forward Guidance (FG), and 3) the surprise change in asset purchases (QE).

The first two identifying assumptions are that QE and FG have no effect on the short term rate. These two zero restrictions can be written as:

$$U'\Lambda_2 = \begin{bmatrix} \cdot \\ 0 \\ 0 \end{bmatrix} \tag{1.22}$$

where Λ_2 denotes the second column of Λ, the loadings of the short term rate (1M OIS) on the three factors F. Letting U_i denote the ith column of U, these restrictions correspond to $\Lambda_2' U2 = 0$ and $\Lambda_2' U_3 = 0$. Effectively, these two restrictions imply that only the first factor has any systematic effect on the short-term rate.

The third identifying restriction is that the variance of the QE factor is as small as possible over the sample from November 2005 to January 2008. The QE factor is given by FU_3, so this restriction amounts to minimizing $U_3'(F^{pre})'U_3(F^{pre})$. This is a constraint on U_3 that does not directly affect U_1 or U_2, except via the orthogonality conditions between the columns of U.

Computationally, I implement these restrictions as follows. Restrictions one and three are implemented in one step as a quadratic minimization problem subject to a linear constraint: I temporarily ignore the unit length requirement on U3 and normalize the third element of U_3 to unity; I then minimize:

$$\begin{bmatrix} u_{13} & u_{23} & 1 \end{bmatrix} (F^{pre})' F^{pre} \begin{bmatrix} u_{13} \\ u_{23} \\ 1 \end{bmatrix} \tag{1.23}$$

subject to $\Lambda_2' \begin{bmatrix} u_{13} & u_{23} & 1 \end{bmatrix}' = 0$ where Λ_2 denotes the first column of Λ. After computing the minimizing vector $\begin{bmatrix} u_{13} & u_{23} & 1 \end{bmatrix}$, I rescale it to have unit length and call the resulting vector U_3.

To implement the second restriction, I then solve equation:

$$\begin{bmatrix} \Lambda_2' \\ U_3' \end{bmatrix} \begin{bmatrix} u_{12} \\ u_{22} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{1.24}$$

for u_{12} and u_{22} which ensures that $\begin{bmatrix} u_{12} & u_{22} & 1 \end{bmatrix}$ satisfies the identifying restriction and is orthogonal to U_3. I then rescale the vector $\begin{bmatrix} u_{12} & u_{22} & 1 \end{bmatrix}$ to have unit length and call the result U_2.

I compute U_1 by an analogous procedure, normalizing the third element of U_1 to unity, then solving the equation:

$$\begin{bmatrix} U_2' \\ U_3' \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{12} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{1.25}$$

and renormalizing $\begin{bmatrix} u_{11} & u_{12} & 1 \end{bmatrix}$ to have unit length. This ensures U_1 is orthogonal to U_2 and U_3.

This uniquely identifies U and $\tilde{F}$ up to a sign and scale. The Jump factor is normalized to have unit effect on the 1-month OIS. The Forward Guidance and QE factors are normalized to have unit effect on the 2-year and on the 10-year yields, respectively. This normalization has no effect on the variance shares and statistical significance of the results.

1.9. APPENDIX D: Proxy FAVAR

In what follows, I lay down the model, give details on the dataset used to run the FAVAR model, run the FAVAR with SDC, QE and LSBP shocks, run the FAVAR considering conventional and Forward Guidance shocks and run robustness checks.

FAVAR

What is the relationship between government bond purchases and DFPA? To address this question, I set up a monthly Factor-Augmented VAR (Bernanke et al., 2005) identified with external instruments. The transition equation for the model can be expressed as follows:

$$\begin{bmatrix} F_t \\ Y_t \end{bmatrix} = \Phi(L) \begin{bmatrix} F_{t-1} \\ Y_{t-1} \end{bmatrix} + \epsilon_t \ , \tag{1.26}$$

where F_t is a $K \times 1$ vector of unobserved factors to be identified, Y_t is a $M \times 1$ vector of observable variables, $\Phi(L)$ is a lag-polynomial of finite order p and $\epsilon_t = (\epsilon_t^F, \epsilon_t^Y)'$ is a $(K + M) \times 1$ vector of independent and identically distributed (iid) white noise.

This model cannot be directly estimated since F_t is unobservable. Therefore, I collected a panel of time series that contains information about real activity, inflation and financial indicators in the euro area. The $N \times 1$ vector X_t denotes this panel and is related to the unobservable factors F_t and the observed variables Y_t by the following observation equation:

$$X_t = \Lambda^f F_t + \Lambda^y Y_t + e_t \ ,$$

where Λ^f is the $N \times K$ matrix of factor loadings, Λ^y is a $N \times M$ matrix of coefficients and $e_t \sim N(0, V)$ is the $N \times 1$ vector of error terms and V is a diagonal matrix. I then estimate the observation equation via principal component method and extract four latent factors, while the observed variables correspond to the policy indicator and the DFPA derived above.

Next, by applying the Wold representation theorem, the FAVAR model can be rewritten as:

$$y_t = \Phi(L)\upsilon_t \,, \tag{1.27}$$

where $\upsilon_t = H\epsilon_t$ is a $N \times 1$ vector of innovations, $\Phi(L) = \phi_0 + \phi_1 L_1 + ... + \phi_p L^p$ is the infinite order vector moving average (VMA) polynomial, Φ_j, $j = 1, ..., \infty$, are $K \times K$ matrices of VMA coefficients and L is the lag operator such that $Ly_t = y_{t-1}$ and H is a matrix of coefficients. VAR models result in an identification problem as the matrix H is not unique. To solve this issue, I follow the so-called external instrument approach (Montiel Olea et al., 2020; Stock and Watson, 2012; Mertens and Ravn, 2013) where information outside the VAR is used to isolate the variation in the reduced-form residuals in the VAR due to monetary policy shocks. The main intuition behind this strategy is that when regressing the instrument z_t on the VAR innovations υ_t, the fitted value of the regression $\Pi\upsilon_t$ identifies the structural shock up to sign and scale.

Following the VAR literature and the notation in Stock and Watson (2012), I model a linear relationship between the VAR innovations υ_t and the structural shocks ϵ_t:

$$\upsilon_t = H\epsilon_t = [H_1 ... H_k] \begin{bmatrix} \epsilon_{1t} \\ \vdots \\ \epsilon_{kt} \end{bmatrix}, \tag{1.28}$$

where H is a matrix of coefficients and H_1 is the first column of H. It follows that $\sum_{\upsilon\upsilon} = H \sum_{\epsilon\epsilon} H'$ where $\sum_{\upsilon\upsilon} = E(\upsilon\upsilon')$ and $\sum_{\epsilon\epsilon} = (\epsilon\epsilon')$. If the system is invertible - a standard assumption in the VAR literature - structural shocks can be expressed as linear combinations of innovations: $\epsilon_t = H^{-1}\upsilon_t$.

An instrument is now employed that is correlated with the structural shock to be uncovered, while being uncorrelated with all other shocks in the system. This corresponds to the standard assumptions of *relevance* and *exogeneity* in the instrumental variables literature. Formally,

1. Relevance: $E(\epsilon_{1t} Z_t) = \alpha \neq 0$

2. Exogeneity: $E(\epsilon_{jt} Z_t) = 0$, $j = 2, \ldots, k$

3. Uncorrelated shocks: $\sum_{\epsilon\epsilon} = D = diag(\sigma^2_{\epsilon k}, \ldots, \sigma^2_{\epsilon k})$,

where D is a $K \times K$ matrix. The last condition is the standard structural VAR assumption that structural shocks are uncorrelated. This assumption does not fix the variance of shocks. From equation Equation 1.28:

$$E(\upsilon_t z_t) = E(H\epsilon_t z_t) = [H_1 \cdots H_k] \begin{bmatrix} E(\epsilon_{1t} Z_t) \\ \vdots \\ E(\epsilon_{kt} Z_t) \end{bmatrix} = H_1 \alpha \,, \tag{1.29}$$

where the last identity follows from the relevance and exogeneity conditions. It follows that H_1 is identified up to scale and sign by the covariance between the VAR innovations and the instrument. To identify the shocks themselves, the third condition on uncorrelated shocks is needed. It implies that the variance-covariance matrix of υ_t can be rewritten as:

$$\sum_{\upsilon\upsilon} = H \sum_{\epsilon\epsilon} H' = HDH' \tag{1.30}$$

Additionally, by defining Π the matrix of coefficients from the OLS population regression of Z_t on υ_t, the fitted value of this regression is:

$$\Pi \upsilon_t = E(z_t \upsilon_t') \sum_{\upsilon\upsilon}^{-1} \upsilon_t \,, \tag{1.31}$$

which, using equation Equation 1.30 and Equation 1.31, can be written as:

$$E(z_t \upsilon_t') \sum_{\upsilon\upsilon}^{-1} \upsilon_t = \alpha H_1'(HDH')^{-1} \upsilon_t \tag{1.32}$$

by simplifying and using equation $\epsilon_t = H^{-1} \upsilon_t$:

$$\alpha(H_1'(H')^{-1})D^{-1}\epsilon_t = \alpha H_1'(HDH')^{-1}\epsilon_t \tag{1.33}$$

Finally, noting that $H^{-1}H_1 = e_1$ where $e_1 = (1, 0, \cdots, 0)'$ implies:

$$\alpha H_1'(HDH')^{-1}\epsilon_t = \alpha D^{-1}\epsilon_t = (\frac{\alpha}{\sigma^2_{\epsilon 1}})\epsilon_{1t} = \alpha\epsilon_{1t} = \Pi\upsilon_t \tag{1.34}$$

This proves the original statement that the fitted value ($\Pi\upsilon_t$) of a regression of the instrument (z_t) on the innovations (υ_t) identifies the structural shock ϵ_{1t} up to a sign and scale.

Following what I just laid down, I now turn to evaluate the choice of the policy indicator and the instruments for policy shocks in the monthly FAVAR. I run the following regression:

$$(\epsilon^i_t)^k = \alpha + \gamma\tilde{F}^j_t + u_t\ , \tag{1.35}$$

where ϵ^i_t is the first stage residual for the policy indicator in the representative VAR with six variables, k indicates respectively EONIA, 1Y OIS, 2Y German and 10Y OIS, $\tilde{F}_t$ is the rotated factor and j is Jump, FG, SDC, QE and LSBP. In short, I consider four policy indicators for every instrument. Table 1.5 summarizes the results.

Table 1.5: Effect of Instruments on the First Stage Residuals of the Representative VAR

	Eonia	Eonia	Eonia	1 year	1 year	1 year
	(1)	(2)	(3)	(1)	(2)	(3)
Jump	0.013***			0.009***		
	(0.003)			(0.0033)		
FG		0.015***			0.006*	
		(0.004)			(0.004)	
LSBP			0.007*			0.009***
			(0.004)			(0.003)
Observations	209	209	136	209	209	136
F Statistic	5.5	**15.13**	3.40	7.26	3.01	5.85

	2 year	2 year	2 year	10 year	10 year	10 year
	(1)	(2)	(3)	(1)	(2)	(3)
Jump	0.014***			0.007		
	(.004)			(0.005)		
FG		0.014***			0.006**	
		(0.003)			(0.003)	
LSBP			0.009**			0.008*
			(0.004)			(0.004)
Observations	209	209	136	209	209	136
F Statistic	1.97	5.55	**12.5**	1.79	4.32	3.48

	2 year	2 year	2 year	10 year	10 year	10 year
	(1)	(2)	(3)	(1)	(2)	(3)
SDC	0.013***					
	(0.002)					
QE		0.015***				
		(0.004)				
Observations	136	136				
F Statistic	**13.11**	**12.91**				

Note: The table shows the effects of instruments on the first stage residuals of the four variable (monthly) VAR described earlier. The columns considered are the first stage regression residual of a particular policy indicator regressed on various instrument sets. The robust F-test assess the relevance of the instrument with respect to the policy indicator. Robust standard errors in parentheses; ***, ** and * denote statistical significance at the 1%, 5% and 10% levels, respectively.

To be confident that a weak instrument problem is not present, Stock et al. (2002) recommend a threshold value of ten for the F-statistic from the first-stage regression. The table shows that the combinations of policy indicator and instrument that provide the best instrument strength are reasonably: EONIA-Jump, FG-EONIA and LSBP-2Y OIS.

Last, in Figure 1.18 I show the results of the representative VAR in industrial production, HICP, CISS Bond Market, real exchange rate, DFPA and 2Y OIS I used to test the relevance of the instrument. I only display impulse responses to LSBP shocks since I got approximately the same results for SDC and QE surprises.

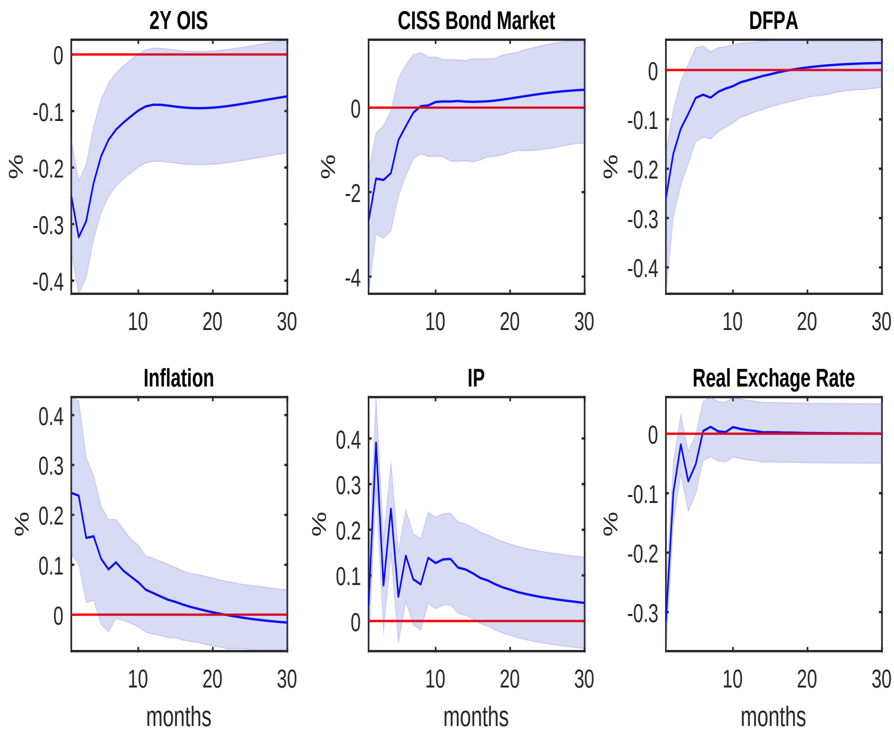


Figure 1.18: Representative VAR Identified with LSBP. The figure plots the impulse responses to a 0.25% decrease in the policy indicator due to a LSBP shock. 90% error bands are displayed.

Even in this representative VAR the results are consistent with the FAVAR and with the main hypotheses of the paper.

Dataset

The table below contains a complete list of the series in the dataset as well as detailed descriptions and information regarding transformations, geographical coverage and sources. Abbreviations and codes are laid out in the following:

Transformation code (T)

1. no transformation
2. logs
3. difference in levels
4. difference in logs

Geography

EA - Euro area

EA12 - Euro area (12 countries)

EA19 - Euro area (19 countries)

EACC - Euro area (changing composition)

$EA11_i$ - 11 individual series for sample countries

Factor analysis (F)

Y - included in data set for principal component analysis

Seasonal adjustment[34]

[34] When the seasonal adjustment was not available, I seasonally adjusted the series in R with X-ARIMA-12.

WDSA - working day and seasonally adjusted

SA - seasonally adjusted (SA)

NA - neither working day nor seasonally adjusted

Description	T	SA	Source	Geography	Start	End	F
Real Activity							
Retail sales volume index	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
Manufacturing turnover index	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
Construction production index	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
Retail sales of non-food production (excl. fuel)	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
Industrial Production (excluding construction)	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
Industrial Production: manufacturing	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
Industrial Production: energy	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
Industrial Production: consumer goods	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
Industrial Production: consumer durables	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
Industrial Production: capital goods	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
Industrial Production: consumer non-durables	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
Industrial Production: intermediate goods	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
New orders: manufacturing	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
New orders: capital goods	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
New orders: consumer durables	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
New orders: consumer non-durables	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
New orders: intermediate goods	4	WDSA	ECB SDW	EA19	2002(1)	2019(12)	Y
Prices							
Harmonised Index of Consumer Prices (HICP)	4	SA	ECB SDW	EACC	2002(1)	2019(12)	Y
HICP: energy	4	SA	ECB SDW	EACC	2002(1)	2019(12)	Y
HICP: food	4	SA	ECB SDW	EACC	2002(1)	2019(12)	Y
HICP: goods	4	SA	ECB SDW	EACC	2002(1)	2019(12)	Y
HICP: excl. energy	4	SA	ECB SDW	EACC	2002(1)	2019(12)	Y
Producer Price Index (PPI)	4	SA	ECB SDW	EACC	2002(1)	2019(12)	Y
Oil Prices (Brent)	4	SA	ECB SDW	EACC	2002(1)	2019(12)	Y
Surveys							
Economic Sentiment Indicator (ESI): industrial confidence	3	SA	EUROSTAT	EA19	2002(1)	2019(12)	Y

Description	T	SA	Source	Geography	Start	End	F
ESI: consumer confidence	3	SA	EUROSTAT	EA19	2002(1)	2019(12)	Y
ESI: economic sentiment	3	SA	EUROSTAT	EA19	2002(1)	2019(12)	Y
ESI: business climate	3	SA	EUROSTAT	EA19	2002(1)	2019(12)	Y
ESI: services confidence	3	SA	EUROSTAT	EA19	2002(1)	2019(12)	Y
ESI: retail confidence	3	SA	EUROSTAT	EA19	2002(1)	2019(12)	Y
Employment							
Unemployment rate: persons under 25	4	WDSA	EUROSTAT	EACC	2002(1)	2019(12)	Y
Unemployment rate (total)	4	WDSA	EUROSTAT	EACC	2002(1)	2019(12)	Y
Unemployment rate: women	4	WDSA	EUROSTAT	EACC	2002(1)	2019(12)	Y
Unemployment rate: persons over 25	4	WDSA	EUROSTAT	EACC	2002(1)	2019(12)	Y
Financial							
VSTOXX volatility index	3	SA	Bloomberg	EA	2002(1)	2019(12)	Y
EU VIX	3	SA	Bloomberg	EA	2002(1)	2019(12)	Y
EUR-USD VIX	3	SA	Bloomberg	EA	2002(1)	2012(12)	Y
CISS: Composite Indicator of Sovereign Stress: GDP weigths	3	SA	ECB SWD	EA	2002(1)	2019(12)	Y
CISS Bond Market: gov and NFC volatility	3	SA	ECB SWD	EA	2002(1)	2019(12)	Y
CISS Money Market: 3M rate volatility	3	SA	ECB SWD	EA	2002(1)	2019(12)	Y
CISS FX Market: euro volatility	3	SA	ECB SWD	EA	2002(1)	2019(12)	Y
CISS FIN Interm: bank volatility	3	SA	ECB SWD	EA	2002(1)	2019(12)	Y
CISS NF equity volatility	3	SA	ECB SWD	EA	2002(1)	2019(12)	Y
Systemic stress indicator	3	SA	ECB SWD	EA	2002(1)	2019(12)	Y
BBB-AAA corporate spread (2Y)	3	SA	Bloomberg	EA	2002(1)	2019(12)	Y
Nominal exchange rate (42 trading partners)	4	SA	ECB SWD	EA	2002(1)	2019(12)	Y
Real eff. exchange rate (42 traing partners, CPI based)	4	SA	Bloomberg	EA	2002(1)	2019(12)	Y
Quarterly (linearly interpolated)							
Fixed Capital Formation	4	SA	ECB SWD	EA19	2002(1)	2019(12)	Y
Real GDP	4	SA	ECB SWD	EA19	2002(1)	2019(12)	Y
Final Consumption Exp. Households	4	SA	ECB SWD	EA19	2002(1)	2019(12)	Y

Description	T	SA	Source	Geography	Start	End	F
Capacity Utilization	4	SA	ECB SWD	EA19	2002(1)	2019(12)	Y
ECB SPF: inflation next year	1	SA	ECB SDW	EA19	2002(1)	2019(12)	Y
ECB SPF: inflation two years ahead	1	SA	ECB SDW	EA19	2002(1)	2019(12)	Y
ECB SPF: GDP growth next year	1	SA	ECB SDW	EA19	2002(1)	2019(12)	Y
ECB SPF: GDP growth two years ahead	1	SA	ECB SDW	EA19	2002(1)	2019(12)	Y
ECB SPF: unemployment rate next year	1	SA	ECB SDW	EA19	2002(1)	2019(12)	Y
ECB SPF: unemployment rate two years ahead	1	SA	ECB SDW	EA19	2002(1)	2019(12)	Y
Government budget balance	3	SA	EUROSTAT	EA19	2002(1)	2019(12)	Y
Government debt/GDP	3	SA	EUROSTAT	EA19	2002(1)	2019(12)	Y
Gov. Net Interest Payments	3	SA	EUROSTAT	EA19	2002(1)	2019(12)	Y
Gov. Total Expenditure	3	SA	EUROSTAT	EA19	2002(1)	2019(12)	Y
Gov. Total Revenue	4	SA	EUROSTAT	EA19	2002(1)	2019(12)	Y
Gov. Investment	4	SA	EUROSTAT	EA19	2002(1)	2019(12)	Y

Results

The sample consists of monthly observations for the period 2002:1-2019:12. Unconventional monetary policy shocks are identified by using SDC, QE and LSBP factors derived above. Since the instruments are available for a shorter period, I first use the full sample to estimate the lag coefficients and obtain the reduced form residuals. Then, I use the factors and residuals for the corresponding period to identify unconventional monetary policy surprises as in Gertler and Karadi (2015). Based on the AIC, the lag length in the FAVAR is set to four. All impulse responses are for a 0.25% decrease in the policy indicator and I show results for a forecast horizon of 30 months. I report 90% confidence intervals computed using a recursive wild bootstrap using 10,000 replications.

Figure 1.19 shows how a selection of macroeconomic time series across the euro area respond to a 0.25% decrease in the policy indicator due to an exogenous SDC (black line), QE (red line) and LSBP (blue line) shock. The model performs in accordance with economic theory. Focusing on the first row, industrial production and manufacturing turnover index rise, as does a soft indicator like industrial confidence. Notably, the model does not suffer from the "price puzzle" since all the variables measuring price pressure correctly increase after the expansionary shock. This result documents that the model can accurately describe the macroeconomic dynamics of the euro area. The remainder of the series also behave as suggested by theory. In particular, risk-related variables significantly drop after the monetary easing, while European stock prices surge. Nominal and real exchange rate depreciate and unemployment contracts after the shock. Upon closer inspection, Figure 1.19 consistently shows that while the effects of SDC and LSBP shocks are often difficult to distinguish, the impact of QE surprises tend to be larger in magnitude.

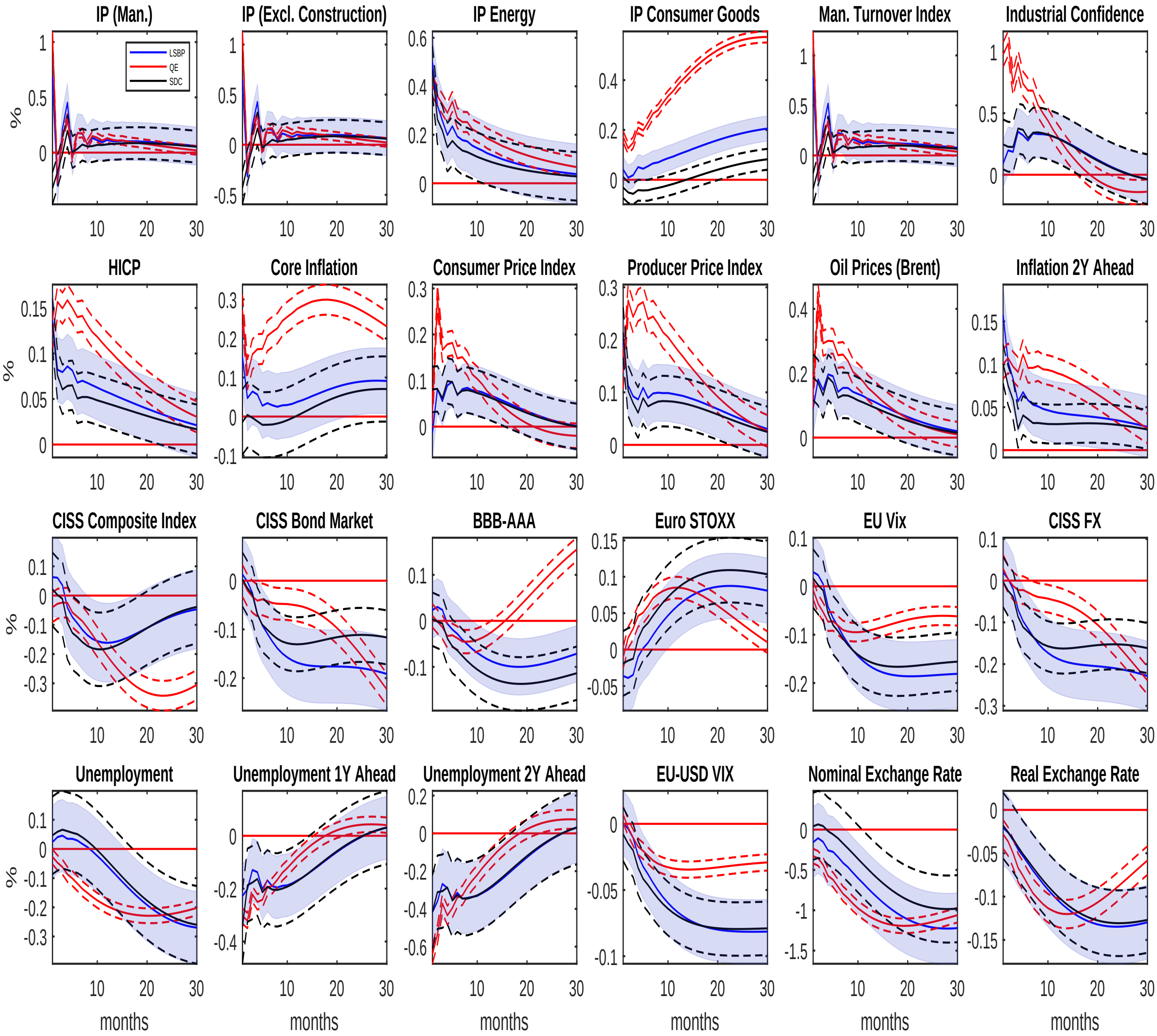


Figure 1.19: Proxy FAVAR. The figure shows the impulse responses to a 0.25% decrease in the policy indicator due to SDC (black line), QE (red line) and LSBP (blue line) shocks. 90% bands are displayed.

