Essays on Central Bank Communication and Monetary Policy

Thesis submitted at the Queen Mary University of London

in partial fulfilment of the requirements of the Degree of Doctor of Philosophy in Economics

by

Leonardo Nogueira Ferreira

School of Economics and Finance

Queen Mary University of London

2022

Declaration

I, Leonardo Nogueira Ferreira, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Name: Leonardo Nogueira Ferreira

Signed: Leonardo Nogueira Ferreira

Date: 1 April 2022

Abstract

This thesis investigates central bank communication and monetary policy from an empirical perspective. Chapters 2 and 3 are devoted to central bank communication, and the last one studies asymmetries in the effects of monetary policy. Specifically, Chapter 2 focusses on the effects of forward guidance and shows that this instrument is an effective policy tool, being at least as strong as conventional monetary policy. Chapter 3 investigates the contribution of information extracted from monetary policy statements to the forecasting of macroeconomic and financial variables. It shows that US central bank communication can improve forecasts, bringing useful informational content even when the benchmark is a traditional large-scale model. Chapter 4 revisits the "string theory", according to which monetary easings have smaller real effects than tightenings, and extends the study of asymmetry to the responses of financial conditions and to the euro area. Financial conditions also respond differently to positive and negative monetary policy shocks, and, overall, the responses found in the euro area are similar to the ones found in the US.

Acknowledgments

I am deeply thankful for all the support I received from my supervisor, Haroon Mumtaz. Our numerous insightful discussions not only improved considerably the quality of my papers but also helped me become an independent researcher. I am fully aware of what a privilege this has been. I am also very grateful to Andrea Carriero for his suggestions and for pushing me until I was able to explain things very clearly. I also would like to thank Emmanuel Guerre, Renato Faccini, and Erich Battistin for always being willing to help, inside and outside the classroom.

Even though the chapters in this thesis are solo-authored, they were written by many hands. The current versions of the papers that comprise this thesis are substantially different than their first drafts thanks to the feedback I received not only from my supervisors but also in many presentations and informal talks. I would like to thank Alistair Macaulay, Marco Bernardini, Nora Lamersdorf, Saulo Benchimol, and Wagner Piazza Gaglianone for their discussions, and Marcelo Medeiros for a very productive talk.

I also thank the participants of many seminars, workshops, and conferences for their instigating comments and fruitful feedback. I am grateful to Refet Gürkaynak and Pascal Paul for generously sharing the high-frequency data. I also would like to thank researchers that post their data and programs on their websites. In addition, I wish to thank Dario Bonciani and Robin Braun for all the support throughout my time at the Bank of England. They helped make my short journey at the Bank more productive and enjoyable at the same time.

This thesis was only possible because I managed to stay sane throughout these difficult years, and for that, I thank the great friends I met in my cohort: Yannis, Akif, Rodrigo, Ali, Ondrej, Sofia, and Aymeric. My friends back in Brazil also played an important role in this endeavour, especially during the pandemic, through several Zoom meetings that helped me keep lucid during the lockdowns. Many thanks for that!

I am also immensely grateful to my family, especially my mother, for always believing in me no matter what. I feel very lucky! Finally, I want to thank Queen Mary and the Central Bank of Brazil for all the support. I hope I was worth it.

Contents

1 Introduction **10**

1.1 Research Themes 10

1.2 Empirical Methodology 11

1.3 Structure of the Thesis 13

1.4 Results and Some Conclusions 15

2 Forward guidance matters: disentangling monetary policy shocks **17**

2.1 Introduction 17

2.2 Econometric Framework 23

2.2.1 Identification 25

2.2.2 Potential Advantages over Proxy SVARs 32

2.3 Results 33

2.3.1 Impulse Responses 33

2.3.2 Informational Sufficiency 37

2.3.3 Alternative Narrative Restrictions and Other Robustness Exercises 38

2.4 Conclusion 39

3 Forecasting with VAR-teXt and DFM-teXt models: exploring the predictive power of central bank communication **40**

3.1 Introduction 40

3.2 Data 44

3.2.1 Some Useful Terms 45

3.2.2 The Corpus of FOMC Statements and the LDA 45

3.3 The VAR-teXt Model 47
3.3.1 Estimation 48
3.4 Forecast Evaluation 50
3.5 Discussion 56
3.6 The DFM-teXt Model 59
3.7 Conclusion 61

4 Monetary policy surprises, financial conditions, and the "string theory" revisited 63

4.1 Introduction 63
4.2 Data 66
4.3 Econometric Framework 67
4.3.1 Linear Local Projections 67
4.3.2 Non-linear Local Projections 68
4.3.3 Estimation 70
4.4 Results 70
4.4.1 Robustness 74
4.5 Euro Area 74
4.5.1 The Euro Area dataset 74
4.5.2 The Euro Area results 75
4.6 Conclusion 76

Bibliography 79

A Appendix to Chapter 2 86

B Appendix to Chapter 3 100

C Appendix to Chapter 4 114

List of Figures

2.1 Forward Guidance Shock for August 2011 30

2.2 Impulse Responses to a Conventional Monetary Policy Shock 34

2.3 Impulse Responses to a Forward Guidance Shock 35

2.4 Difference in Impulse Responses of IP 36

3.1 Cumulative Log-Score Difference over Time 53

3.2 Giacomini and White (2006)'s Decision Rule 54

4.1 Impulse Responses: Linear LP 71

4.2 Impulse Responses: VAR 72

4.3 Impulse Responses: Positive and negative shocks 73

4.4 Impulse Responses: Linear case for the euro area 76

4.5 Impulse Responses: Positive and negative shocks for the euro area 77

A.1 Forecast Error Variance Decomposition 90

A.2 Impulse Responses to a Conventional Monetary Policy Shock: Aug/2011 and Jan/2012 92

A.3 Impulse Responses to a Forward Guidance Shock: Aug/2011 and Jan/2012 . . . 92

A.4 Impulse Responses to a Conventional Monetary Policy Shock: Ludvigsson type . 93

A.5 Impulse Responses to a Forward Guidance Shock: Ludvigsson type 93

A.6 Impulse Responses to a Conventional Monetary Policy Shock: Aug/2011 and Oct/2015 94

A.7 Impulse Responses to a Forward Guidance Shock: Aug/2011 and Oct/2015 . . . 94

A.8 Impulse Responses to a Conventional Monetary Policy Shock: 1993-2012 95

A.9 Impulse Responses to a Forward Guidance Shock: 1993-2012 95
A.10 Impulse Responses to a Conventional Monetary Policy Shock: No LSAP 96
A.11 Impulse Responses to a Forward Guidance Shock: No LSAP 96
A.12 Impulse Responses to a Conventional Monetary Policy Shock 97
A.13 Impulse Responses to a Forward Guidance Shock 97
A.14 Impulse Responses to a Conventional Monetary Policy Shock 98
A.15 Impulse Responses to a Forward Guidance Shock 98
A.16 Impulse Responses to a Conventional Monetary Policy Shock 99
A.17 Impulse Responses to a Forward Guidance Shock 99

B.1 Cumulative Log-Score Difference over Time: 1-year rate 102
B.2 Giacomini and White (2006)'s Decision Rule: 1-year rate 102
B.3 Cumulative Log-Score Difference over Time: Fed funds 103
B.4 Giacomini and White (2006)'s Decision Rule: Fed funds 103
B.5 Relative Frequency of the Top 5 Words for K=15 104
B.6 Cumulative Log-Score Difference over Time: Labelled topics 105
B.7 Giacomini and White (2006)'s Decision Rule: Labelled topics 105
B.8 Cumulative Log-Score Difference over Time: Tone-adjusted 106
B.9 Giacomini and White (2006)'s Decision Rule: Tone-adjusted 106
B.10 Cumulative Log-Score Difference over Time: Endogenous textual factors 107
B.11 Giacomini and White (2006)'s Decision Rule: Endogenous textual factors 107

C.1 US Monetary Policy Surprises 114
C.2 Non-Linear Impulse Responses: No lags of the proxy 116
C.3 Non-Linear Impulse Responses: Pre-ZLB 117
C.4 Non-Linear Impulse Responses: Swanson's factors 118
C.5 Non-Linear Impulse Responses: BAA spread 119
C.6 Non-Linear Impulse Responses: EBP 120
C.7 Updated Jarociński and Karadi (2020)'s MP Shocks 121

List of Tables

2.1 Zero and sign restrictions on responses 26

2.2 P-Values of the Orthogonality F-Test Proposed by Forni and Gambetti (2014) 37

3.1 Log-Scores: VAR-teXt versus benchmark VAR 52

3.2 Forecast Evaluation after 2015M12: VAR-teXt versus benchmark VAR 55

3.3 Forecast Evaluation: Alternative policy indicators 57

3.4 Forecast Evaluation: Labelled VAR-teXt versus benchmark VAR 57

3.5 Forecast Evaluation: Tone-adjusted VAR-teXt versus standard VAR-teXt 58

3.6 Forecast Evaluation: Endogenous textual factors versus benchmark VAR 59

3.7 Forecast Evaluation: DFM-teXt versus benchmark DFM 61

B.1 RMSE: VAR-teXt versus benchmark VAR 101

Chapter 1

Introduction

Monetary policy witnessed several innovations in the last decades. Unprecedented challenges prompted central bankers to employ unconventional instruments, such as forward guidance and large-scale asset purchases. The innovations, however, were not restricted to the instruments in central banks' toolkit. New identification methods, such as narrative sign restrictions and high-frequency identification, contributed to the study of the causal effects of conventional and unconventional monetary policy shocks. New text mining techniques allowed scholars and central bankers to address old issues differently or even to tackle different economic problems.

The papers in this thesis take advantage of such developments, while also introducing new developments, to investigate the use of new and traditional tools, central bank communication and conventional monetary policy respectively, from an empirical perspective.

1.1 Research Themes

The literature on monetary policy is vast. Nonetheless, new events and methods always nudge researchers to consider new lines of thinking (Friedman and Woodford, 2010). The zero lower bound, for instance, prodded central bankers to find new ways to boost the economy and researchers to find new ways of measuring the stance of monetary policy. Wu and Xia (2016) construct a measure that captures both conventional and unconventional monetary policy. Gertler and Karadi (2015) propose the one-year government bond rate as the relevant monetary policy indicator. Other papers, such as Bundick et al. (2017) and Lakdawala (2019), disentangle forward guidance from conventional monetary policy.

The growing use of unconventional instruments such as central bank communication has pushed the literature in a new direction. In fact, central banks have been using communication as a way to interact with markets and society for a long time. However, it assumed a central role in the conduction of monetary policy only in the last decades. Therefore, it has become crucial to understand properly how communication affects the economy. The special issue dedicated to the topic at the *Journal of Monetary Economics* in December 2019 is a testament to the efforts made by the profession in this regard. Two papers in this edition are related to this thesis and illustrate the two main ways used to measure communication: text mining and high-frequency surprises.

Using text mining, Hansen et al. (2019) show that news on economic uncertainty can have increasingly large effects along the yield curve. In order to do that, they employ Latent Dirichlet Allocation (LDA), a probabilistic topic model, and dictionary methods to measure a set of high-dimensional signals based on central bank communication. Chapter 3 is connected to this strand of the literature.

Altavilla et al. (2019) represent the other strand of the literature, which measures central bank communication indirectly, through its effects on financial markets. Together with Gürkaynak et al. (2005), Campbell et al. (2012, 2017), Swanson (2021), and many others, they use surprises around monetary announcements to isolate conventional and unconventional monetary policy from other shocks affecting the economy. In fact, high-frequency surprises are so useful in the identification that they will be used in Chapters 2 and 4 to study both conventional and unconventional monetary policy.

1.2 Empirical Methodology

The papers in the thesis make extensive use of vector autoregressive (VAR) models and local projections. Pioneered by Sims (1980), VAR models can be defined as follows:

$$Y_t = \sum_{p=1}^{P} \beta_p Y_{t-p} + \mu + u_t \tag{1.1}$$

where Y_t is the $N \times 1$ vector of standard macro variables, p denotes the lags, with $p = 1, ..., P$, and u_t are the reduced-form innovations. Sims made the link between the innovations to this linear system and macroeconomic shocks:

$$u_t = A_0 \varepsilon_t \tag{1.2}$$

where ε_t are the structural shocks and A_0 is a decomposition of the covariance matrix Σ such that $Var(u_t) = A_0 A_0' = \Sigma$. Identification then boils down to finding A_0. That is the focus of the literature on identification in Structural VAR models. The most common identification methods used to involve the imposition of timing or sign restrictions.

Recently, there has been a trend that involves the combination of different strategies or extraneous pieces of information to improve inference in VAR models. For instance, Jarociński and Karadi (2020) use sign restrictions on high-frequency surprises, therefore taking advantage of a narrow window around monetary announcements to minimise the probability that their monetary policy and central bank information shocks capture unrelated news announcements. In parallel, Antolín-Díaz and Rubio-Ramírez (2018) use narrative sign restrictions to narrow down the set of admissible models. This is done by constraining the structural shocks and/or the historical decomposition around key historical events, which is particularly convenient in the study of central bank communication since forward guidance is itself a narrative policy instrument.

VAR models and Dynamic Factor Models (DFM), which can be seen as VAR in the factors, have also been widely used to forecast (Granger et al., 2006; Elliott and Timmermann, 2013). At the same time, new machine learning methods have received a lot of attention, and there have been some efforts to connect these two fields. For example, Thorsrud (2018) and Larsen and Thorsrud (2019) use LDA to produce inputs to augment traditional econometric models. Essentially, LDA assembles words into meaningful groups and describes documents in terms of them. LDA can also be seen as a Bayesian factor model in which such groups are factors (Hansen et al., 2018). Thorsrud (2018) and Larsen and Thorsrud (2019) show that the inclusion of text improves the forecasting (or nowcasting) performance of the models, showing that this is an avenue worth pursuing.

Finally, this thesis also makes use of local projections. Proposed by Jordà (2005), a local projection is a method to compute impulse responses that does not require the specification of the underlying multivariate dynamic system. The impulse responses can then be directly estimated from:

$$y_{t+h} = \beta_h \varepsilon_t + \gamma_h x_t + u_{t+h} \tag{1.3}$$

where y_{t+h} is a vector of n variables, with $h = 0, ..., H$ denoting the horizons. ε_t is the measure of a shock/instrument, β_h gives the conventional direct estimates of the impulse responses, x_t collects the controls, and u_{t+h} denotes the residuals.

As pointed out by Ramey (2016), the comparison between this procedure and VAR models has an analogy with direct forecasting versus iterated forecasting, with local projections being analogous to the former and VAR models to the latter. In population, Plagborg-Møller and Wolf (2021) show that local projections and VAR models estimate the same impulse responses. Such equivalence, however, does not hold in the non-linear case. That is why it is necessary to depart from standard VAR models when studying asymmetric effects. A straightforward way of dealing with this form of non-linearity was implemented by Tenreyro and Thwaites (2016):

$$y_{t+h} = \beta_h^- \max\{0, \varepsilon_t\} + \beta_h^+ \min\{0, \varepsilon_t\} + \gamma_h x_t + u_{t+h} \tag{1.4}$$

where positive and negative shocks are allowed to have different effects: β_h^- and β_h^+. They show that have positive shocks have a larger impact on output than negative shocks in line with other studies such as Angrist et al. (2018) and Debortoli et al. (2020).

1.3 Structure of the Thesis

The thesis is set out as follows. Chapters 2 and 3 are devoted to central bank communication. The last chapter delves into potential asymmetries in the effects of conventional monetary policy.

Chapter 2: "Forward Guidance Matters: Disentangling Monetary Policy Shocks"

Chapter 2 assesses the dynamic effects of forward guidance. Central banks have usually employed short-term rates as the main instrument of monetary policy. In the last decades, however, forward guidance has also become a central tool. This paper combines two sources of extraneous information - high-frequency surprises and narrative evidence - with sign restrictions in a Bayesian structural vector autoregressive (VAR) model to disentangle forward guidance from conventional monetary policy.

Chapter 3: "Forecasting with VAR-teXt and DFM-teXt models: exploiting changes in central bank communication"

Chapter 3 explores the complementarity between traditional econometrics and machine learning and applies the resulting model - the VAR-teXt – to central bank communication. The VAR-teXt is a vector autoregressive (VAR) model augmented with information retrieved from text, turned into quantitative data via a Latent Dirichlet Allocation (LDA) model, whereby the number of topics (or textual factors) is chosen based on their predictive performance. A Markov chain Monte Carlo (MCMC) sampling algorithm for the estimation of the VAR-teXt that takes into account the fact that the textual factors are estimates is also provided. The approach is then extended to dynamic factor models (DFM) generating the DFM-teXt.

Chapter 4: "Monetary policy surprises, financial markets, and the "string theory" revisited"

Chapter 4 studies the asymmetric effects of monetary policy. Many are the attempts, by the economists, at testing whether it is true that "you can't push on a string", reputedly John Maynard Keynes's words. Exploiting high-frequency surprises, this paper explores whether the responses of standard macroeconomic variables as well as financial conditions are asymmetric in recent US and euro area samples. In order to do that, I estimate non-linear local projections using a Bayesian version of the procedure proposed by Lusompa (2021).

1.4 Results and Some Conclusions

The analysis brings out a number of important results. Chapter 2 shows that, in contrast with the evidence surveyed by Barakchian and Crowe (2013) and Ramey (2016), the identification scheme leads to the expected responses for output following a conventional monetary policy shock even when the model is estimated for a recent sample: 1993-2017. Results also show that forward guidance has been an effective policy tool. Therefore, forward guidance matters not only to the proper identification of conventional monetary policy shocks but also due to its effect on output and other macroeconomic variables. Specifically, its effects on industrial production are at least as strong as the effects of conventional monetary policy.

Chapter 3 shows that textual factors based on FOMC statements are indeed useful for forecasting in the small-scale VAR and in the DFM-teXt. Specifically, the VAR-teXt outperforms the benchmark VAR in forecasting the consumer price inflation and the interest rate, and this holds under various specifications. Thus, like factors in factor-augmented models, textual factors can increase the forecasting performance of VAR models even without necessarily having a clear meaning. This approach favours replicability since the choice of the number of textual factors is data-driven and does not rely on researchers' interpretability. Another clear advantage of an automated procedure such as this is scalability, so it is easy to apply it to datasets containing many more documents and words. Therefore, Chapters 2 and 3 show that taking into account central bank communication is of paramount importance and can help both in causal investigations as well as in forecasting efforts. As the results show, text matters and the value added is far from negligible.

Chapter 4 presents empirical evidence on the asymmetric effects of monetary policy. Impulse responses from non-linear local projections show that industrial production, unemployment, and the financial conditions index respond more strongly to positive, and CPI responds more strongly to negative shocks. The main findings are similar when the non-linear local projections are estimated using a euro area dataset. This is meaningful because it shows that, although the literature on the asymmetry of dynamic responses has focused on the US, the empirical evidence in favour of asymmetry in the dynamic effects of monetary policy is not US-specific. In addition to its intrinsic value, such analysis based on shocks split into positive and negative values matters in that it shows features that can be masqueraded in the symmetric case.

Chapter 2

Forward guidance matters: disentangling monetary policy shocks[1]

2.1 Introduction

Central banks have usually employed short-term interest rates as the main instrument of monetary policy. The extent to which such instrument is effective depends upon its ability to affect the path of expected future short-term real interest rates since, according to standard macroeconomic theory, such as Woodford (2003) and Galí (2015), consumption and output are driven by the sum of all future short-term real rates: the long-term real rate.[2]

In recent years, a prominent alternative way to affect long-term interest rates has been intensively used: the communication about the likely future course of monetary policy, known as forward guidance. In this framework, if central banks can commit to a future path of interest rates, their communication may affect the economy even in the absence of changes in the short-

[1] A previous version of this chapter is available in the Central Bank of Brazil Working Paper Series as: Ferreira (2020), 'Forward guidance matters: disentangling monetary policy shocks' (No. 530). I have benefited from comments from referees and participants at presentations at the 28th Annual Symposium of the Society for Nonlinear Dynamics and Econometrics, 7th Annual MMF PhD Conference, Central Bank of Brazil, Central Bank of Malta, 8th SIdE Workshop for PhD students in Econometrics and Empirical Economics and Queen Mary University of London. I also thank the Working Paper Series of the Central Bank of Brazil for the Best Economics Paper Award. All errors are mine. The views expressed in this chapter are those of the author and not necessarily reflect those of the Banco Central do Brasil.

[2] To see the monetary transmission mechanism in the New Keynesian model clearly, it is useful to remember that, iterating forward, the Euler equation becomes:

$$\hat{y}_t = -\frac{1}{\sigma} E_t \sum_{i=0}^{\infty} (\hat{\imath}_{t+i} - \pi_{t+i+1})$$

where $\hat{x}$ denotes the percentage deviation of a variable X_t around its steady state, y is the output, i is the nominal interest rate, π is the inflation rate, and $\frac{1}{\sigma}$ governs the intertemporal elasticity of substitution.

term policy rate. Hence, forward guidance (or more broadly, communication) also becomes a policy tool.

These two policy instruments (short-term interest rates and forward guidance) are obviously intrinsically connected. First, forward guidance matters for the identification of the conventional monetary policy shocks in that such shocks cannot be properly recovered unless anticipated changes in the policy rates are taken into account. Second, forward guidance is one reason why, as pointed out by Ramey (2016), estimating the causal effects of conventional monetary policy has become a challenge. With anticipation effects and monetary policy conducted more systematically, finding truly exogenous monetary policy shocks in recent samples has become increasingly difficult.

On the other hand, forward guidance shocks can be a valuable source of not so systematic policy. This tool became prevalent during the zero lower bound (ZLB) period when the use of the conventional policy rate was constrained and episodes of truly exogenous forward guidance shocks can be found. Campbell et al. (2012) and Campbell et al. (2017) show effectiveness of forward guidance shocks in moving long-term government bond rates. But what about the dynamic responses of macroeconomic and financial variables to these shocks?

This paper tackles this question by disentangling forward guidance and conventional monetary policy shocks in an innovative way: combining two sources of extraneous information with sign restrictions in a structural vector autoregressive (VAR) model estimated using data since the 90s, which is when the Federal Open Market Committee (FOMC) started to issue statements immediately after each meeting.

The first source of extraneous information is based on high-frequency futures prices and it builds on Kuttner (2001) and Gürkaynak et al. (2005). The use of high-frequency surprises around FOMC announcements is important to address endogeneity concerns as well as to help in the decomposition of the shocks. Specifically, the vector of variables of the VAR incorporates Gürkaynak et al. (2005)'s target and path factors, which capture surprises in the current and future rates respectively. Their inclusion together with the other variables in the spirit of Jarociński and Karadi (2020) is an alternative to their use as external instruments in Proxy SVARs.

Jarociński and Karadi (2020) combine sign restrictions and high-frequency surprises to identify monetary policy and information shocks, which is their object of study. In this paper, however, this combination will be used to cleanse the shocks of interest from any informational advantage the central bank may have. This is an alternative to the customary use of the Greenbook forecast data, with the advantage of not limiting the sample.[3]

Nonetheless, sometimes sign restrictions may have to be complemented with additional restrictions in order to generate a sufficiently rich shock structure as pointed out by Inoue and Kilian (2013) and Arias et al. (2019). The second source of extraneous, which is the narrative account of some particular episodes, is then used to enhance and refine the identification. The idea was formalised by Antolín-Díaz and Rubio-Ramírez (2018) as narrative sign restrictions.[4] First, sign restrictions consistent with economic theory are placed not only on the standard variables but also on the factors in order to properly isolate the shocks of interest. Then, uncontroversial episodes of forward guidance and conventional monetary policy shocks are used to refine the credible set. This is particularly convenient since forward guidance is itself a narrative policy instrument.

Most importantly, following Uhlig (2005), the sign restrictions are agnostic. Therefore, the VAR model does not place any restriction on the responses of the industrial production and lets the data and the adjacent restrictions "decide" them, avoiding the circularity pointed out by Cochrane (1994). As in Uhlig (2005), the idea is to leave the question of interest open, but using prior information about the behaviour of the other variables through the sign and the narrative sign restrictions.

This agnostic approach is especially important because, notwithstanding the relevance of the topic, there is still a lack of consensus among researchers and policy-makers about the effects of forward guidance. For example, McKay et al. (2016) find that the effect on GDP in models with incomplete markets is much lower than in models with complete markets.[5] Nonetheless, a few quarters of forward guidance is still powerful enough to effectively prevent recessions. In contrast, after adding several features to McKay et al. (2016)'s model to bring it closer to the

[3]The Greenbook forecast data is made available to the public only five years after the end of the year of the forecast.

[4]Narrative information has also been used in different contexts and set-ups, such as in Kilian and Murphy (2014) and in Ben Zeev (2018).

[5]McKay et al. (2016) combine elements from standard New Keynesian models, such as nominal rigidities, with elements from standard incomplete models, such as uninsurable risks and borrowing constraints. See also Del Negro et al. (2012) for a discussion of the forward guidance puzzle.

data, Hagedorn et al. (2019) find that the effects of forward guidance are, in fact, negligible.

VAR models can then help shed some light on New Keynesian models, providing them with some reference and bringing them even closer to the data. Being Bayesian, it also allows for a formal comparison between the effects of forward guidance and the effects of conventional monetary policy in a high posterior density interval (HPDI) sense.

Therefore, following a trend that involves the combination of different strategies to improve inference in SVARs[6], this paper uses different sources of extraneous information in the identification of structural shocks and contributes to the forward guidance literature, showing that this tool is at least as strong as conventional monetary policy. The use of narrative evidence is particularly convenient since forward guidance is itself a narrative policy instrument, and combining it with high-frequency surprises sharpens the results and helps in the disentangling of the shocks.

Specifically, results show that the direction of the effect of conventional monetary policy is the expected even in a recent US sample, in contrast with the evidence revisited by Barakchian and Crowe (2013) and Ramey (2016). Results also show that the effect of forward guidance on industrial production are not different from the effect of conventional monetary policy in a HPDI sense.

Related Literature

The papers most closely related to this one can be divided into two groups. In the first group, forward guidance is mixed with conventional monetary policy. By using futures contracts whose horizon comprises at least the next FOMC meeting, the shocks coined as monetary policy shocks in the next two papers incorporate the impact of forward guidance. Andrade and Ferroni (2018) employ market-based measures of inflation expectations and future interest rates together with sign restrictions to identify Delphic and Odyssean monetary shocks. In a similar endeavour, Jarociński and Karadi (2020) explore the co-movements of interest rates and stock prices around the announcements combined with sign restrictions to identify monetary policy shocks and central bank information shocks. Debortoli et al. (2019) estimate a time-varying VAR that uses the 10-year government bond rate as a policy indicator and find that the responses to different shocks do not present material differences in the ZLB. The corollary is that unconventional monetary

[6]For instance, Braun and Brüggemann (2020) combine sign restrictions with external instruments, Podstawski et al. (2018) combine heteroskedasticity with external instruments, and Ludvigson et al. (2020) combine covariance and sign restrictions with 'external variable inequality constraints'.

policy (including forward guidance) acted as a substitute for conventional monetary policy.[7]

In the second group, forward guidance is isolated from conventional monetary policy. Similar to this paper, D'Amico and King (2015) combine measures of expectations with sign restrictions. Differently, however, they use survey-based measures of macroeconomic variables, which may respond with some delay as pointed out by Coibion and Gorodnichenko (2012, 2015) and may not fully isolate forward guidance shocks from other shocks affecting expectations. In fact, D'Amico and King (2015) acknowledge that any information, not only the shocks generated by forward guidance, which causes agents to change beliefs about the future course of monetary policy, should be captured in their identification. They see it as an advantage as they seem to be interested in overall anticipated monetary policy. Nevertheless, for the purpose of this paper, disentangling forward guidance shocks from conventional monetary shocks or any other kind, this would be a weakness.

Ben Zeev et al. (2019) identify anticipated monetary shocks following the literature on news shocks. The monetary news shock is orthogonal to current policy residual and maximises the sum of contributions to its forecast error variance over a finite horizon. As in D'Amico and King (2015), this captures the effects of forward guidance shocks but not only. Ben Zeev et al. (2019) also acknowledge that their "approach allows for any channels through which changes in expectations may arise". Moreover, by estimating a quarterly VAR they are not able to capture near-term since the next meeting is approximately half of the time within the quarter.

The approach carried out by Ben Zeev et al. (2019) and the one undertaken in this paper can be seen as two extremes. On the one hand, one approach only uses information coming from the meetings. On the other hand, his news shock can capture much more than forward guidance. The choice made in this paper is based on the potential of high-frequency surprises to ameliorate the identification problem and on the fact that, despite using more speeches and other forms of communication, the statements are the main outlet to communicate forward guidance. In this extent, this paper is closer to the following 2 papers.

Lakdawala (2019) uses market-based measures of expectations, specifically the Gürkaynak et al. (2005)'s target and path factors, as external instruments in a VAR to decompose the

[7] In a similar vein, Swanson (2018) shows the Federal Reserve was not very constrained in its ability to influence medium- and longer-term interest rates and the economy due to effective forward guidance and the large-scale asset purchases.

effects of monetary policy. Here, on the other hand, these factors are incorporated into the vector of variables of the VAR, what makes the inference valid even if the VAR without them is not fully or partially invertible. Moreover, Lakdawala (2019)'s sample starts in 1979 and includes the Volcker disinflation period, what may affect the findings, while this paper focuses on a more recent sample, which corresponds to the period experiencing an increase in FOMC communication.

Bundick and Smith (2020) examine the macroeconomic effects of forward guidance at the ZLB using a modified path factor. They order this measure after real activity and the price level but before the 2-year rate in a recursive VAR. A caveat is that by restricting their sample to the ZLB, their estimation disregard numerous episodes of forward guidance that took place in the periods pre or post-ZLB. They work around this issue by also estimating the model over the pre-ZLB period. They find that forward guidance shocks produce similar results.

Hansen and McMahon (2016) follow a different path. They use tools from computational linguistics to extract and measure the information released by the FOMC on the state of economic conditions and on forward guidance, which is inputted in a factor-augmented VAR model identified recursively. Zlobins (2019) studies the effects of ECB's forward guidance and also employs sign, zero and narrative sign restrictions. However, because he covers a period dominated by the ZLB and uses the 3-month EURIBOR rate, which also captures near-term forward guidance, for the identification of conventional monetary policy, shocks may be not properly disentangled.

This paper complements this recent literature by combining the advantages of high-frequency identification with the appeal of narrative sign restrictions to identify the dynamic responses of important macroeconomic and financial variables to conventional monetary policy and forward guidance shocks. The rest of the paper is organised as follows. Section 2.2 describes the econometric approach. Section 2.3 presents the results. Section 2.4 concludes.

2.2 Econometric Framework

The point of departure for the analysis is a VAR model of the form:

$$\begin{pmatrix} m_t \\ y_t \end{pmatrix} = c + \sum_{p=1}^{P} \beta^p \begin{pmatrix} m_{t-p} \\ y_{t-p} \end{pmatrix} + A_0 \varepsilon_t, \tag{2.1}$$

where m_t is a vector of N_m surprises. The monthly series are built by adding up the intra-day surprises occurring in month t on the days of FOMC meetings and letting the series take a value of zero in months without FOMC announcements. y_t is a vector of N_y monthly macroeconomic and financial variables. p denotes the lags, with $p = 1, ..., P$. The structural shocks ε_t are related to the reduced-form innovations u_t via $u_t = A_0\varepsilon_t$ where A_0 is a decomposition of the the covariance matrix Σ such that $Var(u_t) = A_0A_0' = \Sigma$.

The baseline Bayesian VAR is estimated for the US using a flat prior and 5 lags[8] on 7 macroeconomic and financial variables (2 high-frequency variables (m_t) and 5 low-frequency variables (y_t)) spanning the period from 1993M01 to 2017M12. m_t includes the target and path factors. y_t consists of the consumer price index (CPI), the industrial production index (IP), the fed funds rate (FF), the 2-year government bond rate (GS2), and the excess bond premium (EBP) computed by Gilchrist and Zakrajšek (2012). The first two variables of y_t are in log levels.

The target and path factors are constructed based on the methodology of Gürkaynak et al. (2005). Surprises in the prices of fed funds futures and Eurodollar futures are computed for a 30-minute window around 220 scheduled and unscheduled FOMC meetings to estimate these two factors.[9] By construction, the target factor accounts for most of the surprise in the futures rates for the current month (FF1) and the path factor influences only expected future rates.[10]

Gürkaynak et al. (2005) show the path factor is closely related with FOMC statements.[11] Such forward-looking statements provide agents with news on future information about changes

[8] This choice was based on the Akaike information criterion (AIC), which is considered the most accurate criterion for monthly VARs by Ivanov and Kilian (2005).

[9] Following Campbell et al. (2012), however, the outlier meetings in September 2001 (9/11) and March 2009 (QE1) were dropped.

[10] See Gürkaynak et al. (2005) for the constructions of the factors.

[11] Gürkaynak et al. (2018) provide further evidence in this regard. They show there is a close correspondence between the path factor and a latent factor that captures non-headline news.

in short-term interest rates. Furthermore, as a market-based measure of expectations, the path factor is robust to concerns usually associated with survey-based measures of expectations, such as staleness and insufficient skin in the game (Coibion and Gorodnichenko, 2012, 2015).

Due to its characteristics, in order to address the effects of anticipation in monetary policy, the path factor is incorporated into the vector of the variables in the VAR and not used as an external instrument.[12] In a model with news shocks, the inclusion of variables that reflect views on the future path of the economy is even more relevant since their omission can potentially introduce non-fundamentalness, altering the mapping between the true news that agents observe and the identified shocks (Leeper et al., 2013; D'Amico and King, 2015).

The excess bond premium, introduced by Gilchrist and Zakrajšek (2012), is a corporate bond credit spread purged from the default risk, with a high informational content about the economy. As pointed out by Caldara and Herbst (2019), the inclusion of credit spreads is of paramount importance and can result in large differences in the effects found in VAR models. This happens because an increase in credit spreads generates a persistent decrease in real activity and a failure to account for this endogenous reaction induces an attenuation in the response of all variables to monetary shocks.[13]

The use of two policy indicators (the fed funds rate and the 2-year government bond rate) is crucial to the decomposition of monetary policy shocks into conventional and forward guidance shocks. Moreover, the sample does not stop in 2007 or 2008, as it is typical in VAR models of monetary policy due to the turbulence caused by the financial crisis and the following ZLB period, because this would jeopardise the objective of this paper since forward guidance was intensively used during the ZLB period. The 2-year government bond rate was chosen because it is consistent with the horizon of forward guidance. The other variables are standard: CPI and IP.[14]

[12] Section 2.2.2 further elaborates on this issue.

[13] See also Section 9 of Miranda-Agrippino and Ricco (2019).

[14] Results are similar when the sample period ends in 2012 as shown in the Appendix. This span reduces the influence of the zero lower bound in the estimation while it still includes an important episode of forward guidance used in the identification.

2.2.1 Identification

This subsection explains how high-frequency data are combined with sign restrictions and narrative information within this econometric framework to identify the two shocks of interest: forward guidance and conventional monetary policy shocks.

Sign restrictions and high-frequency identification

Following Rubio-Ramirez et al. (2010) and Arias et al. (2018), a candidate A_0 is found by calculating $\tilde{A}_0$, an arbitrary matrix square root of Σ, using Cholesky and multiplying it with a rotation matrix Q. The impulse responses using this candidate structural impact matrix are then checked and kept if the restrictions are satisfied.

There has been increasing concern about the informativeness of priors. Giacomini and Kitagawa (2018), for instance, propose imposing posterior bounds on the impulse response functions that are robust to the choice of priors. An alternative is proposed by Baumeister and Hamilton (2015) who directly draw in the structural parametrisation, requiring the use of Metropolis-Hastings. As the implementation of these procedures would imply additional computational burden and many thousands of draws will be necessary in this paper, the approach of Rubio-Ramirez et al. (2010) and Arias et al. (2018) is preferred.

Sign restrictions are then placed on high- and low-frequency variables. Nevertheless, similar to Uhlig (2005)'s proposal, the procedure is agnostic about the response of the industrial production after both shocks. This is robust to the mixed evidence for the importance of monetary shocks found by Ramey (2016) and compatible with the absence of effects of monetary policy on real activity in regressions run for the Great Moderation period. Following Uhlig (2005), however, to compensate for that agnostic approach, restrictions are applied to a longer period.[15]

It is postulated that an expansionary monetary policy shock decreases the fed funds rate and the EBP and increases the CPI for periods 0 to 5 months. In order to disentangle monetary policy shocks and prevent them from being a combination of other underlying shocks that satisfy the restrictions placed on the low-frequency variables, it is further assumed the target factor

[15] Canova and Paustian (2011) call dynamic sign restrictions into questions. Uhlig (2017) addresses these issues and concludes that sign restrictions beyond the initial impact can make a difference and should be used whenever plausible. Kilian and Lütkepohl (2017) also agree that such restrictions can be useful in restricting further the space of admissible models.

moves down on impact.[16]

An expansionary forward guidance shock is defined as a shock that decreases the 2-year government bond rate and the EBP and increases the CPI for period 0 to 5 months. Because forward guidance shocks are assumed to have no contemporaneous effect on the fed funds rate, the response of the fed funds rate is zero on impact. Furthermore, the path factor goes down impact. Once more, the inclusion of the factor is important to isolate the shock of interest from other shocks that might affect similarly the 2-year government bond rate, the CPI, and the EBP. Table 2.1 summarises the restrictions.[17]

Table 2.1: Zero and sign restrictions on responses

	MP shock	FG shock
target factor	-	
path factor		-
IP		
CPI	+	+
EBP	-	-
fed funds	-	0
2-year rate		-

The restrictions on the lower-frequency variables are standard and motivated by the New Keynesian set-up. Several sources (e.g. Smets and Wouters (2007), Gertler and Karadi (2011), McKay et al. (2016), Hagedorn et al. (2019)) show that Table 2.1 describes the expected responses to monetary and forward guidance shocks.[18] The restrictions on the high-frequency variables are such that the shock of interest is isolated. Because the window around the release is very narrow, it is assumed the surprises are not affected by macroeconomic news other than the announcement.

Moreover, since, as aforementioned, the path factor is closely related to FOMC statements, which telegraph not only forward guidance but also central bank private information, the sign restrictions on the EBP and the CPI are important to cleanse, by construction, the forward guidance shock from any informational advantage the central bank may have. As shown by

[16] Uhlig (2005) restricted the response of the nonborrowed reserves with the same objective.

[17] No zero restrictions were placed on the factors as this would increase the burden on the importance sampling.

[18] Despite the conflicting quantitative results for forward guidance shocks, the different models agree on the direction of the responses.

Jarociński and Karadi (2020)'s results, information shocks have the opposite effect on the EBP and the CPI. The combination of sign restrictions and high-frequency data then lends itself to an ideal way to properly disentangle pure monetary and forward guidance shocks.

However, the set of admissible structural parameters implied by sign restrictions can sometimes be too large with very different or implausible implications for the results. Arias et al. (2019) pointed out this is the case in Uhlig (2005), for instance, in which the posterior probability bands of the impulse responses are very wide and structural parameters incompatible with the systematic response of monetary policy to output are retained.

Narrative information

In order to refine the set of admissible structural parameters, the narrative account of a small number of key and uncontroversial events will be used to motivate further restrictions when estimating sign-identified VAR models as in Antolín-Díaz and Rubio-Ramírez (2018). This approach brings some flavour of the historical case studies pioneered by Friedman and Schwartz (1963), which are seen by Ramey (2016) as the best sources of evidence regarding the effects of monetary policy shocks.

In practice, to check if the narrative sign restrictions are satisfied, evaluate the following inequalities:

$$\varepsilon_{j,t}(\Theta) < 0 \tag{2.2}$$

$$|H_{i,j,t}(\Theta, \varepsilon_t(\Theta))| > \max_{j' \neq j} |H_{i,j',t}(\Theta, \varepsilon_t(\Theta))| \tag{2.3}$$

where Θ collects the values of all structural parameters, the first inequality implies jth shocks must be negative at time t and the second inequality implies the contribution H of the jth shock to variable i at time t must be greater than the contribution of any other shocks to variable i at time t.[19] The full algorithm is described in the Appendix.

Inspired by Ludvigson et al. (2020), an alternative type of shock-based constrained will also

[19]To have positive narrative sign restrictions, just impose equation (2.2) with a negative sign on the left-hand side.

be exploited:

$$\varepsilon_{j,t}(\Theta) < \bar{k} \tag{2.4}$$

This condition requires $\varepsilon_{j,t}(\Theta)$ to be less than $\bar{k}$ standard deviations below zero. Such condition is in-between the restrictions placed by equations (2.2) and (2.3), so it will be useful for the cases whereby a restriction on the historical decomposition would be considered too strong, and a restriction just on the sign of the shock would be considered too weak. A variant of it, however, in which $\bar{k}$ denotes the standard deviations above zero is weaker. It is easy to see that the restriction in equation (2.2) is a special case of the condition in equation (2.4) when $\bar{k} = 0$. One has to choose the type of restriction according to their confidence in the episodes.

For the monetary policy shocks, the main source is Antolín-Díaz and Rubio-Ramírez (2018), who examined in detail episodes that are good candidates to have been conventional monetary policy. The dates that are comprised in the shorter sample period here considered are: February 1994, October 1998, April 2001 and November 2002. Antolín-Díaz and Rubio-Ramírez (2018) also point out it is possible to obtain qualitatively similar results imposing narrative restrictions in February 1994 on its own.[20] Particularly, February 1994 was the month the FOMC began a series of tightening moves and caught the market by surprise.

For the forward guidance shocks, the main references are the site of the Federal Reserve Board as well as Gürkaynak et al. (2005), Campbell et al. (2012), and Borio and Zabai (2018), who scrutinised the FOMC statements and highlighted some important episodes of forward guidance. Some examples are:

August 2011, when the FOMC specified the intended time length of the stimulus and replaced “extended period” with “mid-2013”: “The Committee currently anticipates that economic conditions ... are likely to warrant exceptionally low levels for the federal funds rate at least through mid-2013.”

January 2012, when the FOMC replaced “mid-2013” with “late 2014”: “the Committee ... currently anticipates that economic conditions ... are likely to warrant exceptionally low levels

[20] It is worth noting, however, that when their sample is shortened to 1993-2007, such restriction is no longer sufficient to imply that contractionary monetary policy shocks cause output to fall (not even using a Minnesota prior).

for the federal funds rate at least through late 2014."

September 2012, when "late 2014" was replaced with "mid-2015": "the Committee ... currently anticipates that exceptionally low levels for the federal funds rate are likely to be warranted at least through mid-2015."

December 2012, when forward guidance became based on the state of the economy: "the Committee ... currently anticipates that this exceptionally low range for the federal funds rate will be appropriate at least as long as the unemployment rate remains above 6-1/2 percent, inflation between one and two years ahead is projected to be no more than a half percentage point above the Committee's 2 percent longer-run goal, and longer-term inflation expectations continue to be well anchored."

December 2014, when the FOMC moved a step closer to the beginning of the normalisation: "Based on its current assessment, the Committee judges that it can be patient in beginning to normalize the stance of monetary policy."

October 2015, when the FOMC replaced the clause "how long it will be appropriate to maintain [the target range]" with "whether it will be appropriate to raise the target range at its next meeting."

Because restrictions are placed on both shocks, parsimony is required.[21] So, only the most informative episodes will be selected as restrictions for the benchmark. Alternative combinations will the presented in the robustness subsection. August 2011 marks the change to date-based forward guidance. Before that, the Committee used to be relatively more vague and write expressions such as "for a considerable period" (December 2003), "for some time" (December 2008) or "for an extended period" (March 2009). August 2011 is the announcement Bundick and Smith (2020) use in their model implied-responses and it is also the episode on which Bundick et al. (2017) focus.

In fact, the posterior distribution of the forward guidance shock implied by the VAR identified with sign restrictions during that month is already very concentrated on the left side of the histogram. Figure 2.1 shows the posterior distribution of the forward guidance shock during

[21] The computational aspects of the estimation are described in the Appendix.

that month. Grey represents the posterior distribution with only sign restrictions, and pink represents the distribution after the imposition of narrative sign restrictions.

On the one hand, this implies sign restrictions are being effective in identifying a forward guidance shock in August 2011, on the other hand it means informing the model that a expansionary forward guidance shock occurred in that particular month does not bring much refinement. Nonetheless, combining such restriction with a restriction on the historical decomposition proves to be informative. Therefore, even though most of the distribution already has negative support even before the imposition of narrative sign restrictions, the new restriction increases the weight in the negative region in line with the narrative account. This will be the benchmark restriction.

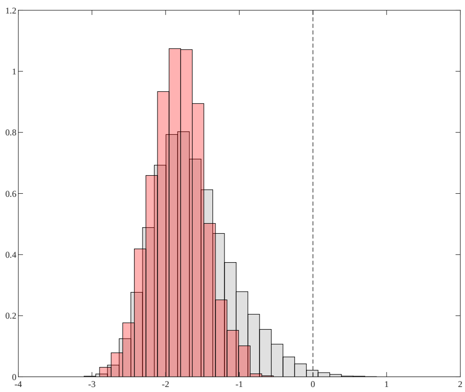

Figure 2.1: Forward Guidance Shock for August 2011

In January 2012, the period of exceptionally low interest rates was extended considerably. This is also considered a strong episode of forward guidance, being even the benchmark for the simulation in Campbell et al. (2012). As in the previous episode, surprises in S&P 500 and in the path factor are consistent with that narrative account. Once more, however, the posterior distribution of the forward guidance shock during that month using only sign restrictions is already concentrated below zero, albeit less than in August 2011. This episode will be explored in the robustness subsection.

September 2012 is considered by Del Negro (2018) an episode of Odyssean forward guidance. The statement declared that the period of low interest rates was going to be further lengthened. Nevertheless, this was also the meeting in which QE3 was announced. So, results could be dominated by large-scale asset purchases (LSAP).[22] In December 2012, there was a switch to state-contingent forward guidance. This episode is harder to interpret in that it represents a change in the nature of forward guidance from calendar-based to one that conditioned the path of interest rates on specific numbers of inflation and unemployment. Also, calendar-based forward guidance is more in line with the experiments conducted in New Keynesian Models for which this paper is the empirical counterpart. Moreover, there was also the announcement of the extension of large scale asset purchases.

In December 2014, the FOMC announced the intention to patiently begin the normalisation. This introduced a message more specific than the contained in the statement from October 2014 in which the FOMC had stated that increases in the target range for the federal funds rate could "occur sooner than currently anticipated" or "later than currently anticipated". However, Fed Chair Janet Yellen said in the press conference that "The committee considers it unlikely to begin the normalization process for at least the next couple of meetings."

In October 2015, the FOMC hinted that a hike might happen in the next meeting. Responding to that the path factor went up and the S&P 500 went down. Even though there is not much literature about this meeting, it will be explored as robustness since, in addition to the market reaction, news at the time are consistent with the narrative account of this episode as a contractionary forward guidance shock. This episode also highlights the importance of the monthly frequency to assess forward guidance since models estimated at the quarterly frequency would aggregate data from October to December, confounding forward guidance and conventional monetary policy shocks.

To sum up, restrictions are based on the confidence in the episodes as well as in their informativeness. After cross-checking news, the market reaction, and the literature, the following episodes were selected as the benchmark restrictions:

Narrative Sign Restriction 1. The monetary policy shock must be positive for the observation corresponding to February 1994.

Narrative Sign Restriction 2. For the period specified by Restriction 1, the monetary policy shock is the most important contributor to the observed unexpected movements in the

[22] The interaction between LSAP and forward guidance will be further explored in the robustness subsection.

federal funds rate. In other words, the absolute value of the contribution of monetary policy shocks is greater than the absolute value of the contribution of any other structural shock.[23]

Narrative Sign Restriction 3. The forward guidance shock must be negative for the observation corresponding to August 2011.

Narrative Sign Restriction 4. For the period specified by Restriction 3, the forward policy shock is the most important contributor to the observed unexpected movements in the 2-year rate.

2.2.2 Potential Advantages over Proxy SVARs

Proxy SVARs rely on external instruments correlated with the shock of interest, and uncorrelated with other structural shocks. Moreover, to address the issue of whether the high-frequency surprises are truly exogenous or just reflect the Fed's private information, the measures or surprises are regressed on measures of the Fed's private information. The results, however, are dependent on the way this measure is built and can be puzzling (Ramey, 2016; Miranda-Agrippino and Ricco, 2019). In the hybrid approach of this paper, however, even when forward guidance shocks are accompanied by information shocks, there is no need to purge the path factor from central bank private information since this is achieved by construction through the sign restrictions as already pointed out.

Another potential advantage is related to the invertibility assumption. Plagborg-Møller and Wolf (2021), Miranda-Agrippino and Ricco (2019) and Paul (2020) show that, under some conditions, the impulse responses obtained with Proxy VARs are equivalent to the ones obtained with a recursive scheme that includes the instrument as an endogenous variable and orders it first. Nonetheless, with news shocks, or more specifically forward guidance shocks, invertibility concerns become even more serious (Ramey, 2016; Plagborg-Møller and Wolf, 2018). Incorporating the path factor into the vector of variables of the VAR makes the inference valid even if the VAR without it is not fully or partially invertible.

[23]Slightly deviating from Antolín-Díaz and Rubio-Ramírez (2018), the restriction placed on the historical decomposition in February 1994 imposes that the forward guidance shock be the most important contributor, not greater than the sum of the contribution of all other shocks. This change is motivated by the fact that a strong restriction would increase the number of required draws without changing the results.

This is in line with and exemplified in D'Amico and King (2015). They show that, for the specific case of forward guidance, measures of expectations should be included in the VAR to avoid misspecification even when there is no special interest in these variables. Including the path factor in the VAR as a variable tackles this issue. Finally, a formal difference pointed out by Arias et al. (2021) is that Jarociński and Karadi (2020)'s strategy assumes that the structural shocks are linear combinations of the proxies while Proxy SVARs only assume that the structural shocks are correlated with linear combinations of the proxies.

2.3 Results

2.3.1 Impulse Responses

Figure 2.2 compares the impulse responses after a conventional monetary policy shock when only sign restrictions are imposed to the case where Narrative Sign Restrictions 1, 2, 3, and 4 are imposed (on top of sign restrictions). Grey and blue represent the results with only sign restrictions, and pink and red represent the narrative sign restrictions. Unless otherwise stated, the estimates discussed in this section refer to the red line.

The impulse responses are computed after a one-standard-deviation expansionary conventional monetary policy shock in the narrative sign restrictions scheme, a decrease of approximately 5 basis points in the fed funds rate (FF). In order to make the impulse responses comparable, the impulse responses of 'sign restrictions only' are normalised so that the initial median impact on FF is the same as in the case with narrative sign restrictions. As in Nakamura and Steinsson (2018) and Jarociński and Karadi (2020), monetary shocks are quite small.[24] Overall, the narrative restrictions narrow down the set of responses substantially.

Industrial production goes up on impact and this effect is persistent. This result is stronger than Antolín-Díaz and Rubio-Ramírez (2018) re-estimated for a post-90 sample and subject to Narrative Sign Restriction 1 and 2 as their VAR would not have found any effect of monetary policy on industrial production for this new specification. The "significant" increase of industrial production also contrasts with the evidence surveyed by Barakchian and Crowe (2013) and Ramey (2016), who find that several specifications and identification schemes do not lead to

[24] Their monetary shocks, however, comprise conventional monetary policy shocks and forward guidance shocks.

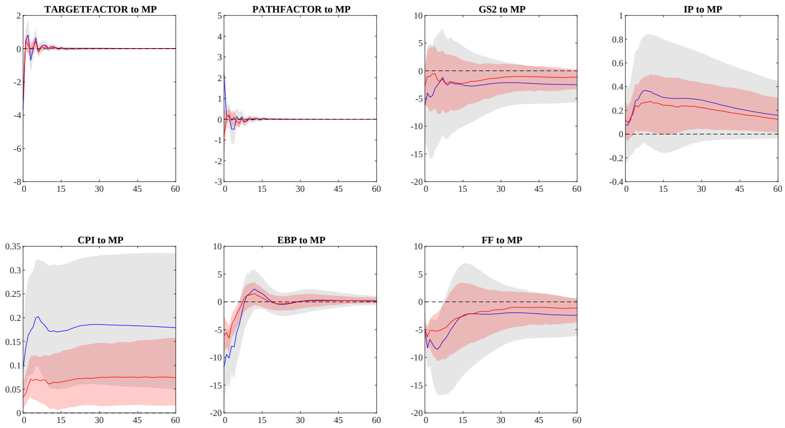

Figure 2.2: Impulse Responses to a Conventional Monetary Policy Shock

Notes: The grey shaded area represents the 68 percent (point-wise) HPD credible sets for the IRFs and the blue lines are the median IRFs using sign restrictions. The pink shaded areas and the red lines display the equivalent quantities for the models that additionally satisfy narrative sign restrictions.

the expected responses when estimated for recent periods. On the other hand, the behaviour of industrial production is similar to Caldara and Herbst (2019), who find a persistent decline in real activity after a contractionary shock for a recent sample after incorporating credit spreads. Coupled with the imposition of restrictions for 6 periods, this leads to rather persistent effects.[25]

Specifically, the posterior median increases 0.2% in the second quarter following the shock, an order of magnitude in line with some studies that employ high-frequency or narrative identification to study the effects of monetary policy, such as Paul (2020) and Gertler and Karadi (2015), after cleansing their measure of policy surprises of the Fed's private information.[26] After some time, however, this movement is attenuated and the effect drops by almost half. CPI increases on impact and slightly more than 0.07% in the long run, a value consistent with previous studies. The excess bond premium goes down 6 basis points and GS2 decreases a little less than half of the magnitude of the initial impact on FF, but 0 is within the interval.[27]

Figure 2.3 compares the impulse responses after a forward guidance shock. The impulse responses are computed to a one-standard-deviation expansionary forward guidance shock in the narrative sign restrictions scheme, a decrease of 8.4 basis points in the 2-year rate (GS2).

[25]In a recent contribution, Jordà et al. (2020) also find that the effects of monetary policy are very persistent.
[26]Holm et al. (2020) also find new evidence of strong effects on industrial production at a monthly frequency for Norway.
[27]Even though factors are built to be unconditionally uncorrelated, their correlation conditional on the other variables in the system differs from zero.

The impulse responses of the case with ‘sign restrictions only’ are then normalised so that the initial median impact on GS2 is the same across the identification schemes. Industrial production increases more than 0.3% after some time, but uncertainty is higher even after the refinement provided by the narrative restriction. Furthermore, the effect is lessened in the medium run. CPI goes up almost 0.1% and the excess bond premium decreases 7 basis points. The fed funds rate goes down in 12 months by a magnitude slightly lower than the initial impact on GS2.

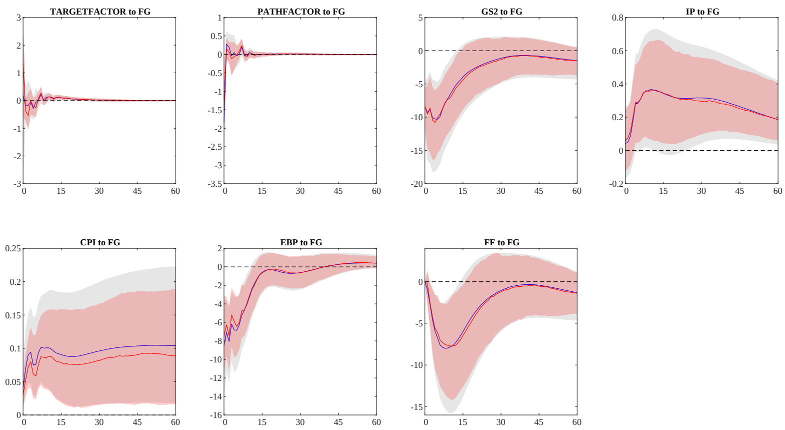

Figure 2.3: Impulse Responses to a Forward Guidance Shock

The credible sets do not shrink as much as in the conventional monetary policy case. This happens because, as aforementioned, most of the posterior for August 2011 was already in negative terrain. Still, the narrative restrictions help reduce the HPDI. Lastly, it should be noted that the response of the industrial production to a forward guidance shock is at least as strong as its response to a conventional monetary policy shock in a HPDI sense as displayed in Figure 2.4.

Benchmarks for the effect of forward guidance shock in VARs are more scarce. Bundick and Smith (2020) find that expansionary forward guidance shocks lead to moderate increases in output and the price level. Nevertheless, because they do not cleanse their modified path factor of information shocks, their estimates of the effects of forward guidance capture only the net-effect of FOMC communication. Despite not being able to formally compare the results with the effects of conventional monetary policy shocks due to their focus on forward guidance shocks, they find, as in here, that forward guidance shocks share many empirical features with conventional monetary policy shocks.

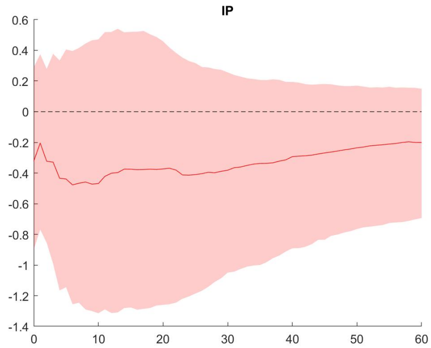

Figure 2.4: Difference in Impulse Responses of IP

Notes: The pink shaded area represents the 68 percent (point-wise) HPD credible sets for the difference between the IRFs of IP after a FG shock and after a MP shock. In order to make the original impulse responses comparable, they are normalised so that the initial impact on GS2 is the same after both shocks: 8.4 basis points.

D'Amico and King (2015) find a significant reduction on impact for both CPI and output after a contractionary shock. They find that responses to the policy-expectations shock are stronger than the responses to the unanticipated shock. Lakdawala (2019) reports an increase in CPI and industrial production following a contractionary shock. Nevertheless, after cleansing the path factor of Fed private information, there is a small but insignificant decline in output while the price puzzle remains. In addition to the distinct sample period, part of the difference in the results derive from the fact that his benchmark VAR does not include the EBP, which induces an attenuation effect also in the response of the variables to forward guidance shocks. The fact that the response of industrial production in Lakdawala (2019) is less puzzling when the EBP is incorporated in his baseline VAR is consistent with that.

To sum up, the results show forward guidance matters for macroeconomic outcomes, including industrial production, being an effective policy tool. In fact, it may be an important part of shocks labelled as monetary policy shocks. Nakamura and Steinsson (2018) and Jarociński and Karadi (2020), for instance, acknowledge that their monetary policy indicator/surprises capture the effects of "forward guidance" whereas here the monetary policy shock captures only the conventional monetary policy shock.

2.3.2 Informational Sufficiency

A common concern about VARs is whether the structural shocks are fundamental. In a model with shocks that may be anticipated by economic agents, such concern is even more important. To check this, the orthogonality F-test proposed by Forni and Gambetti (2014) is conducted. It consists of a regression of the shocks on a large dataset capturing agents' information set and an F-test for the significance of the regression. In practice, the agents' information set is summarised by the past values of principal components of the FRED-MD database (McCracken and Ng, 2016).

The idea is that if the shocks are predicted by past available information, the structural MA representation of the variables included in the VAR is non-fundamental and the VAR is misspecified, in the sense that there is not sufficient information to recover the structural shocks. Such an approach is appealing in that it does not require a well-defined theoretical model of reference.

Table 2.2 presents the results of the test for different combinations of the number of lags and principal components. The hypothesis that the shocks are not predicted by past available information is not rejected for either of shocks or number of lags, when the choice of principal components is based on the Bai and Ng (2002)'s criteria (PC=7). Such result is robust to different numbers of PCs. This orthogonality to the past of the "state variables" associated with a correct identification scheme implies both shocks are indeed the desired object of interest: conventional monetary shocks and forward guidance shocks.