To further inspect the rationale underpinning this result, I report the behavior of some key fiscal variables given the shocks under investigation. Figure 1.20 displays the results.

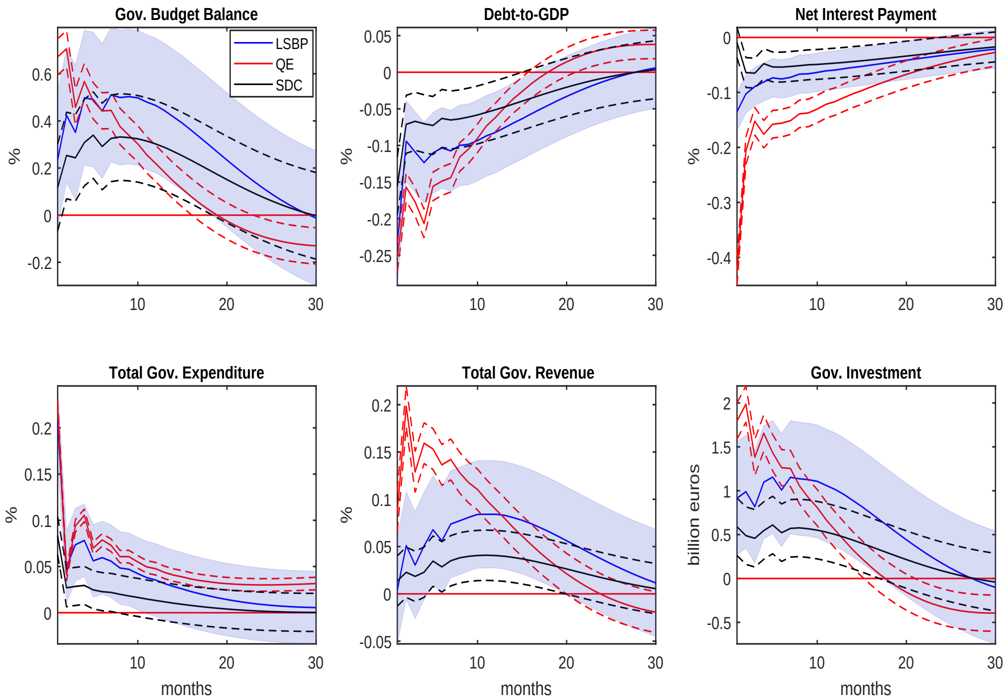


Figure 1.20: Fiscal Variables. Fiscal responses at the euro area level. The figure shows the response to a 0.25% decrease in the policy indicator due to a SDC (black line), QE (red line) and LSBP (blue line) shock. 90% bands are displayed.

The figure shows that government bond purchases have a sizeable windfall for EA countries, that is, government budget improves, debt-to-GDP and net interest payments significantly decline and government revenue increases. Besides, governments internalize the positive fiscal effects of monetary measures and increase government spending and investment, thereby proving, as in Hachula et al. (2020), that the response of government spending behaves countercyclically. Similar to the results in Figure 1.19, the impact of QE shocks is quantitatively larger than LSBP and SDC ones. This time, however, LSBP surprises convincingly exhibit to have a more prominent effect than SDC shocks.

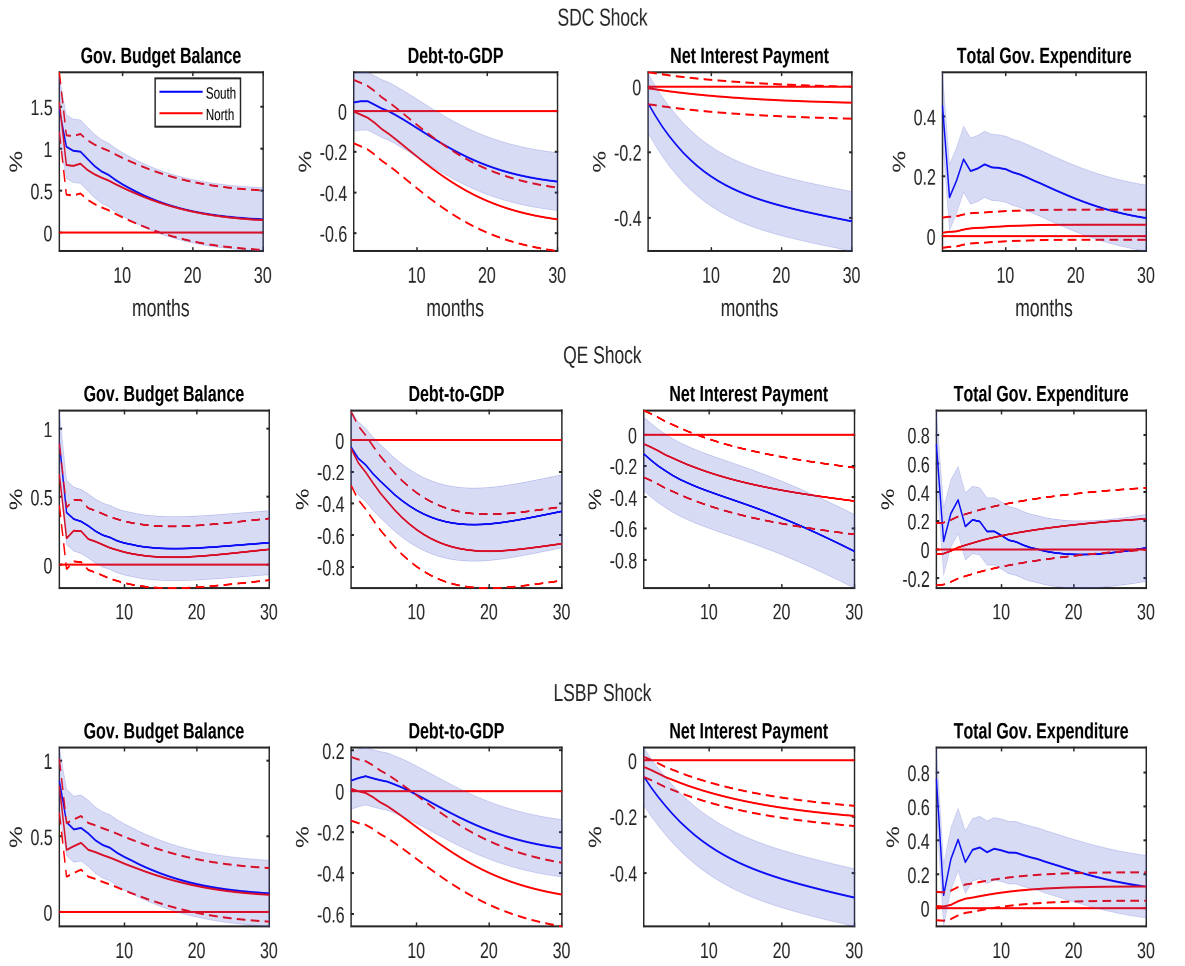


Figure 1.21: Fiscal Variables. Fiscal responses at the euro area, South and North level. The figure shows the response to a 0.25% decrease in the policy indicator due to SDC, QE and LSBP shocks. 90% bands are displayed.

Focusing instead on the difference between North and South blocks, Figure 1.21 illustrates that while debt-to-GDP improves more for the North than the South, the South benefits significantly more than the North in terms of net interesting spending - both results are consistent across the type of shock. Moreover, the gain with regard to government budget is slightly more in favour of Southern countries, whereas the response of total government expenditure is less straightforward to interpret. Throughout the forecast horizon, Southern countries tend to spend significantly more than Northern ones in the presence of SDC and LSBP surprises. Instead, QE shocks tilt Southern expenditure upwards for the first ten months after the shock, while Northern government expenditure displays a muted reaction on impact before rising around 0.2% at 20-30 months after the monetary stimulus.

Additional Results

In this subsection I identify the monthly FAVAR, first, with a conventional monetary policy shock, and then with a Forward Guidance (FG) one. The aim is to understand whether the inverse relationship between DFPA and government bond purchases can be generalized to other shocks.

Figure 1.22 shows the effects of a 0.25% decrease in the policy indicator (EONIA) due to a conventional monetary policy shock (left panel) identified with standard Cholesky decomposition and FG shock (right panel) identified with the FG factor derived above[35]. The sample consists of monthly observations for the period 2003:1-2019:12. Lag-length is set to 3 according to AIC criterion and 90% and 68% confidence intervals computed using a recursive wild bootstrap using 10,000 replications are displayed.

[35] The Cholesky identification was chosen since the Jump factor proved to be rather weak with an F-test for the first stage regression around 5.5 in the representative VAR.

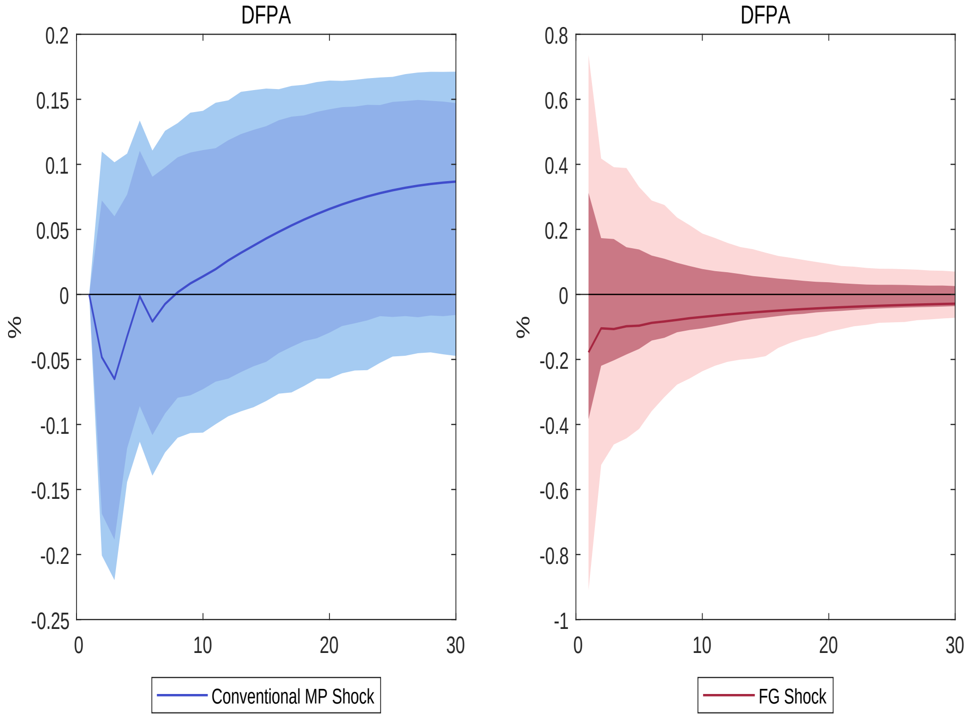


Figure 1.22: Impulse Responses to a Conventional Monetary Policy Shock. The figure shows the impulse response to a 25 basis points expansionary conventional monetary policy shock (left panel) and FG shock (right panel). 90% and 68% error bands are displayed.

Given a conventional rate cut, the fiscal stance of the ECB, first, slightly decreases for 11 months and then rises for the entire forecast horizon. As to FG surprises, DFPA turns marginally negative, while the CBI indicator somewhat picks on impact. In either case, the impulse responses are never statistically significant.

In summary, the results prove that the degree of fiscal policy accommodation of the ECB behaves differently compared to government purchases shocks. While bond purchases drive the ECB to be more fiscally hawkish, conventional monetary policy and FG shocks have no substantial effect on the "fiscal stance" of the European monetary authority. This finding can be explained as follows: even if every monetary policy has fiscal consequences, the magnitude of these consequences significantly differs between conventional and unconventional monetary policy (especially in the case of government bond purchases). It follows that the ECB's DFPA does not react to the negligible effects of the former, but it does respond to the sizeable effects of the latter. Put it differently, the fiscal stance of the ECB reacts proportionally to the size of the fiscal effects of monetary measures.

Sensitivity Analysis

First, I set up the model with different lags (Figure 1.23). Second, I ran the model with different number of factors (Figure 1.24). Third, I ran it using the series derived in the expert survey (Figure 1.25). As the figures below show, in the vast majority of cases the results are robust to lag-length and the selection of the number of factors. I display only the results concerning LSBP surprises since results are similar to SDC and QE ones.

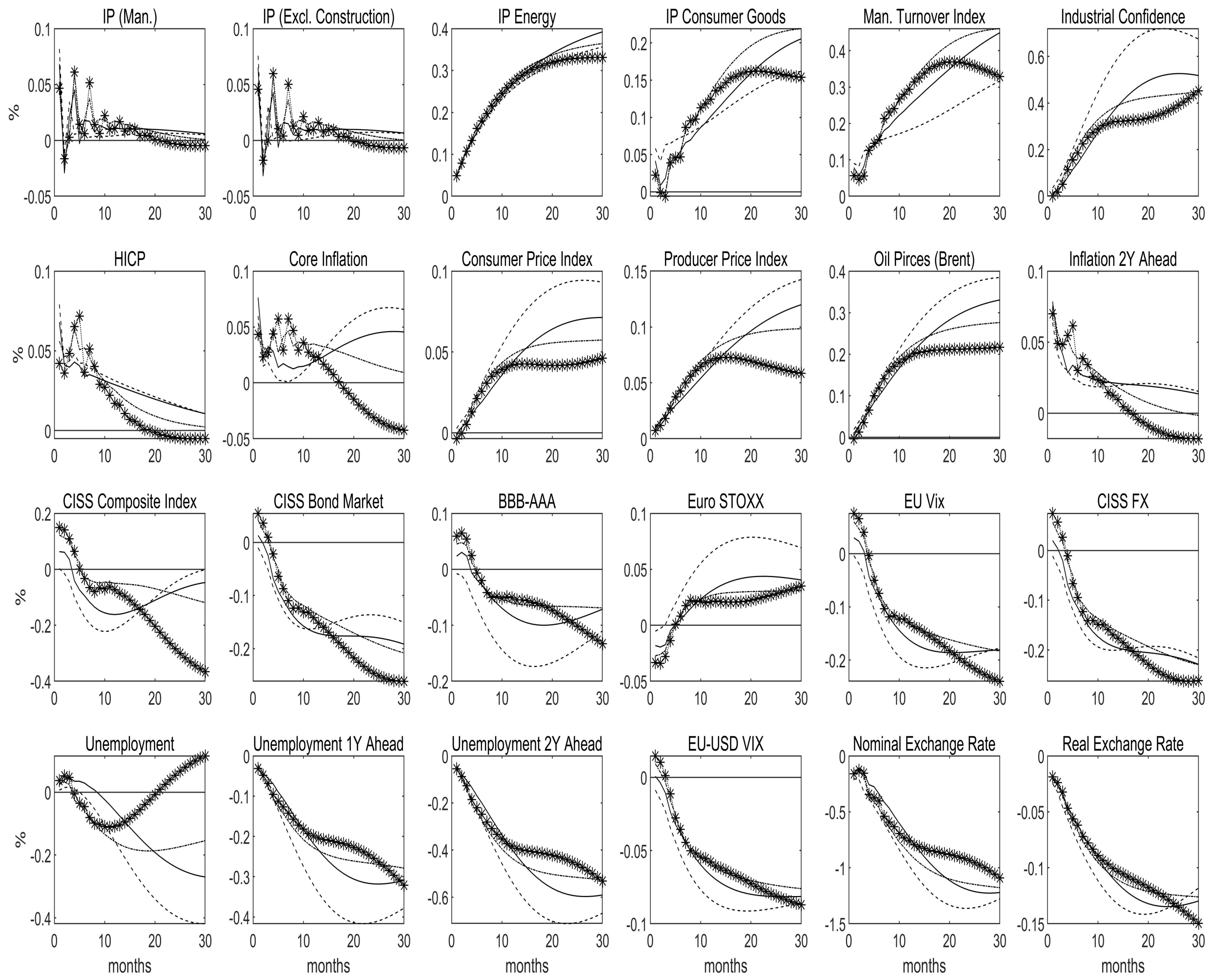


Figure 1.23: Proxy FAVAR. The figure shows the impulse responses of a selection of macroeconomic variables to a 0.25% decrease in the policy indicator due to a LSBP shock. The solid line indicates 3 lags as in the baseline model, the dashed line is for 2 lags, the dash-dot line indicates 4 lags, the dotted line 5 and the asterisk line 6.

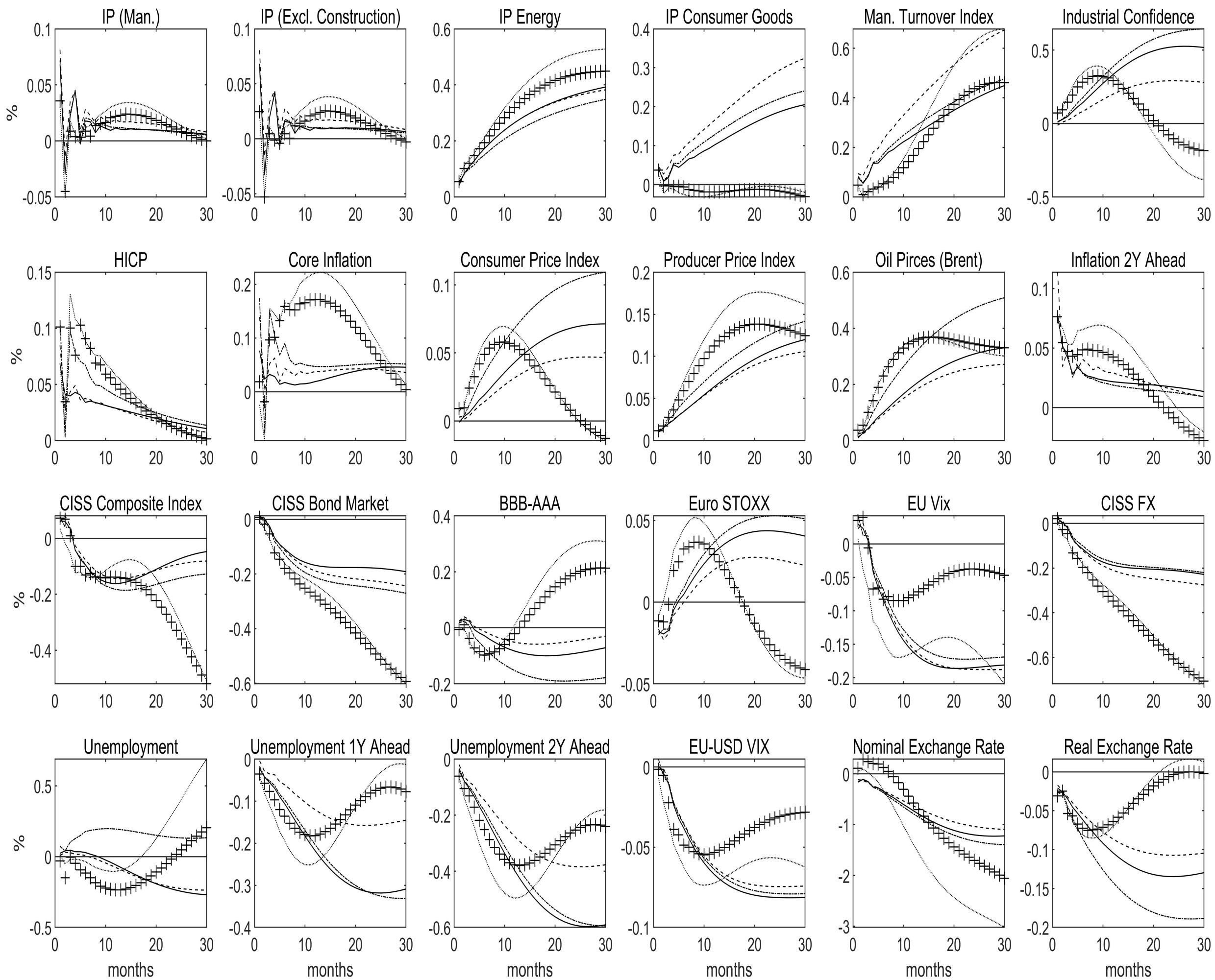


Figure 1.24: Proxy FAVAR. The figure shows the impulse responses of a selection of macroeconomic variables to a 0.25% decrease in the policy indicator due to a LSBP shock. The solid line indicates 4 factors as in the baseline model, the dashed line is for 3 factors, the dash-dot line indicates 5 factors, the dotted line 6 and the plus line 7.

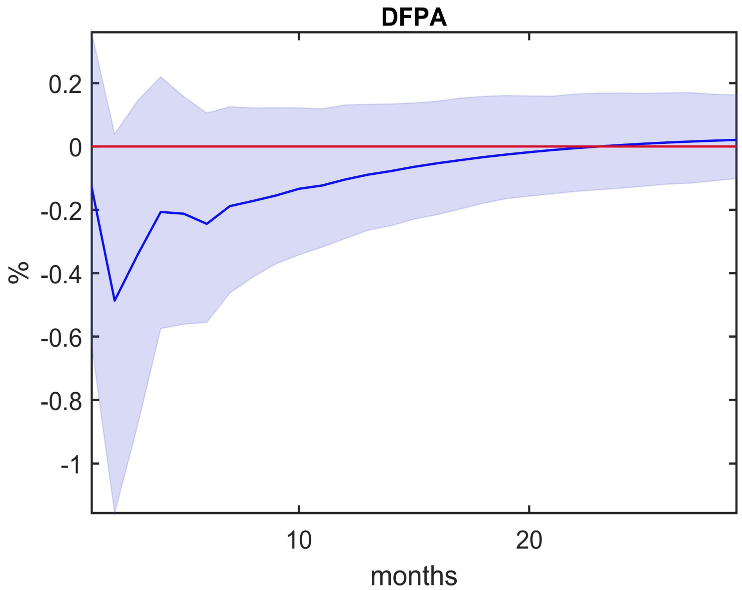


Figure 1.25: DFPA Response to an Expansionary LSBP Shock – Expert Survey. The figure displays the DFPA impulse response to a 0.25% LSBP shock. 90% error bands are shown. The DFPA variable is the time-series of the expert survey in this case.

2. The ECB's Tracker: Nowcasting the Press Conferences of the ECB

2.1. 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. Investors 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 non-conventional 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.

This paper attempts to exploit the entire flow of information that occurs in the timespan between two consecutive ECB press conferences. This implies the construction of a dataset that contains not only conventional data such as financial and macroeconomic releases but also news broadcasted through the media. There are at least two reasons to consider textual data in a macroeconomic forecasting exercise: first, while there is an abundance of high-frequency financial data, the availability of high-frequency macroeconomic variables is much scarcer. Under this regard, textual data can provide real-time information on key macroeconomic indicators. Second, since the importance of central bank communication has grown exponentially in the last decade, news broadcasted through the media have become as relevant as official data. In fact, on the one hand, media channels are increasingly used by central banks to spread out their messages and, on the other hand, they are carefully scrutinized by market participants in search of information. It then follows that textual data shape and reflect the expectations of economic agents[1].

To create a coherent econometric framework that embeds both conventional and textual data, I proceed with the following steps. 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 and finance-related information. Third, I model the total dataset containing around 140 variables as a Dynamic Factor Model (DFM) with flow and stock variables. 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

[1] Between two press conferences, the ECB communicates also through speeches and interviews. Although they can be considered directly relevant for monetary policy, I decided to exclude them from the dataset. The reason is twofold: first, ECB board members not always agree on the optimal path of monetary policy. As a result, considering different signals might create more issues than it would actually solve; second, interviews and speeches are rarely about one single topic. This would result in a complex classification exercise that requires a paper on itw own. Further research can thus explore the predictive power of alternative ECB official releases.

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 over the period 2015-2020.

In light of these results, there are two main contributions to highlight. First, to the best of my knowledge, 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 introduction of a mixed-frequency econometric framework for describing and predicting central banks' reaction function is a novel attempt. Second, the model sheds light on the role of news broadcasted through the media as a fundamental channel of expectation formation. It captures, in fact, the prominent contribution of textual variables in explaining forecast revisions from 2015 on, that is, a period of renewed importance of communication as a policy tool.

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 variables[2]. Thorsrud (2020) and Cimadomo et al. (2020) come also close to this paper: the

[2] 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).

former employs news text to derive a daily business cycle index and predict quarterly 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 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. Similarly, Ardia et al. (2019) test the predictive power of text-based sentiment indices by forecasting the growth in US industrial production using major newspapers from the news database *LexisNexis* over the period January 2001 to December 2016.

The rest of the paper is organized as follows: Section 2.2 describes the datasets, the text-mining techniques and the estimated news topics. Section 2.3 details the models. Section 2.4 illustrates the results and the validation exercises. The last section concludes the paper.

2.2. Data and Textual Methods

Section 2.2.1 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 2.2.2 details the procedural steps to structure and quantify textual data; finally, Section 2.2.3 describes the total dataset used in the model.

2.2.1 Quantifying the Press Conferences of the ECB

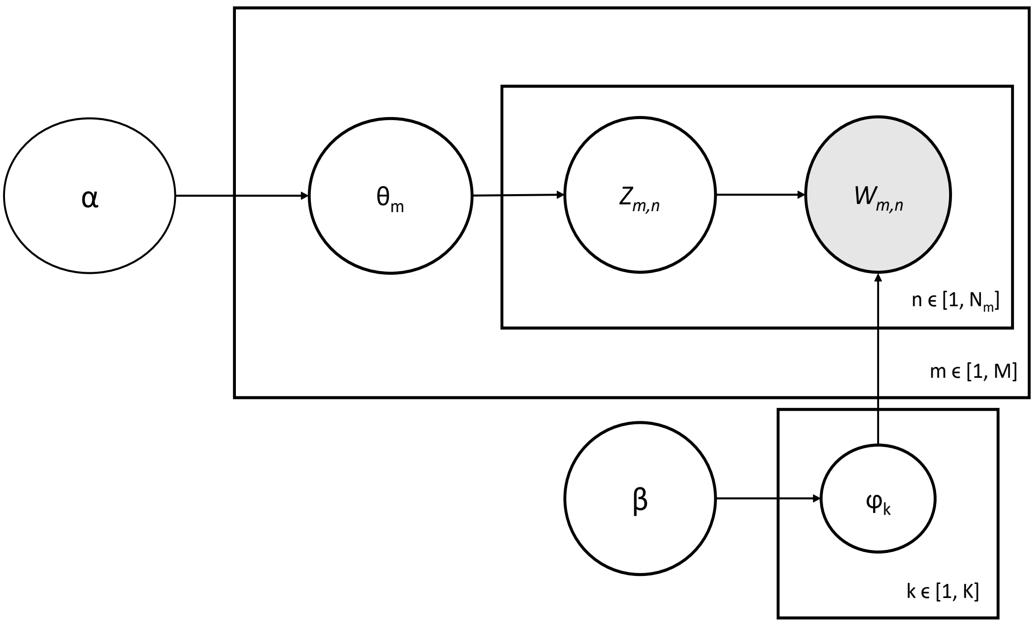
The ECB's press conferences have a pivotal role in transmitting monetary policy information and are therefore suited to extracting macro and policy signals. In particular, the aim is to derive indexes of monetary policy stance, economic and inflation outlook. While the first one is the variable to nowcast, 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, as a first step, 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).

Once the dataset is preprocessed, two challenges arise: on the one hand, the press conferences need to be subsetted to identify only those paragraphs belonging to the topics of interest; on the other hand, the textual content requires to be structured into time-series.

To address the first issue, I apply Latent Dirichlet Allocation (LDA) model (Blei et al., 2003). 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. In other words, LDA assigns to every paragraph the probability to belong to a topic. Paragraphs that belong to the same topic in every press conference are then clustered together in a unique topic-specific document. Figure 2.1 provides a graphical illustration of the model.

Figure 2.1: 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 $\boldsymbol{w} = (w_1, w_2, \ldots, w_N)$ be a document formed by a sequence of words where w_n indicates the n^{th} word and let $D = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_M\}$ 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 $\boldsymbol{z}$ 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 $\boldsymbol{w}$ 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:

$$P(\mathbf{w}_m, \mathbf{z}_m, \boldsymbol{\theta}_m, \boldsymbol{\Phi}; \alpha, \beta) = \prod_{n=1}^{N_m} \underbrace{\underbrace{P(w_{m,n} \mid \boldsymbol{\varphi}_k) P(z_{m,n} \mid \boldsymbol{\theta}_m)}_{\text{word level}} \cdot P(\boldsymbol{\theta}_m; \alpha)}_{\text{document level (1 document)}} \cdot \underbrace{P(\boldsymbol{\Phi}; \beta)}_{\text{topic level}} \tag{2.1}$$

where $\boldsymbol{\Phi} = \{\boldsymbol{\varphi_k}\}_{k=1}^{K}$ is a $K \times N$ matrix. As Equation 2.1 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 $\boldsymbol{\theta_m}$ and $\boldsymbol{\varphi_k}$ out of Equation 2.1 and using Gibbs sampling simulations as in Griffiths and Steyvers (2004). Further technical details are described in Appendix 2.6. The estimated words' probabilities ($\boldsymbol{\varphi_k}$) allow me to identify the three topics of interest (monetary policy stance, economic outlook and inflation outlook)[3]. Then, once the topics are identified, I use the probability of each document to belong to one of these three topics ($\boldsymbol{\theta_m}$) 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 indications[4].

The last step implies the quantification of these datasets. To do so, I use a dictionary approach. Given the specific content of ECB communication, it would be however inappropriate to employ a non-field and non-ECB specific lexicon. A generic dictionary might in fact fail to capture all nuances of central bank communication. Taking the widely-employed dictionary developed by Loughran and McDonald (2011) (LM) as an example, three representative limitations can be highlighted. First, the word "downward" is classified as negative while "upward" is not classified. In contrast, in the ECB's introductory statements the two words are opposites. Second, considering single words (hereafter unigram) rather than a contiguous sequence of n words (hereafter n-grams) might lead to a misclassification of tone. For instance, "lower unemployment" will be classified as negative using the LM dictionary due to the presence of the negative word "unemployment". Third, LM does not control for valence shifters (i.e., negators, amplifiers, de-amplifiers and adversative conjunctions). For example, "recover" and "recover strongly" will have the same tone in the LM dictionary, while "strongly" is meant to amplify the preceding word. Moreover, LM does not have a list of negators and adversative conjuctions that can reverse the tone of a sentence.

To address these limits, I first generate ECB field-specific lexicons and then apply a rule-based algorithm to quantify the ECB press conferences. Starting with the creation of ECB field-specific dictionaries for each topic c, I draw on Picault and Renault (2017)'s methodology. I first subset each dataset PC^c into unigrams and n-grams and manually

[3] Appendix 2.6 provides insights on topics' identification.
[4] Since every document m has a probability $\boldsymbol{\theta_m}$ 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 (${\boldsymbol{\theta_m}}^{MAX}$) exceeds 35%.

classify each of them 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)[5]. 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%. Table 2.1 shows a sample of $n$grams and unigrams for each topic c.

In parallel and unlike Picault and Renault (2017), 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 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} cluster[6]. 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 $1 + [(w_{advcon} - 1) \cdot \eta]$.

5 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.

6 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.

Table 2.1: Sample of N-grams and Unigrams for Each Topic

Monetary Policy	**Economic Outlook**	**Inflation Outlok**
Dovish	Positive	Positive
– favourable financing conditions	– stronger than expected	– tighter labour markets
– additional asset purchases	– growth momentum	– higher than expected
– accommodative stance	– positive outlook	– second round effects
– increase the envelope	– revised upwards	– higher energy price
– reduce interest rates	– strong recovery	– solid wage growth
– heightened alertness	– encouraging	– higher pressure
– extension of credit	– upswing	– higher inflation
– liquidity provision	– impetus	– fast growing
– weakening	– surging	– upside risks
– subdued	– upturn	– build up
Hawkish	Negative	Negative
– higher commodity prices	– negative cyclical momentum	– lower unit labour costs
– increase interest rates	– heightened uncertainty	– higher unemployment
– inflationary pressure	– revised downwards	– lower wage pressures
– strong fundamentals	– downside risks	– unutilised capacity
– counter upside risks	– vulnerabilities	– lower oil prices
– rise in inflation	– decelerating	– disappointing
– end purchases	– disequilibria	– contained
– winding down	– contracting	– fall below
– rising wages	– headwinds	– sluggish
– withdrawal	– softening	– muted

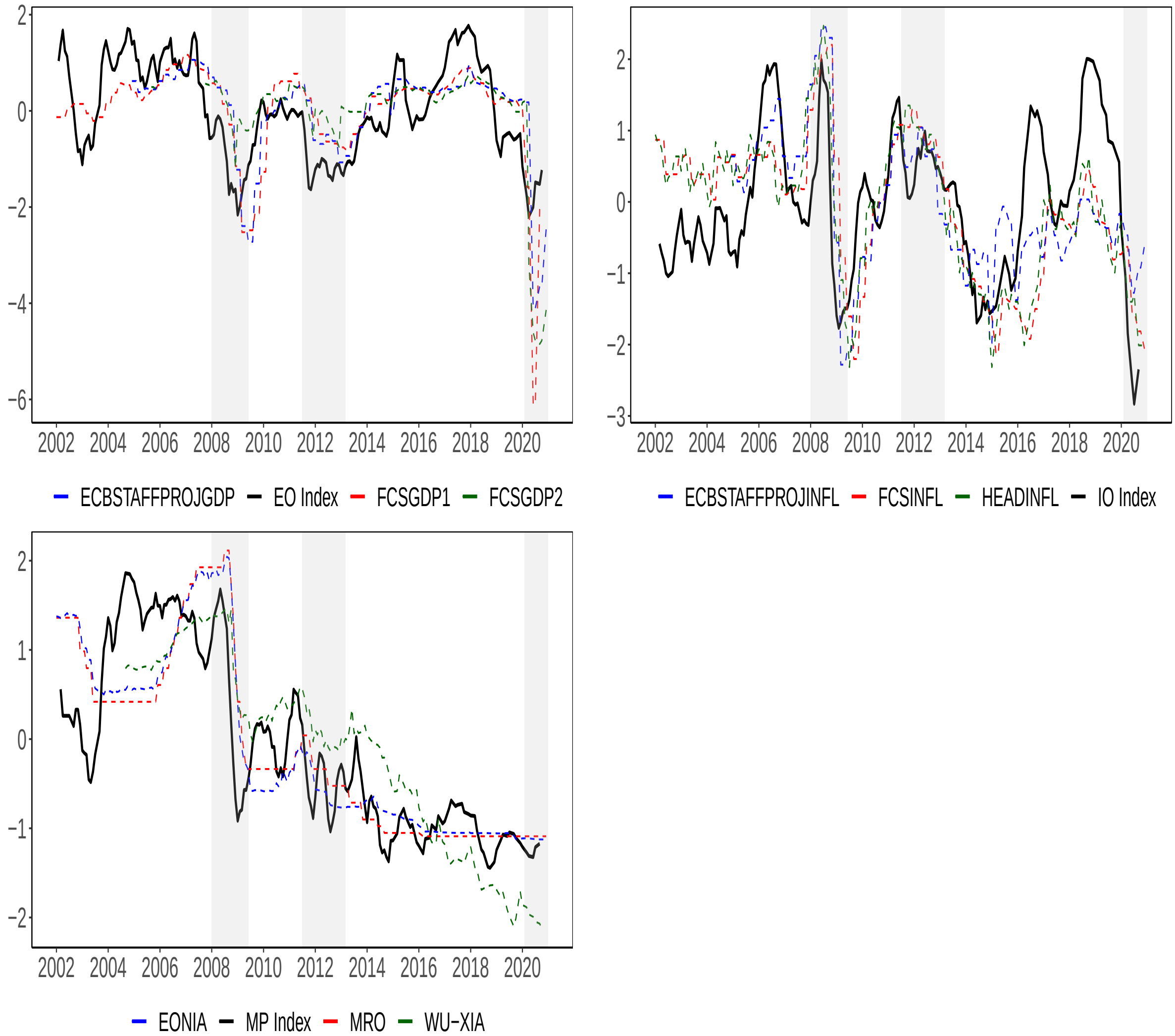
Note: The table shows a sample of n-grams and unigrams for each ECB field-specific dictionary.

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:

$$\psi_\tau = \frac{\gamma^n_{w,\tau}}{\sqrt{\sum_{n=1}^{N} w_n}} \tag{2.2}$$

where $\gamma^n_{w,\tau} = \sum[(1 + w_a + w_d) \cdot w^p_\gamma \cdot w^{neg}_\gamma]$ 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 $PC^d = \frac{1}{n}\sum\psi_\tau$. Figure 2.2 displays the evolution of the economic outlook, inflation outlook and monetary policy stance indexes against relevant macroeconomic data.

Figure 2.2: Indexes Based on the Introductory Statements of the Press Conferences



Note: The figure plots the time series of economic outlook ("EO Index"), inflation outlook ("IO Index") and monetary policy stance ("MP Index") against key macro variables. The variables are standardized to facilitate the comparison. Since the indexes are somewhat jagged, I show the moving average. The gray zones indicate CEPR-based crisis periods. "FCSGDP1" and "FCSGDP2", respectively, stand for the Bloomberg weighted-average of private and official GDP forecasts. While "ECBSTAFFPROJGDP" and "ECBSTAFFPROJINFL" are the ECB staff projections for GDP and inflation, "FCSINFL" is private inflation forecasts and "HEADINFL" is actual headline inflation. "EONIA", "MRO" and "WU-XIA" are self-explanatory.

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. Similarly, the inflation outlook index appears to approximate the development of the ECB (ECBSTAFFPROJ) and private (FCSINFL) inflation forecasts as well as actual headline inflation (HEADINFL). Lastly, the monetary policy stance index closely resembles the dynamics of the EONIA, the marginal refinancing operations (MRO) and the Wu-Xia ECB shadow rate.

2.2.2 Words as Data: Augmenting the Nowcasting Exercise with Textual Data

In addition to 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 2020[7]. This dataset includes newspapers' articles, online websites, magazines, TV news, etc. that explicitly mention the ECB in their content (see Appendix 2.6 for details).

To make this dataset applicable for time series analysis, I follow a similar procedure to the one employed in Section 2.2.1. After pre-processing the dataset, I decompose the textual corpus into news topics using LDA estimated with $K = 80$[8]. Out of 80 topics, I identify and label 60 of them which I reduce to 40 once filtering for monetary policy relevance[9]. 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"[10]. I refer to Appendix 2.6 for further details.

From the perspective of time series analysis, I consider these meta-topics differently. "Financial Crisis", "Eurexit" and "European Banks" enter the model in the form of topic probabilities retrieved from LDA. The intuition for such a choice is that the more a topic is represented in the textual dataset at a given point in time, the more likely it is that this topic signals something meaningful for the economy. Under this regard,

[7] "Internal" does not mean "private". In fact, the textual information is public. ECB staff merely translated every non-English article into English and every non-textual information (radio interviews and podcasts) into text. For further information on the dataset see Appendix 2.6

[8] 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).

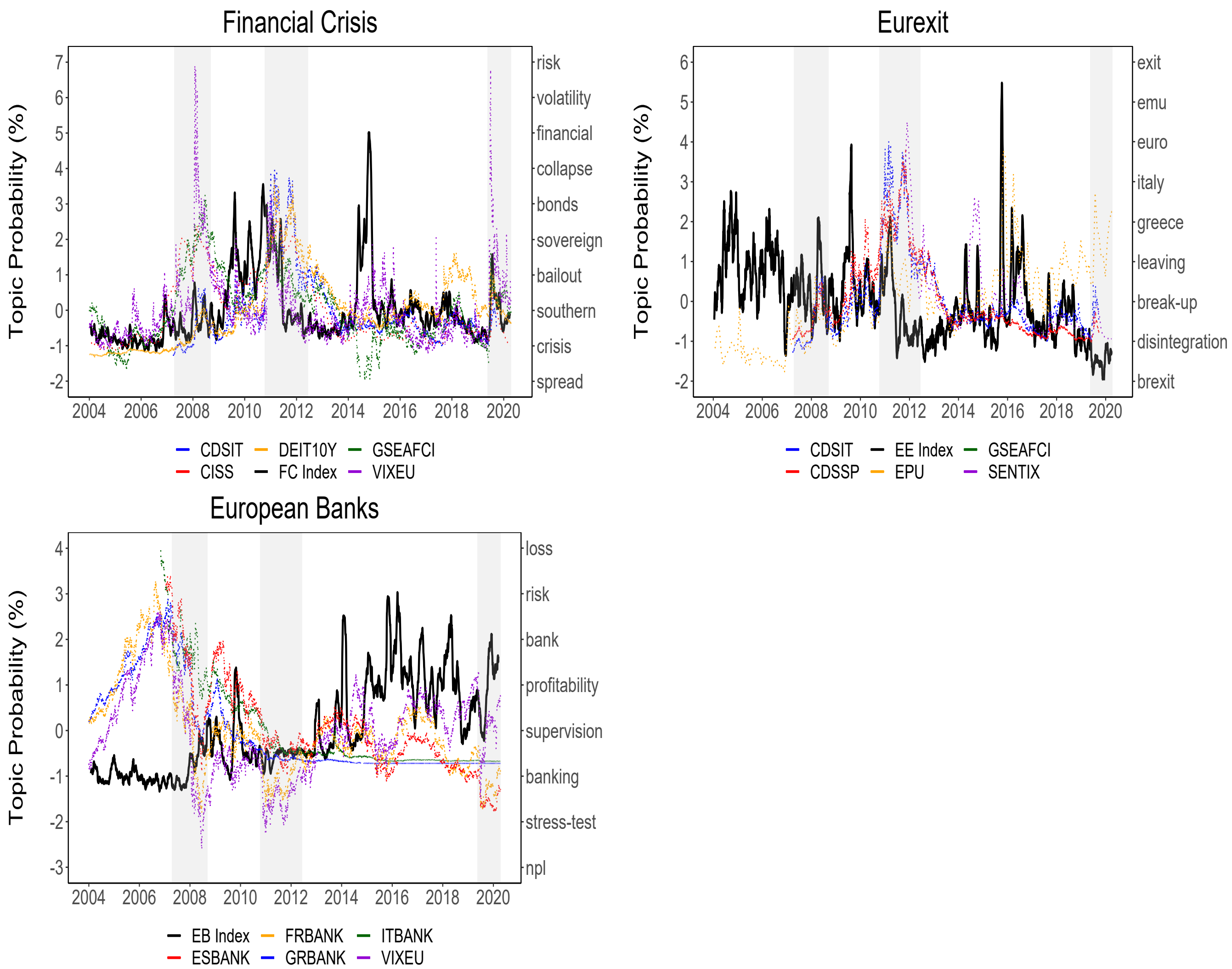
[9] I identify and label only 60 topics out of 80 since the remaining topics were not clearly identifiable.

[10] It should be noted that while the indexes for inflation and economic outlook in Section 2.2.1 represent the ECB internal assessment of those macro variables, the ones derived here reflect the tone of a much broader "audience".

none of these topics gains additional value from being tone adjusted. For example, if the topic probability for "Financial Crisis" increases and to the extent that the textual dataset provides a relevant description of the economy, one would reasonably expect that a financial turmoil might be looming. No further information is needed. Instead, "Inflation Outlook","Economic Outlook","Monetary Policy" and "Fiscal Policy" require a quantification step to be made. It is in fact not sufficient to know how much "inflation outlook" is being discussed to understand whether prices are rising or falling.

Figure 2.3 shows the evolution of the topic probability of "Financial Crisis", "Eurexit" and "European Banks".

Figure 2.3: Topic Probabilities for Daily Indexes

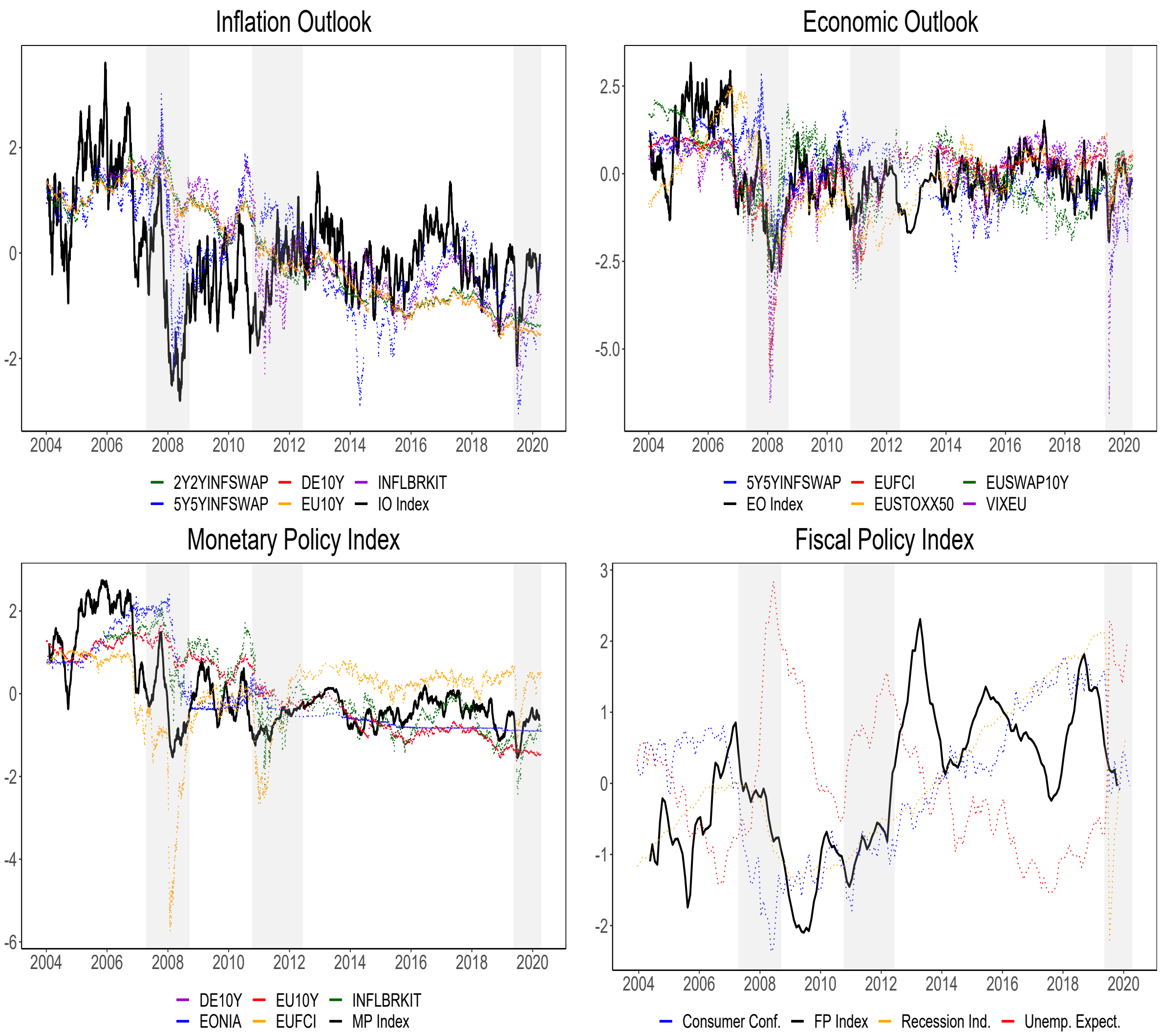


Note: The figure plots the evolution of the topic probabilities derived from LDA for "Financial Crisis" ("FC Index"), "Eurexit" ("EE Index") and "European banks" ("EB Index"). The variables are standardized to facilitate the comparison. The second y-axis shows the words with the highest probability to belong to a certain topic. Since the indexes are somewhat jagged, I display the moving average. The gray zones indicate CEPR-based crisis periods. While "CDSIT" and "CDSSP" indicate the credit default swaps for 10y government bond yields of Italy and Spain, "DEIT10Y" is the spread between the 10y BUND and BTP. "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 and "SENTIX" is the Sentix Investor Confidence. "ESBANK", "FRBANK" "GRBANK" and "ITBANK", respectively, stand for CAC 40, ATHEX Banks Index, IBEX 30 and FTSE-MIB. "CISS" is the composite indicator of systemic stress in the euro area and "VIXEU" the EU volatility Index.

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 indicators.

Moving to "Inflation Outlook","Economic Outlook","Monetary Policy" and "Fiscal Policy", these datasets are translated into tone adjusted time series. The procedure is similar to the one followed in Section 2.2.1. The only difference is that the field-specific dictionaries need to be extended to map a significantly bigger dataset compared to the ECB press conferences. To do so, I apply word embedding (Stoltz and Taylor, 2019); that is, a method to identify patterns of the co-occurrence of words within a "context window" centered around, respectively, "inflation", "economy", "monetary", and "fiscal", where these key words simply correspond to the topics of the field-specific dictionaries. I then compare the vector of unigrams just obtained with the field-specific dictionaries in Section 2.2.1. 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). Once the dictionaries were extended, I apply equation Equation 2.2 to score and aggregate each sentence for each content-specific dataset. As a final step, 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 2.4 presents the results.

Figure 2.4: Quantified Indexes Derived from the ECB Textual Database



Note: The figure plots the quantified indexes of inflation outlook ("IO Index"), economic outlook ("EO Index"), monetary policy index ("MP Index") and fiscal policy index ("FP Index") derived from the ECB textual dataset. The variables are standardized to facilitate the comparison. Since the indexes are somewhat jagged, I display the moving average. The gray zones indicate CEPR-based crisis periods. "DE10Y" and "EU10Y" are, the 10y government bond for Germany and Italy, respectively. While "2Y2YINFSWAP" and "5Y5YINFSWAP" indicate the 2y2y and 5y5y inflation swaps in the euro area, "INFLBRKIT" stands for the Italian 10Y inflation break-even rate. "EUFCI" is the OIS-LIBOR spread, "VIXEU" the EU volatility Index and "EUSWAP10Y" the Euro 10Y swaps. "Consumer Conf.", "Recession Ind." and "Unemp. Expect." are, respectively, the EU Consumer Confidence, the EU Recession Indicator (Household Survey measure of total employment) and the EU Employment Expectation Indicator. "EONIA" is self-explanatory.

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[11]. Instead, the index for fiscal policy appears to have its closest fit with the European Commission's Consumer Confidence Indicator.

2.2.3 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 periodically[12]. 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 2.2)[13]. For a complete breakdown of the dataset I refer to Appendix 2.7. 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 2.2.2[14]. As for the monthly frequency, while *output* spans

[11] Comparing the monetary policy index based on press conferences and the one based on the ECB textual dataset yields a correlation of 52%. The comparison is based on taking the monthly mean of daily observations for the higher frequency index.

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

[13] 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.

[14] To have forecasts in a unique release block, when forecasts were not available at a daily frequency, I

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 Section 2.2.1.

Table 2.2: 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.

filled missing observations with previous values.

2.3. Methodology

Section 2.3.1 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 2.3.2 explains the benchmark model employed to compare the performance of the DFM.