Table 2.2: P-Values of the Orthogonality F-Test Proposed by Forni and Gambetti (2014)

	2 lags			4 lags		
	PC=4	PC=7	PC=10	PC=4	PC=7	PC=10
MP shock	0.98	0.80	0.37	1.00	0.77	0.14
FG shock	0.86	0.27	0.30	0.94	0.52	0.23

2.3.3 Alternative Narrative Restrictions and Other Robustness Exercises

The benchmark results relied on the information provided by August 2011. Other potentially good candidates are January 2012, which was also used by previous literature, and October 2015. A restriction on the sign for the observation corresponding to January 2012 refines just a little the set of admissible parameters. In fact, as shown in the appendix when such restriction is added on top of NSR 1, 2, 3, and 4, results are almost the same. On the other hand, a restriction on the historical decomposition does not work well, being too restrictive and decreasing substantially the effective sample size as a consequence.

That is when the type of restriction inspired by Ludvigson et al. (2020) can be helpful. Making $\bar{k} = 0.3$, jointly with NSR 1 and 2, gives results very similar to Figure 2.3. This restriction is convenient because it brings more information than a restriction on the sign while one has only to be confident that a forward guidance shock occurred regardless of what happened to the other shocks. Results are presented in the appendix. Finally, placing a restriction on the sign of the forward guidance shock in October 2015 does not help reduce identification uncertainty, even though it changes the posterior distribution for this shock in this particular month.

Henceforth, NSR 1, 2, 3, and 4 are always imposed, and the robustness exercises involve other types of modifications. To account for the fact that even the 2-year government bond rate may have been affected by the ZLB, the sample period is also ended in 2012. The results are qualitatively the same. Furthermore, to check whether results are dominated by LSAP dates, the path factor from November 2010 (QE2), September 2011 (MEP[28]), and September 2012 (QE3) are also dropped in addition to September 2001 (9/11) and March 2009 (QE1) that had already been excluded from the benchmark. The impulse responses are similar to the baseline estimates. Overall, the main message is still the same: forward guidance is an effective policy tool.

Exploring alternative sign restrictions can also help understand the drivers of the results. In the first alternative set of restrictions, CPI is left unrestricted while sign restrictions are placed on IP on top of the benchmark restrictions placed on the other variables. In that case, there is no longer an effect after a MP shock and, even though, the impulse response after a FG shock

[28]Maturity Extension Program.

is more concentrated is the positive region, zero is still within the credible set. Second, the sign restrictions are agnostic about CPI and IP. Results do not change much in comparison with the previous impulse response, and there is still an effect on IP after both shocks. This shows the strength of the effects on IP and the reliance of the benchmark results for CPI on the sign restriction on the first horizons, especially after MP shocks.

Finally, Jarociński and Karadi (2020)'s central bank private information shocks are included as an exogenous variable in the VAR. This is important assess whether the sign restrictions are being effective in cleansing the shocks of interest from informational advantages central banks may have. Results are very similar to the benchmark, suggesting that sign restrictions are indeed effective.

2.4 Conclusion

This paper has addressed the identification of conventional monetary policy and forward guidance shocks. In order to do that, two sources of extraneous information – high-frequency surprises and narrative evidence – were combined with sign restriction in a structural VAR. The factors allow for a proper isolation of conventional monetary and forward guidance shocks from other shocks (or a combination of shocks) that satisfy the sign restrictions placed on the low-frequency variables. The narrative restrictions help further refine the credible set.

Results show that, in contrast with the evidence surveyed by Barakchian and Crowe (2013) and Ramey (2016), the identification scheme leads to the expected responses for output following a conventional monetary policy shock even when the model is estimated for a recent sample: 1993-2017. In fact, a strong effect emerges from the refinements in the identification.

Results also show that forward guidance has been an effective policy tool. Therefore, forward guidance matters not only to the proper identification of conventional monetary policy shocks but also due to its effect on output and other macroeconomic variables. Specifically, its effects on industrial production are at least as strong as the effects of conventional monetary policy.

Several robustness exercises show that the results hold under alternative specifications. An important implication of such results is that they provide additional support for the view that the Federal Reserve may not be so constrained even during ZLB periods.

Chapter 3

Forecasting with VAR-teXt and DFM-teXt models: exploring the predictive power of central bank communication[1]

3.1 Introduction

Central bank communication has become an increasingly relevant aspect of monetary policy. The days of "never explain, never excuse", reputedly Sir Montagu Norman's motto[2], were replaced with days of regular communication that takes place via a wide array of formats: statements, minutes, implementation notes, press conferences, and so on.

During crisis times, the role of communication in shaping expectations becomes even more relevant since the short-term policy rate is usually constrained by the zero lower bound (Blot and Hubert, 2018). In these times, as claimed by Gros (2018), central bank communication becomes the policy. In fact, central bank releases not only telegraph likely future actions but also some private information about the future state of the economy (Campbell et al., 2012). Hence, they

[1] This chapter is available in the Central Bank of Brazil Working Paper Series as: Ferreira (2021), 'Forecasting with VAR-teXt and DFM-teXt models: exploring the predictive power of central bank communication' (No. 559). I have benefited from comments from referees and participants at the 29th Annual SNDE Symposium, the Econometric Society European Meeting 2021, the 2021 Latin American Meeting of the Econometric Society, 43rd Meeting of the Brazilian Econometric Society, the 2021 Africa Meeting of the Econometric Society, the European Seminar on Bayesian Econometrics 2021, 25th Spring Meeting of the European Association of Young Economists, the First Annual Southern PhD Economics Conference, the 21st Brazilian Finance Meeting, Universidade Católica de Brasília, and the 5th Workshop of the Research Network of the Banco Central do Brasil. All errors are mine. The views expressed in this chapter are those of the author and not necessarily reflect those of the Banco Central do Brasil.

[2] See Bernanke (2007). Sir Montagu Norman was the Governor of the Bank of England from 1920 to 1944.

can provide information beyond that contained in traditional economic models, being potentially important for forecasting.

This paper investigates this potential by assessing the contribution of information extracted from Federal Open Market Committee (FOMC) statements to the forecasting of macroeconomic and financial variables. In order to do that, text mining and conventional econometric techniques are combined in a VAR-teXt: a vector autoregressive (VAR) model augmented with exogenous variables that capture information retrieved from text.

Taking advantage of the common estimation approach – Gibbs sampling – the transformation of text into quantitative data is done through a Latent Dirichlet Allocation (LDA) model. LDA is a probabilistic topic model developed by Blei et al. (2003) that, essentially, groups words into topics and describes documents in terms of them. As such, it helps reduce documents to low-dimensional space.

Nevertheless, departing from the common approach in economic applications, which involves selecting the number of topics based on their interpretability, and consistently with the aim of this paper, the number of topics (or textual factors) is chosen based on the predictive performance of the VAR-teXt. While some topics will be meaningful, others will not. Still, the full set of time series of statement-specific topic distributions may contain variation that is worth exploring. That is why they will generally be treated as textual factors. This paper then evaluates whether such textual factors are forecast improving by comparing the performance of the VAR-teXt with that of the benchmark VAR.

This agnostic approach is inspired by factor-augmented models, where factors help exploit additional data sources and uncover hidden patterns even without necessarily being identified as specific economic concepts.[3] In fact, both being examples of unsupervised learning tasks, topic modelling and factor analysis share many features. A complementary, and probably more familiar to economists, definition of LDA, attributed to Hansen et al. (2018), highlights this connection: LDA is a Bayesian factor model for discrete data with factors representing topics.

This paper also extends previous literature by proposing a Markov chain Monte Carlo (MCMC) sampling algorithm for the estimation of VAR-teXt models that takes into account the fact that

[3]For instance, in the context of sufficient information in structural VARs, Forni and Gambetti (2014) suggest keeping adding principal components to the VAR until the model is no longer deficient.

textual factors themselves are estimates. By exploring the entire distribution over textual factors, the proposed MCMC sampling algorithm properly captures the estimation uncertainty, leading to more accurate posterior predictive distributions.

Finally, a natural extension of the VAR-teXt is evaluated. With the purpose of using more conventional data in the analysis, the approach is extended to dynamic factor models (DFM). This allows for use of all the information that is easily accessible regardless of its form and gives rise to the DFM-teXt, which is simply a VAR-teXt in the (data) factors.

In summary, the contribution of this paper is twofold. In terms of methodology, the novelty is to explore the complementarity between traditional econometrics and machine learning to estimate more accurately VAR-teXt and DFM-teXt models. The economic contribution is to show that, using the proposed models and algorithms, central bank communication can improve forecasts.[4]

Specifically, results show the VAR-teXt consistently outperforms the benchmark VAR in forecasting the consumer price inflation and the interest rate as measured by log-scores, and such results are robust to several alternative specifications. The DFM-teXt also performs better than the DFM, even for industrial production, as well as many other variables, but only for the 3-month ahead forecasts.

Related Literature

Several papers have explored the link between text and the economy. The closest ones can be divided into two groups. The first one explores central bank communication. Hayo and Neuenkirch (2010), for instance, interpret and sort central bank communications into three categories depending on whether they indicate likely increases (+1), decreases (-1), or no change (0) in the fed funds rate. They then use this indicator in an ordered probit model estimated to predict changes in monetary policy and show it performs very well in their out-of-sample assessment.

While such narrative approaches may capture some nuances, they are prone to subjectivity. Automated approaches, on the other hand, have two meaningful advantages: scalability to larger

[4]Even though this may seem expected, there is also evidence such as Lustenberger and Rossi (2020) showing more communication may increase forecast errors.

corpora and reduction of biases that may unfold when readers overlook patterns that do not conform to prior beliefs (Bholat et al., 2015).

Lucca and Trebbi (2009) address this concern by using an automated approach to extract information from FOMC statements in order to forecast macroeconomic variables with univariate and VAR models. Their objective is rather similar to this paper's, but they differ in that i) they construct semantic orientation scores; and ii) they focus on in-sample predictive power. Results show that changes in communication as measured by the scores help predict future policy rate.

Hansen et al. (2019) use LDA and dictionaries to measure a set of high-dimensional signals based on the Bank of England's Inflation Report. Using these signals in an elastic net regression, they find that, beyond the conventional expectations channel, signals about the expected uncertainty can have important effects along the yield curve, especially in the long run. Their analysis, however, is also in sample.

The second group gathers papers that are closer to forecasting evaluations used in practice. Models are re-estimated recursively and out-of-sample predictions are assessed with log-scores or root mean square errors (RMSE). Nonetheless, they use text from newspapers instead of from central banks. Thorsrud (2018) builds a daily coincident index of the business cycle based on textual data retrieved from a major Norwegian business newspaper. Specifically, he uses tone-adjusted topic time series in a mixed-frequency time-varying factor model and shows the resulting model performs substantially better than simple time series models.

As it is becoming standard in this new literature, topic proportions before tone adjustment are given by the average over LDA draws. In addition to this and the application, Thorsrud (2018) differs from this paper in that topics are added in the panel of variables of the observation equation and not as exogenous variables. This can create an inconsistency between the data generating process for the topics, which rules out time-series dependencies across documents, and their dynamics in the regression.

Larsen and Thorsrud (2019) use LDA on the same newspaper and augment autoregressive (AR) models with tone-adjusted topic time series (AR-X). Unlike Thorsrud (2018), however, they include these series as exogenous variables in the regression. They then compare the predictive power of the AR-X with the benchmark AR and identify which topics improve the forecasts of key economic variables. In particular, they find some topics have predictive power for asset prices, which are forward looking and supposed to reflect all publicly available information.

Kalamara et al. (2020) retrieve information from UK newspapers using several methods and incorporate the resulting measures in AR and factor models. They choose to exclude topic models to avoid identification issues that may arise when the model is re-estimated recursively. The forecasts of most macroeconomic variables are improved when news text is included in the AR model, but not in the factor model. As the papers cited above, however, they do not address the fact that the outputs of the machine learning techniques are estimates.

This paper complements this literature by evaluating the contribution of textual factors based on FOMC statements to the prediction of macroeconomic and financial variables in a recursive exercise, useful for policy makers and forecasters, while taking into account that such factors are estimates, and thus providing more accurate posterior predictive distributions. The rest of the paper is organised as follows. Section 3.2 describes the data. Section 3.3 presents the VAR-teXt model and the MCMC sampling algorithm. Section 3.4 evaluates the performance of this model. Section 3.5 presents some robustness exercises. Section 3.6 introduces the DFM-teXt model and Section 3.7 concludes.

3.2 Data

The traditional dataset consists of 3 macroeconomic and 1 financial variables from 1998M09, when the release of statements after scheduled meetings became more common, to 2020M02. Industrial production growth and consumer price inflation are calculated by taking the first difference of the logarithm of the corresponding indices downloaded from FRED-MD, a large macroeconomic database with 128 time series described in McCracken and Ng (2016). The shadow rate proposed by Wu and Xia (2016) is used as a measure of the stance of the Fed since the zero lower bound affected a considerable part of the sample period. The shadow rate is convenient in that it also summarises unconventional monetary policy. Finally, to capture information about future economic activity and improve the informational content of the regression, the excess bond premium (EBP), introduced by Gilchrist and Zakrajšek (2012), is also incorporated into the vector of variables. The EBP is a corporate bond credit spread purged from the default risk, a useful leading indicator.

3.2.1 Some Useful Terms

Some key terms from the text mining literature will be frequently used in this chapter. They are defined as follows:

- Stopwords: words that are not very informative, such as “a”, “and” and “the”.

- Stemming: the process of reducing words to common linguistic roots. For instance, “increasing”, “increased” and “increases” become “increas”, which is the stem.

- Token: a meaningful unit of text. In this paper, tokens are words.

- Corpus: a collection of documents.

3.2.2 The Corpus of FOMC Statements and the LDA

The text-augmented model also includes information retrieved from the corpus of the 169 FOMC statements that followed scheduled meetings during the period of analysis. The structure of the statements that follow scheduled meetings is fairly comparable over time, making them a fitting choice for the text analysis. The corpus was scraped from the Federal Reserve website and pre-processed before estimation. This involves conversion to lower case, removal of punctuation, white spaces, numbers and stopwords, and stemming.

After this pre-processing, the corpus of 169 statements has 29,903 stems, of which 858 are unique. Nonetheless, in order to make the analysis more granular, the first step of the estimation is conducted at the level of the paragraph, meaning that there are actually 693 documents. LDA can then be used. As explained by Blei et al. (2003), LDA reduces any document to a fixed set of real-valued features on a low-dimensional latent space: the posterior Dirichlet parameters. Given such parameters, it is possible to obtain the statement-specific mixing probabilities $\tilde{\theta}_d$: the textual factors.

It is worth mentioning there are alternatives to the LDA, such as the Dynamic Topic Model (DTM), which introduces time-series dependencies into the data generating process, and the Structural Topic Model (STM), which introduces covariates into a topic model (Blei and Lafferty, 2006; Roberts et al., 2016). Nonetheless, non-conjugacy makes sampling methods more difficult for such models, and the algorithms usually depart from Gibbs sampling, which is more familiar to economists. Therefore, the focus here will be on handling the generated regressors issue while leaving a joint solution for both problems for future research.

LDA is estimated with a collapsed Gibbs sampling algorithm. First, the estimation of word's topic assignment consists of the following steps:[5]

Step 1. Randomly allocate to each token in the corpus a topic assignment drawn uniformly from $\{1, ..., K\}$, where K denotes the number of textual factors.

Step 2. For each token, sequentially draw a new topic assignment via multinomial sampling where

$$P[q_{d,n} = k|Q^{-}, W, \alpha, \eta] \propto \frac{m^{-}_{k,v} + \eta}{\sum_v m^{-}_{k,v} + V\eta}(n^{-}_{d,k} + \alpha) \tag{3.1}$$

where the '-' superscript denotes counts excluding (d, n) term, with d representing documents and n their terms. $q_{d,n} = k$ denotes the topic assignment of (d, n) term and $Q = (\mathbf{q}_1, ..., \mathbf{q}_D)$ the other topic assignments. $W = (\mathbf{w}_1, ..., \mathbf{w}_D)$ is the observed data. α is the prior on the document-specific mixing probabilities and η on the topic-specific term probabilities. $m_{k,v} \equiv \sum_n \sum_d \mathbb{1}(q_{d,n} = k)\mathbb{1}(w_{d,n} = v)$ is the number of times topic k allocation variables generate term v, and $n_{d,k} \equiv \sum_n \mathbb{1}(q_{d,n} = k)$ is the number of words in document d that have topic allocation k. V is the number of unique terms.

Step 3. Repeat Steps 1 and 2 until the required number of draws has been reached.

To fit the LDA, the number of textual factors K and hyperparameters α and η need to be fixed a priori. In applications in economics, K is usually chosen based on interpretability. In this application, however, a grid for K will be explored and the best K will be selected based on the out-of-sample forecasting performance of the VAR-teXt. This helps discipline the choice.[6] Following Griffiths and Steyvers (2004) and Hansen et al. (2018), η is set to $200/V$ and α to

[5] This subsection brings forward part of the estimation so as to show how documents and words are summarised into time series. This follows very closely the Online Appendix of Hansen et al. (2018). For more details see also Murphy (2012).

[6] There are alternative ways to make the selection of the number of topics more objective and disciplined such as the topic coherence proposed by Newman et al. (2010), which uses a co-occurrence measure based on pointwise mutual information over Wikipedia. Nonetheless, selecting the number of topics based on the predictive ability of the VAR-teXt is more consistent with the objective of this paper.

$50/K$.

Next, a process called querying is carried out. Querying allows the document distributions aggregated at the level of the statement to be recovered. It corresponds to running the Gibbs sampling keeping the topic-specific term probabilities φ fixed at their estimated values. This is done by collapsing paragraphs back into the statement level and sequentially sampling from:

$$P[\tilde{q}_{d,n} = k|\tilde{Q}^{-}, \tilde{W}, \alpha, \eta] \propto \hat{\varphi}_{v_{d,n}}(n^{-}_{\tilde{d},k} + \alpha) \tag{3.2}$$

where tilde denotes the new document level. Nonetheless, this gives estimates of each word's topic assignment since the topic proportions θ were integrated out in the derivation of the collapsed Gibbs sampling. In order to recover them, the output of interest, the final step is to compute the statement predictive distributions using:

$$\tilde{\hat{\theta}}^{k}_{\tilde{d}} = \frac{n_{\tilde{d},k} + \alpha}{\sum_{k=1}^{K} n_{\tilde{d},k} + \alpha} \tag{3.3}$$

Such textual factors are incorporated in the VAR as described in the next section.[7]

3.3 The VAR-teXt Model

The point of departure for the analysis is the benchmark VAR model of the form:

$$Y_t = \sum_{p=1}^{P} \beta_p Y_{t-p} + \mu + A_0 \varepsilon_t \tag{3.4}$$

where Y_t is the $N \times 1$ vector of standard macro variables, p denotes the lags, with $p = 1, ..., P$, and A_0 is a decomposition of the covariance matrix Σ such that $Var(u_t) = A_0 A_0' = \Sigma$.

The VAR-teXt, text-augmented VAR, is given by:

$$Y_t = \sum_{p=1}^{P} \beta_p Y_{t-p} + \phi X_{t-1} + \mu + A_0 \varepsilon_t \tag{3.5}$$

[7] The textual factors are transformed into first differences and standardised as in Larsen and Thorsrud (2019). The monthly series are built letting the series take a value of zero in months without scheduled FOMC meetings (the FOMC only holds eight regularly scheduled meetings a year).

where X_{t-1} are the first $K-1$ textual factors.[8]

Because traditional data is not taken into account in the estimation of the textual factors, the feedback from lagged endogenous variables to the textual factors is restricted and they are incorporated into the model as exogenous variables. Hence, the VAR-teXt model is consistent with the data generating process for the textual factors.

Moreover, as in static factor models and approximate static factor models, LDA does not account for the time-series dependence in estimating the textual factors. That is why the textual factors are treated differently and appear in the equation only once. As aforementioned, these issues could be tackled with the use of alternative approaches. This, however, would impair the use of a simple MCMC sampling algorithm.

3.3.1 Estimation

Following Bańbura et al. (2010), equation (3.5) can be written in a more compact way, which is more convenient for the estimation:

$$Y = Z\beta + \varepsilon A_0 \tag{3.6}$$

where $Y = (Y_1, ..., Y_T)'$, $Z = (Z_1, ..., Z_T)'$ with $Z_t = (Y'_{t-1}, ..., Y'_{t-P}, X'_{t-1}, 1)'$, $\varepsilon = (\varepsilon_1, ..., \varepsilon_T)'$, and $\beta = (\beta_1, ..., \beta_P, \phi, \mu)'$ is the $(NP+K) \times N$ matrix containing all the coefficients.

A conjugate prior for the VAR parameters is then introduced by augmenting the vector of variables with the following dummy observations:

$$Y_{d,1} = \begin{pmatrix} \frac{diag(\delta_1\sigma_1,...,\delta_N\sigma_N)}{\tau} \\ 0_{N(P-1)\times N} \\ .... \\ diag(\sigma_1, ..., \sigma_N) \\ .... \\ 0_{1\times K} \end{pmatrix}, \quad Z_{d,1} = \begin{pmatrix} \frac{J_p \otimes diag(\sigma_1,...,\sigma_N)}{\tau} & 0_{NP\times K} \\ & .... \\ 0_{N\times NP} & 0_{N\times K} \\ & .... \\ 0_{K\times NP} & diag(1_{K\times 1}\ c) \end{pmatrix} \tag{3.7}$$

where δ_1 to δ_N denote the prior mean for the coefficients on the first lag, τ controls the overall tightness of the prior distribution on the VAR coefficients, c is the tightness of the prior on the

[8] $K-1$ because, being proportions, textual factors add up to 1 and the model has an intercept.

intercept and the exogenous variables and $J_p = diag(1, ..., P)$ to denote the lags, with $P = 13$. The prior means are chosen as the OLS estimates of the coefficients of an AR(1) regression estimated for each endogenous variable and σ_i's are set using the standard deviation of the error terms from these regressions. As is standard for US data, $\tau = 0.1$ and $c = 1/10^5$ indicating an informative prior on the lags of the endogenous variables but flat for the intercept and the exogenous variables.

As highlighted by Bańbura et al. (2010), the literature suggests the forecasting performance can be improved by adding a prior on the sum of the coefficients of the form:

$$Y_{d,2} = \begin{pmatrix} \frac{diag(\delta_1\mu_1,...,\delta_N\mu_N)}{\lambda} \end{pmatrix}, \quad Z_{d,2} = \begin{pmatrix} \frac{1_{1\times P}\otimes diag(\sigma_1,...,\sigma_N)}{\lambda} & 0_{N\times K} \end{pmatrix} \tag{3.8}$$

where μ_i denotes the sample means of the endogenous variables. Following Bańbura et al. (2010), the tightness of the prior of the sum of coefficients is set to $\lambda = 10\tau$, a loose prior.

These vectors are plugged into the vector with the actual observations. However, there is a difference with respect to conventional VAR-X models: as aforementioned, X assumes a different value in each draw.[9] Given the artificial data, the observables and the product of the LDA, the model is estimated using an MCMC sampler. The algorithm cycles through the following steps:

Step 1. Compute X using the draw of the LDA. Start from the statement predictive distributions as in Equation (3.3):

$$\tilde{\hat{\theta}}^k_{\tilde{d}} = \frac{n_{\tilde{d},k} + \alpha}{\sum_{k=1}^{K} n_{\tilde{d},k} + \alpha} \tag{3.9}$$

X is then given by:

$$X = [\tilde{\hat{\theta}}^1, \ ..., \ \tilde{\hat{\theta}}^{K-1}]$$

where $\tilde{\hat{\theta}}^k$ denotes the time series of $\tilde{\hat{\theta}}^k_{\tilde{d}}$.

Step 2. Draw Σ from the Inverse Wishart:

$$H(\Sigma/Y, X) = IW(S^*, T^*)$$

[9]This is inspired by Bernanke et al. (2005)'s Bayesian FAVAR, even though here, for the aforementioned reasons, textual factors are added to the model as exogenous variables.

where the posterior scale matrix is given by $S^* = (Y^* - Z^*\beta^*)'(Y^* - Z^*\beta^*)$, with $\beta^* = (Z^{*'}Z^*)^{-1}(Z^{*'}Y^*)$, $Y^* = [Y; Y_{d,1}; Y_{d,2}]$ and $Z^* = [Z; Z_{d,1}; Z_{d,2}]$. T^* denotes the posterior degrees of freedom given by the number of rows of Y^*.

Step 3. Draw β from the Normal distribution:

$$H(\beta/\Sigma, Y, X) = N(\beta^*, \Sigma \otimes (Z^{*'}Z^*)^{-1})$$

Step 4. Repeat steps 1 to 3 until the required number of draws has been reached.[10]

In this application, the algorithm is iterated 35,000 times for each data window, with the first 5,000 draws discarded as burn-in. Moreover, in order to reduce correlation across draws and increase efficiency, only every 10th draw from the Markov chain is kept. This gives 3,000 draws which are used to simulate the posterior distributions of β, Σ and X.[11]

3.4 Forecast Evaluation

Models are compared and selected based on their predictive performance as in Geweke and Amisano (2010) and many other applications in forecasting. In particular, the VAR-teXt is estimated with K = 5, 10, 15, ..., 50 and evaluated in comparison with the benchmark VAR. All the models are estimated recursively over an expanding data window, including the entire LDA part. This prevents look-ahead biases that could appear had the LDA estimation been allowed to use the full sample, and therefore future information, to estimate the textual factors whose objective is to predict the future in the first place.[12] As the full sample period is short, due to

[10]This was implemented using a combination of the 'topicmodels' package in Python (available on https://github.com/sekhansen/text-mining-tutorial) and Matlab. As the textual factors enter in the model as exogenous variables, the estimation can be split into two steps. First, the LDA is estimated in Python. Then, all the draws of the topics/textual factors are transferred to Matlab to be used as the time series X in the VAR.

[11]Results are similar with 6,000 draws and a sampling lag of 5.

[12]It is worth noting that topics may change over the expanding window, but this is not a concern here since $K - 1$ textual factors are used in the VAR-teXt. An alternative to fully re-estimate the model every month would involve estimating the topic distributions based on a truncated corpus and using these estimates to obtain the topic time series, selecting only the interpretable topics. This, however, has the caveat of not using all the information available in the regression estimation window to estimate the topic distributions. For example, if the model is estimated for the UK and the term Brexit does not appear in the truncated corpus, it will be overlooked by the LDA even if its number of occurrences in the later period is high. Still, for robustness, this alternative will be explored in the next section.

the availability of statements, only approximately a third of the sample period is evaluated.

Starting from an initial 1998M09-2013M07 window, this results in a set of 77 out-of-sample forecasts.[13] The comparison is conducted based on forecast densities. Given the draws for β and Σ (and X_t in the case of the augmented model), it is straightforward to compute the 1- and 3-step ahead forecast densities by simulating Y_t forward:

$$H(\hat{Y}_{t+h}/Y_t) = \int H(\hat{Y}_{t+h}/Y_t, \Gamma) \times H(\Gamma/Y_t, \tilde{W}_t) d\Gamma \tag{3.10}$$

where $h = 1, 2, 3, ...$ and $\Gamma = \{\beta, \Sigma, X_t\}$.

As the proposed MCMC sampling is designed to properly capture textual factor estimation uncertainty leading to more accurate posterior predictive distributions, the analysis will focus on log-scores.[14] Log-scores are the (log) likelihood the model assigns to the actual observations Y_{t+1} given data up to t:

$$LS^i_{t,h} = \ln H(Y^i_{t+h}/Y_t) \tag{3.11}$$

where both Y^i_{t+h} and Y_t are actual data. Following Alessandri and Mumtaz (2017), these (log) predictive densities are estimated using kernel methods.

Table 3.1 reports the differences in log-scores over the entire evaluation period relative to the benchmark: positive values favour the VAR-teXt. The table also shows the p-values of Giacomini and White (2006)'s tests of unconditional and conditional predictive ability.[15] Overall, the inclusion of text improves the forecasting performance. In particular, the forecasts of the consumer price inflation and the interest rate are the ones that benefit the most from the information retrieved from the statements.

The improvement for $h = 1$ is only marginal, suggesting it may take some time for the economy to react in line with the literature of monetary policy transmission. Moreover, approx-

[13]Models perform poorly in forecasting the shadow rate in 2013M07 due to a spike in this time series, and the resulting relative performance at this month is such an outlier that it would determine the ranking of the average performance had it been included in the evaluation period. That is why the initial window ends in 2013M07 and not earlier in that year. Nonetheless, starting the forecast evaluation at the beginning of 2013 and dropping 2013M07 from the average gives very similar results.

[14]Point forecasts are presented in the appendix.

[15]Such p-values, however, are only indicative since the distributions of the tests are derived based on fixed rolling window estimators. The same applies to the decision rule, which will soon be described. As in Giacomini and White (2006) and Alessandri and Mumtaz (2017) the conditional test is based on the same information set used to generate the forecasts.