2.3.1 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 2.2.3 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:

$$\begin{bmatrix} Y_t^{i,f} \\ Y_t^{m,f} \\ Y_t^{m,s} \\ Y_t^{d} \end{bmatrix} = \begin{bmatrix} \tilde{\Lambda}^{i,f} & 0 & 0 & 0 \\ 0 & \tilde{\Lambda}^{m,f} & 0 & 0 \\ 0 & 0 & \Lambda^{m,s} & 0 \\ 0 & 0 & 0 & \Lambda^{d} \end{bmatrix} \begin{bmatrix} \tilde{F}_t^{i,f} \\ \tilde{F}_t^{m,f} \\ F_t^{m,s} \\ F_t^{d} \end{bmatrix} + \begin{bmatrix} E_t^{i,f} \\ E_t^{m,f} \\ E_t^{m,s} \\ E_t^{d} \end{bmatrix}, \tag{2.3}$$

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 $E_{i,t} = i.i.d.N(0, \sum_E)$ where $\sum_E$ is diagonal. I leave technical details on aggregators in Appendix 2.8. The transition equation with time-varying coefficients can then be written as:

$$\begin{bmatrix} I_{2r} & 0 & 0 & \mathcal{W}_t^{i,f} \\ 0 & I_{2r} & 0 & \mathcal{W}_t^{m,f} \\ 0 & 0 & I_r & \mathcal{W}_t^{m,s} \\ 0 & 0 & 0 & I_r \end{bmatrix} \begin{bmatrix} \tilde{F}_t^{i,f} \\ \tilde{F}_t^{m,f} \\ F_t^{m,s} \\ F_t^{d} \end{bmatrix} = \begin{bmatrix} \mathcal{I}_t^{i,f} & 0 & 0 & 0 \\ 0 & \mathcal{I}_t^{m,f} & 0 & 0 \\ 0 & 0 & \mathcal{I}_t^{m,s} & 0 \\ 0 & 0 & 0 & \Xi \end{bmatrix} \begin{bmatrix} \tilde{F}_{t-1}^{i,f} \\ \tilde{F}_{t-1}^{m,f} \\ F_{t-1}^{m,s} \\ F_{t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ U_t \end{bmatrix}, \tag{2.4}$$

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 $U_{i,t} \sim i.i.dN(0, \sum_U)$ where $\sum_U$ is a diagonal matrix.

The model is then estimated using the methodology developed by Bańbura and Modugno (2014) that adapts from **?** 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 2.8).

Equation 2.3 and Equation 2.4 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:

$$P(y_{t+h} = j|\hat{X}_{t+h}) = \Phi(\alpha + \beta' \hat{X}_{t+h})\,. \tag{2.5}$$

Equation 2.5 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 $t + h$[15]; 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 $t+1$[16]. This bridge equation thus yields updated nowcasts of

[15] Appendix 2.8 provides details on how I turned a continuous variable into a categorical one.

[16] Although the DFM makes predictions of all variables in the dataset, I narrow the selection to the following: core inflation, industrial production, M3, EONIA, 10Y BUND-BTP spread, trade-weighted exchange rate and forecasts of inflation and GDP. This choice is mainly to avoid overparametrization and estimation issues of the bridge equation.

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 available[17]. 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 as[18]:

$$\underbrace{\mathbb{E}(y_t^{ECB}|\Omega_t)}_{\text{new forecast}} = \underbrace{\mathbb{E}(y_t^{ECB}|\Omega_{t-1})}_{\text{old forecast}} + \mathbb{E}(y_t^{ECB}|\underbrace{I_{new}}_{\text{news}}) \,. \tag{2.6}$$

Note that in Equation 2.6, $\mathbb{E}(y_t^{ECB}|I_{new}) \neq 0$ only if $I_{new} = y_t^j - \mathbb{E}(y_t^j|\Omega_{t-1}) \neq 0$, that is, the nowcast is updated only if the new observations in y_t^j are different from the forecasts made at time $t-1$[19]. 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 follows[20]:

[17] 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.

[18] 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 horthogonality holds, the revision is also uncorrelated to the old forecast.

[19] 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

[20] 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.

$$\mathbb{E}(y_t^{ECB}|I_{new}) = \mathbb{E}(y_t^{ECB} I'_{new})\,\mathbb{E}(I_{new}I'_{new})^{-1} I_{new}\,, \tag{2.7}$$

where, for the state space model written as in Equation 2.3 and Equation 2.4:

$$\begin{aligned}
\mathbb{E}(y_t^{ECB} I'_{new}) &= \Lambda^{k,\cdot}\,\mathbb{E}[(F_t^{k,n} - F_{t|\Omega_{t-1}}^{k,n})(F_t^{k,n} - F_{t|\Omega_{t-1}}^{k,n})']\Lambda_j^{k,\cdot\prime} \\
\mathbb{E}(I_{new}I'_{new}) &= \Lambda_j^{k,\cdot}\,\mathbb{E}[(F_{tj}^{k,n} - F_{tj|\Omega_{t-1}}^{k,n})(F_{tj}^{k,n} - F_{tj|\Omega_{t-1}}^{k,n})']\Lambda_j^{k,\cdot\prime} + \sum\nolimits_{\mathrm{j}}\,,
\end{aligned} \tag{2.8}$$

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 $A = (a_1, \ldots, a_J)$ such that:

$$\underbrace{\mathbb{E}(y_t^{ECB}|\Omega_t)}_{\text{new forecast}} - \underbrace{\mathbb{E}(y_t^{ECB}|\Omega_{t-1})}_{\text{old forecast}} = A(\underbrace{y_t^j - y_t^j|\Omega_{t-1}}_{\text{news}})\,. \tag{2.9}$$

Equation 2.9 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 2.2.

2.3.2 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 2.2.1. 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)[21]. I then estimate the model in the following specification:

$$ECB_t = \alpha + \phi_\pi(\pi_t - \pi^*) + \phi_\gamma(\gamma_t - \gamma_t^*) + \phi_{\pi,h}\mathbb{E}_t\pi_{t+h} + \phi_{\gamma,h}\mathbb{E}_t\gamma_{t+h} + \epsilon_t \, , \tag{2.10}$$

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 ($h = 4$)[22]; 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.

2.4. Empirical Results

Section 2.4.1 presents the in-sample and out-of-sample results, while Section 2.4.2 attempts to validate the robustness of the model.

2.4.1 The Model at Work

The DFM is parametrized with one factor per frequency and variable type and one lag[23]. 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 perspectives.

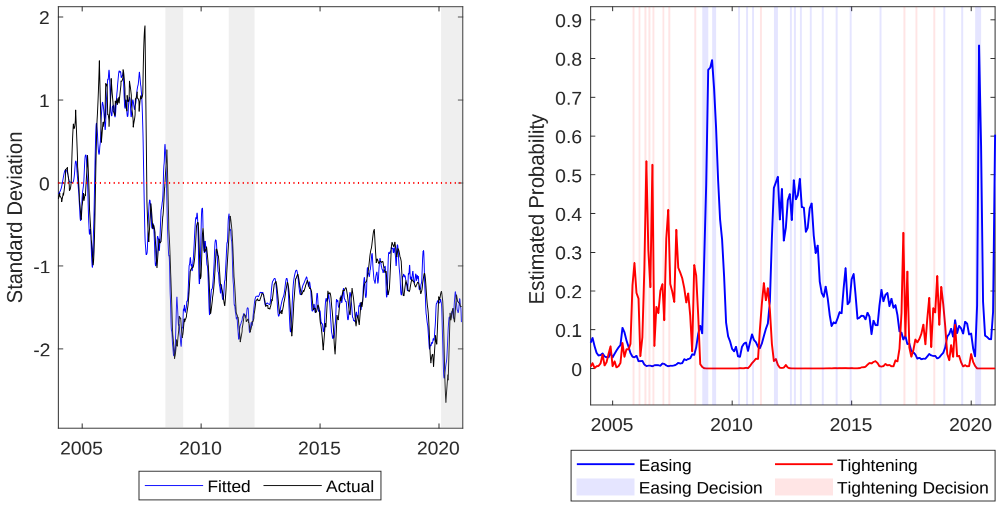
[21] Since the literature abounds with different Taylor rule specifications, the choice to focus on one single "benchmark model" may be excessively discretionary. For this reason, in Appendix 2.8 I compare the performance of the DFM with alternative benchmark models.

[22] 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-Prescrott (HP) filter.

[23] 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.

First, Figure 2.5 reports 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 2.5: 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). The gray zones indicate CEPR-based crisis periods.

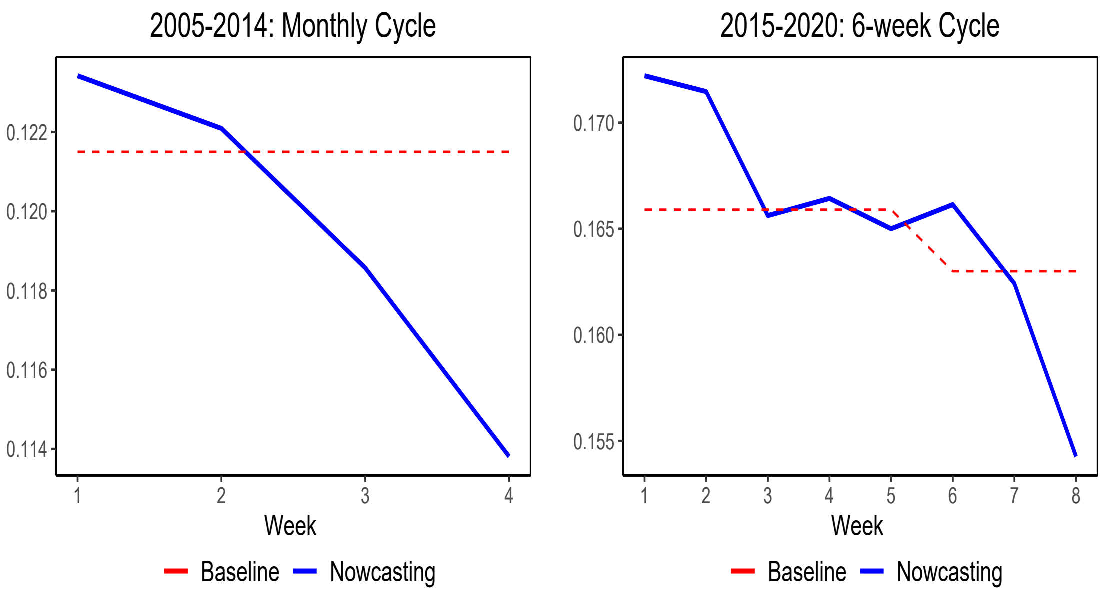
As the figure shows, the fitted values track closely the actual index of monetary policy stance. Besides, the estimated probability of tightening and easing significantly increase in correspondence of introductory statements where such decisions were taken. More precisely, the bridge equation correctly predicts 86% of tightening decisions and 90% of easing decisions[24].

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.

[24] To understand how these percentages were calculated see Appendix 2.8.

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 2.3.2. Since the frequency of the ECB press conferences is approximately monthly until 2014 before moving to a six-week cycle, I calculate the RMSFE for each subperiod. Figure 2.6 reports the result.

Figure 2.6: Root Mean Squared Forecast Error (RMSFE)



Note: The figure shows the RMSFE of the nowcasting model (solid blue lines) compared to the baseline model (dashed red lines) outlined in Section 2.3.2. 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 behavior, 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 2.3.2 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 2.3. The DFM, therefore, maintains an informational advantage compared to the benchmark model due to the ablity to process mixed-frequency data in a unified framework.

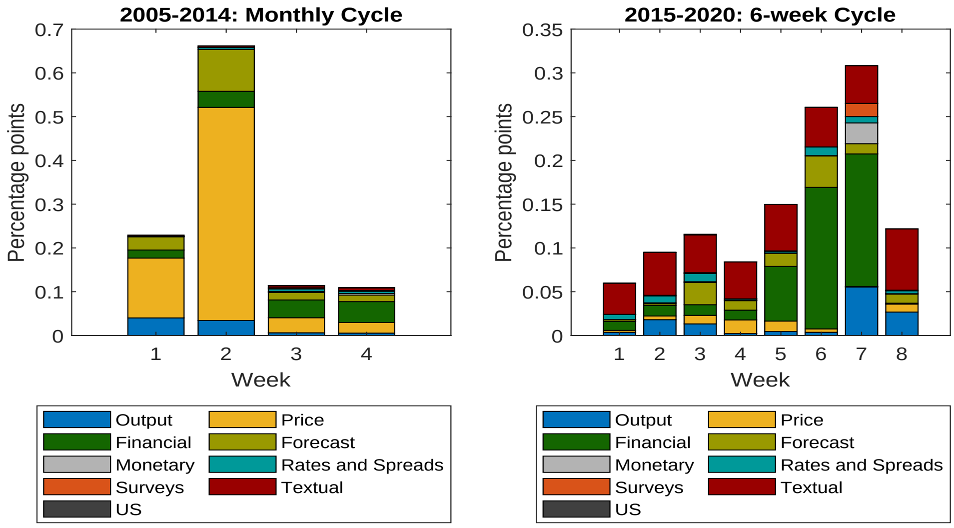
Table 2.3: Forecast Evaluation

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 2.7 documents the average absolute weekly impact of each block on the updated nowcast computed in real time. In other words, Figure 2.7 sheds light on the reaction function of the ECB.

Figure 2.7: 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, a period of renewed attention to ECB communication due to its systematic leverage on unconventional tools.

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 2.4. 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 not only bring new evidence to the usefulness of text-data in forecasting purposes but they also shed new light on the reaction function of the ECB. The model is in fact able to capture the prominent role of news broadcasted through the media in a period in which many unconventional monetary policy decisions have been launched. News is therefore proven to be an essential pass-through of expectation formation in the feedback loop between the ECB and the market.

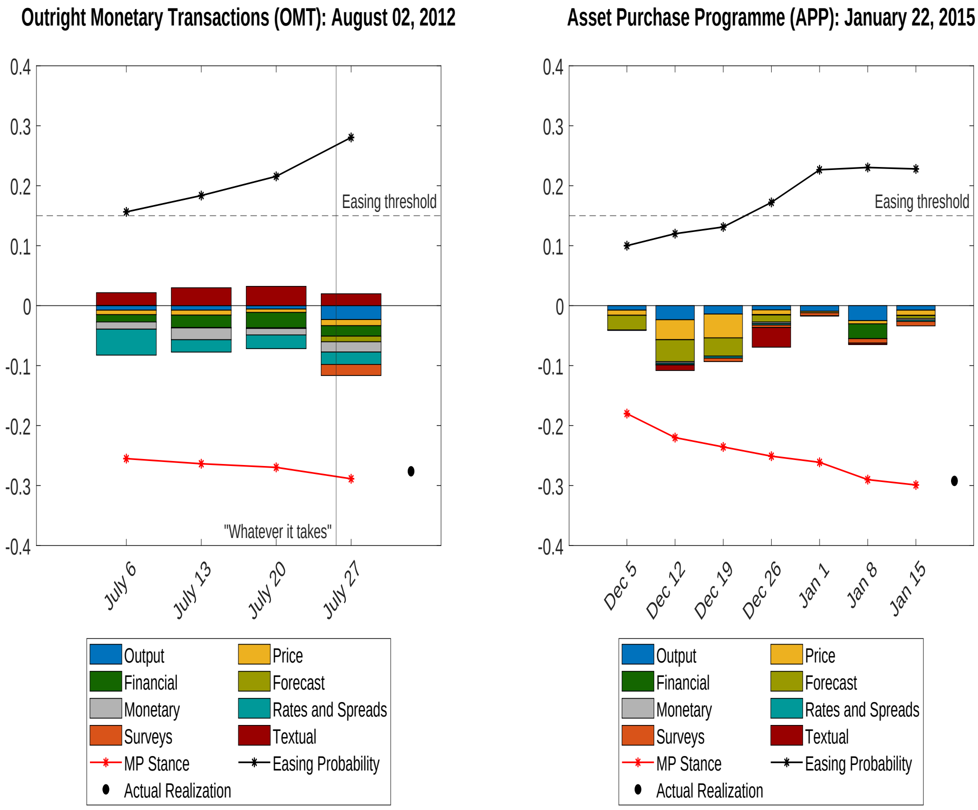
Table 2.4: Do Textual Variables Matter?

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 2.8 and Figure 2.9 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 2.8: 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. The vertical black line indicates then-President Draghi's "whatever it takes" at the Global Investment Conference in London on July 26 2012.

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 2.8), 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 differentials[25]. 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 data[26].

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 [$\cdots$] 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"[27]. Coherently, the first column of Figure 2.9 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 2.8), 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"[28], the second column of Figure 2.9 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.

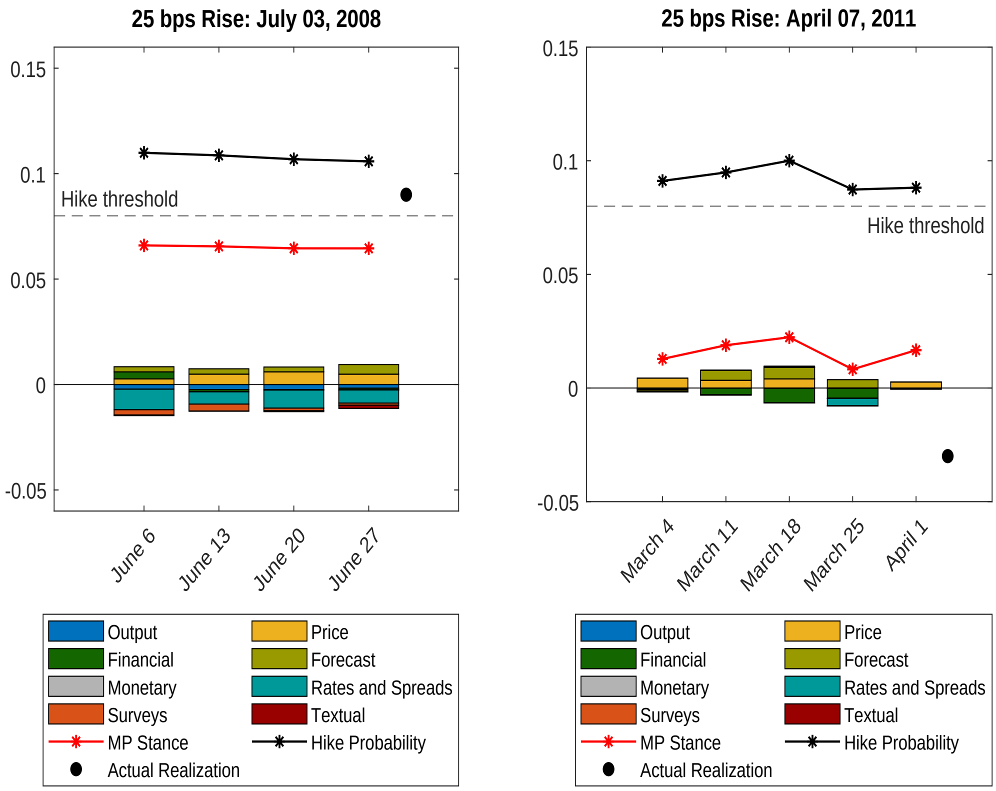
[25] The ECB Governing Council justified the OMTs stating that "the severe malfunctioning in the price formation process in the bond markets of euro area countries [$\cdots$] need to be addressed in a fundamental manner". *Introductory statement to the press conference (with Q&A)*, Mario Draghi, President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

[26] The ECB decided to launch the APP as a result of "inflation dynamics [$\cdots$] [being] weaker than expected [$\cdots$] [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.

[27] *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.

[28] *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 2.9: Nowcasting Historical Episodes - Tightening



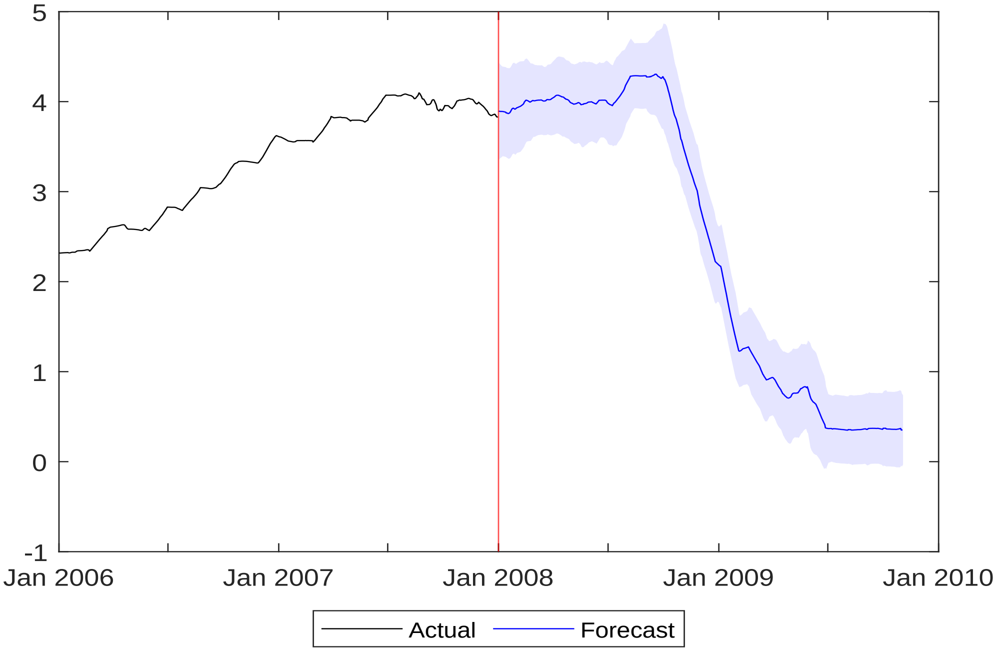
Note: The figure shows the evolution of the nowcast for the announcements of the 25bps rate increase on July 3, 2008 and on April 7, 2011. While the black line represents the nowcast of the expected probability of an easing decision, the red line indicates the nowcast of the monetary policy stance. News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

2.4.2 Validating the DFM

To validate the DFM, three exercises are carried out: first, nowcasting EONIA rates throughout the Great Recession; second, nowcasting the EU shadow rate (Wu and Xia, 2016) at different time windows in 2015; third, forecasting the EONIA, the EU shadow rate and a textual measure of monetary policy stance derived using Loughran and McDonald (2011)'s dictionary at three different points in time in 2005. 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.

Starting with the first exercise, I forecast, 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 variables[29]. 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 2.10 graphs what would have been the forecast for the EONIA rates conditional on the information available at the time of the analysis.

Figure 2.10: Forecasting EONIA



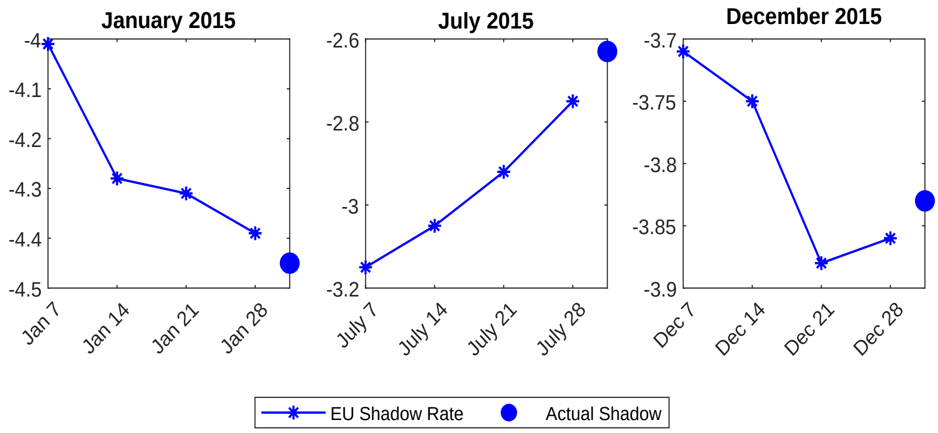
Note: The figure shows the actual values of EONIA rates (black line) and the out-of-sample nowcast with its 90% confidence intervals (blue line).

As the figure shows, the model proves to be valuable 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.

[29] 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.

Next, I forecast the EU shadow rate in January, July and December 2015. In such a period, in fact, the EU shadow rate was significantly different from the EONIA. Figure 2.11 displays the results.

Figure 2.11: Forecasting EU Shadow Rates



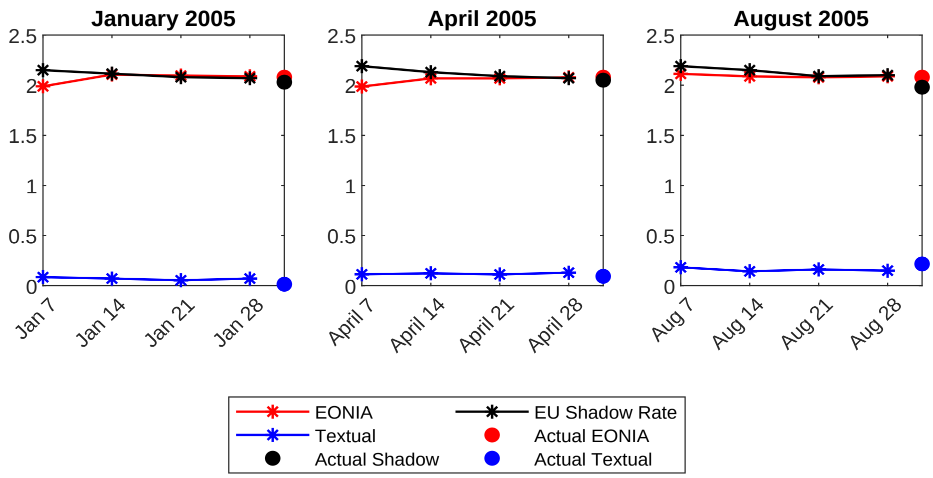
Note: The figure shows the out-of-sample nowcasts of EU shadow rates in Jnuary, July and December 2015.

The model appears to perform accurately also in tracking the developments of EU shadow rates at different windows in 2015.

The last exercise consists in forecasting out-of-sample EONIA rates (red line), EU shadow rates (black line) and an alternative textual measure of ECB monetary policy stance (blue line) in January, April and August 2005[30]. Figure 2.12 shows the results.

[30] The alternative textual measure should not be taken as representative of the ECB monetary policy stance since it suffers from the flaws outlined in Section 2.2.2

Figure 2.12: Additional Validation Exercises



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.

As in the previous exercises, the DFM proves its forecasting accuracy across the indexes under investigation.

It can therefore be concluded that, overall, the validation exercise 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, appears to track accurately the actual realizations of alternative monetary policy stance indicators.

2.5. 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 modelled as a dynamic factor model 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 from 2015 onwards, a period of renewed attention to central bank communication.

2.6. APPENDIX A: Text Mining

In what follows, I provide details on the estimation of the LDA model; compare the indexes derived from the press conferences with alternative dictionaries; assess the selection and identification of the topics; describe the textual dataset and shed additional light on the quantification of the time series extracted from the textual dataset.

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 $D = \{\boldsymbol{w}_1, \ldots, \boldsymbol{w_M}\}$ 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, $\boldsymbol{z}$ 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 $\boldsymbol{w}$ 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:

$$P(\mathbf{z}, \mathbf{w}; \alpha, \beta) = \frac{P(\mathbf{w},\mathbf{z};\alpha,\beta)}{P(\mathbf{w};\alpha,\beta)} \tag{2.11}$$

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:

$$P(\mathbf{z}, \mathbf{w}, \boldsymbol{\Theta}, \boldsymbol{\Phi}; \alpha, \beta) = \prod_{k=1}^{K} P(\boldsymbol{\varphi_k}; \beta) \prod_{m=1}^{M} P(\boldsymbol{\theta_m}; \alpha) \prod_{n=1}^{N} P(z_{m,n}|\boldsymbol{\theta_m}) P(w_{m,n}|\boldsymbol{\varphi_k}) \tag{2.12}$$

Integrating out the parameters $\boldsymbol{\varphi}$ and $\boldsymbol{\theta}$:

$$P(\mathbf{z}, \mathbf{w}; \alpha, \beta) = \int_{\Theta} \int_{\Phi} P(\mathbf{z}, \mathbf{w}, \boldsymbol{\Theta}, \boldsymbol{\Phi}; \alpha, \beta) d\boldsymbol{\Phi} d\boldsymbol{\Theta}$$

$$= \int_{\Phi} \prod_{k=1}^{K} P(\boldsymbol{\varphi_k}; \beta) \prod_{m=1}^{M} \prod_{n=1}^{N} P(w_{m,n}|\boldsymbol{\varphi_k}) d\boldsymbol{\Phi} \int_{\Theta} \prod_{m=1}^{M} P(\boldsymbol{\Theta_m}; \alpha) \prod_{n=1}^{N} P(z_{m,n}|\boldsymbol{\Theta_m}) d\boldsymbol{\Theta} \tag{2.13}$$

In Equation 2.13, the first integral does not include $\boldsymbol{\theta}$ nor the second integral contains $\boldsymbol{\varphi}$. As a result, $\boldsymbol{\varphi}$ and $\boldsymbol{\theta}$ can be solved separately. Thus, drawing on the properties of the conjugate Dirichlet distribution it can be shown that:

$$\int_{\Phi} \prod_{k=1}^{K} P(\boldsymbol{\varphi_k}; \beta) \prod_{m=1}^{M} \prod_{n=1}^{N} P(w_{m,n}|\boldsymbol{\varphi_k}) d\boldsymbol{\Phi} = \frac{\Gamma(\sum_{k=1}^{K} \alpha_k)}{\prod_{k=1}^{K} \Gamma(\alpha_k)} \frac{\prod_{k=1}^{K} \Gamma(n_m^{(k)} + \alpha_k)}{\Gamma(\sum_{k=1}^{K} n_m^{(k)} + \alpha_k)} \tag{2.14}$$

and

$$\int_{\Theta} \prod_{m=1}^{M} P(\boldsymbol{\Theta_m}; \alpha) \prod_{n=1}^{N} P(z_{m,n}|\boldsymbol{\Theta_m}) d\boldsymbol{\Theta} = \prod_{k=1}^{K} \frac{\Gamma(\sum_{t=1}^{V} \beta_t)}{\prod_{t=1}^{V} \Gamma(\beta_t)} \frac{\prod_{m=1}^{M} \Gamma(n_k^{(t)} + \beta_t)}{\Gamma(\sum_{t=1}^{V} n_k^{(t)} + \beta_t)} \tag{2.15}$$

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 2.11 $P(\mathbf{w}; \alpha, \beta)$ 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:

$$P(z_{(m,n)}|\mathbf{z}_{\text{-}(m,n)}, \mathbf{w}; \alpha, \beta) = \frac{P(z_{(m,n)}, \mathbf{z}_{\text{-}(m,n)}, \mathbf{w}; \alpha, \beta)}{P(\mathbf{z}_{\text{-}(m,n)}, \mathbf{w}; \alpha, \beta)} \tag{2.16}$$

where $z_{(m,n)}$ denotes the hidden variable of the n^{th} word token in the m^{th} document, and $z_{\text{-}(m,n)}$ denotes all z but $z_{(m,n)}$. Denoting the index of a word token by $\iota = (m, n)$, and using Equation 2.14 and Equation 2.15, cancellation of terms (and some extra manipulations exploiting the properties of the gamma function) yields:

$$P(z_\iota = k|\mathbf{z}_{(\iota)}, \mathbf{w}; \alpha, \beta) \propto (n^{(k)}_{m,\text{-}\iota} + \alpha_k) \frac{n_{k,\text{-}\iota} + \beta_t}{\sum_{t=1}^{V} n^{(t)}_{k,\text{-}\iota} + \beta_t} \tag{2.17}$$

where the counts $n^{(\cdot)}_{\cdot,\text{-}\iota}$ indicate that token ι is excluded from the corresponding document or topic. Therefore, sampling topic indexes using Equation 2.17 for each word in a document and across documents until convergence allows us to approximate the posterior distribution given by Equation 2.11.

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 exercise[31]. 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 jut described, but

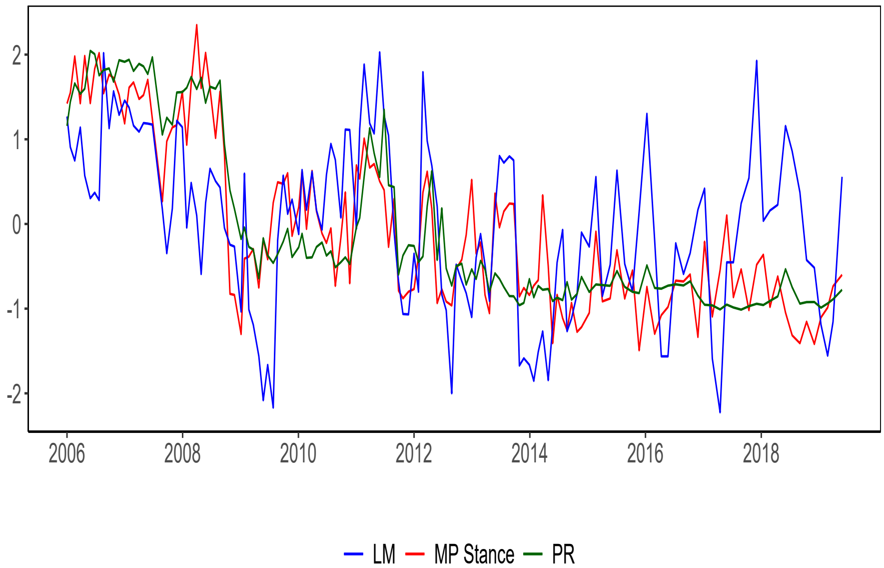
[31]Updating the topic model recursively is infeasible, as one estimation round takes approximately 10 days using my infrastructure. Thus, at the end of the evaluation period, I am essentially using topic distributions based on 10-year old estimates. Another caveat with re-estimating the LDA recursively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

with the difference that the sampler is run with the estimated word distributions (from the training sample) held constant.

Alternative Tone's Measures

Figure 2.13 compares the monetary policy stance index derived in Section 2.2.1 with alternative measures available in the literature.

Figure 2.13: Alternative Tone's Measures



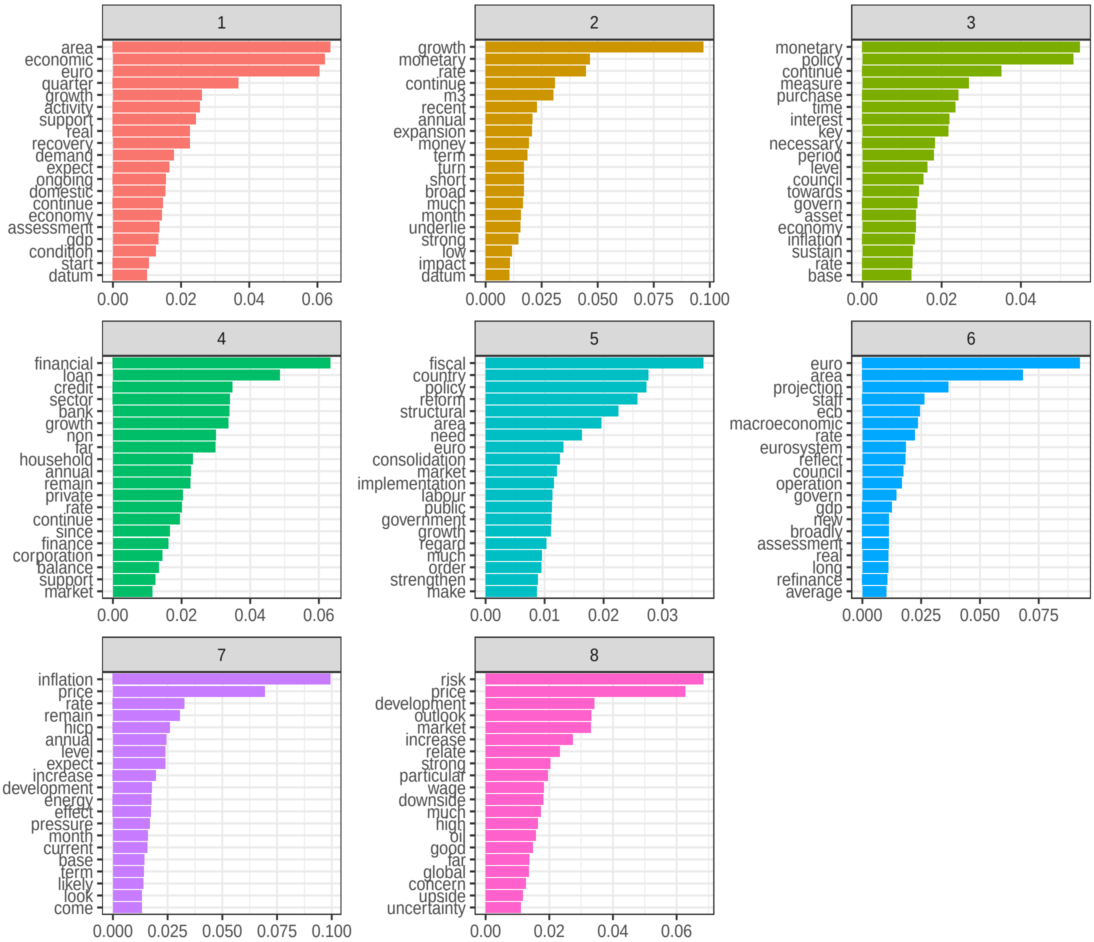
Note: The figure shows the evolution of the monetary policy stance (red) index against alternatives specifications available in the literature.

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.

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 2.14 illustrates the probability that each word belongs to a specific topic[32]:

Figure 2.14: 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 2.14, I label the topics as follows[33]:

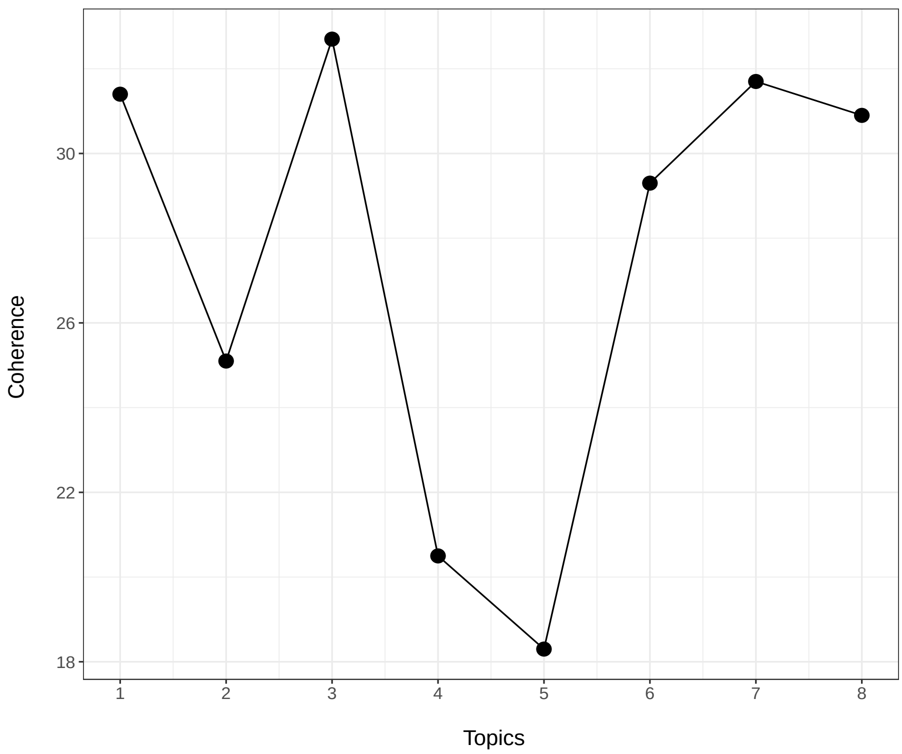
[32] Besides quantitative and graphical evidence, I checked qualitatively the plausibility of the identification strategy. In fact, once the probability of a topic to belong to a paragraph is retrieved, it is straightforward to investigate whether the content of the paragraph matches the labeling. Upon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

[33] 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 2.15.

Figure 2.15: Coherence within Topics

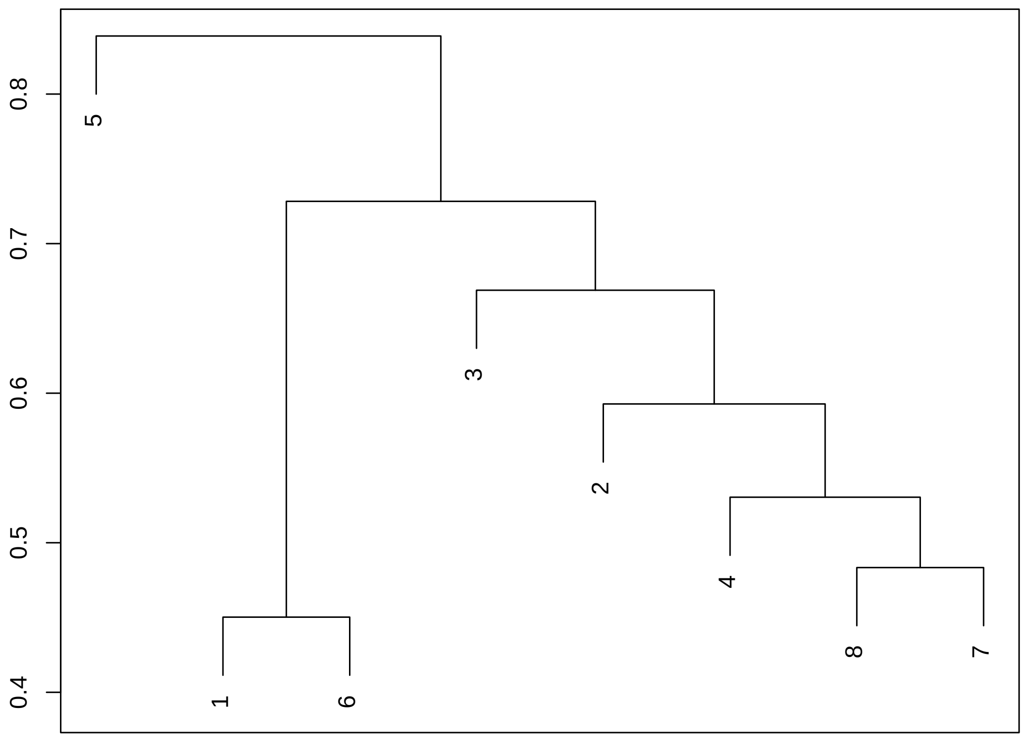


Note: The figure shows the coherence score for each topic.

Figure 2.15 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"[34].

Finally, Figure 2.16 shows a cluster dendogram that uses Hellinger distance (distance between two probability vectors) to decide whether the topics are closely related.

Figure 2.16: Cluster Dendogram



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

The dendogram, 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,

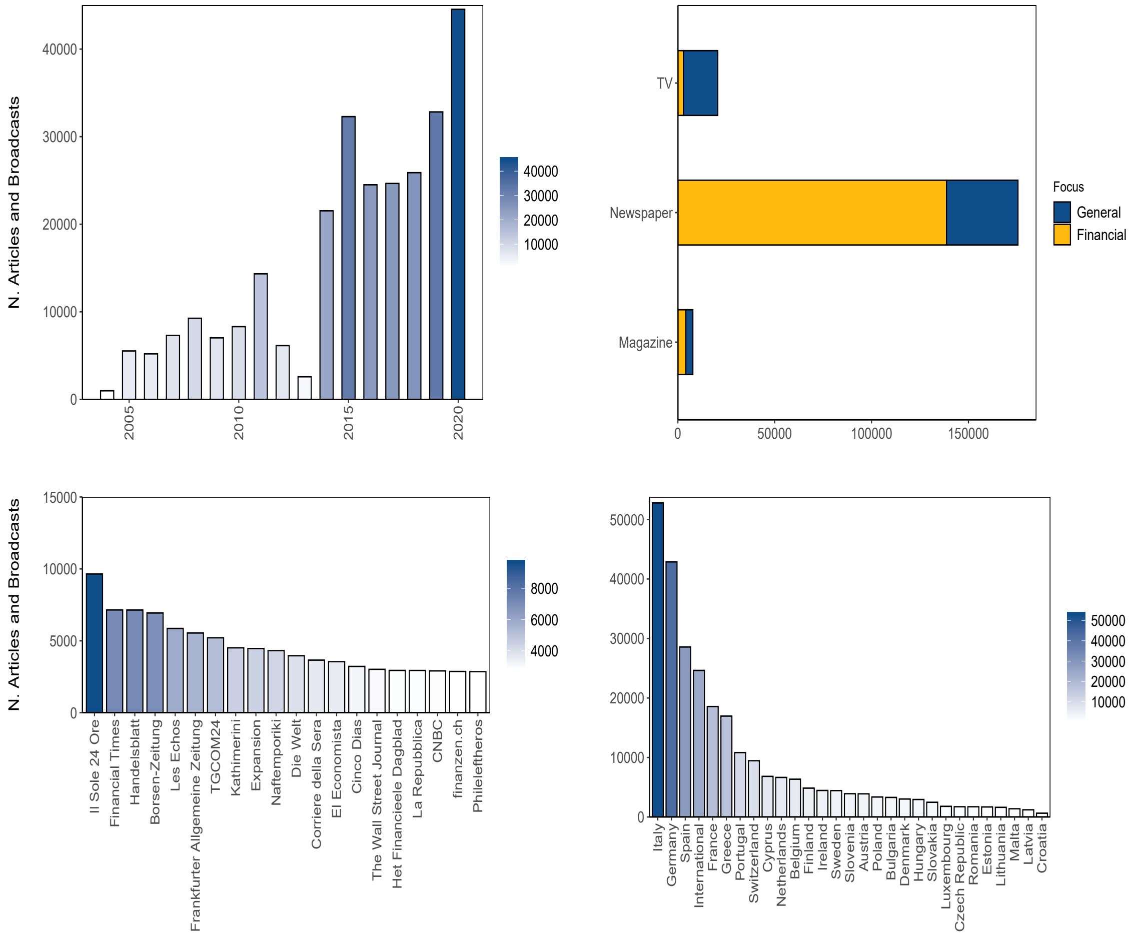
[34] It is not surprising that Topic 4 and Topic 5 have lower score. In fact, the words listed in "Financial Conditions" span from market-related words to households and banks. Similarly, "Fiscal and Structural Reforms" intersect both fiscal and labour related issues.

likewise Topic 1 and 6.

The Textual Dataset

In Figure 2.17, I provide some descriptive statistics of the textual dataset collected in December 2020 from an ECB internal database. The database consists of any communication event (mainly but not only newspapers) that mentioned the ECB in its content. The database is prepared and constantly updated by ECB staff who translates and summarizes any non-English article, TV news etc. Therefore, the database contains original English articles as well as detailed translated summaries of non-English documents. For example, if the Financial Times publishes an article on the ECB, this article is stored in the database in its original form since its first language is English. On the other hand, if Il Sole 24 Ore releases an article on the ECB in Italian, ECB staff will summarize it in English and store it in the database.

Figure 2.17: 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.

The database 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.

The main strength of the database is that it contains only articles that explicitly refer to ECB matters. Moreover, a non negligible advantage from a practical perspective is the possibility to have access to only English documents despite articles may be originating from countries where English is not the first language. There are also a few shortcomings that are worth highlighting. First, the volume of documents is, so to speak, "heteroskedastic". In fact, it has a visible structural break in 2015. Second, non-English documents are translated and summarized by ECB staff. This implies that a translation, for as much as it can be accurate, is necessarily different from the original source.

However, I don't expect the weaknesses of the database to have biased the analysis for two reasons: first, I didn't directly use volumes in the paper. As long as the number of textual documents is large enough (an assumption that seems to hold except for 2004 and supposedly 2014), results should have avoided any sampling bias. Second, considering the results in Figure 2.3 and Figure 2.4, it does not seem that the mediation of ECB staff on the news played a negative role.

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 2.2.1. However, in Section 2.2.1 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 neighbor" 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':

$$WMD_{ij} = \min_{\mathbf{T} \geq 0} \sum_{i,j=1}^{n} T_{ij} c(i,j) \tag{2.18}$$

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:

$$\sum_{i=1}^{n} T_{ij} = d_i \text{ }, \forall_i \in \{1, \dots, n\} \tag{2.19}$$

$$\sum_{j=1}^{n} T_{ij} = d'_j \text{ }, \forall_i \in \{1, \dots, n\} \tag{2.20}$$

where Equation 2.19 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 2.20 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 2.19 removed, and then once with just the constraint from Equation 2.20 removed. The RWMD for each i, j pair is then calculated twice following Equation 2.18, and the final reported RWMD score for the i, j pair is:

$$RWMD_{ij} = \max(\min_{\mathbf{T}\geq 0} \sum_{i,j=1}^{n} T_{ij}c(i,j), \min_{\mathbf{T}\geq 0} \sum_{i,j=1}^{n} T'_{ij}c(i,j)) \tag{2.21}$$

In the paper, for example, I applied Equation 2.21 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 2.2.1. These words are then manually classified as “positive” or “negative” using discretion with regard to the topic of interest.

2.7. APPENDIX 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
HWWI 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

Series	**Transformation**	**Publication Lag**	**Frequency**
Spain CLIFS	1	0	M
Portugal CLIFS	1	0	M
Business Climate Indicator	1	0	M
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
BARCLAYSMBSFIXXRATE	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

Series	Transformation	Publication Lag	Frequency
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
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

Series	Transformation	Publication Lag	Frequency
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
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

Series	**Transformation**	**Publication Lag**	**Frequency**
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
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

Series	Transformation	Publication Lag	Frequency
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
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

Series	**Transformation**	**Publication Lag**	**Frequency**
Treasury yields 10Y	2	0	D

2.8. APPENDIX C: State Space Representation and Estimation

In what follows, I explain the temporal aggregation implemented in the Dynamic Factor Model, provide details on the Expectation Maximization (EM) algorithm, shed light on the bridge equation and consider alternative benchmark models to compare the DFM with.

Temporal Aggregation

Letting bold-font variables indicate the vector version of the variables, I write the model in Equation 2.3 and Equation 2.4 compactly as follows:

$$\mathbf{Y_t} = \boldsymbol{\Lambda}_t \mathbf{F_t} + \mathbf{e_t} \tag{2.22}$$

$$\mathbf{F_t} = \mathbf{A_t} \mathbf{F_{t_1}} + \mathbf{u_t} \tag{2.23}$$

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 **?** for an overview). For a generic variable Y_t^k, time aggregation from higher to lower frequency is restricted as follows:

$$\begin{aligned} y_t^k &= \log(v_{1,t}^k) - \log(v_{1,t-k}^k) \\ &\approx \log(\sum_{i=0}^{k-1} v_{1,t_i}^k) - \log(\sum_{i=k}^{2k-1} v_{1,t-i}^k) \\ &\approx \sum_{i=0}^{k-1} \log(v_{1,t_i}^k) - \sum_{i=0}^{2k-1} \log(v_{1,t-i}^k) \\ &= \sum_{i=0}^{2k-2} \omega_i^k y_{1,t-i} \, , t = k, 2k, \dots , \end{aligned} \tag{2.24}$$

where y_t^k is the observed low-frequency growth rate, v_t^k its level, and $\omega_i^k = i+1$ for $i = 1, \ldots, k-1$; $\omega_i^k = 2k - i - 1$ for $i = k, \ldots, 2k-2$; and $\omega_i^k = 0$ otherwise. It follows from Equation 2.24 that imposing a common factor structure to y_t^k gives:

$$y_t^k = \sum_{i=0}^{2k-2} \omega_i^k y_{1,t-i} = \sum_{i=0}^{2k-2} \omega_i^k (\Lambda F_{t-i}^d + e_{t-i}) \tag{2.25}$$

However, Equation 2.25 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:

$$F_t^k = \sum_{i=0}^{2k-2} \omega_i^k F_{t-i} \quad t = k, 2k, \ldots \tag{2.26}$$

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:

$$\tilde{F}_t^{k,f} = \begin{pmatrix} \bar{F}_t^{k,f} \\ F_t^{k,f} \end{pmatrix} = \begin{cases} \begin{pmatrix} \bar{F}_t^{k,f} + \omega_{k-1}^{k,f} F_t \\ 0 \end{pmatrix}, & t = 1+k+1, 2k+1, \ldots \\ \begin{pmatrix} F_{t-1}^{k,f} + \omega_{R(k-t,k)}^{k,f} F_t \\ 0 \end{pmatrix}, & \text{otherwise.} \end{cases} \tag{2.27}$$

For the stock variables only single aggregator variable is necessary:

$$F_t^{k,s} = \begin{cases} \omega_{k-1}^{k,s} F_t \, , & t = 1+k+1, 2k+1 \\ F_{t-1}^{k,s} + \omega_{k-1}^{k,s} F_t \, , & \text{otherwise.} \end{cases} \tag{2.28}$$

This is implemented via the transition Equation 2.4 with the following weight vector $\mathcal{W}_t^k$ and selection matrix $\mathcal{I}_t^k$:

$$\mathcal{W}_t^{k,f} = \begin{cases} \begin{pmatrix} -\omega_{k-1}^{k,f} \\ 0 \end{pmatrix}, & t = 1 + k + 1, 2k + 1, \dots \\ \begin{pmatrix} -\omega_{R(k-t,k)}^{k,f} \\ -\omega_{R(k-t,k)+k}^{k,f} \end{pmatrix}, & \text{otherwise.} \end{cases} \tag{2.29}$$

$$\mathcal{I}_t^{k,f} = \begin{cases} \begin{pmatrix} 0 & I_r \\ 0 & 0 \end{pmatrix}, & t = 1 + k + 1, 2k + 1, \dots \\ I_{2r}\ , \text{otherwise,} \end{cases} \tag{2.30}$$

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, Equation 2.27-Equation 2.30 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:

$$z_t^k = \frac{k}{k_t} \sum_{i=0}^{k_t - 1} z_{t-i}\ , t = k_1, k_1 + k_{k_1+1}, \dots, \tag{2.31}$$

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:

$$\gamma_t^k = k\left(\sum_{i=0}^{k_t-1} \tfrac{i+1}{k_t}\gamma_{t-i} + \sum_{i=k_t}^{k_t+k_{t-k_t}-2} \tfrac{k_t+k_{t-k}-i-1}{k_{t-k}}\gamma_{t-i}\right), t = k_1, k_1 + k_{k_1+1}, \ldots, \tag{2.32}$$

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 $i = 0, 1, \ldots, k_{t-1}$, $\omega_{t,i}^{k,f} = k\frac{k_t+k_{t-k_t}-i-1}{k_{t-k_t}}$ for $i = k_t, k_{t+1}, \ldots, k_t + k_{t-k_t} - 2$ and $\omega_{t,i}^{k,f} = 0$ otherwise.