Table 3.1: Log-Scores: VAR-teXt versus benchmark VAR

	1M				3M			
K	y	π	r	s	y	π	r	s
5	-0.98	2.26	0.06	0.29	-0.88	2.22	-0.93	-0.08
	(0.228)	*(0.022)*	*(0.954)*	*(0.794)*	*(0.348)*	*(0.001)*	*(0.466)*	*(0.915)*
	(0.102)	*(0.068)*	*(0.460)*	*(0.818)*	*(0.606)*	*(0.004)*	*(0.232)*	*(0.982)*
10	-1.60	2.13	0.51	-1.12	-0.59	3.84	0.73	0.16
	(0.223)	*(0.132)*	*(0.731)*	*(0.301)*	*(0.634)*	*(0.000)*	*(0.517)*	*(0.842)*
	(0.319)	*(0.322)*	*(0.334)*	*(0.410)*	*(0.183)*	*(0.000)*	*(0.062)*	*(0.727)*
15	-2.45	2.07	1.21	-0.08	-0.54	4.14	1.51	0.77
	(0.149)	*(0.172)*	*(0.369)*	*(0.938)*	*(0.663)*	*(0.000)*	*(0.256)*	*(0.385)*
	(0.284)	*(0.421)*	*(0.651)*	*(0.558)*	*(0.877)*	*(0.000)*	*(0.007)*	*(0.098)*
20	-1.99	1.68	1.46	-1.89	0.07	4.83	2.25	1.99
	(0.207)	*(0.291)*	*(0.352)*	*(0.355)*	*(0.966)*	*(0.000)*	*(0.116)*	*(0.041)*
	(0.319)	*(0.582)*	*(0.318)*	*(0.459)*	*(0.656)*	*(0.000)*	*(0.007)*	*(0.024)*
25	-1.37	1.98	0.33	-1.00	-0.57	5.62	2.09	1.73
	(0.418)	*(0.317)*	*(0.839)*	*(0.411)*	*(0.781)*	*(0.000)*	*(0.208)*	*(0.097)*
	(0.625)	*(0.647)*	*(0.804)*	*(0.681)*	*(0.959)*	*(0.000)*	*(0.01)*	*(0.017)*
30	-2.28	1.24	0.10	-3.3	-0.61	6.73	2.43	1.10
	(0.255)	*(0.575)*	*(0.958)*	*(0.133)*	*(0.747)*	*(0.000)*	*(0.197)*	*(0.309)*
	(0.396)	*(0.714)*	*(0.869)*	*(0.298)*	*(0.89)*	*(0.000)*	*(0.031)*	*(0.105)*
35	-1.62	1.24	0.11	-2.95	-0.04	7.44	2.65	1.38
	(0.447)	*(0.667)*	*(0.956)*	*(0.294)*	*(0.985)*	*(0.000)*	*(0.173)*	*(0.262)*
	(0.27)	*(0.382)*	*(0.955)*	*(0.493)*	*(0.951)*	*(0.000)*	*(0.035)*	*(0.386)*
40	-3.02	1.23	0.39	-5.05	-0.4	7.64	2.95	1.41
	(0.230)	*(0.688)*	*(0.873)*	*(0.201)*	*(0.858)*	*(0.000)*	*(0.099)*	*(0.299)*
	(0.201)	*(0.556)*	*(0.736)*	*(0.387)*	*(0.968)*	*(0.000)*	*(0.036)*	*(0.229)*
45	-2.71	0.84	-1.49	-4.06	-0.79	8.99	2.11	1.91
	(0.384)	*(0.784)*	*(0.612)*	*(0.189)*	*(0.784)*	*(0.000)*	*(0.353)*	*(0.252)*
	(0.221)	*(0.559)*	*(0.846)*	*(0.324)*	*(0.779)*	*(0.000)*	*(0.046)*	*(0.347)*
50	-2.87	0.76	-2.7	-8.19	-0.16	9.44	1.32	3.07
	(0.433)	*(0.825)*	*(0.392)*	*(0.23)*	*(0.953)*	*(0.000)*	*(0.605)*	*(0.085)*
	(0.311)	*(0.503)*	*(0.785)*	*(0.494)*	*(0.779)*	*(0.000)*	*(0.24)*	*(0.232)*

Notes: The table shows the average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 1-month and the 3-month ahead forecasts over 2013M08–2020M02. The p-values of Giacomini and White (2006)'s test of unconditional (conditional) predictive ability are in the first (second) parenthesis. The variables are industrial production growth (y), inflation (π), shadow rate (r), and EBP (s).

imately half of the meetings are not followed by a meeting in the next month, so changes in the statements that may indicate changes in interest rate will take more time to be fulfilled. If one is mostly interested in forecasting inflation, $K = 5$ gives the largest difference in log-scores: 2.26%. However, for the other variables, it is hard to distinguish between the augmented and the benchmark models.

The improvement is stronger for the 3-month ahead forecasts, a horizon that always comprises the next FOMC meeting. For this horizon, the observations and the forecasts of the variables in first differences (y and π) are cumulated, so the performance is evaluated based on the industrial production growth and the inflation over the following quarter. The model with the largest gains for inflation is the one with K=50 and for the interest rate is the one with K=40. $K = 45$

delivers a good performance for both.[16] In this model, there is an improvement of 8.99% for inflation and 2.11% for the shadow rate. For the excess bond premium, $K = 50$ is the best choice, and there is no significant difference for industrial production.

Since the predictive power of textual factors and, consequently, models can change, henceforth, the forecast evaluation will be conducted over time. Figure 3.1 shows the model selection over time using the cumulative difference in log-scores between the augmented model with $K = 45$ and the benchmark model.

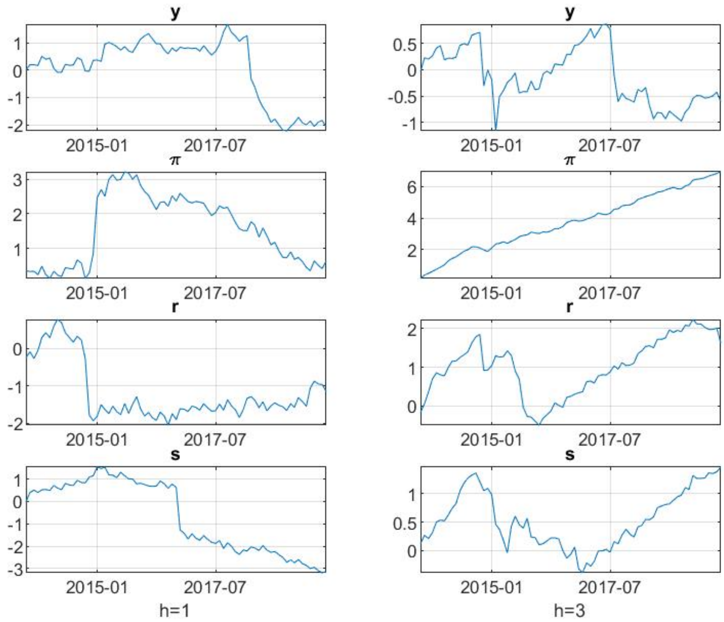

Figure 3.1: Cumulative Log-Score Difference over Time

Notes: The lines show the cumulative difference in log-scores between the VAR-text and the VAR models for horizons 1 and 3.

For $h = 1$, the plots are more erratic and confirm the average performance displayed in the table. For $h = 3$, the plot of industrial production confirms that the textual factors do not contribute to its forecasts. Inflation exhibits an upward trend indicating that the evidence in favour of the augmented model is gradually built over time but with some reversals in the relative performance as indicated by the changes in the slope. As for the shadow rate, there is also an upward trend, especially from 2016 on. Finally, the behaviour over time of the forecasts

[16] While some textual factors are interpretable topics, most of them are not.

of the excess bond premium is more volatile. Overall, ex-post, it is possible to associate some of the changes in the slope with the statements. At time t, however, it would have been hard to anticipate the effects of particular changes in the statements on the performance.

The decision rule proposed by Giacomini and White (2006) provides an alternative way to evaluate the forecasting performance over time. It uses current information to select at every t the model that is expected to work better in $t+h$. Figure 3.2 shows that, using the log-scores up to time t as the decision criterion, the VAR-teXt is predicted to perform better in the future to forecast inflation and the interest rate apart from few troughs at the beginning of the evaluation period.

In particular, and consistent with the previous table and graph, the augmented model is predicted to perform around 10% better in relation to the benchmark in terms of the log-scores of inflation for $h = 3$ throughout almost the entire evaluation period, sometimes even more than 15%. For the interest rate, the augmented model is expected to perform around 5% better most of the time. The other plots display a more erratic behaviour.

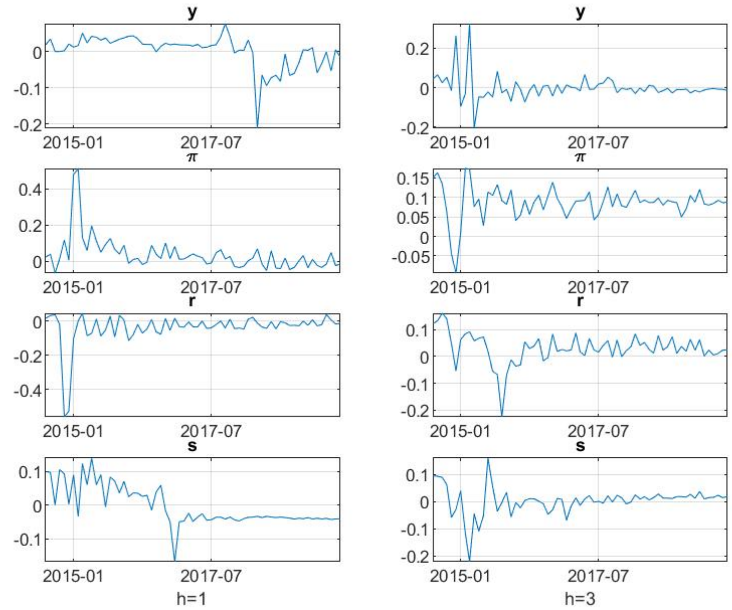

Figure 3.2: Giacomini and White (2006)'s Decision Rule

Notes: Positive values indicate that the VAR-teXt is expected to work better in the future and should be selected.

It is worth noting at this point that the first 2 years of the evaluation period are within the ZLB period. This does not invalidate the previous exercise because textual factors can capture unconventional monetary policy through pattern changes in the statements and the shadow rate also summarises this unconventional policy. Nevertheless, the results are also presented next for the post-ZLB period. This reduces the set of forecasts but it is a useful exercise since it allows for the reaction of the Federal Reserve to economic and financial conditions to be properly captured during the ZLB period and for the evaluation of the forecast of the actual fed funds rate, which replaces the shadow rate after 2015M12 in Wu and Xia (2016)'s time series.[17]

Table 3.2: Forecast Evaluation after 2015M12: VAR-teXt versus benchmark VAR

	1M				3M			
K	y	π	r	s	y	π	r	s
5	-1.61	0.71	1.54	-1.32	0.26	1.92	0.64	0.22
	(0.129)	*(0.339)*	*(0.131)*	*(0.361)*	*(0.800)*	*(0.010)*	*(0.394)*	*(0.751)*
	(0.037)	*(0.614)*	*(0.137)*	*(0.710)*	*(0.493)*	*(0.040)*	*(0.294)*	*(0.443)*
10	-1.99	0.33	2.24	-2.68	-0.54	3.15	1.37	0.79
	(0.192)	*(0.728)*	*(0.042)*	*(0.079)*	*(0.715)*	*(0.000)*	*(0.196)*	*(0.327)*
	(0.296)	*(0.355)*	*(0.129)*	*(0.180)*	*(0.756)*	*(0.000)*	*(0.109)*	*(0.690)*
15	-3.49	0.11	2.63	-1.60	-0.82	3.57	2.85	1.05
	(0.091)	*(0.924)*	*(0.037)*	*(0.200)*	*(0.551)*	*(0.000)*	*(0.006)*	*(0.159)*
	(0.211)	*(0.676)*	*(0.067)*	*(0.443)*	*(0.646)*	*(0.000)*	*(0.018)*	*(0.281)*
20	-2.83	-0.39	3.63	-4.83	-0.60	4.61	4.13	2.12
	(0.130)	*(0.745)*	*(0.012)*	*(0.115)*	*(0.759)*	*(0.000)*	*(0.001)*	*(0.024)*
	(0.217)	*(0.804)*	*(0.039)*	*(0.292)*	*(0.920)*	*(0.000)*	*(0.003)*	*(0.094)*
25	-2.95	-0.57	1.45	-2.99	-0.12	5.45	3.71	2.26
	(0.125)	*(0.654)*	*(0.343)*	*(0.043)*	*(0.957)*	*(0.000)*	*(0.004)*	*(0.027)*
	(0.217)	*(0.619)*	*(0.409)*	*(0.101)*	*(0.702)*	*(0.000)*	*(0.009)*	*(0.052)*
30	-4.19	-1.82	2.10	-6.34	-0.65	5.87	4.26	0.88
	(0.100)	*(0.277)*	*(0.261)*	*(0.045)*	*(0.752)*	*(0.000)*	*(0.007)*	*(0.470)*
	(0.167)	*(0.160)*	*(0.157)*	*(0.149)*	*(0.654)*	*(0.000)*	*(0.020)*	*(0.395)*
35	-4.07	-3.42	1.52	-6.44	-0.16	6.44	4.24	1.54
	(0.135)	*(0.097)*	*(0.495)*	*(0.123)*	*(0.941)*	*(0.000)*	*(0.007)*	*(0.228)*
	(0.113)	*(0.216)*	*(0.089)*	*(0.231)*	*(0.668)*	*(0.000)*	*(0.023)*	*(0.255)*
40	-5.30	-4.78	1.88	-9.75	-0.52	6.70	4.50	2.25
	(0.109)	*(0.033)*	*(0.436)*	*(0.104)*	*(0.835)*	*(0.000)*	*(0.004)*	*(0.137)*
	(0.070)	*(0.037)*	*(0.352)*	*(0.244)*	*(0.812)*	*(0.000)*	*(0.014)*	*(0.111)*
45	-5.57	-4.79	0.70	-8.43	-0.38	8.12	3.86	1.84
	(0.194)	*(0.056)*	*(0.786)*	*(0.059)*	*(0.892)*	*(0.000)*	*(0.033)*	*(0.284)*
	(0.103)	*(0.044)*	*(0.174)*	*(0.123)*	*(0.700)*	*(0.000)*	*(0.072)*	*(0.235)*
50	-5.66	-5.02	-0.03	-14.72	-1.13	8.73	3.75	2.93
	(0.278)	*(0.071)*	*(0.992)*	*(0.161)*	*(0.697)*	*(0.000)*	*(0.064)*	*(0.110)*
	(0.146)	*(0.077)*	*(0.356)*	*(0.387)*	*(0.675)*	*(0.000)*	*(0.087)*	*(0.023)*

Notes: The table shows the average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 1-month and the 3-month ahead forecasts over 2015M12–2020M02. The p-values of Giacomini and White (2006)'s test of unconditional (conditional) predictive ability are in the first (second) parenthesis. The variables are industrial production growth (y), inflation (π), shadow rate (r), and EBP (s).

Table 3.2 reports the relative log-scores averaged over 2015M12-2020M02. For $h = 3$, gains

[17] Remember that the shadow rate is given by the minimum value between the fed funds rate and the product of a shadow rate term structure model, so in normal times it is simply the fed funds rate.

are significant for both the inflation and the interest rate for all K. In particular, the forecast of the interest rate improves 3.86% when K=45 and 4.50% when K=40, suggesting the model is better at forecasting the fed funds rate after the ZLB than the shadow rate during it. There are also relevant gains for the inflation forecasts: 8.12% when K=45 and 8.73% when K=50. The VAR-teXt also performs better in forecasting the excess bond premium, and this is significant for some $K's$.

Overall, using this approach, text improves the forecast of the consumer price inflation and the fed funds rate, and sometimes of the excess bond premium, as summarised by the average performance and further detailed in the graphs.

3.5 Discussion

This section raises some issues found during the analysis.[18] First, the shadow rate is not the only possible choice of policy indicator. Other candidates are the 1-year government bond rate and the fed funds rate. The 1-year rate has been commonly used by papers that include the ZLB, such as Gertler and Karadi (2015) and Jarociński and Karadi (2020), and the fed funds rate is ultimately the variable of interest although the lack of substantial variation during the ZLB may compromise estimation.

Table 3.3 shows the main results hold when these variables replace the shadow rate in the model.[19] Results with the alternative policy indicators are as good as the benchmark results. In fact, the average difference in predictive log-scores for r is even higher when the fed funds rate is used.

Furthermore, in order to check whether only some interpretable topics are driving the results while the other textual factors are just introducing noise, topic distributions are now estimated based on a truncated corpus (1998M09-2012M12) and these estimates are used to obtain the topic time series as it is customary in text mining applications in forecasting. The point of departure is Hansen and McMahon (2016), who examine the causal effects of FOMC statements using 15 topics. Their sample period is very similar, and this choice leads indeed to a good fit in terms of interpretability. For details, see the appendix. Five topics related to the economic situation and

[18] For more details, the reader is referred to the appendix.
[19] Note that the benchmark models to which the new VAR-teXts are compared are different.

Table 3.3: Forecast Evaluation: Alternative policy indicators

		1M				3M			
r	K	y	π	r	s	y	π	r	s
GS1	45	-2.52	0.08	0.41	-4.94	-0.53	8.52	-0.06	2.69
		(0.382)	*(0.979)*	*(0.883)*	*(0.170)*	*(0.831)*	*(0.000)*	*(0.969)*	*(0.040)*
		(0.285)	*(0.465)*	*(0.335)*	*(0.142)*	*(0.757)*	*(0.000)*	*(0.032)*	*(0.110)*
FF	45	-1.75	0.04	0.21	-4.70	0.24	8.69	3.95	2.35
		(0.548)	*(0.990)*	*(0.934)*	*(0.081)*	*(0.925)*	*(0.000)*	*(0.015)*	*(0.065)*
		(0.326)	*(0.499)*	*(0.093)*	*(0.005)*	*(0.964)*	*(0.000)*	*(0.049)*	*(0.159)*

Notes: The table shows the average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 1-month and the 3-month ahead forecasts over 2013M08–2020M02. The p-values of Giacomini and White (2006)'s test of unconditional (conditional) predictive ability are in the first (second) parenthesis. The variables are industrial production growth (y), inflation (π), policy indicator (r), and EBP (s).

prospects are selected. Table 3.4 reports the log-scores of this VAR-teXt with labelled topics relative to the benchmark VAR. The gains are smaller when this specification is implemented. Therefore, results are not driven only by the interpretable textual factors.[20]

Table 3.4: Forecast Evaluation: Labelled VAR-teXt versus benchmark VAR

	1M				3M			
K	y	π	r	s	y	π	r	s
15	-0.13	1.58	-0.95	1.82	-0.35	1.39	0.58	2.30
	(0.915)	*(0.060)*	*(0.518)*	*(0.131)*	*(0.746)*	*(0.030)*	*(0.652)*	*(0.028)*
	(0.491)	*(0.133)*	*(0.769)*	*(0.343)*	*(0.709)*	*(0.096)*	*(0.921)*	*(0.042)*

Notes: The table shows the average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 1-month and 3-month ahead forecasts over 2013M08–2020M02. The p-values of Giacomini and White (2006)'s test of unconditional (conditional) predictive ability are in the first (second) parenthesis. The variables are industrial production growth (y), inflation (π), shadow rate (r), and EBP (s).

Still in line with the usual approach of text mining applications in the economics literature, a combination of topic models and dictionary methods is explored. Following Thorsrud (2018), this is done by i) counting the number of positive and negative words in the statement[21], ii) normalising such statistics by the total number of words in each statement such that each document observation reflects the fraction of positive and negative words, iii) and taking the difference between the positive and negative fractions, resulting in an index for the mood of each

[20] 15 is also the number of topics in the grid that gives the lowest perplexity. Therefore, the goodness-of-fit of the LDA does not seem correlated with the forecasting performance of the VAR-teXt, and criteria should be used according to the objective of the model.

[21] The list from the 'topicmodels' package developed by Stephen Hansen was used.

statement: $M_{\tilde{d}}$. Each textual factor is then adjusted as follows:

$$\tilde{\hat{\theta}}^{k}_{\tilde{d},\pm} = \tilde{\hat{\theta}}^{k}_{\tilde{d}} M_{\tilde{d}} \tag{3.12}$$

This new index can be decomposed into two dimensions: the textual factor, which extracts information from the way words are grouped over time, and the mood, which extracts information about the tone used in the text. By comparing the tone-adjusted VAR with the standard VAR-teXt, the table shows the valued added by the inclusion of tone in the analysis. There are significant improvements for the 3-month ahead forecasts of inflation and the excess bond premium, indicating that the sentiment of the text matters.

Table 3.5: Forecast Evaluation: Tone-adjusted VAR-teXt versus standard VAR-teXt

	1M				3M			
K	y	π	r	s	y	π	r	s
45	-1.52 *(0.640)* *(0.870)*	0.33 *(0.960)* *(0.053)*	-2.15 *(0.632)* *(0.300)*	3.01 *(0.547)* *(0.681)*	-7.64 *(0.180)* *(0.015)*	4.43 *(0.090)* *(0.005)*	0.14 *(0.955)* *(0.643)*	2.98 *(0.154)* *(0.002)*

Notes: The table shows the average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 1-month and 3-month ahead forecasts over 2013M08–2020M02. The p-values of Giacomini and White (2006)'s test of unconditional (conditional) predictive ability are in the first (second) parenthesis. The variables are industrial production growth (y), inflation (π), shadow rate (r), and EBP (s).

In fact, for inflation, there is an additional improvement of 4.43%. Therefore, compared to the benchmark VAR, the inflation forecasts are 13.42% (=8.99% + 4.43%) better, and the mood dimension contributes to approximately 1/3 of the gains. As expected, however, the multiplication of $\tilde{\hat{\theta}}^{k}_{\tilde{d}}$ by $M_{\tilde{d}}$ increases the volatility of the new index, and this affects the comparison over time and the real-time decision between models.[22]

Finally, the assumption of exogeneity of the text-data is relaxed. Under weak exogeneity of the text-data, a more internally consistent version of the algorithm is provided. However, such assumption can be seen as too strong in some applications. To see how it affects the case of this paper, textual factors are included as endogenous variables in the VAR. In general, one should use the option that better suits their own application.

In practice, when textual factors are treated as exogenous, the vector X_{t-1} is appended with

[22]The cumulative difference in log-scores over time and Giacomini and White (2006)'s decision rule are shown in the appendix.

the values for the textual factors in t so as to allow the generation of the 1-month ahead forecasts. In the case of the 3-month ahead forecasts, the values for the textual factors in t are repeated in the vector. When textual factors are treated as endogenous, their future values will be given by simulating the system of equations forward as usual.

Table 3.6: Forecast Evaluation: Endogenous textual factors versus benchmark VAR

	1M				3M			
K	y	π	r	s	y	π	r	s
15	0.65	6.79	6.62	4.77	4.76	14.07	5.25	5.18
	(0.813)	*(0.002)*	*(0.005)*	*(0.028)*	*(0.251)*	*(0.000)*	*(0.059)*	*(0.194)*
	(0.624)	*(0.011)*	*(0.017)*	*(0.004)*	*(0.100)*	*(0.000)*	*(0.000)*	*(0.000)*

Notes: The table shows the average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 1-month and 3-month ahead forecasts over 2013M08–2020M02. The p-values of Giacomini and White (2006)'s test of unconditional (conditional) predictive ability are in the first (second) parenthesis. The variables are industrial production growth (y), inflation (π), shadow rate (r), and EBP (s).

When the textual factors are included in the regression as endogenous variables, the ranking reverses and lower numbers of K are the best performing ones. Table 3.6 shows the comparison between the VAR augmented with endogenous textual factors when $K = 15$ and the benchmark VAR. Once more, the text-augmented VAR shows a good performance in forecasting the interest rate and, mainly, inflation, with gains even in the 1-month ahead forecasts.

As pointed out before, the text mining literature has advanced considerably in recent years, and other methods could also be explored. However, the main takeaway from the results is that even a standard LDA can produce textual factors that are useful for forecasting.

3.6 The DFM-teXt Model

The DFM-teXt model is a natural extension of the VAR-teXt since the transition equation of dynamic factor models is simply a VAR in the factors, and observables are expressed as a weighted average of factors via the observation equation. The idea behind the DFM-teXt model is to use all available information regardless of its form: quantitative data or text. This is helpful in understanding whether the results from previous sections are driven by the fact that the benchmark VAR is a small-scale model and teXt is simply providing information that could have been found in traditional time series. The point of departure for the analysis is a DFM model of the form:

$$W_{i,t} = \Gamma_i F_t + e_{i,t} \tag{3.13}$$

$$F_t = \sum_{p=1}^{P} \phi_p F_{t-p} + \mu + A_0 \varepsilon_t \tag{3.14}$$

where $W_t = (W_{1,t}, ..., W_{N_W,t})$ is a panel of N_W variables, $F_t = (F_t^1, ..., F_t^R)$ denotes the R latent factors, $W_{i,t}$ is related to the factors via the factor loadings Γ_i, $e_{i,t}$ is the *i.i.d.* idiosyncratic component in the observation equation, and ε_t is the error term in the transition equation. Given the transition equation (3.14) and Γ_i from the observation equation (3.13), it is possible to forecast all the variables in the panel.

As in the VAR-teXt, the DFM is augmented with textual factors:

$$F_t = \sum_{p=1}^{P} \phi_p F_{t-p} + \phi_X X_{t-1} + \mu + A_0 \varepsilon_t$$

The algorithm for DFM-teXt is presented in the appendix and is similar to the algorithm for VAR-teXt, but, being the DFM-teXt a state-space model, one has to carry out additional steps to sample factor loadings and factors. The panel of variables is composed of the 128 series downloaded from the FRED-MD dataset, the excess bond premium and the shadow rate. Following Bai and Ng (2002)'s criteria, 7 factors are used.[23] Furthermore, a flat prior is used in the estimation of the observation equation and the same prior specification of the VAR is imposed for the transition equation: 13 lags and dummy observations. The DFM-teXt is estimated with K=45, and then evaluated as the VAR-teXt was.[24]

Table 3.7 shows that, in contrast with Kalamara et al. (2020), the added value of text does not degrade in the DFM, although the gains are concentrated in the 3-month ahead forecasts. In particular, even the forecast of industrial production is enhanced using this model. The augmented model continues to be very good at forecasting inflation, but its ability to forecast the interest rate decreases. This is related to the fact that the interest rate enters the DFM in first differences. The appendix reports the results for the other variables.

[23]Even though the estimation is fully Bayesian, this criterion is chosen because it is less computationally intensive.

[24]As the DFM-teXt takes much longer to run, only $K = 45$ is explored.

Table 3.7: Forecast Evaluation: DFM-teXt versus benchmark DFM

	1M				3M			
K	y	π	r	s	y	π	r	s
45	-2.09	0.62	-0.96	-3.58	5.00	8.15	0.16	3.03
	(0.324)	*(0.730)*	*(0.475)*	*(0.010)*	*(0.045)*	*(0.000)*	*(0.901)*	*(0.013)*
	(0.484)	*(0.150)*	*(0.167)*	*(0.018)*	*(0.131)*	*(0.000)*	*(0.751)*	*(0.028)*

Notes: The table shows the average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 1-month and the 3-month ahead forecasts over 2013M08–2020M02. The p-values of Giacomini and White (2006)'s test of unconditional (conditional) predictive ability are in the first (second) parenthesis. The variables are industrial production growth (y), inflation (π), shadow rate (r), and EBP (s).

In summary, for $K = 45$, 97 variables present a positive average difference in predictive log-scores, of which 53 (47) are significant at the 10% level according to the unconditional (conditional) test. The forecasts of only 9 (8) variables become significantly worse when the model is augmented with text. This shows that teXt brings information on top of the contained in traditional time series and can improve the informational content even of a large-scale model.

3.7 Conclusion

This paper has explored the complementarity between traditional econometrics and machine learning and applied the resulting models – the VAR-teXt and the DFM-teXt – to central bank communication. It is becoming a consensus that text can provide information beyond that contained in traditional economic models, being potentially important for forecasting. This paper adds to this literature.

In order to deal with the fact that the LDA outputs are estimates, an MCMC sampling algorithm that tackles the generated regressors issue is also provided. Moreover, unlike previous economic applications, the number of topics is selected based on the predictive performance, which is consistent with the objective of this paper and helps discipline the choice.

Results show that textual factors based on FOMC statements are indeed useful for forecasting in the small-scale VAR and in the DFM-teXt, especially at the 3-month horizon as it seems that the information retrieved from FOMC statements takes some time to affect the forecasts. Specifically, the VAR-teXt outperforms the benchmark VAR in forecasting the consumer price inflation and the interest rate, and this holds under various specifications.

As for the DFM-teXt, gains are more general, despite also being even more concentrated in the 3-month horizon. Thus, like factors in factor-augmented models, textual factors can increase the forecasting performance of VAR models even without necessarily having a clear meaning. As a consequence, this approach also favours replicability, since the choice of the number of textual factors is data-driven and does not rely on researchers' interpretability.

Another clear advantage of an automated procedure such as this is scalability, so it is easy to apply it to datasets containing many more documents and words. This approach can also be easily extended to incorporate more than one corpus. Provided that the aggregate number of textual factors is low, just add them as exogenous variables in the regression; otherwise, one has only to run the LDA for each corpus and extract principal components from the pooled textual factor estimates in each draw. Then add the principal components, rather than the textual factors, as exogenous variables at each iteration of the VAR.

In this vein, a straightforward extension is to apply the model to other types of central bank communication, such as minutes and speeches. As text is ubiquitous in many other branches of economics, there are also many potential applications outside monetary policy. Equipped with a supercomputer, one could also depart from previous literature in setting the values of the hyperparameters and explore a grid for α, η and K, selecting the triplet with the best out-of-sample performance. In terms of methodology, an avenue for future work is to make the estimation of textual factors also depend on traditional data.

Chapter 4

Monetary policy surprises, financial conditions, and the "string theory" revisited[1]

4.1 Introduction

Many are the attempts, by economists, at testing whether it is true that "you can't push on a string", reputedly John Maynard Keynes's words, according to which monetary easings have smaller real effects than tightenings. Ravn and Sola (2004), Tenreyro and Thwaites (2016), Angrist et al. (2018), and Barnichon and Matthes (2018) are only some of the examples. If changes in the policy rate are less powerful in any of the directions, central banks have to internalise this feature and act accordingly.