The Expectation Maximization Algorithm

To describe the Expectation Maximization (EM) algorithm for the daily model in Equation 2.3 and Equation 2.4, I follow once again Bańbura and Modugno (2014). Assuming one lag in the factor VAR, I can group the parameters as follows: $\theta = (\mu, \Lambda, \Xi, \sum_E, \sum_U)$, where the only restriction is that $\sum_E$ 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 $Y_t^{k,n} = (Y_1^{k,n}, Y_2^{k,n}, \ldots, Y_{Tv}^{k,n})$ and $F_t^{k,n} = (F_1^{k,n}, F_2^{k,n}, \ldots, F_{Tv}^{k,n})$ 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:

$$\begin{aligned} E - step &: L(\theta, \theta(j)) = E_\theta[l(Y_t^{k,n}, F_t^{k,n}; \theta)|\Omega_v], \\ M - step &: \theta_{j+1} = \underset{\theta}{\operatorname{argmax}}\, L(\theta, \theta(j)). \end{aligned} \tag{2.33}$$

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:

$$\begin{aligned} vec(\Lambda^{k,\cdot}(j+1)) &= \left(\sum_{t=1}^{Tv} E_{\theta(j)}[Y_t^{k,n} F_t^{k,n'}|\Omega_v] \otimes \mathcal{S}_t\right)\left(\mathcal{S}_t \sum_{t=1}^{Tv} E_{\theta(j)}[Y_t^{k,n} F_t^{k,n'}|\Omega_v]\right)^{-1}, \\ \Phi(j+1) &= \left(\sum_{t=1}^{Tv} E_{\theta(j)}[F_t^{k,n} F_{t-1}^{k,n'}|\Omega_v]\right)\left(\sum_{t=1}^{Tv} E_{\theta(j)}[F_{t-1}^{k,n} F_{t-1}^{k,n'}|\Omega_v]\right)^{-1}. \end{aligned} \tag{2.34}$$

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:

$$\begin{aligned}\sum_E(j+1) = diag(&\sum_{t=1}^{Tv}(\mathcal{S}_t Y_t^{k,n} Y_t^{k,n'} \mathcal{S}_t' - \mathcal{S}_t Y_t^{k,n} E_{\theta(j)}[F_t^{k,n'}|\Omega_v]\Lambda^{k,\cdot}(j+1)'\mathcal{S}_t \\ &- \mathcal{S}_t\Lambda^{k,\cdot}(j+1)E_{\theta(j)}[F_t^{k,n}|\Omega_v]Y_t^{k,n'}\mathcal{S}_t + \mathcal{S}_t\Lambda^{k,\cdot}(j+1)E_{\theta(j)}[F_t^{k,n}F_t^{k,n'}|\Omega_v]\Lambda^{k,\cdot}(j+1)'\mathcal{S}_t \\ &+ (I_N - \mathcal{S}_t)\sum_{E(j)}(I_N - \mathcal{S}_t))),\end{aligned} \tag{2.35}$$

and

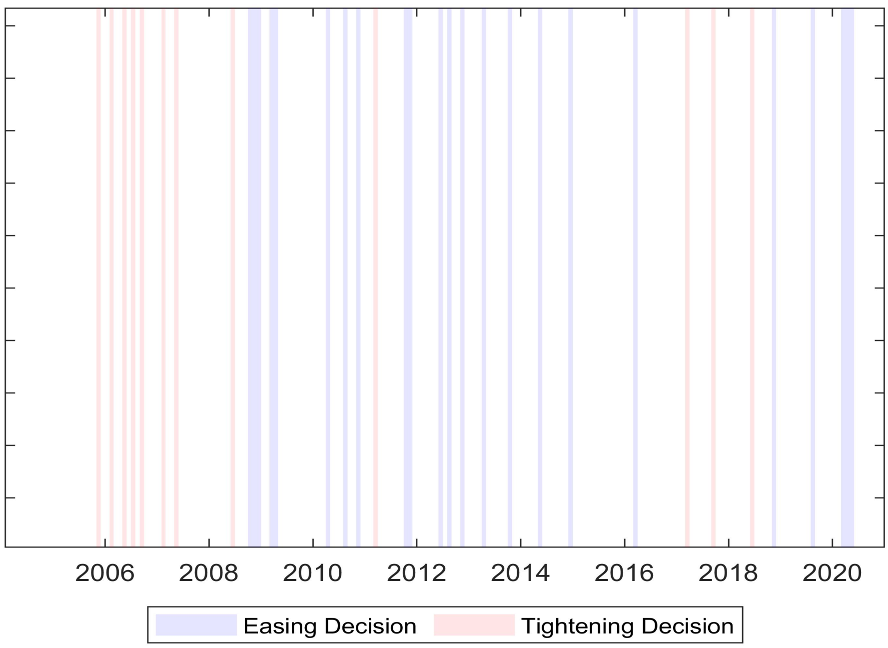
$$\sum\nolimits_U(j+1) = \tfrac{1}{T}\left(\sum_{t=1}^{Tv} E_{\theta(j)}[F_t^{k,n}F_t^{k,n'}|\Omega_v] - \Xi(j+1)\sum_{t=1}^{Tv} E_{\theta(j)}[F_t^{k,n}F_t^{k,n'}|\Omega_v]\right) \tag{2.36}$$

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 $F_t^{k,n}$.

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 2.18 displays the tightening monetary policy decisions (shaded red) and the easing ones (shaded blue).

Figure 2.18: 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:

$$MAE = \frac{1}{T}\sum_{t=1}^{T}(|y_t - P(y_t = j)|) \tag{2.37}$$

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:

$$R^2 = \frac{n_{tt}}{n_t} \cdot \frac{n_t}{n} + \frac{n_{ee}}{n_e} \cdot \frac{n_e}{n} + \frac{n_{cc}}{n_c} \cdot \frac{n_c}{n} = \frac{n_{tt}+n_{ee}+n_{cc}}{n} \tag{2.38}$$

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 $D : [0, 1] \rightarrow \{0, 1\}$, $\hat{y}_t = D[P(y_t = j)]$, 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 $p_N \vee p_T \approx 8\%$.

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 models[35]. 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:

$$ECB_t = \rho ECB_{t-1} + (1-\rho)(r^* + \pi^* + \beta_\pi(\mathbb{E}_t[\pi_{t+h}] - \pi^*)) + \beta_y[y_{t+k} - y_{t+k-1}] \tag{2.39}$$

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 slack[36].

Based on Equation 2.39, I estimate in total 1,290 policy rules, exploiting all plausible combinations of 43 measures of economic activity and 30 indicators of inflation expectations (see below)[37].

[35] "Pseudo" because my dependent variable is not "interest rate" but the ECB monetary policy stance series derived above.

[36] y^* is estimated, as in the main text, with the HP filter.

[37] 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 Prodution 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)

Series
Foreign industrial production growth forecast (US, JP, UK)
Next Exports
Gross Fixed Capital Formation
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

Series
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)
Consumer Price Index (CPI)
CONSUMER PRICES OF PETROLEUM PRODUCTS (PUMP PRICES)
HWWI 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 2.8 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 decision[38].

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

Table 2.8: 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	-0.51	1.20	0.90	0.2017	0.9516
			(0.03)	(0.34)	(0.39)	(0.20)		
Forward-looking	46	29.9%	0.91	-1.05	2.11	1.10	0.2134	0.9598
			(0.04)	(0.38)	(0.86)	(0.19)		
Backward-looking	51	32.4%	0.92	-0.36	0.79	0.92	0.2001	0.9511
			(0.03)	(0.32)	(0.14)	(0.25)		
Mixed Rules	60	38.2%	0.91	-0.61	1.52	0.96	0.2211	0.9522
			(0.04)	(0.39)	(0.61)	(0.20)		

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 2.8. Table 2.9 reports the results of the Diebold-Mariano test statistic (Diebold and Mariano, 1995).

Table 2.9: Forecast Evaluation for Alternative Benchmark Models

2005-2014				
H	1	2	3	4
$DFM\backslash Eonia_1$	0.75	1.22	1.90*	2.61***
$DFM\backslash Eonia_2$	1.14	1.45	2.01**	2.37**
$DFM\backslash Eonia_3$	1.22	1.38	1.79*	3.11***
$DFM\backslash Eonia_4$	0.98	1.13	1.59	2.41**
$DFM\backslash Eonia_5$	0.67	0.94	1.55	2.78***
$DFM\backslash Eonia_{all}$	1.55	1.58	1.73*	2.26**
$DFM\backslash Eonia_{for}$	1.41	1.49	1.65*	1.88*
$DFM\backslash Eonia_{back}$	1.05	1.33	2.11**	2.90***
$DFM\backslash Eonia_{mix}$	1.15	1.34	2.23**	3.08***
2015-2020				
H	1	4	6	8
$DFM\backslash Shadow_1$	1.172	1.19	1.71*	3.10***
$DFM\backslash Shadow_2$	1.16	1.26	1.90*	3.22***
$DFM\backslash Shadow_3$	1.01	0.163	1.34	3.11***
$DFM\backslash Shadow_4$	1.15	1.16	1.45	2.99***
$DFM\backslash Shadow_5$	1.15	1.26	2.33**	2.65***
$DFM\backslash Shadow_{all}$	1.10	1.16	2.55**	2.76***
$DFM\backslash Shadow_{for}$	0.78	1.11	1.56	3.01***
$DFM\backslash Shadow_{back}$	0.98	1.32	2.89***	3.44***
$DFM\backslash Shadow_{mix}$	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 2.9 shows that the DFM, in the vast majority of cases, outperforms different versions of the (pseudo) Taylor rule at longer horizons.

This exercise can be repeated replacing the textual dependent variable I developed above with EONIA and EU shadow rates. Table 2.10 and Table 2.11 show the results.

Table 2.10: Forecast Evaluation with EONIA Models

2005-2014				
H	1	2	3	4
$DFM\backslash Eonia_1$	0.81	0.92	1.93*	2.91***
$DFM\backslash Eonia_2$	1.24	1.11	2.11**	2.45**
$DFM\backslash Eonia_3$	1.29	1.17	1.87*	3.01***
$DFM\backslash Eonia_4$	1.18	1.23	1.49	2.11**
$DFM\backslash Eonia_5$	1.17	0.94	1.55	2.81***
$DFM\backslash Eonia_{all}$	1.35	1.36	1.92**	1.68*
$DFM\backslash Eonia_{for}$	1.22	0.99	1.75*	2.12**
$DFM\backslash Eonia_{back}$	0.95	1.22	1.66*	2.91***
$DFM\backslash Eonia_{mix}$	1.34	1.14	2.23**	3.01***
2015-2020				
H	1	4	6	8
$DFM\backslash Eonia_1$	0.72	1.19	1.51	3.21***
$DFM\backslash Eonia_2$	0.96	1.26	1.33	3.02***
$DFM\backslash Eonia_3$	1.11	0.163	1.34	2.11**
$DFM\backslash Eonia_4$	1.05	1.16	1.45	2.09**
$DFM\backslash Eonia_5$	1.05	1.26	1.23	2.05**
$DFM\backslash Eonia_{all}$	0.90	1.16	1.55	1.76*
$DFM\backslash Eonia_{for}$	0.76	1.11	1.56	3.11***
$DFM\backslash Eonia_{back}$	1.18	1.32	2.63***	3.21***
$DFM\backslash Eonia_{mix}$	1.21	1.44	2.81***	3.34***

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. Both in the DFM and stylized Taylor rule scenario, the dependent variable is the EONIA. $*$, $**$, and $***$ denote, respectively, the 10%, 5%, and 1% significance level.

Table 2.11: Forecast Evaluation with EU Shadow Rate Models

2005-2014				
H	1	2	3	4
$DFM\backslash Shadow_1$	0.75	1.22	1.90*	2.61***
$DFM\backslash Shadow_2$	1.14	1.45	2.01**	2.37**
$DFM\backslash Shadow_3$	1.22	1.38	1.79*	3.11***
$DFM\backslash Shadow_4$	0.98	1.13	1.59	2.41∗∗
$DFM\backslash Shadow_5$	0.67	0.94	1.55	2.78***
$DFM\backslash Shadow_{all}$	1.55	1.58	1.73*	2.26**
$DFM\backslash Shadow_{for}$	1.41	1.49	1.65*	1.88*
$DFM\backslash Shadow_{back}$	1.05	1.33	2.11**	2.90***
$DFM\backslash Shadow_{mix}$	1.15	1.34	2.23**	3.08***
2015-2020				
H	1	4	6	8
$DFM\backslash Shadow_1$	1.11	0.92	1.11	1.68*
$DFM\backslash Shadow_2$	1.11	1.16	0.90	1.22
$DFM\backslash Shadow_3$	1.00	0.76	1.34	1.81*
$DFM\backslash Shadow_4$	1.13	1.46	1.55	3.09***
$DFM\backslash Shadow_5$	1.12	1.16	2.23**	2.81***
$DFM\backslash Shadow_{all}$	1.01	1.15	2.45**	2.82***
$DFM\backslash Shadow_{for}$	0.88	1.01	1.56	3.00***
$DFM\backslash Shadow_{back}$	1.08	1.12	2.99***	3.30***
$DFM\backslash Shadow_{mix}$	1.30	1.44	2.89***	3.11***

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. Both in the DFM and stylized Taylor rule scenario, the dependent variable is the EU shadow rate. $*$, $**$, and $***$ denote, respectively, the 10%, 5%, and 1% significance level.

These results show that, in the vast of majority of cases, a “pseudo” Taylor rule with a mixed-frequency framework produces significantly smaller forecast errors compared to stylized Taylor rules with conventional dependent variables such as EONIA and EU shadow rates.

3. To Agree or not to Agree? The Effect of ECB Governing Council Disagreements in the Euro Area

3.1. Introduction

Macroeconomists have long studied the effects of monetary policy on the economy and financial markets. More recently, scholars have taken macroeconomic non-linearities into account and explored the effectiveness of monetary policy during cycles of recessions and expansions. Yet, the business cycle is not the only variable that matters for the transmission of monetary policy. Monetary policy decisions, for example, can be reached either unanimously or after heated disagreement. However, no study has investigated whether monetary policy is more powerful in a regime of agreement or disagreement among central bankers.

I investigate this hypothesis by focusing on the European Central Bank (ECB). More in detail, let me assume two regimes: one in which the Governing Council (GC) of the ECB announces an expansionary monetary policy decision achieved unanimously and another one where it announces a monetary policy easing acknowledging significant disagreement among its members. The hypothesis of the paper is that while in the first regime monetary policy propagates as documented by the literature (see, for example, Romer and Romer, 2004), in the second regime the impact of a monetary policy easing might vanish due to second-round effects through financial markets. In fact, the market may interpret disagreement as a fragile commitment that may lead the monetary authority to unwind monetary policy support more quickly than actually warranted.

To test this hypothesis, I proceed in two steps: first, since the ECB does not publish voting records, I use a dictionary approach based on the ECB's press conferences to measure the level of agreement/disagreement within the GC over monetary policy decisions. I call this proxy "consensus index" (CI); second, I use smooth transition-local projection model (STLPM) employed in Auerbach and Gorodnichenko (2012) and Ramey and Zubairy (2014) to analyze fiscal policy and Tenreyro and Thwaites (2016) to study the effects of a monetary policy easing conditional on the agreement within the GC.

The results support the hypothesis that monetary policy easing significantly differs depending on whether members of the GC agree or disagree over a monetary policy decision. While, in the consensus regime, a 1% monetary easing leads to results in line with the literature, the same shock in a disagreement regime has no statistically significant effect on macroeconomic variables − if anything, it has contractionary effects. One important considerations in evaluating the results is on the role of endogeneity. In fact, the disagreement measure obtained in the paper may reflect the fact that there is divergence in the assessment of macroeconomic conditions. The results may therefore be biased and require further investigation and robustness checks before drawing "causal" conclusions.

To the best of my knowledge, this is one of the first attempts to study the transmission of monetary policy conditional on a novel measure of agreement/disagreement within a central bank.

This paper refers to two streams of literature. The first one regards state-dependent papers on monetary policy. Tenreyro and Thwaites (2016) apply the STLPM to examine the impact of monetary policy shocks in expansions and recessions. The authors find that the effects of monetary policy are less powerful in recessions, especially for durables expenditure and business investment. Matthes and Barnichon (2015) propose a method to identify the non-linear effects of structural shocks by using Gaussian basis functions to parametrize impulse response functions. The finding is that a contractionary monetary policy is always more potent than its expansionary counterpart. Pellegrino (2021) estimates a nonlinear vector autoregression (VAR) model to assess whether the real effects of monetary policy shocks depend on the level of uncertainty. He finds that monetary policy shocks are about 50-75% more powerful during tranquil times than during uncertain times. Ascari and Haber (2021) document that large monetary policy shocks yield proportionally larger initial responses of the price level. Moreover, in a high trend inflation regime, the response of the price level to monetary policy shocks is larger and real effects smaller. Bernstein (2021) shows that when stockholders' incomes are more pro-cyclical than non-stockholders', output responds less to monetary policy in recessions, and contractionary monetary policy is more powerful than expansionary policy. Rüth (2017) finds that monetary policy impacts macroeconomic, housing, and financial variables stronger and more persistently when financial frictions are high.

The second stream concerns the literature on central banks' disagreement. The paper is most closely related to Falck et al. (2018) that investigate how disagreement about inflation expectations interacts with the efficacy of monetary policy. In times of high disagreement, they estimate that a 100 basis points (bps) increase in the U.S. policy rate leads to a significant short-term increase in inflation and in inflation expectations of up to 1 percentage point, whereas in times of low disagreement they find a significant decline of close to 1 percentage point. The paper is also very close to Madeira and Madeira (2019) that show that stock prices increase after monetary policy announcements when votes are unanimous in the Federal Open Market Committee (FOMC) but fall when dissent (which typically is due to preference for higher interest rates) occurs. Moreover, Seelajaroen et al. (2020) studies how monetary policy transparency of the Bank of England has information content in reducing disagreement about interest rate forecasts. The authors find that disagreement among the Monetary Policy Committee in policy rate decisions is associated with lower disagreement among professional forecasters on interest rate outlook, whereas neither announcement of changes in policy rates nor publication of inflation reports affects forecast disagreement. Similarly, Jitmaneeroj et al. (2019) provide evidence that greater transparency surrounding monetary policy reduces uncertainty of interest rates and inflation, primarily by reducing uncertainty that is common to agents rather than disagreement between agents.

The remainder of this paper is structured as follows. Section 3.2 derives the measure of agreement/disagreement over monetary policy decisions. Section 3.3 documents the empirical method and describes the dataset. Section 3.4 sets out the main results. I conduct sensitivity analysis in Section 3.5. Section 3.6 concludes with some thoughts for future research.

3.2. Measuring Consensus in the ECB's Governing Council

The first challenge of the paper is to measure agreement/disagreement in the ECB's GC. In fact, unlike other major central banks, the ECB doesn't provide voting records of monetary policy meetings or other data that can proxy the variable of interest. To solve this problem, I apply natural language process (NLP) techniques to the ECB's press conferences[1]. In particular, the procedure involves two steps. First, I create a dictionary of polarized unigrams (single words) and n-grams (sequences of words) denoting either agreement or disagreement. A representative list of words is shown in Table 3.1.

Table 3.1: Representative Polarized Dictionary

Agreement	Disagreement
agree*	disagree*
unanim*	dissent*
consens*	abstain*
majority	against
collegial	resign*
unite*	criticism

Note: The table shows the polarized words denoting agreement or disagreement in the Governing Council. * indicates any further declension of the unigram or n-gram under consideration (e.g. agree* = agrees, agreeing, agreed, etc.).

Second, following Rinker (2019) I develop a rule-based algorithm that takes into account the nuances of natural language. 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] I exclude journalists' questions from the press conferences since they can inflate the word-count and hence the final results.

sentence τ is then compared to the dictionaries of polarized words just displayed. 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} cluster[2]. 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:

$$\psi_\tau = \frac{\gamma_{w,\tau}^n}{\sqrt{\sum_{n=1}^N w_n}} \tag{3.1}$$

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 $PC^d = \frac{1}{n}\sum \psi_\tau$. Figure 3.1 displays the evolution of the index against relevant inflation data.

[2] 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.

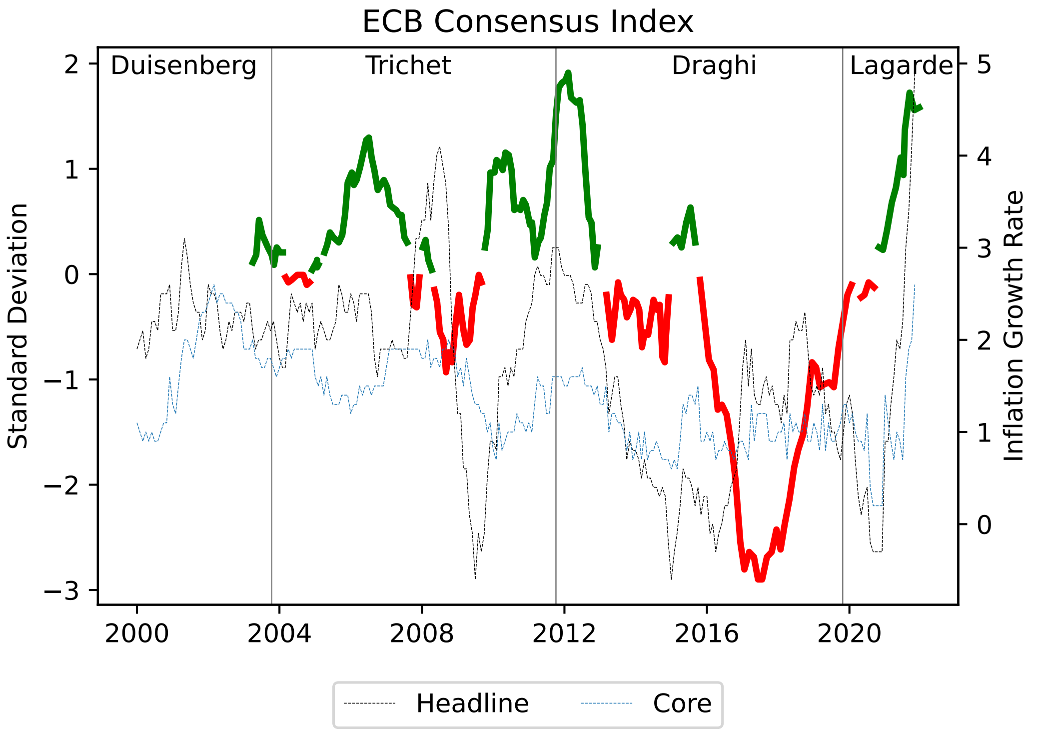


Figure 3.1: The figure shows the ECB consensus index alongside core (blue dotted line) and headline (solid black line) inflation.

The results show a significant correlation coefficient between the consensus index and, respectively, headline (34%) and core inflation (35%). This means that the level of agreement in the GC tends to be higher when inflation is high and lower when inflation is lower.

3.3. The Econometric Framework

In this section I first set out the specification of the econometric model used in this paper alongside the approach to statistical inference. Then, I describe the data, including the state variable and the identified monetary policy shocks.

3.3.1 Specification and Inference

The econometric model closely resembles the smooth transition-local projection model (STLPM) employed in Auerbach and Gorodnichenko (2012) and Ramey and Zubairy

(2014) to analyze fiscal policy and in Tenreyro and Thwaites (2016) to investigate monetary policy. The impulse response of variable y_t at horizon $h \in \{0, H\}$ in state $j \in \{a, d\}$[3] to a shock ε_t is estimated as the coefficient β_h^j in the following regression:

$$y_t = \tau_t + F(z_t)(\alpha_h^a + \beta_h^a \varepsilon_t + \gamma^{a'} x_t) + (1 - F(z_t))(\alpha_h^d + \beta_h^d \varepsilon_t + \gamma^{d'} x_t) \quad (3.2)$$

where τ_t is a linear time trend, α_h^j is a constant and x_t are controls[4]. $F(z_t)$ is a smooth increasing function of the indicator of the state of the agreement in the GC z_t. Following Granger and Teräsvirta (1993), I employ the logistic function:

$$F(z_t) = \frac{exp(\theta \frac{z_t - c}{\sigma_z}))}{1 + exp(\theta \frac{(z_t - c)}{\sigma_z})} \quad (3.3)$$

where c is a parameter that controls what proportion of the sample the GC spends in either state and σ_z is the standard deviation of the state variable z. The parameter θ determines how violently the GC switches from agreement to disagreement when z_t changes.