Given the policy relevance of this question, I revisit this topic and extend it in some dimensions. First, I exploit high-frequency changes around monetary announcements. High-frequency surprises have become a standard method of measuring monetary policy shocks because they can refine identification (Kuttner, 2001; Gürkaynak et al., 2005; Ramey, 2016; Kaminska et al., 2021). Since the window around the release is very narrow, it is assumed the surprises are not affected by macroeconomic news other than the announcement.

Second, I estimate the model for a recent sample, what may lead to different conclusions for the real variables as reported by Ramey (2016). Furthermore, the study of potential asymmetries is not only relevant by itself, but it can also help dissect what is behind the results in the

[1]I have benefited from comments from participants at the presentations at the Bank of England and at the 6th Workshop of the Central Bank of Brazil Research Network for useful comments. All errors are mine.

symmetric case. It could be the case that positive shocks are driving the results for some variables, while negative shocks are driving the results for others. Third, I apply the same methodology to the euro area to investigate whether the results are general or US-specific.

Fourth, I also study the responses of financial conditions. The empirical literature on credit and financial conditions is growing fast. Brave and Butters (2011) produce financial condition indices that provide a timely assessment of how tightly or loosely financial markets are and contain information on future economic activity beyond that found in non-financial measures of economic activity.[2] Gilchrist and Zakrajšek (2012) show the component of credit spreads not explained by expected defaults has considerable predictive power. Alessandri and Mumtaz (2017) find credit and financial conditions are useful in forecasting.

Turning to the interaction with monetary policy, in a linear vector autoregressive (VAR) model, Gertler and Karadi (2015) find monetary shocks lead to enhanced movements in credit costs and the excess bond premium is one of the channels. Caldara and Herbst (2019) show the failure to account for the endogenous reaction of spreads causes attenuation in the response of all variables to monetary shocks. Carriero et al. (2020) highlight the role of credit market conditions as a source of asymmetry in the effects of monetary policy. They do so by employing a smooth transition model, and find asymmetric effects are explained by how easings and tightenings affect credit conditions and the probabilities of regime changes differently.

The role played by financial conditions in the transmission of monetary policy is also highlighted in the official Fed communication. In the press conference that followed the July 2022 Federal Open Market Committee (FOMC) Meeting[3], for instance, Fed Chair Powell explained: "we set our policy, and financial conditions react, and then financial conditions are what affects the economy." In fact, the dynamic responses of financial conditions to monetary policy shocks are crucial in the understanding of the monetary transmission. It remains to be more explored, however, whether the responses of financial conditions to monetary tightenings and loosenings are symmetric.

In order to investigate such issues, I use the local projection method of Jordà (2005). This

[2] Many other papers have explored the link between financial conditions and economic activity. For instance, Aramonte et al. (2017) find predictive power of a selection of financial condition indices, especially if the financial crisis is included in the analysis. Hatzius et al. (2010) also build a new financial index that shows a tighter link with future economic activity.

[3] https://www.federalreserve.gov/mediacenter/files/FOMCpresconf20220727.pdf.

approach is suitable in that it allows for the dynamic effects to be asymmetric in a simple way. Differently, however, I estimate a Bayesian version of the Generalised Least Squares (GLS) procedure proposed by Lusompa (2021). Lusompa (2021) shows that the autocorrelation process of local projections is known and that estimating the model with GLS is more efficient than standard estimation with heteroskedasticity- and autocorrelation-consistent (HAC) standard errors. A Bayesian approach is convenient in that the sampling already takes into account the fact that estimates are used in the Feasible GLS (FGLS) transformation.

Furthermore, as pointed out by Stock and Watson (2018), high-frequency measures typically have measurement error, which can lead to bias if the measure is treated as the true shock. Hence, the surprises are treated as instruments in a Local Projection Instrumental Variable (LP-IV) set-up. As shown by Plagborg-Møller and Wolf (2021), this can be done by simply computing the ratio of reduced-form to first-stage coefficients. Estimating the model with Two-Stage Least Squares (2SLS) as it is usually done would allow only the first stage to be asymmetric while the dynamics would still be the same after both shocks.

Results show there is evidence of asymmetry for all the variables. Industrial production, unemployment, and FCI respond more strongly to positive shocks than to negative shocks and the differences are 'significant' in a high posterior density interval (HPDI) sense. On the other hand, CPI responds more weakly to positive shocks than to negative shocks, especially at the beginning, due to downward nominal rigidities. The main findings are similar when the methodology is applied to the euro area, so the asymmetry is not a specific feature of the US.

Therefore, this paper complements the literature by exploring the dynamic responses of financial conditions as well as revisiting the evidence for traditional macroeconomic variables in the US. The use of high-frequency surprises around policy announcements in the LP-IV also represents an advance in comparison with previous studies that implicitly rely on the strong identifying assumption of selection on observables. Equally important is the investigation of the effects in the euro area since, to my knowledge, this is the first paper to study potential asymmetries in the dynamic responses of the ECB's monetary policy.

The algorithm for the (non-)linear reduced-form BLP can also be applied to many other economics questions involving asymmetry, or not, provided that a measure of shock or an instrument is available. The rest of the paper is organised as follows. Section 4.2 describes the data. Section

4.3 introduces the econometric approach. Section 4.4 presents the results for the US, followed by evidence on the euro area in Section 4.5. Section 4.6 concludes.

4.2 Data

The model is estimated using US data on the fed funds rate (FFR), the consumer price index (CPI), the industrial production index (IP), the unemployment rate, and the Financial Condition Index (FCI). FFR, FCI, and unemployment are in levels and the remaining variables are in log levels. The data is monthly and runs from 1987M11 to 2020M02.

The FCI used here is constructed by the Chicago Fed using a dynamic factor model and represents a single common factor that captures financial conditions in money markets, debt and equity markets, and the traditional and "shadow" banking systems.[4] Therefore, in addition to allowing the study of the responses of financial conditions, its inclusion in the set of variables increases the informational content of the model, essentially turning it into a factor-augmented local projection.

To identify all these effects, I use monetary policy surprises. Following Paul (2020), high-frequency surprises are extracted from the current-month fed funds futures and adjusted for the remaining days within a month, as suggested by Kuttner (2001).[5] The surprises are calculated based on a 30-minute window around scheduled FOMC meetings and are relatively balanced between positive and negative values: 42% of them are positive. In practice, they are virtually zero during the zero lower bound period, but keeping this episode in the sample allows estimating the effects in the period following it as in Paul (2020).[6]

[4]The index is constructed to have an average value of zero and a standard deviation of one over a sample period extending back to 1971. More information can be found on the Chicago Fed website (https://www.chicagofed.org/research/data/nfci/background), Brave and Butters (2011), and Brave and Butters (2012).

[5]This measure was later coined MP1 by Gürkaynak et al. (2005). Paul (2020) presents detailed evidence that suggests it provides a strong instrument to identify monetary policy shocks and is less likely to be contaminated with information effects, especially when only scheduled meetings are used.

[6]The series of monetary surprises is presented in the appendix.

4.3 Econometric Framework

4.3.1 Linear Local Projections

The point of departure for the analysis is the local projection method of Jordà (2005):

$$y_{t+h} = \beta_h \varepsilon_t + \gamma_h x_t + u_{t+h} \tag{4.1}$$

where y_{t+h} is a vector of n monthly macroeconomic and financial variables, with $h = 0, ..., H$ denoting the horizons. ε_t are the high-frequency surprises, β_h gives the conventional direct estimates of the impulse responses, and x_t collects the controls: P lags of the endogenous variables and of the proxy, the intercept, and a linear trend. Finally, u_{t+h} denotes the residuals.

High-frequency surprises ε_t are useful to address endogeneity concerns, but typically have measurement error that can lead to bias if treated as the true shock (Stock and Watson, 2018). Hence, the monetary surprises are treated as instruments in a Local Projection Instrumental Variable (LP-IV) approach, whereby equation (4.1) represents both the reduced-form and the first-stage regressions. As shown by Plagborg-Møller and Wolf (2021), the LP-IV approach can be implemented by simply calculating the ratio of reduced-form to first-stage coefficients $\beta_{LPIV,h} \equiv \beta_h/\beta_{FS}$ where β_{FS} is the term in β_0 from the equation for FFR, i.e. the response of FFR on impact.

The residuals u_{t+h} are serially correlated since they are a combination of one-step-ahead forecast errors. Previous studies (e.g. Jordà (2005); Miranda-Agrippino and Ricco (2021); Ramey (2016)) have treated this autocorrelation process as unknown and dealt with this issue by incorporating corrections for serial correlation such as Newey and West (1987). However, Lusompa (2021) shows that the autocorrelation process of local projections is known and proposes a more efficient correction. Taking advantage of the fact that the residuals u_{t+h} are VMA(h) even if the true model is not a VAR, he introduces a consistent GLS estimator that involves transforming the data and estimating the regressions as follows:

1. Estimate the horizon 0 LP

$$y_t = \beta_0 \varepsilon_t + \gamma_0 x_t + u_t$$

which is equivalent to a VAR, implying $u_t = \epsilon_t$, where ϵ_t denotes the VAR forecast error terms.

2. Using the estimates of ϵ_t and $\gamma_0^{(1)}$ (the elements of γ_0 associated with the first lag), do the GLS transformation

$$\tilde{y}_{t+1} = y_{t+1} - \hat{\gamma}_0^{(1),OLS} \hat{\epsilon}_t \tag{4.2}$$

and estimate the horizon 1 LP replacing y_{t+1} with $\tilde{y}_{t+1}$ in equation (4.1).

3. For each horizon $h = 2, ..., H$, repeat this recursive procedure and estimate the local projections using the transformed $\tilde{y}_{t+h}$ that can be generalised as

$$\tilde{y}_{t+h} = y_{t+h} - \hat{\gamma}_{h-1}^{(1),GLS} \hat{\epsilon}_t - ... - \hat{\gamma}_0^{(1),OLS} \hat{\epsilon}_{t+h-1} \tag{4.3}$$

This procedure cleanses the left-hand variables from $\hat{\epsilon}_{t+h}$, eliminating the autocorrelation in local projections (see Lusompa (2021) for details).

4.3.2 Non-linear Local Projections

In order to investigate asymmetry, I also apply this GLS transformation to the local projections implemented by Tenreyro and Thwaites (2016):

$$y_{t+h} = \beta_h^- \max\{0, \varepsilon_t\} + \beta_h^+ \min\{0, \varepsilon_t\} + \gamma_h x_t + u_{t+h} \tag{4.4}$$

or, equivalently,

$$y_{t+h} = \tilde{\beta}_h \varepsilon_t + \tilde{\beta}_h^+ \max\{0, \varepsilon_t\} + \gamma_h x_t + u_{t+h} \tag{4.5}$$

where positive and negative shocks are allowed to have different effects: β_h^- and β_h^+, with $\beta_h^- = \tilde{\beta}_h$ and $\beta_h^+ = \beta_h + \tilde{\beta}_h^+$. Differently, however, I adjust these conventional estimates of the impulse responses using the correction for non-linearities proposed by Gonçalves et al. (2021). They show that local projection estimators currently used in the literature, which effectively ignore the non-linearity of the impulse responses are invalid. This can be seen by applying the definition of non-linear impulse response to equation (4.5)

$$IRF_{h,\delta} = E(y_{t+h}(\delta) - y_{t+h})$$

where δ is the size of the shock, and

$$y_{t+h}(\delta) - y_{t+h} = \tilde{\beta}_h^{+}[\max\{0, \varepsilon_t + \delta\} - \max\{0, \varepsilon_t\}] + \tilde{\beta}_h[\varepsilon_t + \delta - \varepsilon_t]$$

which can be written as follows:

$$y_{t+h}(\delta) - y_{t+h} = \tilde{\beta}_h^{+}[\max\{0, \varepsilon_t + \delta\} - \max\{0, \varepsilon_t\}] + \tilde{\beta}_h\delta$$

It is easy to see that when there is no non-linear term, this expression boils down to $\tilde{\beta}_h\delta$ (or just $\tilde{\beta}_h$ when $\delta = 1$) since the square brackets multiplying this term becomes simply δ. However, that is not the case with the first square brackets, where the non-linearity appears. Therefore, existing studies that interpret the raw β's as the impulse response in the non-linear case are implicitly assuming $[\max\{0, \varepsilon_t + \delta\} - \max\{0, \varepsilon_t\}] = \delta$, being, therefore, unable to recover the population response functions even asymptotically.

Nevertheless, when ε_t is i.i.d, the correction assumes a very simple form and the construction of the modified LP estimator amounts to estimating the term in brackets and adjusting the standard impulse response as follows:

1. Obtain an estimate of $A_{0,\delta} \equiv E[\max\{0, \varepsilon_t + \delta\} - \max\{0, \varepsilon_t\}]$ as[7]

$$\hat{A}_{0,\delta} \quad = \tfrac{1}{T}\textstyle\sum_{t=1}^{T}(\max\{0, \varepsilon_t + \delta\} - \max\{0, \varepsilon_t\})$$

2. Then compute

$$IRF_{h,\delta^+}^{LP} \quad = \tilde{\beta}_h\delta + \tilde{\beta}_h^{+}\hat{A}_{0,\delta}$$

The adjustment for non-linearity depends on $\hat{A}_{0,\delta}$, which is the sample average of the difference between the non-linear functions $\max(\cdot)$ evaluated at '$\varepsilon_t + \delta$' and 'ε_t', and enables the consistent estimation of the population impulse responses as demonstrated by Gonçalves et al.

[7]Note that this is just the non-linear counterpart of the impulse given in the traditional linear IRF which is given by δ since: $\frac{1}{T}\sum_{t=1}^{T}(f(\varepsilon_t + \delta) - f(\varepsilon_t)) = \frac{1}{T}\sum_{t=1}^{T}((\varepsilon_t + \delta) - \varepsilon_t) = \delta$.

(2021).

4.3.3 Estimation

Modelling the autocorrelation process allows for a fully Bayesian approach.[8] Specifically, a Monte Carlo sampler with a conjugate flat prior (described in the appendix) is employed.[9] The baseline model is estimated with 12 lags. It is crucial that the number of lags be large enough so that the residuals from the first LP are uncorrelated (Lusompa, 2021).

A Bayesian approach is particularly convenient in that i) LP-IV inference becomes straightforward and only requires the computation of the ratio of reduced-form to first-stage coefficients for each draw; and ii) the sampling naturally takes into account the fact that estimates are used in the FGLS transformation. By using draws of ϵ_t, ..., ϵ_{t+h-1} and of $\gamma_0^{(1)}$, ..., $\gamma_{h-1}^{(1)}$ in the transformation described in equations (4.2) and (4.3), the algorithm properly captures the uncertainty arisen from the fact that $\hat{\epsilon}_t$, ..., $\hat{\epsilon}_{t+h-1}$ are used in place of the unobserved ϵ_t, ..., ϵ_{t+h-1} and $\hat{\gamma}_0^{(1)}$, ..., $\hat{\gamma}_{h-1}^{(1)}$ in place of $\gamma_0^{(1)}$, ..., $\gamma_{h-1}^{(1)}$. The posterior is simulated based on 1,000 draws.

4.4 Results

Figure 4.1 displays the impulse responses after a monetary policy shock, normalised to have an impact of 25 basis points on the federal funds rate. The grey shaded area represents the 68 percent HPD credible sets, and the blue line represents the median.

CPI goes down after some time, but the effect is only temporary. Differently, industrial production displays a persistent decrease. The unemployment rate increases, but the credible set becomes very wide, so there is high uncertainty for longer horizons. FCI goes up indicating tighter financial conditions, and this fades away only in approximately one year. Overall, results are consistent with the literature (Gertler and Karadi, 2015; Paul, 2020; Miranda-Agrippino and Ricco, 2021).

Albeit more erratic as it is usually the case with local projections, such responses are very

[8]Miranda-Agrippino and Ricco (2021) are agnostic about the source of model misspecification and replace the scale matrix in the inverse-Wishart posterior with a HAC-corrected covariance matrix.

[9]Even though, in principle, an informative prior based on a training sample could also be used, this would make the remaining sample period even shorter.

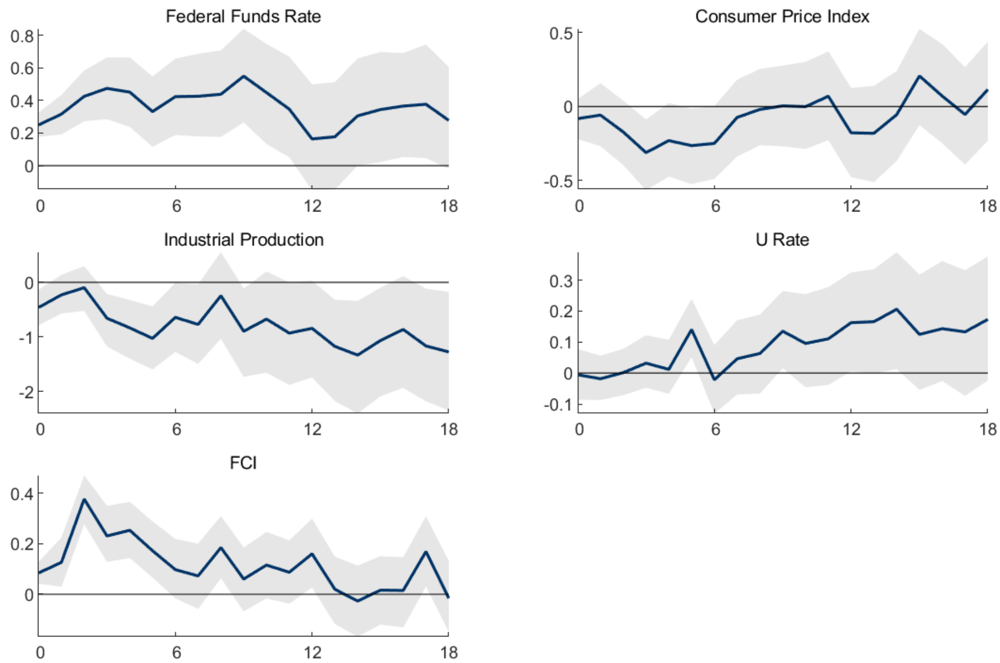

Figure 4.1: Impulse Responses: Linear LP

Notes: The shaded area represents the 68 percent HPD credible sets, and the blue line represents the median. The monetary policy shock has been normalised to have an impact of 25 basis points on the federal funds rate.

similar to the ones produced by a Proxy VAR presented in Figure 4.2. This is not surprising since Plagborg-Møller and Wolf (2021) show that local projections and VARs estimate the same impulse responses in population. Such equivalence, however, does not hold in the non-linear case.

Figure 4.3 compares the impulse responses after positive and negative shocks based on local projections. Specifically, it shows the IRFs after a tightening in the first column, the flipped IRFs after a loosening in the second column, and the difference between them in the last column. As in the linear case, the impact on the fed funds rate is normalised on impact. Nevertheless, the response after a loosening shows more persistence than the response to a tightening.

CPI goes down after a positive shock, but this is only borderline “significant”. On the other hand, there is a stronger and steadier response to a negative shock. This more persistent behaviour of CPI after a loosening can be at least partially accounted for the persistence in the fed funds ratio. Given the uncertainty, however, this does not produce a difference in an HPDI sense for the longer horizons.

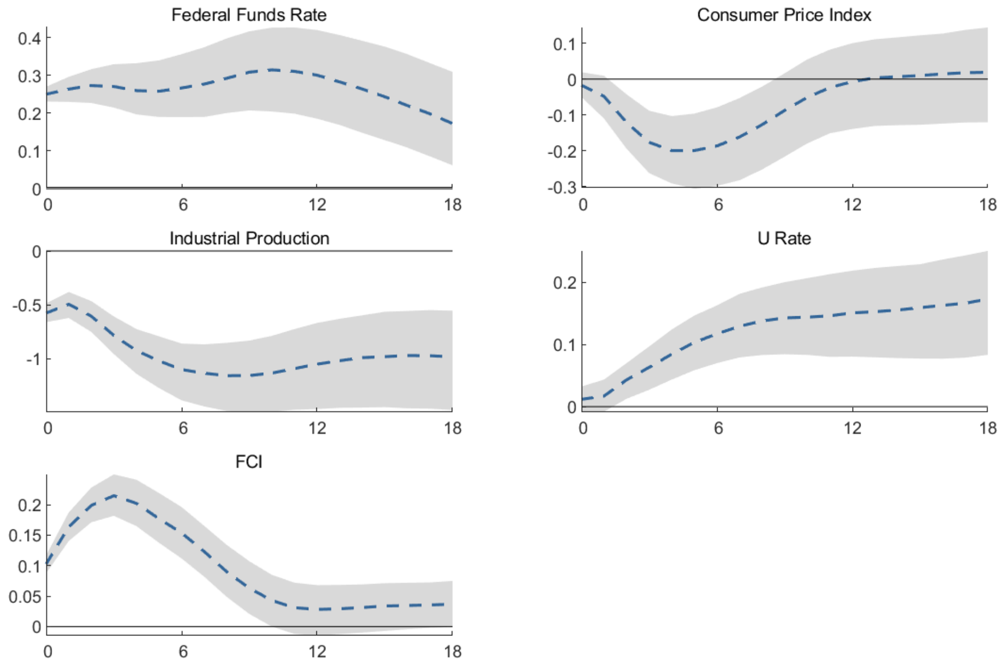

Figure 4.2: Impulse Responses: VAR

Notes: The shaded area represents the 68 percent HPD credible sets, and the blue line represents the median. The monetary policy shock has been normalised to have an impact of 25 basis points on the federal funds rate.

The main source of the difference in the right column is the fact that prices do not react on impact to a tightening, what is consistent with theories of downward nominal rigidities. It is also worth noting that the version without lags of the proxy displays a price puzzle after a positive shock, highlighting the importance of controlling for lags of ε and $f(\varepsilon)$ as suggested by Gonçalves et al. (2021).[10]

The response of industrial production after a tightening is very similar to the one displayed in Figure 4.1. In fact, it seems that positive shocks are the main drivers of the response in the linear case as industrial production does not react to negative shocks in this sample. As a consequence, there is a large difference between them. The responses of the unemployment rate mirror the ones of industrial production. The unemployment rate increases persistently after a tightening without any significant reaction after a loosening. Accordingly, the difference in an HPDI sense is also persistent and “significant”.

Despite the difference in terms of sample and identification, such responses are in line with

[10]See the appendix.

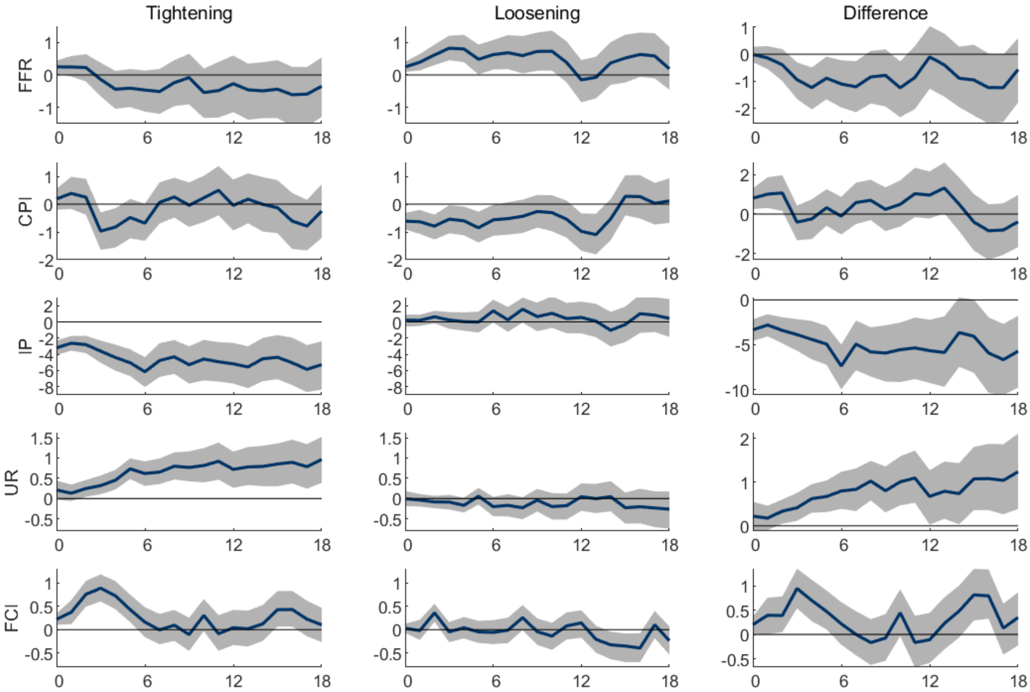

Figure 4.3: Impulse Responses: Positive and negative shocks

Notes: The shaded area represents the 68 percent HPD credible sets, and the blue line represents the median. The left panel presents impulse responses to positive shocks. The central panel shows the impulse responses to negative shocks. They are flipped to facilitate comparison. The right panel shows their difference. The monetary policy shock has been normalised to have an impact of 25 basis points on the federal funds rate.

previous studies, such as Angrist et al. (2018) and Debortoli et al. (2020) among others. Theoretical macro models encompassing downward nominal rigidities, menu costs and other types of frictions also set the ground for asymmetry in prices and real variables, with monetary easings have smaller real effects than tightenings in line with the words attributed to Keynes.

The behaviour of financial conditions is also asymmetric. FCI responds more strongly to positive shocks than to negative shocks, displaying a difference in an HPDI sense. This behaviour, coupled with the "balance sheet" channel reviewed by Bernanke and Gertler (1995), amplifies the movements and, therefore, the asymmetry in the industrial production and the unemployment rate.

Empirical benchmarks for financial and credit conditions in a set-up that allows for asymmetry are not so easy to find. Carriero et al. (2020) explore different sizes of shocks and different regimes, and find effects that are sometimes not different than zero and sometimes only borderline "significant".

4.4.1 Robustness

Some additional robustness checks are conducted, and similar results can be obtained when estimating different specifications as shown in the appendix. First, the sample ends before the zero lower bound period (in 2007M12). In this case, impulse responses are less precisely estimated given the large reduction in the sample period. There is a small price puzzle following a loosening and this affects the comparison. Apart from that, overall conclusions are maintained.

Second, Swanson (2021)'s factors are used in the regression. This offers an opportunity both to check if results are robust to the use of an alternative instrument (FFR factor) and to control for forward guidance (FG factor) and Fed's asset purchases (LSAP factor). Even using an LP-IV, given the scale of these interventions and the frequency in the use of these tools, it is important to control for unconventional monetary policy as a robustness check. Except for CPI, whose difference is now only slight - positive at the beginning as in the benchmark, but positive at some later horizons -, the main conclusions remain the same.

Finally, FCI is replaced with the excess bond premium and the BAA spread. Even though they are narrower measures since they are constructed using exclusively spreads on non-financial corporate bonds, they are broadly used in this literature Gertler and Karadi (2015); Caldara and Herbst (2019). Once more, results are similar.

4.5 Euro Area

This section employs the same methodology to euro area data. This helps understand whether the results presented in previous sections are US-specific or more general.

4.5.1 The Euro Area dataset

The local projections are estimated on the German 1-year government bond yield in order to capture the safest one-year interest rate as in Jarociński and Karadi (2020), the harmonised index of consumer prices (HICP), the industrial production index (IP), the unemployment rate, and the spread non-financial corporate euro area with respect to the German yield, built by

Gilchrist and Mojon (2018), to capture financial conditions.[11] They show that credit spreads provide substantial predictive content for a variety of real activity and lending measures for the euro area.

The sample runs from 1999M01 to 2020M02, and to identify the effects for the euro area, I use the updated conventional monetary policy shocks of Jarociński and Karadi (2020).[12] They are based on the co-movement of the high-frequency surprises of interest rates and stock prices around policy announcements, which allow the disentangling of monetary policy and the central bank's assessment of the economy. As the authors show, measures free of information effects are crucial to achieving unbiased inference.

4.5.2 The Euro Area results

Figure 4.4 displays the impulse responses for the euro area in the linear case. Overall, they are qualitatively similar to the results from the US estimation. Prices and industrial production go down as expected. However, while the effect on prices is more long-lasting than the effect found in the US, the effect on industrial production is much less persistent. The unemployment rate, on the other hand, does not react to a monetary tightening whereas the spread goes slightly up.

Figure 4.5 shows the impulse responses after positive and negative shocks, and their differences. Once more, the response of prices after a loosening is slightly stronger than the response to a tightening although the difference is much smaller. As in the US case, once shocks are split into positive and negative, only positive shocks have an impact on industrial production, resulting in a reasonable difference in terms of HPDI. In fact, the negative shocks are the ones causing the responses in the linear case to be so small. This is a good example of how splitting the shocks into positive and negative can help us understand what is behind the results in the symmetric case.

Differently, however, the unemployment rate displays no difference. More than that, in this sample, neither positive nor negative shocks seem to have strong effects on the unemployment

[11] This choice differs from the US baseline specification because the FCI for the euro area constructed by Petronevich et al. (2019) is available only since 2008.

[12] This series is available on Marek Jarocinsky's webpage. An alternative closer to the US specification would be to use Altavilla et al. (2019)'s rate factor. However, besides reducing the sample period even further, this would bring little information for the estimation of the impulse responses since, in contrast to the US, the euro area has been in the zero lower bound for a large portion of the sample period.

Figure 4.4: Impulse Responses: Linear case for the euro area

Notes: The shaded area represents the 68 percent HPD credible sets, and the blue line represents the median. The monetary policy shock has been normalised to have an impact of 25 basis points on the German 1-year rate.

rate in the euro area. Spreads, on the other hand, are in line with the evidence found for the US, with positive shocks having stronger effects. Once more, the weak effects found in the linear case are caused by the fact that spreads do not react after a loosening in the euro area.

In general, the evidence of asymmetry found in the euro area is very similar to the one found in the US. This is relevant because it shows that asymmetric responses are not a particularity of the US and that the predictions of theoretical macro models also find support in the euro area.

4.6 Conclusion

This paper has investigated whether the effects of monetary tightenings and loosenings on standard macro and financial variables have been asymmetric recently. This is relevant because recent samples may lead to different conclusions for the real variables as reported by Ramey (2016). Moreover, the responses of financial conditions are very important for monetary transmission.