For each variable, I estimate the $H + 1$ equations of the impulse-response function (IRF) at horizon $0, \ldots, H$ as a system of apparently unrelated regression equations. By Kruskal's theorem, this results in the same point estimates of the regression coefficients as equation-by-equation ordinary least squares (OLS) because the explanatory variables are the same in each equation.

I follow Tenreyro and Thwaites (2016) in conducting inference on the estimated impulse response functions. I calculate standard errors analytically, allowing for the possibility of serially correlated residuals within equations and across equations. To capture this, I follow Ramey and Zubairy (2014) and use the Driscoll and Kraay (1998) method to adjust standard errors for the possibility of correlation in the residuals across dates t and horizons h. This amounts to estimating the parameters of the equations separately and then averaging the moment conditions across horizons h when calculating Newey-West

[3] a denotes agreement, while d disagreement

[4] In the baseline specification, x_t contains one lag each of the dependent variable and EONIA rates.

standard errors. Following Jordà (2005), I set the maximum autocorrelation lag $L = h+1$, where h is the maximum horizon of the impulse response function.

The specification just outlined might hide some drawbacks, of which the most significant one is endogeneity. In fact, the disagreement in the Governing Council of the ECB about monetary policy decisions may reflect divergent views on macroeconomic conditions. Under this regard, the index derived with textual techniques may hide macroeconomic confounding variables such as inflation, growth, etc. Although in the following sections, I will provide some robustness checks, results must be interpreted with caution.

3.3.2 Data

I run the model above with six variables at monthly frequency. While industrial production, Harmonized Index of Consumer Price (HICP), EURO STOXX 50 and unemployment enters in logarithmic form, Composite Indicator of Sovereign Stress (CISS) (Kremer et al., 2012) and EONIA are left unchanged in level. The period of analysis goes from January 2002 to December 2021 and the forecast horizon is 30 months.

I define the state variable z_t as a twelve month moving average of the consensus index derived in Section 3.2. Following Ramey and Zubairy (2014), and in contrast to Auerbach and Gorodnichenko (2012), the moving average term z_t is a lagging rather than centered moving average, so that future values of response variables do not appear on the right-hand side of the regression. Higher values of θ indicate that $F(z_t)$ spends more time close to the $\{0, 1\}$ bounds of the process, moving the model closer to a discrete regime-switching setup. Instead, smaller values of θ denote that more of the observations are taken to contain some information about behavior in both regimes. I follow Auerbach and Gorodnichenko (2012) and calibrate rather than estimate the parameters of the smooth transition model, for the same reasons they cite − it is difficult in practice to identify the curvature and location of the transition function in the data − and given the need for distributional assumptions on the error term when estimating by maximum likelihood. I set $\theta = 3$ to give an intermediate degree of intensity to the regime switching, and $c = 52\%$ which is the time spent in an "agreement" period in the sample. Figure 3.2 shows the probability of ending up in an agreement versus disagreement regime (orange line) versus the evolution of the consensus index (blue line).

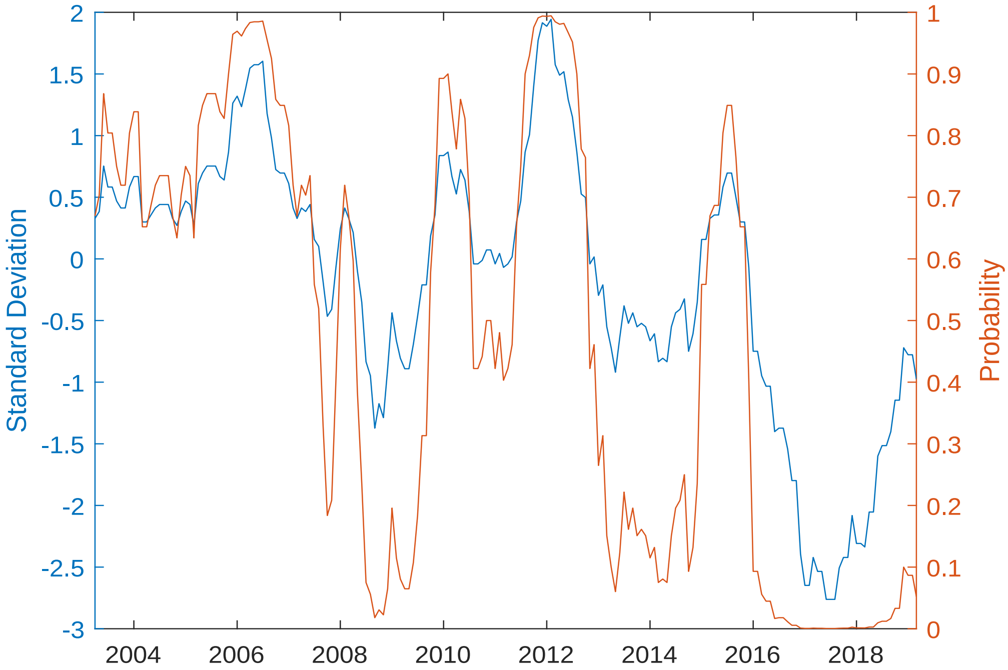


Figure 3.2: The figure shows the probability of ending up in an agreement versus disagreement regime against the ECB consensus index (blue line).

Finally, I identify monetary policy surprises ε_t in Equation 3.2 following Altavilla et al. (2019)[5]. The focus of the paper is on conventional monetary policy as well as asset purchases shocks.

3.4. Results

Figure 3.3 shows the impulse responses of unemployment, HICP, industrial production, CISS, EURO STOXX 50 and EONIA to an identified conventional monetary policy shock that generates an initial 1 percentage point decline in the EONIA. The first row displays linear impulse responses (black lines), while the second row shows impulse responses in

[5] Altavilla et al. (2019) proposes the identification of four types of monetary policy shocks: Timing, Target, Forward Guidance and QE. In this paper, I only focus on conventional monetary policy shocks (Target) and asset purchases shocks (QE). For a full derivation of the shocks I refer to Appendix 1.9.

a regime of agreement (red lines) and disagreement (blue lines). The solid lines indicate central tendencies while dashed lines denote 90% confidence intervals for the models under investigation.

Focusing on the linear impulse responses, the results show a familiar pattern: given a 1% exogenous decline in the interest rate, unemployment significantly declines after a year from the shock, HICP rises throughout the forecast horizon and industrial production increases up to 0.3% before declining from a year and a half after the shock. Moreover, the conventional monetary policy easing propagates through fast-moving market variables in accordance with economic theory: while the European stock market rallies after the shock, sovereign systemic risk significantly drops.

The results however vary across the regimes under investigation. In the agreement regime, impulse responses closely resemble linear impulse response where inflation, industrial production and EURO STOXX 50 increase while CISS and unemployment rightly decrease. Instead, in the disagreement regime, unemployment appears to rise for more than a year after the shock before gradually declining. Inflation stays put for 20 months before dropping into negative territory. Similarly, industrial production marginally rises for the first 12 months before turning negative. Notably, no result appears to be statistically significant.

The difference between the two regimes can be explained by investigating the transmission of the surprise through the market variables: CISS increases and the index of the European stock market mildly drops. This is evidence in support of the hypothesis of the paper: in a regime of heightened disagreement, the market looks through the expansionary monetary policy since it interprets disagreement as a fragile commitment that may lead the monetary authority to unwind monetary policy support sooner than actually warranted.

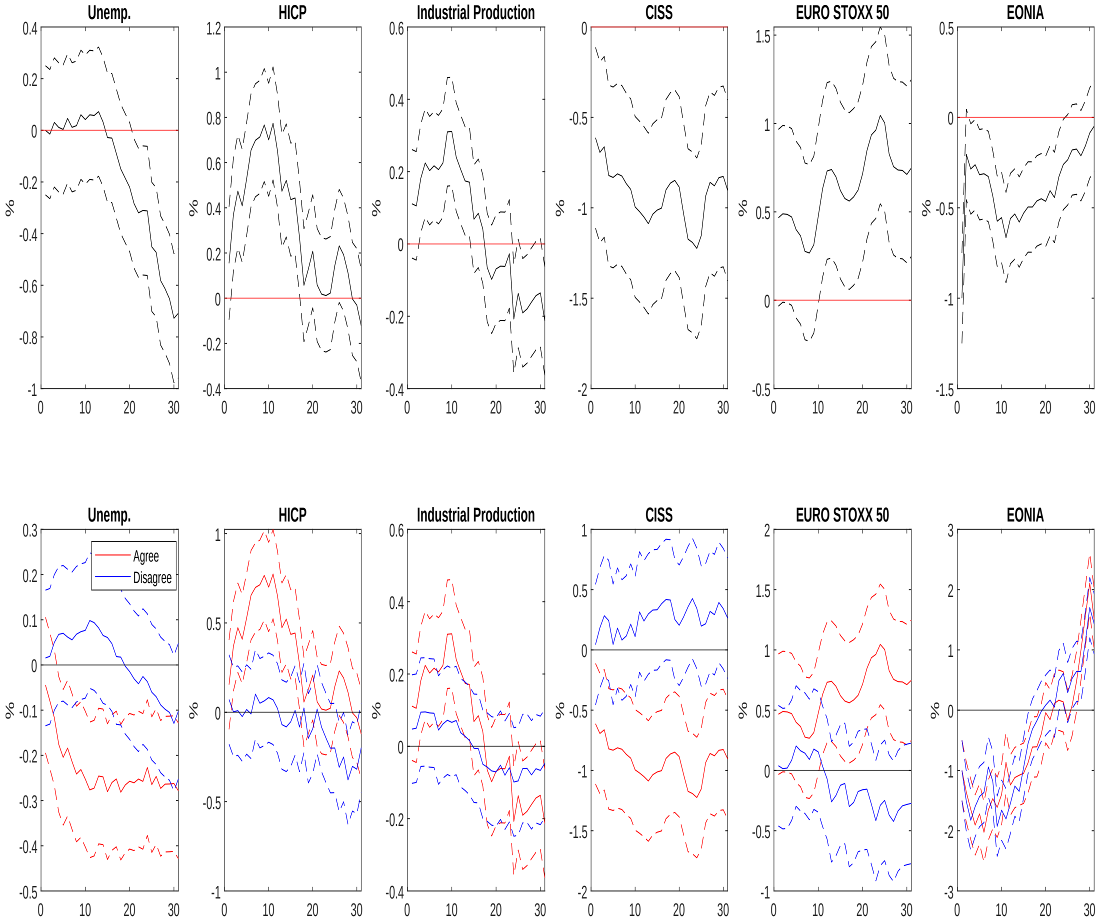


Figure 3.3: The figure shows the impulse responses to a conventional monetary policy shock. The first row displays linear impulse responses while the second one illustrates impulse responses in a regime of agreement (red lines) and disagreement (blue lines).

Figure 3.4 shows the linear and non-linear impulse responses to an unconventional QE surprise. Linear impulse responses are similar to the ones observed in Figure 3.3. As before, the variation occurs between the two regimes under study. While in a regime of agreement, impulse responses approximate linear impulse responses, in a regime of disagreement a QE shock has contractionary effects on the macro variables. In particular, unemployment rises and inflation and industrial production marginally drop. The transmission channel is once again the market: sovereign systemic risk increases, while EUR STOXX 50 significantly declines. Moreover, EONIA appears to recover much faster in the disagreement regime, turning positive a year after the shock.

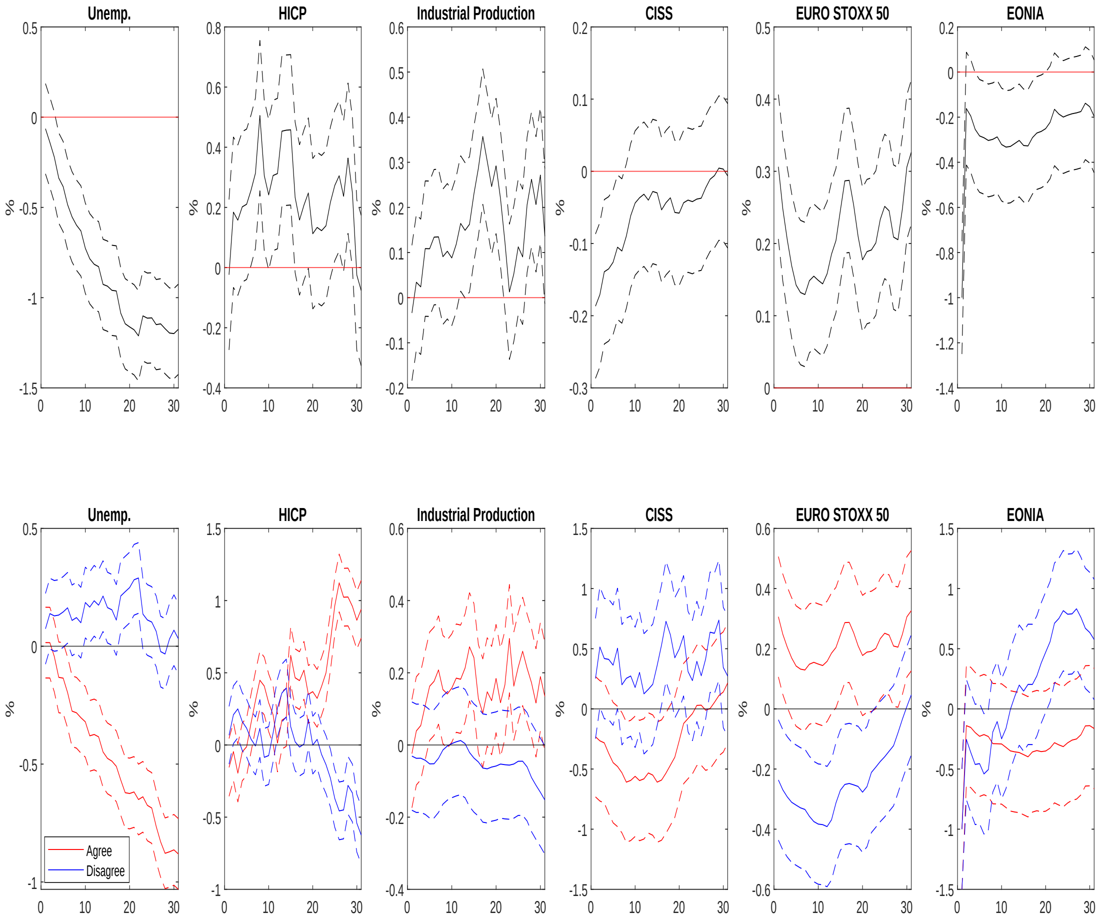


Figure 3.4: The figure shows the impulse responses to an unconventional QE shock. The first row displays linear impulse responses while the second one illustrates impulse responses in a regime of agreement (red lines) and disagreement (blue lines).

Overall, these results provide evidence in support of the main hypothesis of the paper: the transmission of monetary policy is affected by the level of agreement within the ECB Governing Council. While in a regime of agreement monetary policy easing retains well-known expansionary effects with regard to macro and financial variables, in a regime of disagreement those effects at best vanish and at worst become contractionary. The rationale for this result is that financial markets interpret disagreement as a fragile commitment that may drive the monetary authority to withdraw monetary policy support earlier than actually warranted.

There are however two concerns that must be raised with these results. The first and most important one relates to endogeneity. The state variable - the GC disagreement over monetary policy - may reflect divergent views on macroeconomic variables. Therefore, the effects could capture states of the world that are related to high/low inflation or recession/expansion. Second, there can be times where it is optimal to signal more consensual views to influence the market despite the real level of disagreement on monetary policy decisions. Such a dimension would be disregarded in the current set up and simulations.

In the next section, I will provide some robustness checks to relax some of the concerns just raised, in particular the one related to endogeneity.

3.5. Sensitivity Analysis

The following section examines the robustness of the findings to alternative choices of lags and trends in the regression equation (Section 3.5.1), the state variable z_t (Section 3.5.2) and additional specifications such as the phase shift of the state variable, the intensity of regime switching and the proportion of sample in a regime of agreement (Section 3.5.3). To save space, I carry out robustness tests only for conventional monetary policy shocks. Results also hold for QE surprises.

3.5.1 Trends and Lags in the Regression Equation

I estimated the model with a log-linear trend and one lag of both the policy and endogenous variable. In this subsection I examine the robustness of the results to both choices.

Figure 3.5 illustrates impulse responses estimated with a regression model identical to Equation 3.2 except for the omission of a time trend. Qualitatively, the results retain the message that the effects of monetary policy easing vanish in a regime of disagreement among central bankers. No impulse response appears to display an anomalous shape. Therefore, the model is robust to the introduction of a log-linear trend.

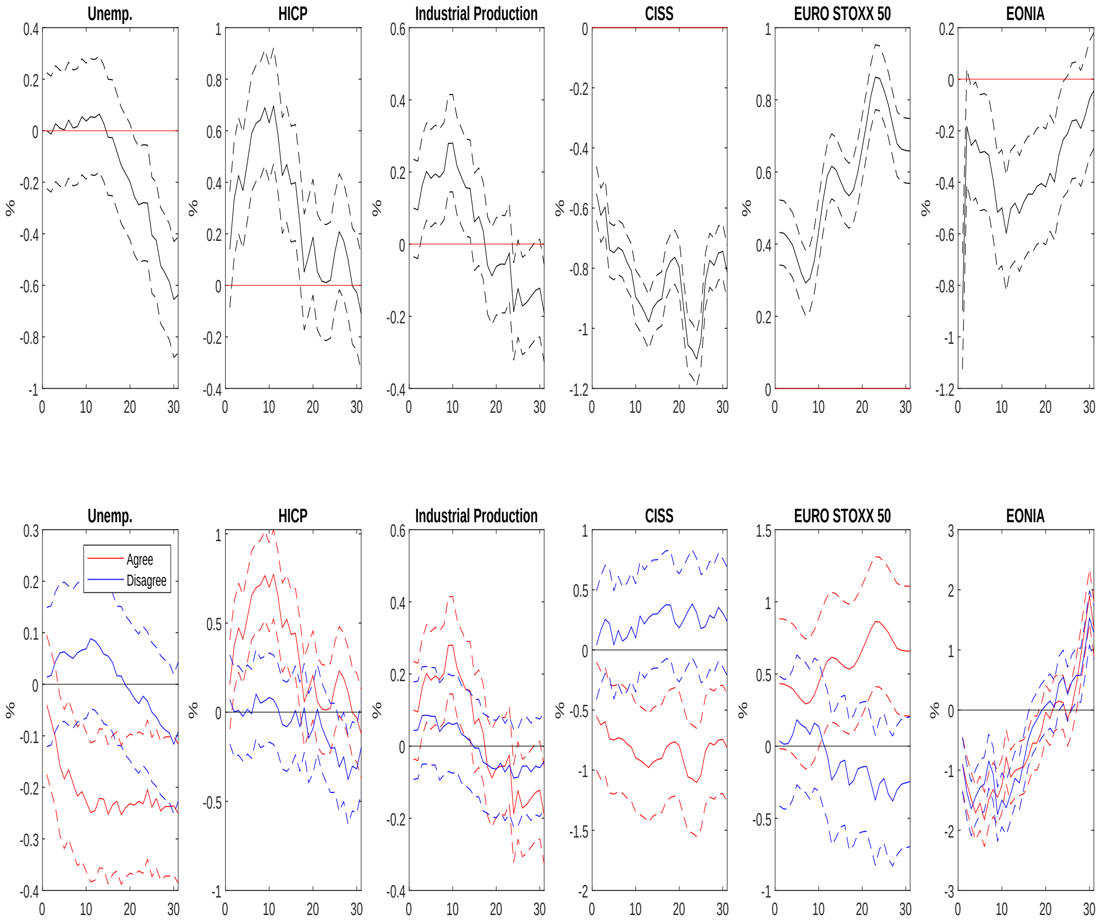


Figure 3.5: The figure shows the impulse responses to a conventional monetary policy shock. The first row displays linear impulse responses while the second one illustrates impulse responses in a regime of agreement (red lines) and disagreement (blue lines). The model is estimated without a log-linear trend.

On the other hand, the model uses one lag each of the dependent variable and the EONIA. The rationale for using such a lag structure is indicated by the Schwartz Bayesian Criterion (SBC), given by:

$$-2ln(\hat{L}) + kln(n) \tag{3.4}$$

where $\hat{L}$ is the maximized value of the likelihood function, k is the number of regressors and n is the number of observations in the sample. As a robustness test, I also calculated optimal lag length with the Akaike Information Criterion (AIC) given by:

$$2k - 2ln(\hat{L}) \tag{3.5}$$

This criterion indicates that two lags are optimal both for the dependent variable and for the policy variable. Figure 3.6 shows the impulse responses estimated with a regression model identical to the baseline but with the optimal lag structure denoted by AIC. The overall picture is unchanged. Hence, I conclude that the model is robust to different lag specifications

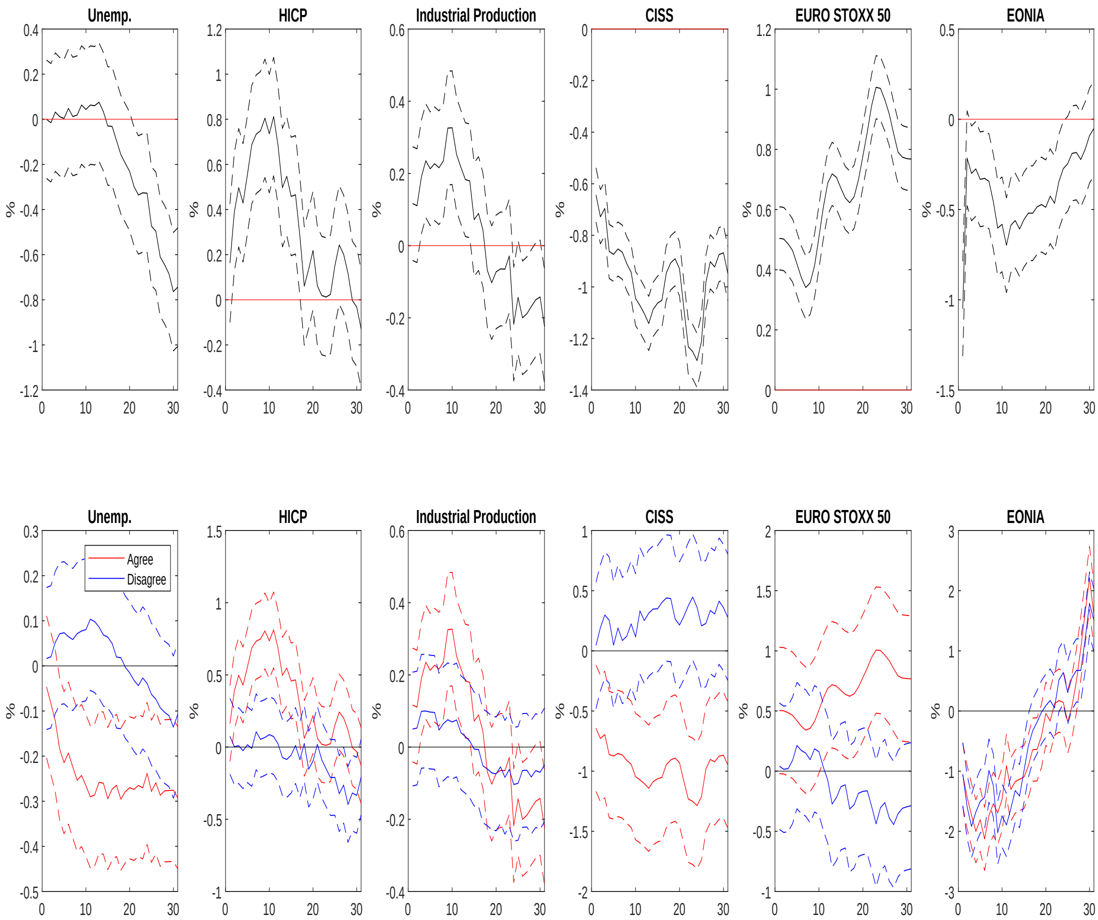


Figure 3.6: The figure shows the impulse responses to a conventional monetary policy shock. The first row displays linear impulse responses while the second one illustrates impulse responses in a regime of agreement (red lines) and disagreement (blue lines). The model is estimated with the optimal lag structure denoted by AIC.

3.5.2 The State Variable

The baseline results employ a measure of agreement/disagreement in the GC to which there are not many reasonable alternatives. Notwithstanding, I examine the sensitivity of the results to an additional specification[6]. I fit a logistic regression where the dependent variable − derived manually from the press conferences where 1 stands for a unanimous decision and 0 for a debated decision − is regressed on one-year ahead inflation rate and output growth. I then use the fitted values of the logistic regression as the state variable z_t. Such a methodology aims to relax the endogeneity concern raised above. Figure 3.7 reports the results:

[6] I also examine the sensitivity of the results to different smoothing transformations. Results remain consistent across different smoothed specifications.