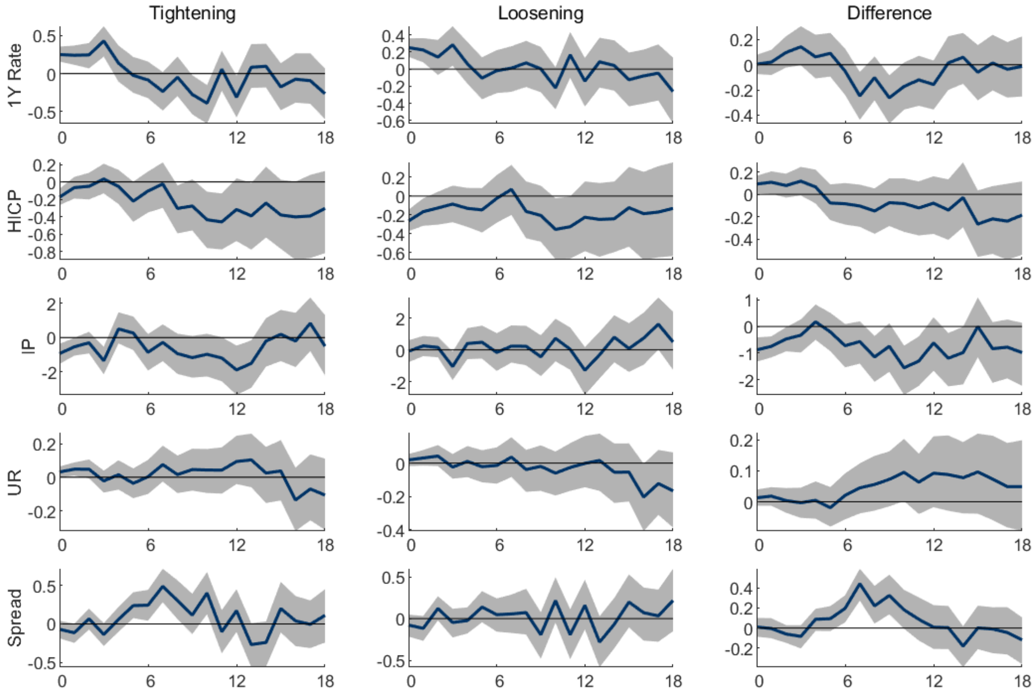

Figure 4.5: Impulse Responses: Positive and negative shocks for the euro area

Notes: The shaded area represents the 68 percent HPD credible sets, and the blue line represents the median. The left panel presents impulse responses to positive shocks. The central panel shows the impulse responses to negative shocks. They are flipped to facilitate comparison. The right panel shows their difference. The monetary policy shock has been normalised to have an impact of 25 basis points on the German 1-year rate.

In order to do that, a Bayesian version of the Generalised Least Squares (GLS) procedure proposed by Lusompa (2021) is employed. The LP-IV estimand is then calculated based on the ratio of reduced-form to first-stage coefficients as pointed out by Plagborg-Møller and Wolf (2021). This is suitable because it allows the dynamics to be asymmetric. The use of high-frequency surprises around policy announcements as instruments in the LP-IV approach is important to cope with selection on unobservables, ameliorating the identification problem.

All these refinements lead to better estimates of the dynamic responses to monetary tightenings and loosenings, which, in turn, can improve the understanding of the transmission of monetary policy. Results show that there is evidence of asymmetry for all variables. Industrial production, unemployment, and FCI respond more strongly to positive shocks while CPI responds more weakly to positive shocks, especially at the beginning, due to downward nominal rigidities.

Most importantly, these results show that the usual linear impulse responses of industrial production and unemployment rate following monetary policy shocks are driven by contractionary shocks, highlighting the importance of taking into account the direction of the intervention in the study of its effects, with meaningful implications for policy.

The main findings are similar when the non-linear local projections are estimated using a euro area dataset. This is important because it shows that, although the literature on the asymmetry of dynamic responses has focused on the US, the empirical evidence in favour of asymmetry in the dynamic effects of monetary policy is not US-specific.

Bibliography

Alessandri, P. and Mumtaz, H. (2017). Financial conditions and density forecasts for US output and inflation. *Review of Economic Dynamics*, 24:66–78.

Altavilla, C., Brugnolini, L., Gürkaynak, R. S., Motto, R., and Ragusa, G. (2019). Measuring euro area monetary policy. *Journal of Monetary Economics*, 108:162–179.

Andrade, P. and Ferroni, F. (2018). Delphic and Odyssean monetary policy shocks: Evidence from the Euro Area. FRB of Chicago Working Paper No. WP-2018-12.

Angrist, J. D., Jordà, Ò., and Kuersteiner, G. M. (2018). Semiparametric estimates of monetary policy effects: string theory revisited. *Journal of Business & Economic Statistics*, 36(3):371–387.

Antolín-Díaz, J. and Rubio-Ramírez, J. F. (2018). Narrative sign restrictions for SVARs. *American Economic Review*, 108(10):2802–29.

Aramonte, S., Rosen, S., and Schindler, J. W. (2017). Assessing and combining financial conditions indexes. *International Journal of Central Banking*, 47.

Arias, J. E., Caldara, D., and Rubio-Ramirez, J. F. (2019). The systematic component of monetary policy in svars: An agnostic identification procedure. *Journal of Monetary Economics*, 101:1–13.

Arias, J. E., Rubio-Ramirez, J. F., and Waggoner, D. F. (2014). Inference based on SVARs identified with sign and zero restrictions: Theory and applications. International Finance Discussion Papers 1100, Board of Governors of the Federal Reserve System (U.S.).

Arias, J. E., Rubio-Ramírez, J. F., and Waggoner, D. F. (2018). Inference based on structural vector autoregressions identified with sign and zero restrictions: Theory and applications. *Econometrica*, 86(2):685–720.

Arias, J. E., Rubio-Ramírez, J. F., and Waggoner, D. F. (2021). Inference in bayesian proxy-svars. *Journal of Econometrics*.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.

Bańbura, M., Giannone, D., and Reichlin, L. (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1):71–92.

Barakchian, S. M. and Crowe, C. (2013). Monetary policy matters: Evidence from new shocks data. *Journal of Monetary Economics*, 60(8):950–966.

Barnichon, R. and Matthes, C. (2018). Functional approximation of impulse responses. *Journal of Monetary Economics*, 99:41–55.

Baumeister, C. and Hamilton, J. D. (2015). Sign restrictions, structural vector autoregressions, and useful prior information. *Econometrica*, 83(5):1963–1999.

Ben Zeev, N. (2018). What can we learn about news shocks from the late 1990s and early 2000s boom-bust period? *Journal of Economic Dynamics and Control*, 87:94–105.

Ben Zeev, N., Gunn, C., and Khan, H. (2019). Monetary news shocks. *Journal of Money, Credit and Banking*.

Bernanke, B. (2007). Federal reserve communications. Speech at the Cato Institute 25th Annual Monetary Conference, Washington, DC, November 14.

Bernanke, B. S., Boivin, J., and Eliasz, P. (2005). Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. *The Quarterly Journal of Economics*, 120(1):387–422.

Bernanke, B. S. and Gertler, M. (1995). Inside the black box: the credit channel of monetary policy transmission. *Journal of Economic perspectives*, 9(4):27–48.

Bholat, D., Hans, S., Santos, P., and Schonhardt-Bailey, C. (2015). *Text mining for central banks*. Number 33 in Handbooks. Centre for Central Banking Studies, Bank of England.

Blei, D. M. and Lafferty, J. D. (2006). Dynamic topic models. In *Proceedings of the 23rd International Conference on Machine learning*, pages 113–120.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan):993–1022.

Blot, C. and Hubert, P. (2018). Central bank communication during normal and crisis times. *European Parliament Directorate-General for Internal Policies, Monetary Dialogue*.

Borio, C. and Zabai, A. (2018). Unconventional monetary policies: a re-appraisal. In *Research Handbook on Central Banking*. Edward Elgar Publishing.

Braun, R. and Brüggemann, R. (2020). Identification of SVAR models by combining sign restrictions with external instruments. University of Konstanz Department of Economics Working Paper Series.

Brave, S. and Butters, R. A. (2012). Diagnosing the financial system: financial conditions and financial stress. *International Journal of Central Banking*, 29.

Brave, S. A. and Butters, R. A. (2011). Monitoring financial stability: A financial conditions index approach. *Economic Perspectives*, 35(1):22.

Bundick, B., Herriford, T., and Smith, A. (2017). Forward guidance, monetary policy uncertainty, and the term premium. *Monetary Policy Uncertainty, and the Term Premium (July 2017)*.

Bundick, B. and Smith, A. L. (2020). The dynamic effects of forward guidance shocks. *The Review of Economics and Statistics*, 102(5):946–965.

Caldara, D. and Herbst, E. (2019). Monetary policy, real activity, and credit spreads: Evidence from Bayesian proxy SVARs. *American Economic Journal: Macroeconomics*, 11(1):157–92.

Campbell, J. R., Evans, C. L., Fisher, J. D., Justiniano, A., Calomiris, C. W., and Woodford, M. (2012). Macroeconomic effects of federal reserve forward guidance [with comments and discussion]. *Brookings Papers on Economic Activity*, pages 1–80.

Campbell, J. R., Fisher, J. D., Justiniano, A., and Melosi, L. (2017). Forward guidance and macroeconomic outcomes since the financial crisis. *NBER Macroeconomics Annual*, 31(1):283–357.

Canova, F. and Paustian, M. (2011). Business cycle measurement with some theory. *Journal of Monetary Economics*, 58(4):345–361.

Carriero, A., Galvao, A. B., and Marcellino, M. (2020). Credit conditions and the asymmetric effects of monetary policy shocks.

Carter, C. K. and Kohn, R. (1994). On gibbs sampling for state space models. *Biometrika*, 81(3):541–553.

Cochrane, J. H. (1994). Shocks. In *Carnegie-Rochester Conference series on public policy*, volume 41, pages 295–364. Elsevier.

Coibion, O. and Gorodnichenko, Y. (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy*, 120(1):116–159.

Coibion, O. and Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, 105(8):2644–78.

D'Amico, S. and King, T. B. (2015). What does anticipated monetary policy do? Federal Reserve Bank of Chicago Working Paper.

Debortoli, D., Forni, M., Gambetti, L., and Sala, L. (2020). Asymmetric effects of monetary policy easing and tightening.

Debortoli, D., Gali, J., and Gambetti, L. (2019). On the empirical (ir) relevance of the zero lower bound constraint. In *NBER Macroeconomics Annual 2019, volume 34*. University of Chicago Press.

Del Negro, M. (2018). Discussion of Miranda-Agrippino & Ricco: "The transmission of monetary policy shocks". BIS Research Network Discussion Slides.

Del Negro, M., Giannoni, M. P., and Patterson, C. (2012). The forward guidance puzzle. FRB of New York Staff Report.

Elliott, G. and Timmermann, A. (2013). *Handbook of Economic Forecasting*. Elsevier.

Ferreira, L. N. (2020). Forward guidance matters: disentangling monetary policy shocks. Central Bank of Brazil Working Paper Series.

Ferreira, L. N. (2021). Forecasting with VAR-teXt and DFM-teXt models: exploring the predictive power of central bank communication. Central Bank of Brazil Working Paper Series.

Forni, M. and Gambetti, L. (2014). Sufficient information in structural VARs. *Journal of Monetary Economics*, 66:124–136.

Friedman, B. M. and Woodford, M. (2010). *Handbook of Monetary Economics*. Elsevier.

Friedman, M. and Schwartz, A. J. (1963). *A monetary history of the United States, 1867-1960.* Princeton University Press.

Galí, J. (2015). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications.* Princeton University Press.

Gertler, M. and Karadi, P. (2011). A model of unconventional monetary policy. *Journal of Monetary Economics*, 58(1):17–34.

Gertler, M. and Karadi, P. (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics*, 7(1):44–76.

Geweke, J. and Amisano, G. (2010). Comparing and evaluating Bayesian predictive distributions of asset returns. *International Journal of Forecasting*, 26(2):216–230.

Giacomini, R. and Kitagawa, T. (2018). Robust bayesian inference for set-identified models. Cemmap working paper.

Giacomini, R. and White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74(6):1545–1578.

Gilchrist, S. and Mojon, B. (2018). Credit risk in the euro area. *The Economic Journal*, 128(608):118–158.

Gilchrist, S. and Zakrajšek, E. (2012). Credit spreads and business cycle fluctuations. *American Economic Review*, 102(4):1692–1720.

Gonçalves, S., Herrera, A. M., Kilian, L., and Pesavento, E. (2021). Impulse response analysis for structural dynamic models with nonlinear regressors. *Journal of Econometrics.*

Granger, C. W. J., Elliott, G., and Timmermann, A. (2006). *Handbook of Economic Forecasting*, volume 1. Elsevier.

Griffiths, T. L. and Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National academy of Sciences*, 101(suppl 1):5228–5235.

Gros, D. (2018). When communication becomes the policy. *European Parliament Directorate-General for Internal Policies, Monetary Dialogue.*

Gürkaynak, R. S., Kısacıkoğlu, B., and Wright, J. H. (2018). Missing events in event studies: Identifying the effects of partially-measured news surprises. National Bureau of Economic Research.

Gürkaynak, R. S., Sack, B., and Swansonc, E. T. (2005). Do actions speak louder than words? The response of asset prices to monetary policy actions and statements. *International Journal of Central Banking.*

Hagedorn, M., Luo, J., Manovskii, I., and Mitman, K. (2019). Forward guidance. *Journal of Monetary Economics*, 102(C):1–23.

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

Hansen, S., McMahon, M., and Prat, A. (2018). Transparency and deliberation within the FOMC: a computational linguistics approach. *The Quarterly Journal of Economics*, 133(2):801–870.

Hansen, S., McMahon, M., and Tong, M. (2019). The long-run information effect of central bank communication. *Journal of Monetary Economics*, 108:185–202.

Hatzius, J., Hooper, P., Mishkin, F. S., Schoenholtz, K. L., and Watson, M. W. (2010). Financial conditions indexes: A fresh look after the financial crisis. Technical report, National Bureau of Economic Research.

Hayo, B. and Neuenkirch, M. (2010). Do federal reserve communications help predict federal funds target rate decisions? *Journal of Macroeconomics*, 32(4):1014–1024.

Holm, M. B., Paul, P., and Tischbirek, A. (2020). The transmission of monetary policy under the microscope. Federal Reserve Bank of San Francisco.

Inoue, A. and Kilian, L. (2013). Inference on impulse response functions in structural var models. *Journal of Econometrics*, 177(1):1–13.

Ivanov, V. and Kilian, L. (2005). A practitioner's guide to lag order selection for var impulse response analysis. *Studies in Nonlinear Dynamics & Econometrics*, 9(1).

Jarociński, M. and Karadi, P. (2020). Deconstructing monetary policy surprises—the role of information shocks. *American Economic Journal: Macroeconomics*, 12(2):1–43.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American economic review*, 95(1):161–182.

Jordà, Ò., Singh, S. R., and Taylor, A. M. (2020). The long-run effects of monetary policy. National Bureau of Economic Research.

Kalamara, E., Turrell, A., Redl, C., Kapetanios, G., and Kapadia, S. (2020). Making text count: economic forecasting using newspaper text. Bank of England Working Paper.

Kaminska, I., Mumtaz, H., and Šustek, R. (2021). Monetary policy surprises and their transmission through term premia and expected interest rates. *Journal of Monetary Economics*, 124:48–65.

Kilian, L. and Lütkepohl, H. (2017). *Structural vector autoregressive analysis*. Cambridge University Press.

Kilian, L. and Murphy, D. P. (2014). The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics*, 29(3):454–478.

Kuttner, K. N. (2001). Monetary policy surprises and interest rates: Evidence from the fed funds futures market. *Journal of Monetary Economics*, 47(3):523–544.

Lakdawala, A. (2019). Decomposing the effects of monetary policy using an external instruments SVAR. *Journal of Applied Econometrics*, 34(6):934–950.

Larsen, V. H. and Thorsrud, L. A. (2019). The value of news for economic developments. *Journal of Econometrics*, 210(1):203–218.

Leeper, E. M., Walker, T. B., and Yang, S.-C. S. (2013). Fiscal foresight and information flows. *Econometrica*, 81(3):1115–1145.

Lucca, D. O. and Trebbi, F. (2009). Measuring central bank communication: an automated approach with application to FOMC statements. National Bureau of Economic Research.

Ludvigson, S. C., Ma, S., and Ng, S. (2020). Shock restricted structural vector-autoregressions. National Bureau of Economic Research.

Lusompa, A. (2021). Local projections, autocorrelation, and efficiency. Federal Reserve Bank of Kansas City Working Paper.

Lustenberger, T. and Rossi, E. (2020). Does central bank transparency and communication affect financial and macroeconomic forecasts? *International Journal of Central Banking.*

McCracken, M. W. and Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4):574–589.

McKay, A., Nakamura, E., and Steinsson, J. (2016). The power of forward guidance revisited. *American Economic Review*, 106(10):3133–58.

Miranda-Agrippino, S. and Ricco, G. (2019). Identification with external instruments in structural vars under partial invertibility.

Miranda-Agrippino, S. and Ricco, G. (2021). The transmission of monetary policy shocks. *American Economic Journal: Macroeconomics*, 13(3):74–107.

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective.* MIT press.

Nakamura, E. and Steinsson, J. (2018). High-frequency identification of monetary non-neutrality: the information effect. *The Quarterly Journal of Economics*, 133(3):1283–1330.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

Newman, D., Lau, J. H., Grieser, K., and Baldwin, T. (2010). Automatic evaluation of topic coherence. In *Human language technologies: The 2010 annual conference of the North American chapter of the association for computational linguistics*, pages 100–108.

Paul, P. (2020). The time-varying effect of monetary policy on asset prices. *Review of Economics and Statistics*, 102(4):690–704.

Petronevich, A., Sahuc, J.-G., et al. (2019). Un nouvel indice banque de france des conditions financières pour la zone euro. *Bulletin de la Banque de France*, (223):1–7.

Plagborg-Møller, M. and Wolf, C. K. (2018). Instrumental variable identification of dynamic variance decompositions. *Unpublished paper: Department of Economics, Princeton University.*

Plagborg-Møller, M. and Wolf, C. K. (2021). Local projections and VARs estimate the same impulse responses. *Econometrica*, 89(2):955–980.

Podstawski, M., Schlaak, T., and Rieth, M. (2018). Monetary policy, external instruments and heteroskedasticity. *DIW Berlin Discussion Paper.*

Ramey, V. A. (2016). Macroeconomic shocks and their propagation. In *Handbook of Macroeconomics*, volume 2, pages 71–162. Elsevier.

Ravn, M. O. and Sola, M. (2004). Asymmetric effects of monetary policy in the united states. *Review-Federal Reserve Bank of Saint Louis*, 86:41–58.

Roberts, M. E., Stewart, B. M., and Airoldi, E. M. (2016). A model of text for experimentation in the social sciences. *Journal of the American Statistical Association*, 111(515):988–1003.

Rubio-Ramirez, J. F., Waggoner, D. F., and Zha, T. (2010). Structural vector autoregressions: Theory of identification and algorithms for inference. *The Review of Economic Studies*, 77(2):665–696.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, pages 1–48.

Smets, F. and Wouters, R. (2007). Shocks and frictions in US business cycles: A bayesian dsge approach. *American Economic Review*, 97(3):586–606.

Stock, J. H. and Watson, M. W. (2018). Identification and estimation of dynamic causal effects in macroeconomics using external instruments. *The Economic Journal*, 128(610):917–948.

Swanson, E. T. (2018). The federal reserve is not very constrained by the lower bound on nominal interest rates. National Bureau of Economic Research.

Swanson, E. T. (2021). Measuring the effects of federal reserve forward guidance and asset purchases on financial markets. *Journal of Monetary Economics*, 118:32–53.

Tenreyro, S. and Thwaites, G. (2016). Pushing on a string: US monetary policy is less powerful in recessions. *American Economic Journal: Macroeconomics*, 8(4):43–74.

Thorsrud, L. A. (2018). Words are the new numbers: A newsy coincident index of the business cycle. *Journal of Business & Economic Statistics*, pages 1–17.

Uhlig, H. (2005). What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics*, 52(2):381–419.

Uhlig, H. (2017). Shocks, sign restrictions, and identification. *Advances in Economics and Econometrics*, 2:95.

Woodford, M. (2003). *Interest and prices: Foundations of a theory of monetary policy*. Princeton University Press.

Wu, J. C. and Xia, F. D. (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit and Banking*, 48(2-3):253–291.

Zlobins, A. (2019). Macroeconomic Effects of the ECB's Forward Guidance. Working Papers 2019/03, Latvijas Banka.

Appendix A

Appendix to Chapter 2

A. Conjugate Priors and Posteriors

Given a multivariate normal inverse Wishart (NIW) distribution (conjugate prior) of the form NIW$(v_0, \Psi_0, \beta_0, S_0)$:

$$\begin{aligned} \Sigma &\sim IW(S_0, v_0) \\ \beta|\Sigma &\sim \mathcal{N}(\beta_0, \Sigma \otimes \Psi_0) \end{aligned}$$

where S_0 is the prior scale matrix, v_0 the degrees of freedom and Ψ_0 is a diagonal matrix with common elements to all equations.

The posterior distribution over the reduced-form parameters is NIW$(v_1, \Psi_1, \beta_1, S_1)$:

$$\begin{aligned} \Sigma|y &\sim IW(S_1, v_1) \\ \beta|y, \Sigma &\sim \mathcal{N}(\beta_1, \Sigma \otimes \Psi_1) \end{aligned}$$

where, for the general case:

$$\begin{aligned} v_1 &= v_0 + T \\ \Psi_1 &= (X'X + \Psi_0^{-1})^{-1} \\ \beta_1 &= \Psi_1(X'Y + {\Psi_0}^{-1}\beta_0) \\ S_1 &= Y'Y + S_0 + \beta_0'\Psi_0^{-1}\beta_0 - \beta_1'\Psi_1^{-1}\beta_1 \end{aligned}$$

and, for the flat (Jeffreys) prior, simply:

$$
\begin{aligned}
v_1 &= T \\
\Psi_1 &= (X'X)^{-1} \\
\beta_1 &= \Psi_1(X'Y) = \hat{\beta}^{OLS} \\
S_1 &= \hat{S}^{OLS}
\end{aligned}
$$

B. Algorithm

The algorithm follows very closely Arias et al. (2018) and Antolín-Díaz and Rubio-Ramírez (2018). As customary, the starting point in identification by sign restrictions is to characterise the set admissible models by drawing Q, where Q $\in O(n)$, the set of all orthogonal $n \times n$ matrices.

Arias et al. (2018) use an alternative parametrisation to emphasise the role of the orthogonal matrix Q, which they call orthogonal reduced-form parametrisation. This parametrisation is characterised by the reduced-form parameters β and Σ together with Q. The NIW(v, Ψ, β, S) then becomes UNIW(v, Ψ, β, S), the uniform-normal-inverse-Wishart distribution over the orthogonal reduced-form parametrisation. They show that independent draws of β, Σ and Q from UNIW(v, Ψ, β, S) are independent draws from a normal-generalised-normal distribution over the structural parametrisation, denoted by NGN(v, Ψ, β, S).[1]

Nevertheless, when zero restrictions are also imposed, additional sub-steps are necessary to properly achieve the objective of drawing from the correct distribution. Arias et al. (2018) argue that the distribution over the structural parametrisation conditional on the zero restrictions is no longer equal to the NGN(v, Ψ, β, S). They then suggest the computation of its density and its use as a proposal distribution for an importance sampler to draw from the NGN(v, Ψ, β, S) distribution over the structural parametrisation conditional on the zero restrictions.

Similarly, Antolín-Díaz and Rubio-Ramírez (2018) show that it is not correct to simply discard the draws that do not satisfy the narrative sign restrictions. Such a procedure would give high posterior probability to draws that are more likely to satisfy the narrative restrictions and would

[1] See Arias et al. (2018) for a detailed description.

deviate from drawing from the UNIW(v, Ψ, β, S) distribution. Therefore, importance weights inversely proportional to the probability of satisfying the narrative restrictions are computed and the draws are re-sampled accordingly. As before, this implies making independent draws from the NGN(v, Ψ, β, S). The following algorithm describes the whole procedure. In practice, the algorithm starts with an educated guess of the number of iterations necessary to achieve the required number of independent draws.[2]

ALGORITHM: This algorithm makes independent draws from the NGN(v, Ψ, β, S) distribution over the structural parametrisation conditional on the zero, traditional and narrative sign restrictions.

1. Independently draw (β, Σ) from the NIW(v, Ψ, β, S) distribution.

2. For $1 \leq j \leq N$, draw $\mathbf{x}_j \in \mathbb{R}^{N+1-j-z_j}$ independently from a standard normal distribution and set $\mathbf{w}_j$=$\mathbf{x}_j/||\mathbf{x}_j||$, where z_j is the number of zero restrictions associated with the jth structural shock.

3. Define Q = [q_1 ... q_N] recursively by $q_j = K_j w_j$ for any matrix K_j whose columns form an orthonormal basis for the null space of the (j – 1 + z_j) × N matrix

$$M_j = [q_1 \ \ ... \ \ q_{j-1} \ \ (Z_j F(f_h^{-1}(\beta, \Sigma, I_n)))]$$

where Z_j defines the zero restrictions on the jth structural shock for $1 \leq j \leq N$, f_h^{-1} is the function that transforms draws over the orthogonal reduced-form parametrisation into draws from the structural parametrisation, and F is a function of the structural parameters defined as a matrix that vertically stacks the impulse responses over which the restrictions will be imposed.[3]

[2] This number is calibrated in order for the algorithm to generate the desired number of draws that satisfy the zero, traditional and narrative sign restrictions.

[3] For instance, make L_0=chol($\hat{\Sigma}$) an initial guess of the structural impact matrix multiplier. That implies L_1=βL_0 and F=[L_0; L_1]. For a numerical example in more detail, see Arias et al. (2014) – working paper version – or Kilian and Lütkepohl (2017), pages 475-482.

4. Check if the sign restrictions are satisfied. If they are, compute the importance weights. Otherwise, discard the draw.

5. Return to Step 1 until the required number of draws satisfying the zero and sign restrictions has been obtained.

6. Re-sample with replacement using the importance weights.

7. Check if the narrative sign restrictions are satisfied:

$$\varepsilon_{j,t}(\Theta) < 0 \tag{A.1}$$

$$|H_{i,j,t}(\Theta, \varepsilon_t(\Theta))| > \max_{j' \neq j} |H_{i,j',t}(\Theta, \varepsilon_t(\Theta))| \tag{A.2}$$

where Θ collects the values of all structural parameters, equation (A.1) implies that jth shocks must be negative at time t and equation (A.2) implies that the contribution of the jth shock to variable i at time t must be greater than the contribution of any the other shock to variable i at time t. If the restrictions for the particular case are satisfied, approximate the new importance weights as the inverse of the probability of satisfying the narrative restrictions as in Antolín-Díaz and Rubio-Ramírez (2018). Otherwise, discard the draw.

8. Re-sample with replacement using the new importance weights.

C. Computational Aspects

The importance samplers are the most onerous part of the algorithm. For the latter, however, it is necessary to compute the weights only for the draws that satisfy the zero, traditional and narrative sign restrictions, which usually are hundreds of draws. On the other hand, the first importance sampler is computationally much more demanding because it involves computing the weights for all the candidate draws that satisfy the zero and sign restrictions. This number of draws can reach hundreds of thousands depending on the number of narrative sign restrictions imposed and how restrictive they are.

For the benchmark case, the number of draws satisfying the zero and sign restrictions was 100,000 and the effective sample size from Step 5 is 48,960. The number of draws satisfying the narrative sign restrictions was 3,888 and the effective sample size from Step 8 calculated based on 100,000 weights is 1,737. Because two importance samplers are used, it is also useful to keep track of the number of unique draws which survive until the end. So it is possible to be sure that the second importance sampler is not dominated by only a few surviving draws from the first one. The number of unique draws that satisfy the narrative sign restrictions is 1,293.

D. Variance Decomposition

The figure below presents the forecast error variance decomposition of the VAR along 10 years, measuring the reduction in the forecast variance resulting from knowing future realisations of each shock. As aforementioned, the estimated shocks are quite small. This is a by-product of the high-frequency identification and has consequences for the variance decomposition.

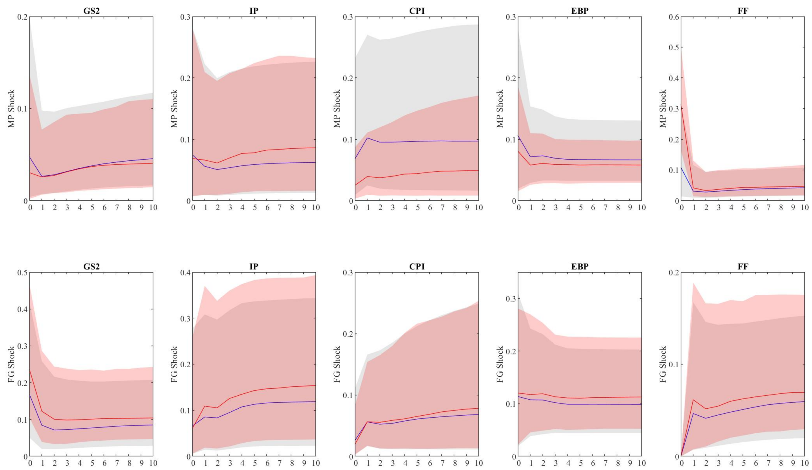

Figure A.1: Forecast Error Variance Decomposition

Notes: The grey shaded area represents the 68 percent (point-wise) HPD credible sets and the blue lines are the medians using sign restrictions. The pink shaded areas and the red lines display the equivalent quantities for the models that additionally satisfy narrative sign restrictions.

The conventional monetary policy shock explains a negligible fraction of short-run movements in CPI and IP in line with previous literature even though the number for the contribution to the CPI is lower than usual. The conventional monetary policy shock explains very little of the forecast error variance of the excess bond premium and of the 2-year government bond rate. The contribution to the fed funds changes heavily after the imposition of the narrative restrictions and the percentage on impact as well as its decay over time are in line with a trivariate VAR identified with standard sign restrictions.

Forward guidance shocks also explain a negligible fraction of short-run movements in CPI and IP, albeit the contributions are slightly larger than the ones of MP shocks. On the other hand, forward guidance shocks explain a reasonable percentage of the forecast error variance of GS2, especially on impact and with some degree of persistence if compared to the behaviour of the other variables. The reverse happens to the fed funds, to whose variance there is no contribution on impact by construction, and the contribution arises over time. For the excess bond premium, the percentage doubles when compared to the MP shock. When comparing, however, one should bear in mind that posterior uncertainty is large.

E. Other Dates and Robustness

i. FG shock in August 2011 and January 2012

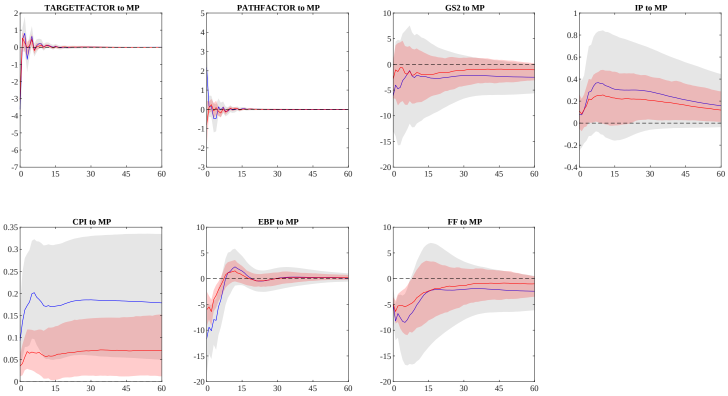

Figure A.2: Impulse Responses to a Conventional Monetary Policy Shock: Aug/2011 and Jan/2012

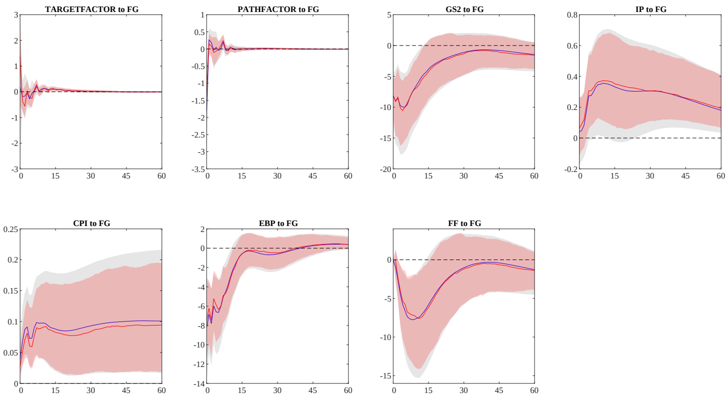

Figure A.3: Impulse Responses to a Forward Guidance Shock: Aug/2011 and Jan/2012

Notes: The grey shaded area represents the 68 percent (point-wise) HPD credible sets for the IRFs and the blue lines are the median IRFs using sign restrictions. The pink shaded areas and the red lines display the equivalent quantities for the models that additionally satisfy narrative sign restrictions.

ii. Ludvigsson type: January 2012

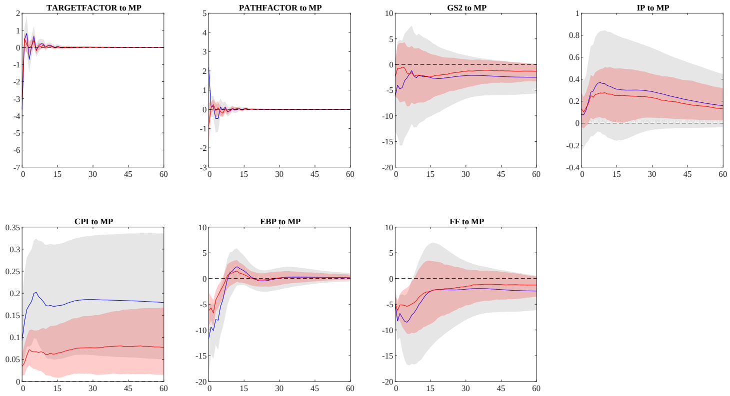

Figure A.4: Impulse Responses to a Conventional Monetary Policy Shock: Ludvigsson type

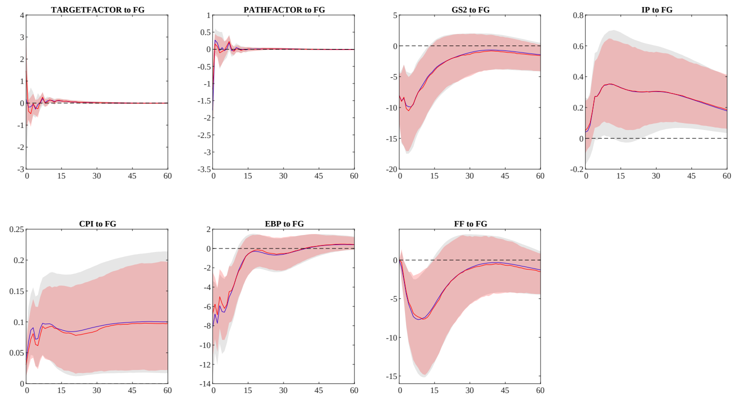

Figure A.5: Impulse Responses to a Forward Guidance Shock: Ludvigsson type

Notes: The grey shaded area represents the 68 percent (point-wise) HPD credible sets for the IRFs and the blue lines are the median IRFs using sign restrictions. The pink shaded areas and the red lines display the equivalent quantities for the models that additionally satisfy narrative sign restrictions.

iii. FG shock in August 2011 and October 2015

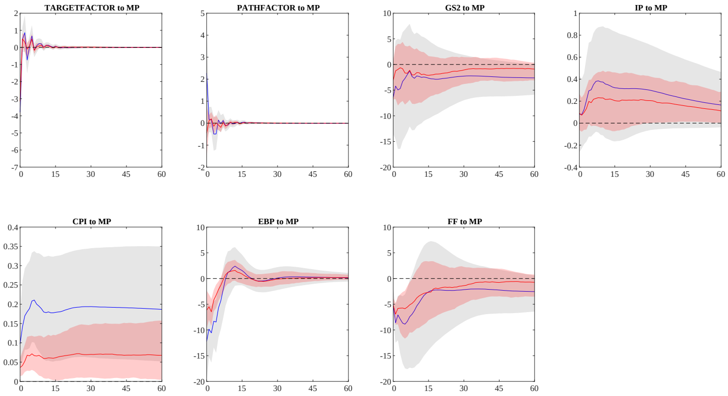

Figure A.6: Impulse Responses to a Conventional Monetary Policy Shock: Aug/2011 and Oct/2015

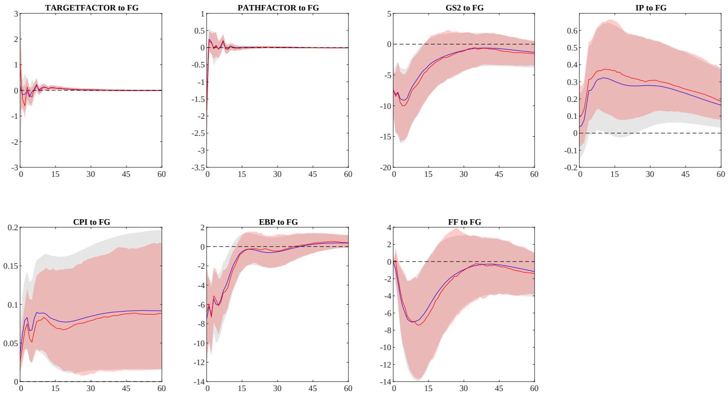

Figure A.7: Impulse Responses to a Forward Guidance Shock: Aug/2011 and Oct/2015

Notes: The grey shaded area represents the 68 percent (point-wise) HPD credible sets for the IRFs and the blue lines are the median IRFs using sign restrictions. The pink shaded areas and the red lines display the equivalent quantities for the models that additionally satisfy narrative sign restrictions.

iv. Sample Period: 1993-2012

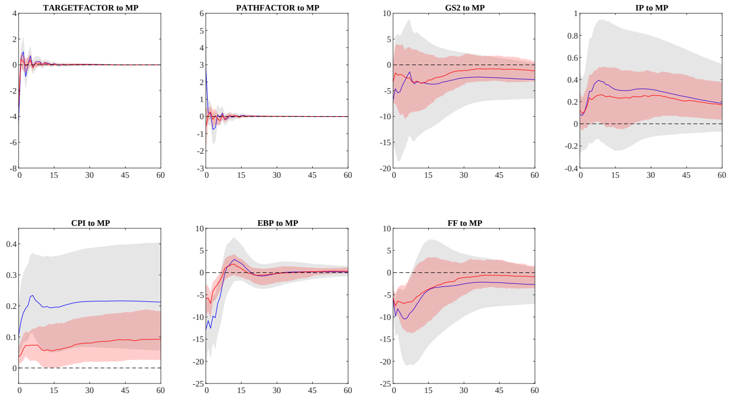

Figure A.8: Impulse Responses to a Conventional Monetary Policy Shock: 1993-2012

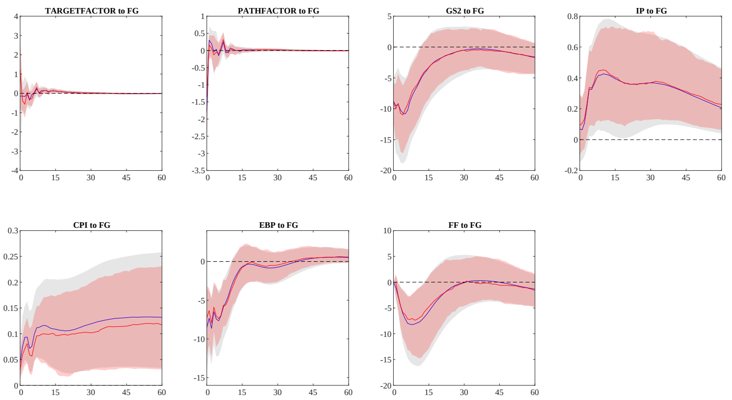

Figure A.9: Impulse Responses to a Forward Guidance Shock: 1993-2012

Notes: The grey shaded area represents the 68 percent (point-wise) HPD credible sets for the IRFs and the blue lines are the median IRFs using sign restrictions. The pink shaded areas and the red lines display the equivalent quantities for the models that additionally satisfy narrative sign restrictions.

v. Removing LSAP dates

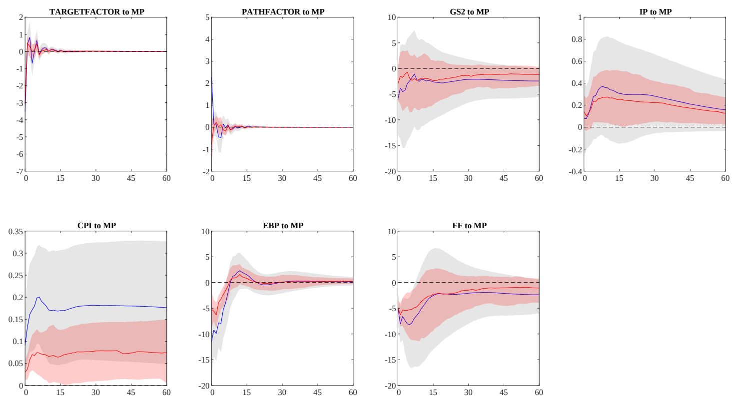

Figure A.10: Impulse Responses to a Conventional Monetary Policy Shock: No LSAP

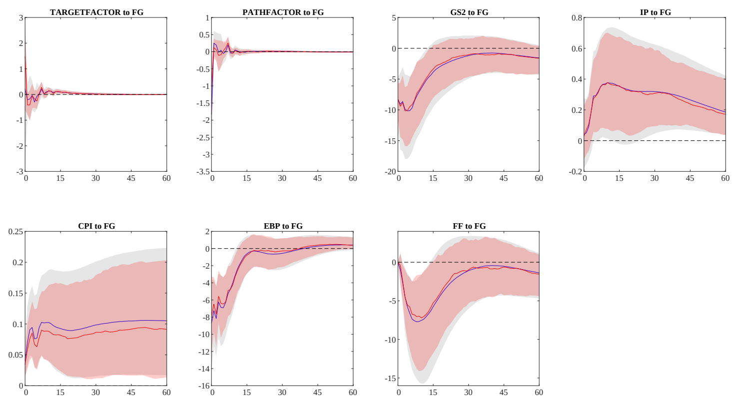

Figure A.11: Impulse Responses to a Forward Guidance Shock: No LSAP

Notes: The grey shaded area represents the 68 percent (point-wise) HPD credible sets for the IRFs and the blue lines are the median IRFs using sign restrictions. The pink shaded areas and the red lines display the equivalent quantities for the models that additionally satisfy narrative sign restrictions.

vi. Agnostic about CPI

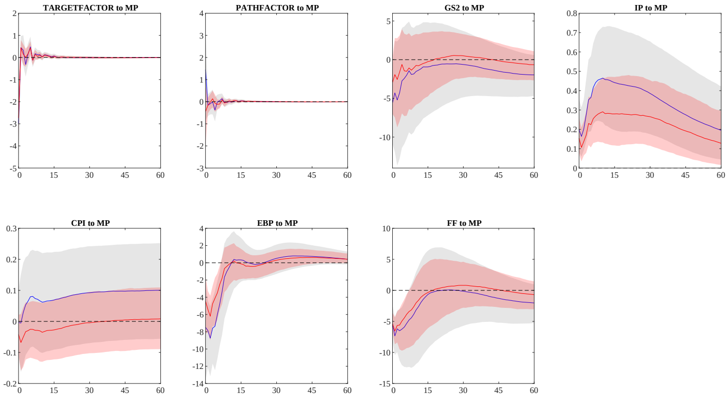

Figure A.12: Impulse Responses to a Conventional Monetary Policy Shock

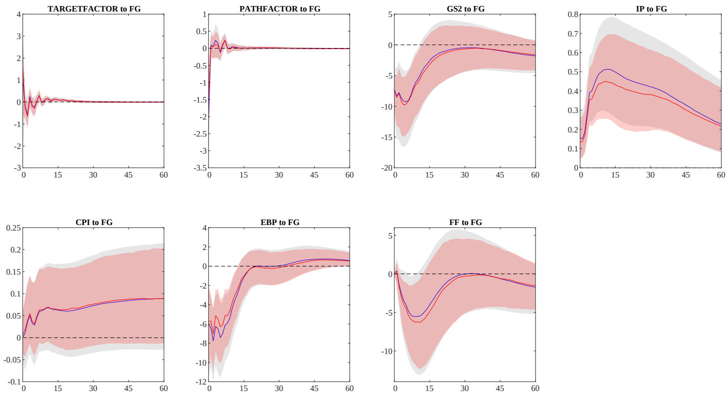

Figure A.13: Impulse Responses to a Forward Guidance Shock

Notes: The grey shaded area represents the 68 percent (point-wise) HPD credible sets for the IRFs and the blue lines are the median IRFs using sign restrictions. The pink shaded areas and the red lines display the equivalent quantities for the models that additionally satisfy narrative sign restrictions.

vii. Agnostic about IP and CPI

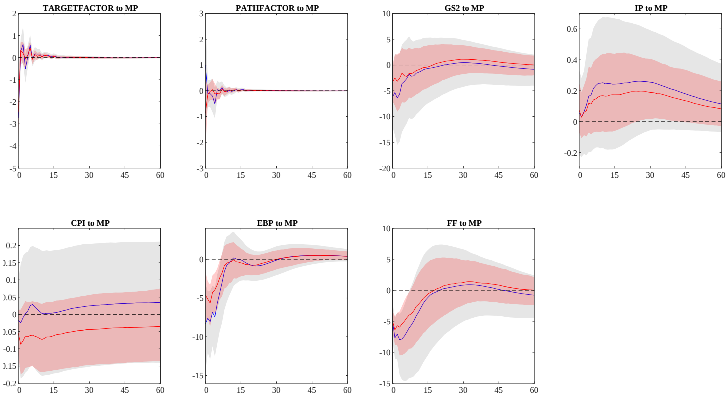

Figure A.14: Impulse Responses to a Conventional Monetary Policy Shock

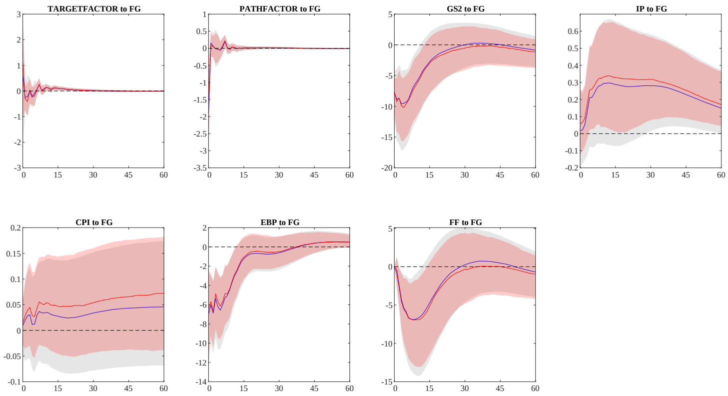

Figure A.15: Impulse Responses to a Forward Guidance Shock

Notes: The grey shaded area represents the 68 percent (point-wise) HPD credible sets for the IRFs and the blue lines are the median IRFs using sign restrictions. The pink shaded areas and the red lines display the equivalent quantities for the models that additionally satisfy narrative sign restrictions.

viii. Controlling for Jarociński and Karadi (2020)'s central bank private information shocks

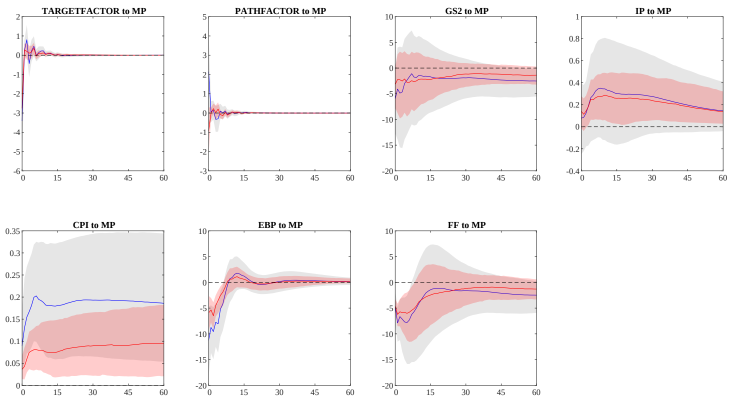

Figure A.16: Impulse Responses to a Conventional Monetary Policy Shock

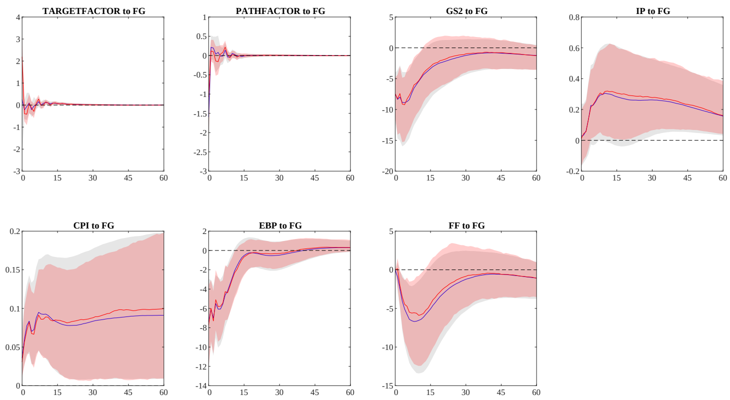

Figure A.17: Impulse Responses to a Forward Guidance Shock

Notes: The grey shaded area represents the 68 percent (point-wise) HPD credible sets for the IRFs and the blue lines are the median IRFs using sign restrictions. The pink shaded areas and the red lines display the equivalent quantities for the models that additionally satisfy narrative sign restrictions.

Appendix B

Appendix to Chapter 3

A. Forecast evaluation

Given a pair of models *(a,b)* the forecast evaluation criterion is calculated for each variable x_t as:

$$BF_t^{a,b} = \sum_{t=1}^{T} [log(LS_t^a) - log(LS_t^b)] \tag{B.1}$$

The conditional tests are constructed as in Giacomini and White (2006) and Alessandri and Mumtaz (2017) where the conditioning information set coincides with the information set that generates the forecasts.

These (log) predictive densities are estimated using kernel methods. This is done to address the point that the textual factors are generated regressors. As in Alessandri and Mumtaz (2017), densities are estimated over evenly-spaced grids of 100 points, and this is carried out using Matlab function *ksdensity*.

B. Point forecasts

Point forecasts are given by the root mean square errors (RMSE) based on the arithmetic mean of the draws of the simulated forecasts:

$$RMSE_{t,h}^i = \sqrt{(\hat{Y}_{t+h}^i(M) - Y_{t+h}^i)^2} \tag{B.2}$$

where $\hat{Y}^i_{t+h}(M)$ denotes the average over the forecast density produced by model M for variable i and Y^i_{t+h} is the actual data.

The table below reports the ratios of the RMSE: values lower than 1 favour the VAR-teXt.

Table B.1: RMSE: VAR-teXt versus benchmark VAR

	1M				3M			
K	y	π	r	s	y	π	r	s
5	1.01	0.99	1.03	0.99	1.00	1.00	1.01	1.00
	(0.211)	*(0.291)*	*(0.255)*	*(0.601)*	*(0.208)*	*(0.680)*	*(0.481)*	*(0.601)*
	(0.129)	*(0.623)*	*(0.461)*	*(0.356)*	*(0.148)*	*(0.718)*	*(0.101)*	*(0.740)*
10	1.02	0.99	1.00	1.00	1.01	0.99	0.99	1.00
	(0.122)	*(0.580)*	*(0.822)*	*(0.767)*	*(0.106)*	*(0.063)*	*(0.399)*	*(0.829)*
	(0.270)	*(0.391)*	*(0.778)*	*(0.967)*	*(0.144)*	*(0.061)*	*(0.005)*	*(0.966)*
15	1.02	0.98	1.01	0.99	1.01	0.99	0.98	1.00
	(0.213)	*(0.252)*	*(0.517)*	*(0.597)*	*(0.068)*	*(0.152)*	*(0.259)*	*(0.933)*
	(0.383)	*(0.569)*	*(0.456)*	*(0.891)*	*(0.115)*	*(0.339)*	*(0.008)*	*(0.790)*
20	1.02	0.98	1.00	0.99	1.01	0.99	0.98	1.00
	(0.290)	*(0.209)*	*(0.829)*	*(0.646)*	*(0.246)*	*(0.026)*	*(0.164)*	*(0.882)*
	(0.518)	*(0.350)*	*(0.224)*	*(0.920)*	*(0.282)*	*(0.063)*	*(0.006)*	*(0.947)*
25	1.01	0.97	1.00	0.98	1.00	0.98	0.98	0.99
	(0.673)	*(0.089)*	*(0.953)*	*(0.303)*	*(0.631)*	*(0.018)*	*(0.235)*	*(0.587)*
	(0.829)	*(0.186)*	*(0.337)*	*(0.599)*	*(0.446)*	*(0.073)*	*(0.015)*	*(0.606)*
30	1.02	0.97	1.01	0.99	1.01	0.98	0.98	0.99
	(0.421)	*(0.157)*	*(0.671)*	*(0.695)*	*(0.369)*	*(0.025)*	*(0.306)*	*(0.424)*
	(0.686)	*(0.347)*	*(0.572)*	*(0.874)*	*(0.434)*	*(0.102)*	*(0.024)*	*(0.313)*
35	1.01	0.97	1.01	0.99	1.00	0.98	0.98	1.00
	(0.572)	*(0.120)*	*(0.569)*	*(0.774)*	*(0.566)*	*(0.028)*	*(0.254)*	*(0.784)*
	(0.754)	*(0.290)*	*(0.617)*	*(0.914)*	*(0.277)*	*(0.055)*	*(0.058)*	*(0.403)*
40	1.01	0.97	1.01	0.99	1.01	0.98	0.98	0.99
	(0.531)	*(0.154)*	*(0.637)*	*(0.765)*	*(0.437)*	*(0.029)*	*(0.241)*	*(0.637)*
	(0.789)	*(0.173)*	*(0.559)*	*(0.937)*	*(0.549)*	*(0.146)*	*(0.032)*	*(0.451)*
45	1.02	0.96	1.01	0.98	1.01	0.98	0.98	0.99
	(0.475)	*(0.100)*	*(0.624)*	*(0.547)*	*(0.501)*	*(0.014)*	*(0.315)*	*(0.428)*
	(0.753)	*(0.052)*	*(0.615)*	*(0.772)*	*(0.555)*	*(0.080)*	*(0.038)*	*(0.347)*
50	1.01	0.97	1.02	0.99	1.00	0.98	0.98	0.99
	(0.680)	*(0.208)*	*(0.424)*	*(0.733)*	*(0.635)*	*(0.028)*	*(0.464)*	*(0.381)*
	(0.751)	*(0.090)*	*(0.983)*	*(0.858)*	*(0.574)*	*(0.134)*	*(0.112)*	*(0.261)*

Notes: The table shows the ratio of the average RMSEs, relative to the benchmark for the 1-month and the 3-month ahead forecasts over 2013M08–2020M02. The p-values of Giacomini and White (2006)'s test of unconditional (conditional) predictive ability are in the first (second) parenthesis. The variables are industrial production growth (y), inflation (π), shadow rate (r), and EBP (s).

According to the table, it is hard to distinguish between the augmented and the benchmark models in terms of point forecasts, apart from the 3-month ahead forecast of inflation and interest rate and the 1-month ahead forecast of inflation. For $K = 45$, the magnitude of the improvement in the former is lower, while the latter displays larger gains than when the metric of comparison is log-scores.

C. Alternative Specifications

i. 1-year rate

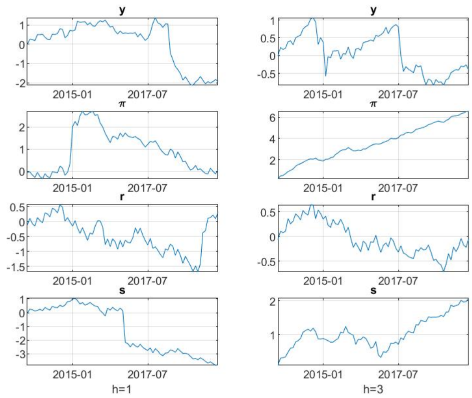

Figure B.1: Cumulative Log-Score Difference over Time: 1-year rate

Notes: The lines show the cumulative difference in log-scores between the VAR-text and the VAR models for horizons 1 and 3.

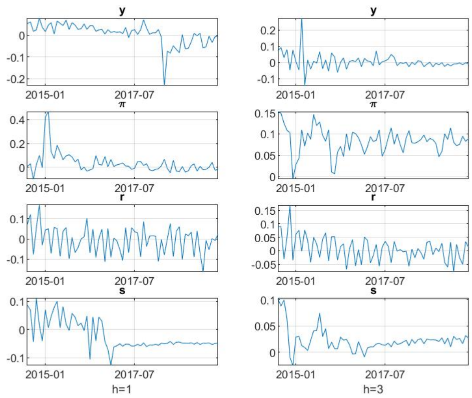

Figure B.2: Giacomini and White (2006)'s Decision Rule: 1-year rate

Notes: Positive values indicate the VAR-teXt is expected to work better in the future and should be selected.

ii. Fed funds

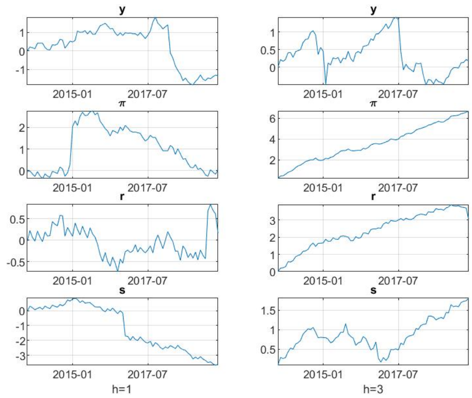

Figure B.3: Cumulative Log-Score Difference over Time: Fed funds

Notes: The lines show the cumulative difference in log-scores between the VAR-text and the VAR models for horizons 1 and 3.

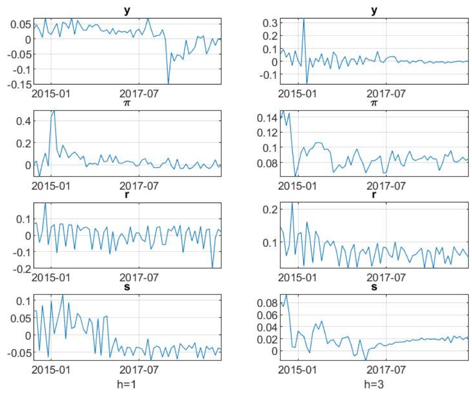

Figure B.4: Giacomini and White (2006)'s Decision Rule: Fed funds

Notes: Positive values indicate the VAR-teXt is expected to work better in the future and should be selected.

D. Labelled Topics

5 topics are selected: topic 3 (economic conditions and growth prospect), topics 8 and 9 (inflation), topic 10 (economic outlook) and topic 13 (interest rates).