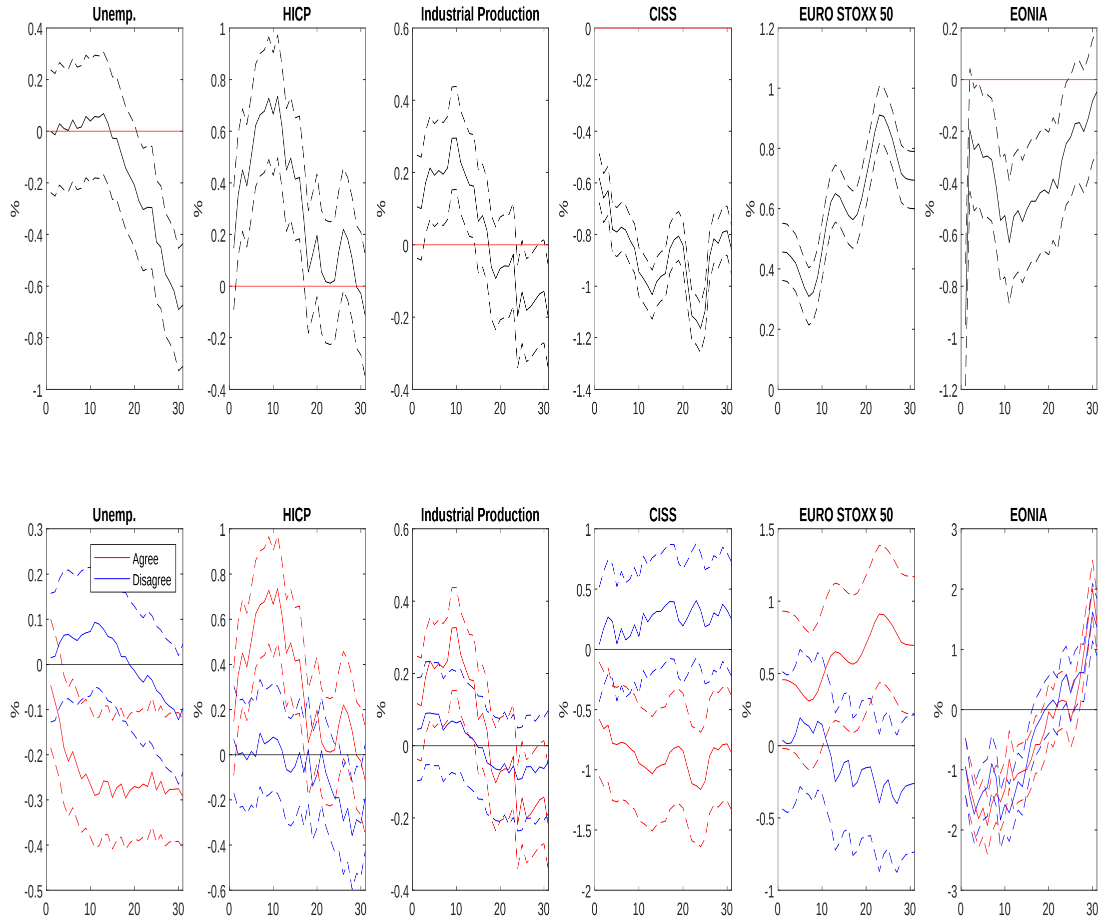


Figure 3.7: The figure shows the impulse responses to a conventional monetary policy shock. The first row displays linear impulse responses while the second one illustrates impulse responses in a regime of agreement (red lines) and disagreement (blue lines). The model is estimated with a state-variable derived from a logistic regression.

The results prove to be robust to different specifications of the state variable. In fact, similar to the baseline model while in the regime of high agreement the effects of monetary policy easing are very similar to the linear specification, in the regime of low agreement the monetary policy surprise fails to stimulate the economy – if anything it appears to have contractionary effects.

3.5.3 Additional Checks

I run additional robustness checks with regard to the phase shift of the state variable, the intensity of regime switching (θ) and the proportion of sample in an agreement state (c)[7].

To begin with the phase shift of the state variable, while in the baseline specification I follow Ramey and Zubairy (2014) in setting z_t as a lagging moving average, I now rely on Auerbach and Gorodnichenko (2012) and set z_t as a centered moving average of the consensus index. Results are consistent across the phase shift of the state variable. Hence, the results do not appear to be an artifact of using a centered moving average to calculate the state of the agreement within the ECB Governing Council.

With regard to the intensity of regime switching, I also experiment the model by setting θ equal to one and ten, respectively. Again, the qualitative message of the earlier analysis holds.

Finally, the impulse responses to an expansionary monetary policy shock are robust to changing from 20 to 40 percent the proportion of the sample judged to be in a regime of agreement.

Overall, the results prove to be robust to different model specifications. Therefore, the main result – the effects of monetary policy easing vanish in a regime of disagreement among central bankers – does not rely on peculiar modelling choices.

There is one further robustness check that can be run in further releases to take care of endogeneity. As the consensus index may reflect the fact that there is divergence in economic conditions, this problem could be taken into account by introducing measures of national economic divergences, i.e. divergencies in inflation and unemployment across member states. One way to find out how important national divergences in economic conditions are is to compute national Taylor rules, using national observations on inflation and output gap. The consensus index can be then filtered out on the fitted values of the national central banks.

[7] To save space and since the results resembles the impulse responses just shown, I don't display the results.

3.6. Conclusion

This paper explored whether monetary policy transmission is affected by regimes of agreement or disagreement among central bankers. The hypothesis is that when a monetary policy easing is accompanied by significant disagreement among central bankers, the expansionary effects of monetary policy are diminished. This is because financial markets might interpret disagreement as a fragile commitment that may drive central bankers to unwind monetary policy decisions sooner than actually warranted.

To test this hypothesis, I first created a proxy for the level of agreement/disagreement in the ECB's GC and then I set up a non-linear model to study the impulse responses due to a monetary policy easing in a regime of agreement and in a regime of disagreement within the GC. The results support the hypothesis of the paper: when an expansionary monetary surprise occurs in a regime of disagreement, the macroeconomic effects are muted, that is, monetary policy fails to stimulate the economy − if anything those effects are contractionary. The results hold for conventional and unconventional (QE) monetary policy. Further robustness checks are required to filter endogeneity issues from the consensus index that may still reflect divergent views on macroeconomic fundamentals.

To the best of my knowledge, this is one of the first attempts to investigate the impact of central bankers' disagreement on the transmission of monetary policy. There are at least two additional avenues of research. On the one hand, one may extend the current framework to different types of unconventional monetary policy shocks such as forward guidance or Targeted Long-Term Refinancing Operations (TLTRO). On the other hand, one may explore the interaction between agreement/disagreement among central bankers and macroeconomic variables for the conduct of monetary policy.

Conclusion

This thesis investigates the communication of the ECB from different perspectives. In the first paper, I study how the communicated degree of fiscal policy accommodation of the ECB affects the macroeconomy of the euro area and how it interacts with government bond surprises. The main results are the following. First, I document the existence of a quantifiable communicated ECB fiscal stance. Second, communicating a fiscally hawkish stance might lower output and inflation. In particular, the fiscal stance of the euro area tends to react to the fiscal recommendations of the ECB by cutting expenditures and raising taxes. This response however varies across blocks and conditional on the level of interest rates. On the one hand, southern countries appear to be more responsive than northern ones to the ECB's fiscal reprimands. On the other hand, both southern and northern blocks are more responsive when their fiscal capacity is constrained by an environment of higher rates. Third, historical decomposition provides empirical evidence that output and inflation would have been higher during the Great Recession as well as the sovereign-debt crisis, had the ECB held its communicated "fiscal stance" constant. Fourth, I find a statistically significant inverse relationship between government bond purchases and the "fiscal stance" of the ECB: the monetary authority becomes more fiscally conservative given a bond-buying easing. Further research is warranted to make the results more robust to endogeneity problems stemming from the absence of a measure of discretionary fiscal policies in the various models.

In the second paper, I provide an econometric framework which enables interested parties to track systematically 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. The 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 over the period 2015-2020.

In the third paper, I investigate whether monetary policy is more powerful in a regime of agreement or disagreement among ECB GC members. The results document that monetary policy easing significantly differs depending on whether members of the GC agree or disagree over a monetary policy decision. While, in the consensus regime, a 1% monetary easing leads to results in line with the literature, the same shock in a disagreement regime has no statistically significant effect on macroeconomic as well as financial variables. As in the first paper, further robustness checks are needed to confirm the "causal" claims as the consensus index may hide divergent views on key macroeconomic variables.

The three papers provide methodological innovations as well as empirical contribution. Methodologically, the first paper provides a method to quantify the "fiscal stance" of a central bank over the polarized policy dimension "hawkish"/"dovish". Similarly, it is also novel the endeavor to identify DFPA shocks. The second paper illustrates a mixed-frequency econometric framework for describing and predicting central banks' reaction function. The third paper proposes a method to construct a new index to measure the agreement within the GC of the ECB. Empirically, the first paper assesses that a DFPA shock brings inflation and output down. It also documents an inverse relationship between government bond purchases and the fiscal stance of the ECB. The second paper sheds light on the role of news broadcasted through the media as a fundamental channel of expectation formation. It captures, in fact, the prominent contribution of textual variables in explaining forecast revisions from 2015 on, that is, a period of renewed importance of communication as a policy tool. The third paper finally studies the transmission of monetary policy conditional on a novel measure of agreement/disagreement within a central bank and finds that monetary policy easing has no effect in a regime of heightened disagreement among central bankers.

References

Laver, M.; Benoit, K.; Garry, J. Extracting Policy Positions from Political Texts Using Words as Data. *American Political Science Review* **2003**, *97*, 311–331.

Miranda-Agrippino, S.; Ricco, G. The impact of the ECB monetary policy decision on the yield curve. *American Economic Journal: Macroeconomics* **2021**, *13*, 74–107.

Altavilla, C.; Brugnolini, L.; Gürkaynak, R. S.; Motto, R.; Ragusa, G. Measuring Euro Area Monetary Policy. *Journal of Monetary Economics* **2019**, *108*, 162–179.

Auerbach, A. J.; Gorodnichenko, Y. *Fiscal Policy after the Financial Crisis*; NBER Chapters; National Bureau of Economic Research, Inc, 2012; pp 63–98.

Ramey, V.; Zubairy, S. *Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data*; NBER Working Papers 20719, 2014.

Tenreyro, S.; Thwaites, G. Pushing on a String: US Monetary Policy Is Less Powerful in Recessions. *American Economic Journal: Macroeconomics* **2016**, *8*, 43–74.

Hansen, S.; McMahon, M. Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics* **2016**, *99*, S114–S133, 38th Annual NBER International Seminar on Macroeconomics.

Hansen, S.; McMahon, M.; Prat, A. Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach. *The Quarterly Journal of Economics* **2017**, *133*, 801–870.

Picault, M.; Renault, T. Words are not all created equal: A new measure of ECB communication. *Journal of International Money and Finance* **2017**, *79*, 136–156.

Bennani, H.; Fanta, N.; Gertler, P.; Horvath, R. Does central bank communication signal future monetary policy in a (post)-crisis era? The case of the ECB. *Journal of International Money and Finance* **2020**, *104*, 102167.

Baranowski, P.; Bennani, H.; Doryń, W. Do the ECB's introductory statements help predict monetary policy? Evidence from a tone analysis. *European Journal of Political Economy* **2021**, *66*, 101964.

Gorodnichenko, Y.; Pham, T.; Talavera, O. *The Voice of Monetary Policy*; Discussion Papers 21-02, 2021.

Marozzi, A. *The ECB's tracker: Nowcasting the press conferences of the ECB*; ECB Working Paper 2609, 2021.

Brand, C.; Buncic, D.; Turunen, J. The Impact of ECB Monetary Policy Decisions and Communication on the Yield Curve. *Journal of the European Economic Association* **2010**, *8*, 1266–1298.

Altavilla, C.; Giannone, D.; Lenza, M. *The financial and macroeconomic effects of OMT announcements*; ECB Working Paper 1707, 2014.

Hansen, S.; McMahon, M. Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics* **2016**, *99*, S114–S133.

Jarociński, M.; Karadi, P. Deconstructing monetary policy surprises-The role of information shocks. *American Economic Journal: Macroeconomics* **2020**, *12*, 1–43.

Swanson, E. T. Measuring the Effects of Federal Reserve Forward Guidance and Asset Purchases on Financial Markets. *Journal of Monetary Economics* **2021**, *118*, 32–53.

Allard, J.; Catenaro, M.; Vidal, J.-P.; Wolswijk, G. Central bank communication on fiscal policy. *European Journal of Political Economy* **2013**, *30*, 1–14.

Ferrara, F. M. The battle of ideas on the euro crisis: evidence from ECB inter-meeting speeches. *Journal of European Public Policy* **2020**, *27*, 1463–1486.

Giannone, D.; Reichlin, L.; Small, D. Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics* **2008**, *55*, 665–676.

Bok, B.; Caratelli, D.; Giannone, D.; Sbordone, A. M.; Tambalotti, A. Macroeconomic Nowcasting and Forecasting with Big Data. *Annual Review of Economics* **2018**, *10*, 615–643.

Bańbura, M.; Modugno, M. Maximum Likelihood Estimation of Factor Models on Datasets with Arbitrary Pattern of Missing Data. *Journal of Applied Econometrics* **2014**, *21*, 1–21.

Bańbura, M.; Giannone, D.; Modugno, M.; Reichlin, L. In *Handbook of Economic Forecasting*; G. Elliott, C. G., Timmermann, A., Eds.; Elsevier, 2013; Vol. 2; Chapter 0, pp 195–237.

Ghysels, E.; Santa-Clara, P.; Valkanov, R. *The MIDAS Touch: Mixed Data Sampling Regression Models*; CIRANO Working Papers, 2004.

Ghysels, E.; Sinko, A.; Valkanov, R. MIDAS regressions: Further results and new directions. *Econometric Reviews* **2007**, *26*, 53–90.

Clements, M.; Beatriz, A., Galvao Macroeconomic Forecasting With Mixed-Frequency Data. *Journal of Business & Economic Statistics* **2008**, *26*, 546–554.

Kuzin, V. N.; Marcellino, M.; Schumacher, C. *MIDAS versus mixed-frequency VAR: nowcasting GDP in the euro area*; Discussion Paper Series 1: Economic Studies 7, 2009.

Marcellino, M.; Schumacher, C. Factor MIDAS for Nowcasting and Forecasting with Ragged-Edge Data: A Model Comparison for German GDP. *Oxford Bulletin of Economics and Statistics* **2010**, *72*, 518–550.

Thorsrud, L. A. Words are the New Numbers: A Newsy Coincident Index of the Business Cycle. *Journal of Business and Economic Statistics* **2020**, *38*, 393–409.

Cimadomo, J.; Giannone, D.; Lenza, M.; Sokol, A.; Monti, F. *Nowcasting with large Bayesian vector autoregressions*; Working Paper Series 2453, 2020.

Taylor, J. B. Discretion versus policy rules in practice: two critical points. A comment. *Carnegie-Rochester Confer. Series on Public Policy* **1993**, *39*, 215–220.

Wieland, V.; Wolters, M. In *Handbook of Economic Forecasting*; G. Elliott, C. G., Timmermann, A., Eds.; Elsevier, 2013; Vol. 2; Chapter 0, pp 239–325.

Sturm, J. E.; de Haan, J. 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* **2011**, *147*, 41–58.

Bennani, H.; Neuenkirch, M. The (home) bias of European central bankers: new evidence based on speeches. *Applied Economics* **2017**, *49*, 1114–1131.

Falck, E.; Hoffmann, M.; Hürtgen, P. *Disagreement and Monetary Policy*; 2018 Meeting Papers 655, 2018.

Seelajaroen, R.; Budsaratragoon, P.; Jitmaneeroj, B. Do monetary policy transparency and central bank communication reduce interest rate disagreement? *Journal of Forecasting* **2020**, *39*, 368–393.

Jitmaneeroj, B.; Lamla, M. J.; Wood, A. The implications of central bank transparency for uncertainty and disagreement. *Journal of International Money and Finance* **2019**, *90*, 222–240.

Leeper, E. Equilibria under 'active' and 'passive' monetary and fiscal policies. *Journal of Monetary Economics* **1991**, *27*, 129–147.

Sargent, T. J.; Wallace, N. Some unpleasant monetarist arithmetic. *Quarterly Review* **1981**, *5*.

Sims, C. A Simple Model for Study of the Determination of the Price Level and the Interaction of Monetary and Fiscal Policy. *Economic Theory* **1994**, *4*, 381–99.

Woodford, M. Price-level determinacy without control of a monetary aggregate. *Carnegie-Rochester Conference Series on Public Policy* **1995**, *43*, 1–46.

Woodford, M. Fiscal Requirements for Price Stability. *Journal of Money, Credit and Banking* **2001**, *33*, 669–728.

Cochrane, J. H. Long-Term Debt and Optimal Policy in the Fiscal Theory of the Price Level. *Econometrica* **2001**, *69*, 69–116.

Debrun, X.; Masuch, K.; Ferrero, G.; Vansteenkiste, I. *Monetary-fiscal policy interactions in the euro area*; Occasional Paper Series 273, 2021.

de Haan, J.; Berger, H.; Jansen, D.-J.; de Haan, J. *The End of the Stability and Growth Pact?*; CESifo Working Paper Series 1093, 2003.

Sacher, M. Avoiding the inappropriate: The European Commission and sanctions under the stability and growth pact. *Politics and Governance* **2021**, *9*, 163–172.

King, M. A. Monetary policy implications of greater fiscal discipline (commentary). *Proceedings - Economic Policy Symposium - Jackson Hole* **1995**, 171–183.

Ferdinandusse, M.; Checherita-Westphal, C.; Attinasi, M. G.; Lalouette, L.; Bańkowski, K.; Palaiodimos, G.; Campos, M. *Euro area fiscal stance*; Occasional Paper Series 182, 2017.

Cimadomo, J.; Hauptmeier, S.; N., L.-K. Reflections on the Euro Area Fiscal Stance. *Intereconomics* **2017**, *52*, 125–131.

Blei, D. M.; Ng, A. Y.; Jordan, M. I. Latent Dirichlet Allocation. *Journal of Machine Learning Research* **2003**, *3*, 993–1022.

Hachula, M.; Piffer, M.; Rieth, M. Unconventional Monetary Policy, Fiscal Side Effects, and Euro Area (Im)balances. *Journal of the European Economic Association* **2020**, *18*, 202–231.

Kremer, M.; Lo Duca, M.; Holló, D. *CISS - a composite indicator of systemic stress in the financial system*; Working Paper Series 1426, 2012.

Wu, J. C.; Xia, F. D. Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound. *Journal of Money, Credit and Banking* **2016**, *48*, 253–291.

Jordà, Ò. Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review* **2005**, *95*, 161–182.

Christiano, L. J.; Eichenbaum, M.; Evans, C. L. Monetary policy shocks: What have we learned and to what end? *Handbook of macroeconomics* **1999**, *1*, 65–148.

Christiano, L. J.; Eichenbaum, M.; Evans, C. L. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of political Economy* **2005**, *113*, 1–45.

Bjørnland, H. C.; Leitemo, K. Identifying the interdependence between US monetary policy and the stock market. *Journal of Monetary Economics* **2009**, *56*, 275–282.

Kilian, L.; Lee, T. K. Quantifying the Speculative Component in the Real Price of Oil: The Role of Global Oil Inventories. *Journal of International Money and Finance* **2014**, *42*, 71–87.

Wright, J. H. Comment on "Measuring Euro Area Monetary Policy" by Carlo Altavilla, Luca Brugnolini, Refet Gürkaynak, Giuseppe Ragusa and Roberto Motto. *Journal of Monetary Economics* **2019**, *108*, 180–184.

Bernanke, B. S.; Boivin, J.; Eliasz, P. Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach. *Quarterly Journal of Economics* **2005**, *120*, 387–422.

Gertler, M.; Karadi, P. Monetary Policy Surprises, Credit Costs,and Economic Activity. *American Economic Journal: Macroeconomics* **2015**, *7*, 44–76.

Miranda-Agrippino, S.; Ricco, G. *Bayesian Direct Methods*; Mimeo, 2015.

Kadiyala, K. R.; Karlsson, S. Numerical Methods for Estimation and Inference in Bayesian VAR-Models. *Journal of Applied Econometrics* **1997**, *12*, 99–132.

Banbura, M.; Giannone, D.; Reichlin, L. Large Bayesian vector auto regressions. *Journal of Applied Econometrics* **2010**, *25*, 71–92.

Müller, U. K. Risk of Bayesian Inference in Misspecified Models, and the Sandwich Covariance Matrix. *Econometrica* **2013**, *81*, 1805–1849.

Giannone, D.; Lenza, M.; Primiceri, G. E. Priors for the Long Run. *Journal of the American Statistical Association* **2019**, *114*, 565–580.

Lafferty, D.; Blei, M. D. Correlated Topic Models. Advances in Neural Information Processing Systems. 2006; pp 147–155.

Heinrich, G. Parameter estimation for text analysis (based on Gibbs sampling). *Technical Report, Fraunhofer IGD* **2009**,

Cao, J.; Xia, T.; Li, J.; Zhang, Y.; Tang, S. A density-based method for adaptive LDA model selection. *Neurocomputing* **2009**, *72*, 1775–1781.

Ehrmann, M.; Fratzscher, M. Explaining Monetary Policy in Press Conferences. *International Journal of Central Banking* **2009**, *5*, 42–84.

Cour-Thimann, P.; Winkler, B. The ECB Non-Standard Monetary Policy Measures: the Role of Institutional Factors and Financial Structure. *Oxford Review of Economic Policy* **2012**, *28*, 765–803.

Cragg, J. G.; Donald, S. G. Inferring the Rank of a Matrix. *Journal of Econometrics* **1997**, *76*, 223–250.

Gürkaynak, R. S.; Sack, B.; Swanson, E. Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements. *International Journal of Central Banking* **2005**, *1*.

Montiel Olea, J. L.; Stock, J. H.; Watson, M. W. Inference in Structural Vector Autoregressions Identified With an External Instrument. *Journal of Econometrics* **2020**,

Stock, J. H.; Watson, M. Disentangling the Channels of the 2007-09 Recession. *Brookings Papers on Economic Activity* **2012**, *43*, 81–156.

Mertens, K.; Ravn, M. O. The dynamic effects of personal and corporate income tax changes in the United States. *American Economic Review* **2013**, *109*, 2679–2691.

Stock, J. H.; Wright, J. H.; Yogo, M. A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business & Economic Statistics* **2002**, *20*, 518–529.

Lucca, D. O.; Trebbi, F. *Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements*; NBER Working Papers 15367, 2009.

Shapiro, A. H.; Wilson, D. J. *Taking the Fed at its Word: A New Approach to Estimating Central Bank Objectives Using Text Analysis*; Working Paper Series 2019-02, 2019.

Gentzkow, M.; Kelly, B.; Taddy, M. Text as Data. *Journal of Economic Literature* **2019**, *57*, 535–74.

Ke, Z. T.; Kelly, B. T.; Xiu, D. *Predicting Returns With Text Data*; NBER Working Papers 26186, 2019.

Bybee, L.; Kelly, B. T.; Manela, A.; Xiu, D. *The Structure of Economic News*; NBER Working Papers 26648, 2020.

Babii, A.; Ghysels, E.; Striaukas, J. Machine Learning Time Series Regressions With an Application to Nowcasting. *Journal of Business & Economic Statistics* **2021**, *0*, 1–23.

Ardia, D.; Bluteau, K.; Boudt, K. Questioning the news about economic growth: Sparse forecasting using thousands of news-based sentiment values. *International Journal of Forecasting* **2019**, *35*, 1370–1386.

Griffiths, T. L.; Steyvers, M. Finding Scientific Topics. *Pnas* **2004**, *101*, 5228–5235.

Loughran, T.; McDonald, B. When is a liability not a liability? Textual Analysis, Dictionaries, and 10-Ks. *Journal of Finance* **2011**, *66*, 35–65.

Rinker, T. W. sentimentr: Calculate Text Polarity Sentiment. 2019; version 2.8.0.

Stoltz, D. S.; Taylor, M. A. Concept Mover's Distance: measuring concept engagement via word embeddings in texts. *Journal of Computational Social Science* **2019**, *2*, 293–313.

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

Modugno, M. Now-casting inflation using high frequency data. *International Journal of Forecasting* **2013**, *29*, 664–675.

Jansen, D.-J.; De Haan, J. Has ECB communication been helpful in predicting interest rate decisions? An evaluation of the early years of the Economic and Monetary Union. *Applied Economics* **2009**, *41*, 1995–2003.

Gerlach, S. Interest Rate Setting by the ECB, 1999-2006: Words and Deeds. *International Journal of Central Banking* **2007**, *3*, 1–46.

Gorter, J.; Jacobs, J.; De Haan, J. Taylor rules for the ECB using expectations data. *Scandinavian Journal of Economics* **2008**, *110*, 473–488.

Diebold, F.; Mariano, R. Comparing Predictive Accuracy. *Journal of Business & Economic Statistics* **1995**, *13*, 253–63.

Hansen, S.; McMahon, M.; Prat, A. Deliberation within the FOMC: a Computational Linguistics Approach. *The Quarterly Journal of Economics* **2018**, *133*, 801–870,.

Green, W. H. *Econometric Analysis*, 7th ed.; 2012.

Blattner, T. S.; Margaritov, E. *Towards a robust monetary policy rule for the euro area*; Working Paper Series 1210, 2010.

Granger, C.; Jeon, Y. Thick modeling. *Economic Modelling* **2004**, *21*, 323–343.

Brand, C.; Bielecki, M.; Penalver, A. *The natural rate of interest: estimates, drivers, and challenges to monetary policy JEL Classification: E52, E43*; Occasional Paper Series 217, 2018.

Romer, C. D.; Romer, D. H. A New Measure of Monetary Shocks: Derivation and Implications. *American Economic Review* **2004**, *94*, 1055–1084.

Matthes, C.; Barnichon, R. *Measuring the Non-Linear Effects of Monetary Policy*; 2015 Meeting Papers 49, 2015.

Pellegrino, G. Uncertainty and monetary policy in the US: A journey into nonlinear territory. *Economic Inquiry* **2021**, *59*, 1106–1128.

Ascari, G.; Haber, T. Non-linearities, state-dependent prices and the transmission mechanism of monetary policy. *The Economic Journal* **2021**,

Bernstein, J. A model of state-dependent monetary policy. *Journal of Monetary Economics* **2021**, *117*, 904–917.

Rüth, S. K. State-dependent monetary policy transmission and financial market tensions. *Economics Letters* **2017**, *157*, 56–61.

Madeira, C.; Madeira, J. The Effect of FOMC Votes on Financial Markets. *The Review of Economics and Statistics* **2019**, *101*, 921–932.

Granger, C.; Teräsvirta, T. *Modelling Non-Linear Economic Relationships*; Oxford University Press, 1993.

Driscoll, J. C.; Kraay, A. C. Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *The Review of Economics and Statistics* **1998**, *80*, 549–560.