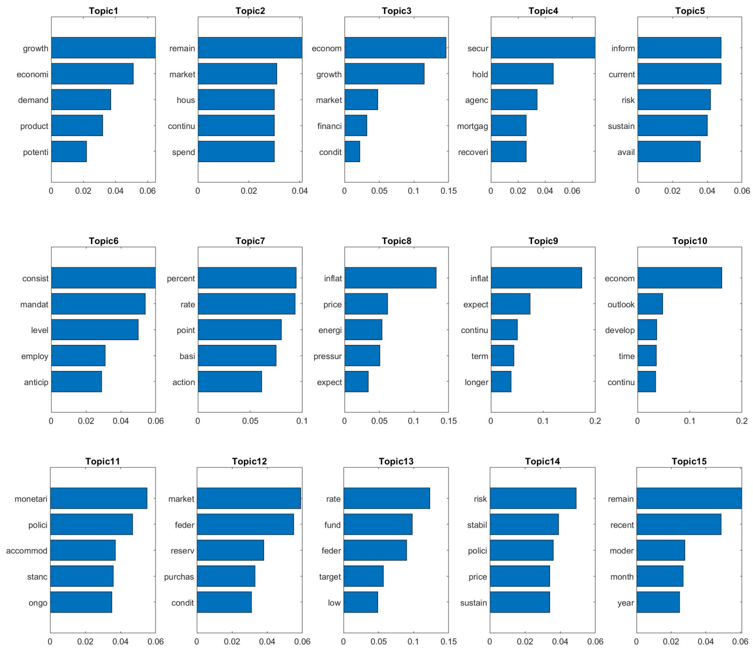

Figure B.5: Relative Frequency of the Top 5 Words for K=15

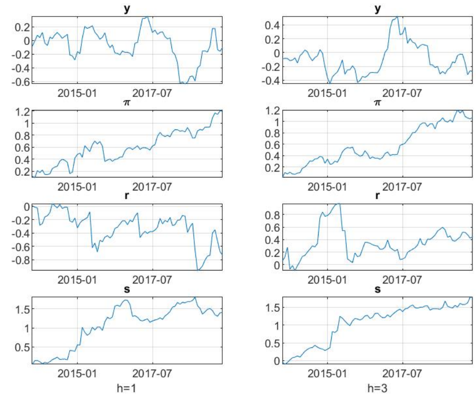

Figure B.6: Cumulative Log-Score Difference over Time: Labelled topics

Notes: The lines show the cumulative difference in log-scores between the VAR-text and the VAR models for horizons 1 and 3.

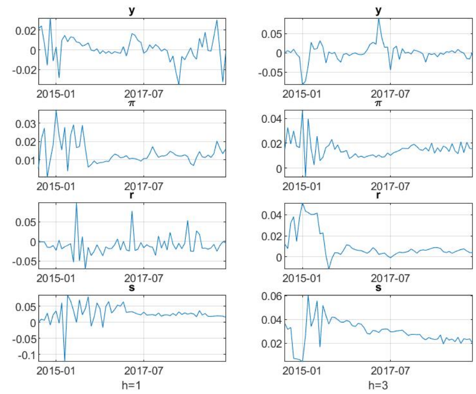

Figure B.7: Giacomini and White (2006)'s Decision Rule: Labelled topics

Notes: Positive values indicate the VAR-teXt is expected to work better in the future and should be selected.

E. Tone-adjusted VAR-teXt

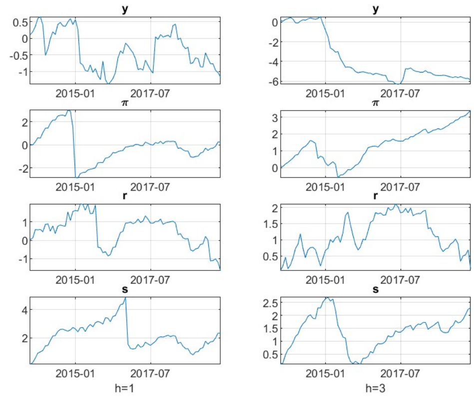

Figure B.8: Cumulative Log-Score Difference over Time: Tone-adjusted

Notes: The lines show the cumulative difference in log-scores between the Tone-adjusted VAR-teXt and the standard VAR-teXt models for horizons 1 and 3.

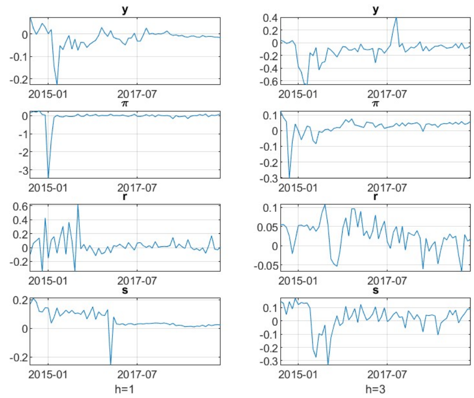

Figure B.9: Giacomini and White (2006)'s Decision Rule: Tone-adjusted

Notes: Positive values indicate the Tone-adjusted VAR-teXt is expected to work better in the future and should be selected.

F. Endogenous textual factors

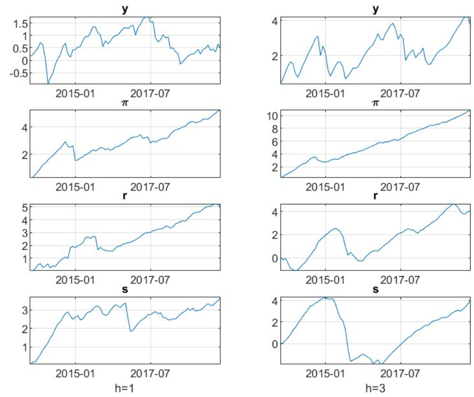

Figure B.10: Cumulative Log-Score Difference over Time: Endogenous textual factors

Notes: The lines show the cumulative difference in log-scores between the VAR with endogenous textual factors and the VAR models for horizons 1 and 3.

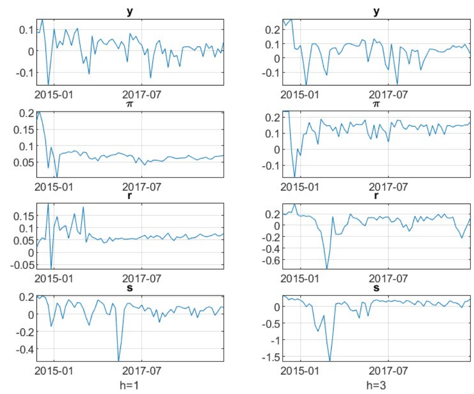

Figure B.11: Giacomini and White (2006)'s Decision Rule: Endogenous textual factors

Notes: Positive values indicate the VAR with endogenous textual factors is expected to work better in the future and should be selected.

G. Algorithm for DFM-teXt

The additional steps in comparison with the algorithm for the VAR-teXt are highlighted:

Step 1. Substep LDA as in Subsection 3.2.2.

Step 2. Sample factor loadings and the variance of the idiosyncratic components.

Conditional on the factors and Ω_{ii}, the factor loadings are sampled from their normal conditional distributions:

$$\Gamma_i \sim N(\bar{\Gamma}_i, \Omega_{ii}\bar{M}_i^{-1})$$

where $\bar{\Gamma}_i$ represents the OLS estimate and $\bar{M}_i = (F'_{i,t}F_{i,t})$.

Conditional on the factors and the factor loadings, the variance of the error term of the observation equation is sampled from:

$$\Omega_{ii} \sim IG\left(\frac{T}{2}, \frac{e'_F e_F}{2}\right) \tag{B.3}$$

where $e_F = (W_{i,t} - \Gamma_i F_t)$.

Step 3. Sample VAR-X coefficients and covariance as in Subsection 3.3.1.

Step 4. Prepare matrices for the state-space treating textual factors as exogenous observables and sample factors via the Carter and Kohn (1994) algorithm.

For details see Bernanke et al. (2005). The only modification is that, as the textual factors enter in the model as exogenous variables, they are treated as part of the intercept in the Kalman filter.

Step 5. Repeat steps 1 to 4 until the required number of draws has been reached.

Because estimating the DFM-teXt recursively takes much longer to run, the algorithm is iterated only 7,000 times for each data window, using every 10th draw from the LDA. With a burn-in of 4,000 draws, this results in 3,000 draws.

G. DFM-teXt Results

Variables	1M	GW-UC	GW-C	3M	GW-UC	GW-C
Real Personal Income	-0.53	0.509	0.868	0.74	0.332	0.279
RPI ex Transfers	-1.26	0.111	0.152	0.68	0.399	0.141
Real Consumption	-0.50	0.463	0.673	-0.79	0.210	0.375
Real Sales	-0.84	0.481	0.074	1.28	0.289	0.339
Retail Sales	-2.04	0.020	0.062	0.62	0.505	0.095
Industrial Production	-2.09	0.324	0.484	5.00	0.045	0.131
IP Products	-2.97	0.296	0.515	6.26	0.018	0.044
IP Finished Goods	-2.20	0.429	0.684	5.99	0.012	0.024
IP Consumer Goods	-0.85	0.742	0.453	4.88	0.002	0.007
IP Cons Durables	-0.27	0.915	0.971	5.91	0.001	0.005
IP Cons Nondurables	0.29	0.876	0.779	2.06	0.086	0.235
IP Bus Equipment	-0.84	0.723	0.889	4.39	0.035	0.059
IP Materials	-1.55	0.361	0.388	3.88	0.043	0.038
IP Durable Materials	-1.79	0.329	0.312	5.71	0.008	0.015
IP Nondurable Materials	-1.60	0.501	0.663	5.42	0.004	0.020
IP Manufacturing	-1.68	0.551	0.803	6.88	0.012	0.041
IP Residential Utilities	-1.05	0.465	0.527	0.35	0.733	0.874
IP Fuels	1.14	0.206	0.434	1.41	0.091	0.127
Cap Util: Manufacturing	-0.82	0.795	0.920	6.69	0.011	0.040
Help Wanted Index	2.06	0.179	0.454	-0.65	0.565	0.513
Help to Unemployment Ratio	2.46	0.069	0.163	-1.13	0.418	0.366
Civilian Labor Force	1.89	0.123	0.304	0.16	0.845	0.281
Civilian Employment	0.51	0.738	0.851	-0.68	0.612	0.336
Unemployment Rate	0.34	0.793	0.859	-3.92	0.022	0.062
U Mean Duration	-0.98	0.517	0.317	-5.24	0.002	0.012
Unemployed <5 weeks	-2.47	0.032	0.099	0.36	0.596	0.750
Unemployed 5-14 weeks	-0.35	0.651	0.878	0.08	0.903	0.784
Unemployed >15 weeks	-2.40	0.024	0.075	-1.99	0.139	0.255
Unemployed 15-26 weeks	-0.34	0.741	0.883	0.45	0.660	0.863
Unemployed >27 weeks	-1.42	0.091	0.175	-2.14	0.041	0.067
Initial Claims	2.85	0.240	0.170	-0.05	0.953	0.808
Non Farm Payroll Employment	-2.80	0.102	0.156	-1.25	0.590	0.791

Notes: The table shows average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 1-month and the 3-month ahead forecasts over 2013M08–2020M02. The GW-UC and GW-C columns display the p-values of Giacomini and White (2006)'s test of unconditional and conditional predictive ability respectively. The variables are described as in the FRED-MD database in addition to the excess bond premium and the shadow rate.

Variables	1M	GW-UC	GW-C	3M	GW-UC	GW-C
NFP Goods	-4.31	0.042	0.050	0.03	0.989	0.749
NFP Mining	9.26	0.169	0.142	35.68	0.340	0.577
NFP Construction	-4.13	0.012	0.027	-4.49	0.060	0.135
NFP Manufacturing	-1.35	0.545	0.370	3.29	0.111	0.198
NFP Durables	-0.20	0.925	0.580	6.51	0.001	0.003
NFP Nondurables	-6.54	0.020	0.048	-10.16	0.003	0.014
NFP Services	-1.13	0.433	0.660	-0.39	0.848	0.929
NFP TT&U	-1.52	0.297	0.553	0.13	0.945	0.358
NFP Wholesale Trade	2.09	0.185	0.315	5.06	0.004	0.021
NFP Retail Trade	-0.91	0.523	0.649	0.73	0.679	0.601
NFP Financial	-2.45	0.065	0.150	-6.04	0.004	0.012
NFP Government	-0.48	0.368	0.029	2.12	0.003	0.017
Average Weekly Hours Goods	-9.22	0.001	0.002	-12.41	0.002	0.010
Overtime Weekly Hours Mfg	1.23	0.183	0.477	3.23	0.005	0.007
Average Weekly Hours Mfg	-10.49	0.000	0.000	-12.52	0.001	0.004
Housing Starts	-1.07	0.562	0.633	0.93	0.597	0.565
HS Northeast	-31.57	0.318	0.381	-44.67	0.219	0.198
HS Midwest	52.60	0.326	0.393	23.32	0.286	0.258
HS South	-0.49	0.727	0.932	0.29	0.804	0.269
HS West	-0.73	0.594	0.863	-1.33	0.461	0.674
Building Permits	-2.73	0.317	0.069	1.54	0.361	0.479
BP Northeast	74.85	0.291	0.372	-81.69	0.353	0.353
BP Midwest	1.11	0.470	0.522	2.15	0.140	0.411
BP South	-2.24	0.127	0.227	1.70	0.211	0.118
BP West	-1.46	0.494	0.685	-1.62	0.374	0.247
Orders Consumer Goods	-2.80	0.185	0.014	0.07	0.966	0.529
Orders Durable Goods	11.62	0.266	0.339	-0.02	0.989	0.629
Orders Nondefense Capital Goods	-7.75	0.634	0.386	1.33	0.187	0.414
Unfilled Orders Durable Goods	7.85	0.091	0.117	2.13	0.379	0.089
Business Inventories	1.84	0.079	0.290	5.98	0.004	0.015
Inventories to Sales Ratio	-2.76	0.102	0.022	2.45	0.231	0.518
M1 Money Stock	0.82	0.360	0.451	1.98	0.002	0.011
M2 Money Stock	0.60	0.398	0.299	1.38	0.037	0.137
M2 Real	0.30	0.780	0.829	-0.17	0.877	0.949
Monetary Base	0.93	0.521	0.401	1.97	0.010	0.032

Notes: The table shows average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 1-month and the 3-month ahead forecasts over 2013M08–2020M02. The GW-UC and GW-C columns display the p-values of Giacomini and White (2006)'s test of unconditional and conditional predictive ability respectively. The variables are described as in the FRED-MD database in addition to the excess bond premium and the shadow rate.

Variables	1M	GW-UC	GW-C	3M	GW-UC	GW-C
Total Reserves	-0.29	0.570	0.229	1.31	0.033	0.111
Nonborrowed Reserves	-3.42	0.418	0.662	2.47	0.561	0.109
Business Loans	0.81	0.278	0.484	0.57	0.486	0.525
Real Estate Loans	-0.27	0.682	0.767	0.86	0.129	0.331
Total Nonrevolving Credit	-6.46	0.222	0.444	1.87	0.086	0.124
Credit to Income Ratio	-6.53	0.259	0.293	0.14	0.860	0.484
S&P 500	-3.32	0.235	0.455	0.30	0.885	0.943
S&P Industrials	-3.06	0.249	0.488	-0.17	0.933	0.883
S&P Dividend Yield	-3.62	0.128	0.265	0.51	0.788	0.992
S&P PE Ratio	-2.17	0.148	0.321	3.75	0.065	0.150
Effective FFR	-0.79	0.611	0.185	3.68	0.005	0.011
3M Commercial Paper	-0.74	0.612	0.524	2.26	0.073	0.081
3M T-Bill	-0.84	0.633	0.507	4.06	0.005	0.003
6M T-Bill	-1.02	0.608	0.419	3.61	0.014	0.004
1Y T-Bond	-0.47	0.804	0.281	3.50	0.012	0.005
5Y T-Bond	-0.51	0.688	0.088	0.84	0.484	0.754
10Y T-Bond	-1.11	0.333	0.128	1.32	0.176	0.129
Aaa Corporate Bond Yield	0.14	0.864	0.624	1.32	0.219	0.247
Baa Corporate Bond Yield	0.05	0.958	0.157	0.39	0.726	0.277
CP-FFR Spread	-0.73	0.498	0.036	4.45	0.000	0.000
3M-FFR Spread	-1.79	0.232	0.093	3.95	0.006	0.003
6M-FFR Spread	-1.23	0.427	0.151	5.15	0.000	0.000
1Y-FFR Spread	-0.73	0.683	0.040	5.33	0.000	0.000
5Y-FFR Spread	-4.76	0.023	0.054	0.03	0.984	0.994
10Y-FFR Spread	-7.94	0.003	0.012	-1.25	0.526	0.116
Aaa-FFR Spread	-4.22	0.071	0.132	1.47	0.404	0.097
Baa-FFR Spread	-2.05	0.279	0.098	2.78	0.065	0.202
Trade Weighted Exchange Rate	-0.10	0.927	0.994	-0.13	0.921	0.284
FX Rate CHF	-0.37	0.685	0.810	1.53	0.055	0.149
FX Rate JPY	-0.44	0.791	0.735	-3.69	0.222	0.271
FX Rate GBP	-1.92	0.114	0.155	-0.60	0.533	0.606
FX Rate CAD	0.96	0.641	0.832	-3.74	0.043	0.036
PPI Final Goods	0.10	0.943	0.814	5.76	0.000	0.000
PPI Consumer Goods	0.29	0.837	0.971	5.85	0.000	0.000
PPI Intermediate Material	-0.41	0.740	0.049	5.65	0.000	0.000

Notes: The table shows average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 1-month and the 3-month ahead forecasts over 2013M08–2020M02. The GW-UC and GW-C columns display the p-values of Giacomini and White (2006)'s test of unconditional and conditional predictive ability respectively. The variables are described as in the FRED-MD database in addition to the excess bond premium and the shadow rate.

Variables	1M	GW-UC	GW-C	3M	GW-UC	GW-C
PPI Crude Material	-0.67	0.468	0.709	2.94	0.000	0.002
Crude Oil WTI Price	-0.99	0.456	0.276	1.69	0.035	0.082
PPI Commodities	-0.39	0.567	0.790	0.78	0.212	0.152
CPI All	0.62	0.730	0.150	8.15	0.000	0.000
CPI Apparel	-1.89	0.061	0.247	0.49	0.456	0.342
CPI Transport	0.33	0.857	0.062	8.22	0.000	0.000
CPI Medical	-1.06	0.373	0.493	-1.59	0.138	0.363
CPI Commodities	0.10	0.958	0.263	8.17	0.000	0.000
CPI Durables	-1.04	0.301	0.427	2.06	0.004	0.016
CPI Services	0.58	0.606	0.414	0.67	0.282	0.081
CPI ex Food	0.42	0.819	0.150	8.56	0.000	0.000
CPI ex Shelter	0.39	0.833	0.247	8.34	0.000	0.000
CPI ex Medical	0.47	0.800	0.209	8.30	0.000	0.000
PCE Deflator	0.81	0.629	0.741	7.88	0.000	0.000
PCE Durables	0.83	0.389	0.539	0.71	0.343	0.611
PCE Nondurables	0.01	0.995	0.202	8.26	0.000	0.000
PCE Services	-1.25	0.081	0.204	-0.12	0.847	0.486
Average Earnings Goods	-10.44	0.325	0.413	-0.31	0.657	0.699
Average Earnings Construction	6.82	0.765	0.952	-0.23	0.750	0.570
Average Earnings Manufacturing	0.84	0.281	0.352	1.03	0.139	0.315
Consumer Sentiment	-2.05	0.088	0.138	1.75	0.143	0.260
MZM Money Stock	-0.39	0.641	0.661	2.10	0.005	0.026
CL Motor Vehicles	3.35	0.181	0.401	2.44	0.001	0.003
Consumer Loans	-11.99	0.757	0.375	6.61	0.121	0.181
Securities in Bank Credit	0.61	0.274	0.580	1.45	0.060	0.273
VXO	-5.72	0.000	0.000	-2.25	0.143	0.277
GZ Excess Bond Premium	-3.58	0.010	0.018	3.03	0.013	0.028
Shadow Rate	-0.96	0.475	0.167	0.16	0.901	0.751

Notes: The table shows average difference in predictive log-scores multiplied by 100, relative to the benchmark for the 1-month and the 3-month ahead forecasts over 2013M08–2020M02. The GW-UC and GW-C columns display the p-values of Giacomini and White (2006)'s test of unconditional and conditional predictive ability respectively. The variables are described as in the FRED-MD database in addition to the excess bond premium and the shadow rate.

Appendix C

Appendix to Chapter 4

A. MP Suprises

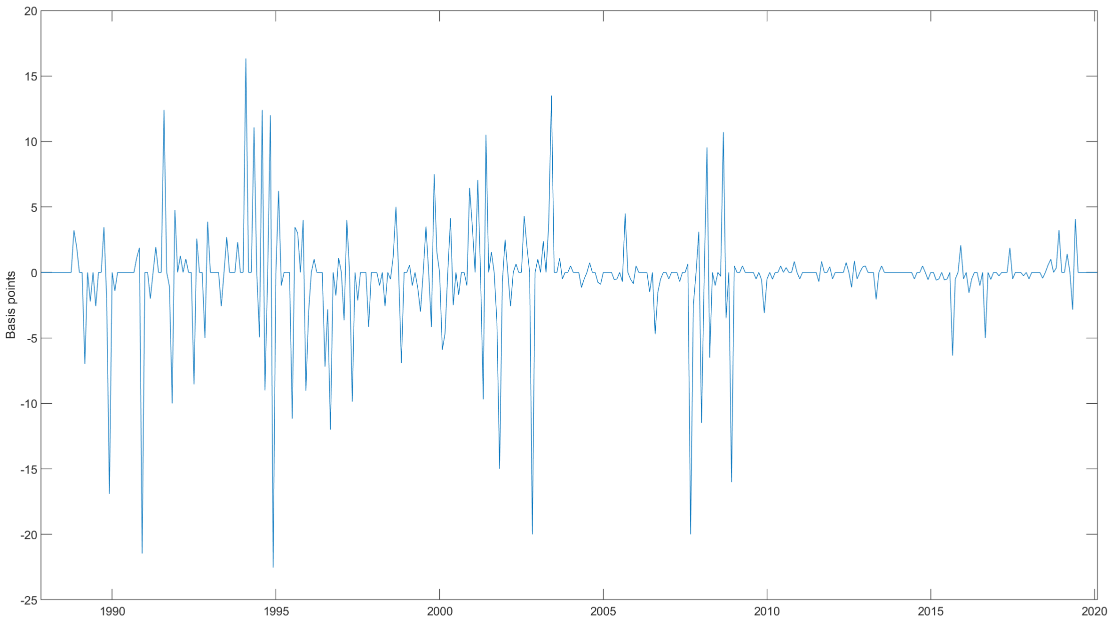

Figure C.1: US Monetary Policy Surprises

B. Conjugate Priors and Posteriors

Given a multivariate normal inverse Wishart (NIW) distribution (conjugate prior) of the form NIW($v_0, \Psi_0, \beta_0, S_0$):

$$\begin{aligned}\Sigma &\sim IW(S_0, v_0) \\ \beta|\Sigma &\sim \mathcal{N}(\beta_0, \Sigma \otimes \Psi_0)\end{aligned}$$

where S_0 is the prior scale matrix, v_0 the degrees of freedom and Ψ_0 is a diagonal matrix with common elements to all equations.

The posterior distribution over the reduced-form parameters is NIW($v_1, \Psi_1, \beta_1, S_1$):

$$\begin{aligned}\Sigma|y &\sim IW(S_1, v_1) \\ \beta|y, \Sigma &\sim \mathcal{N}(\beta_1, \Sigma \otimes \Psi_1)\end{aligned}$$

where, for the general case:

$$\begin{aligned}v_1 &= v_0 + T \\ \Psi_1 &= (X'X + \Psi_0^{-1})^{-1} \\ \beta_1 &= \Psi_1(X'Y + \Psi_0{}^{-1}\beta_0) \\ S_1 &= Y'Y + S_0 + \beta_0'\Psi_0^{-1}\beta_0 - \beta_1'\Psi_1^{-1}\beta_1\end{aligned}$$

and, for the flat (Jeffreys) prior, simply:

$$\begin{aligned}v_1 &= T \\ \Psi_1 &= (X'X)^{-1} \\ \beta_1 &= \Psi_1(X'Y) = \hat{\beta}^{OLS} \\ S_1 &= \hat{S}^{OLS}\end{aligned} \qquad \begin{aligned}v_1 &= T \\ \Psi_1 &= (X'X)^{-1} \\ \beta_1 &= \Psi_1(X'\tilde{Y}) = \hat{\beta}^{GLS} \\ S_1 &= \hat{S}^{GLS}\end{aligned}$$

according to the horizon.

C. Robustness

i. No lags of the proxy

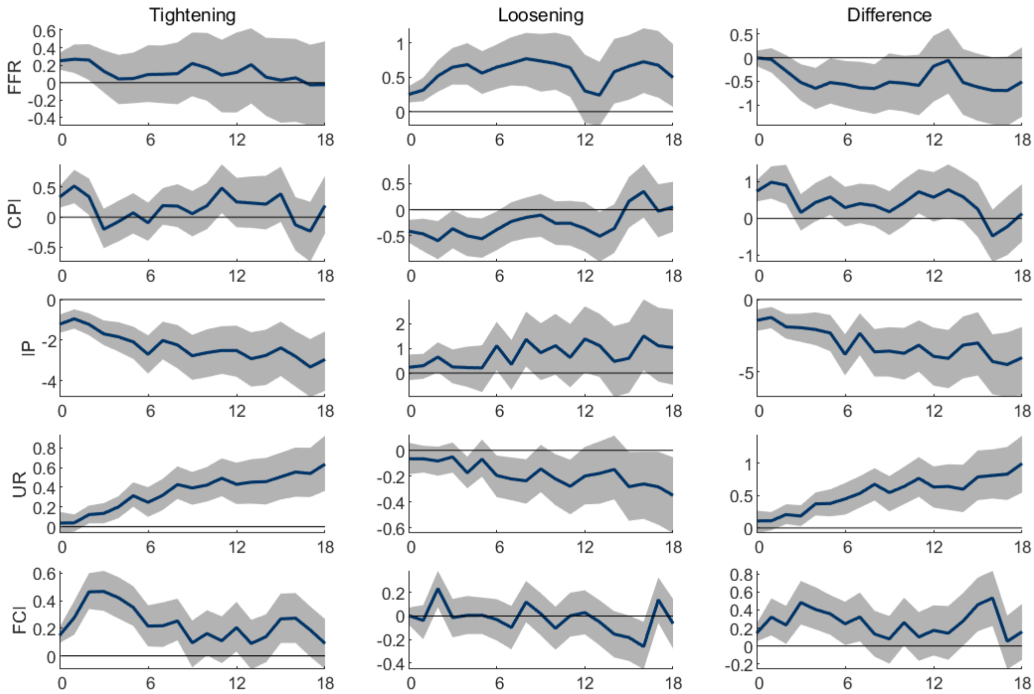

Figure C.2: Non-Linear Impulse Responses: No lags of the proxy

Notes: The shaded area represents the 68 percent HPD credible sets, and the blue line represents the median. The left panel presents impulse responses to positive shocks. The central panel shows the impulse responses to negative shocks. They are flipped to facilitate comparison. The right panel shows their difference. The monetary policy shock has been normalised to have an impact of 25 basis points on the federal funds rate.

ii. Pre-ZLB: 1987M11-2007M12

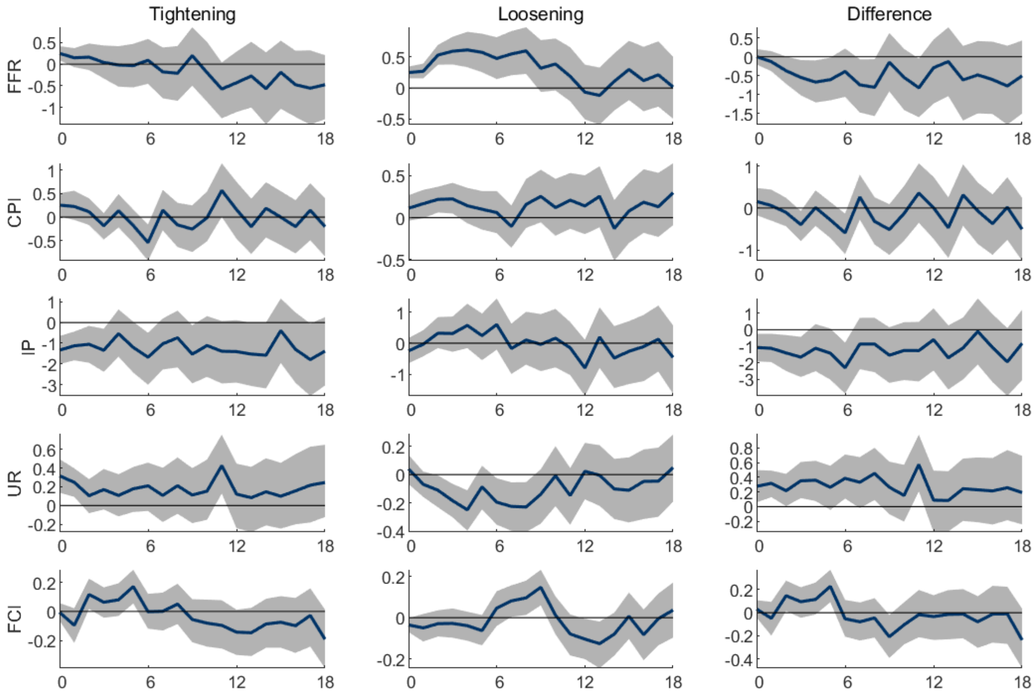

Figure C.3: Non-Linear Impulse Responses: Pre-ZLB

Notes: The shaded area represents the 68 percent HPD credible sets, and the blue line represents the median. The left panel presents impulse responses to positive shocks. The central panel shows the impulse responses to negative shocks. They are flipped to facilitate comparison. The right panel shows their difference. The monetary policy shock has been normalised to have an impact of 25 basis points on the federal funds rate.

iii. Swanson's factors

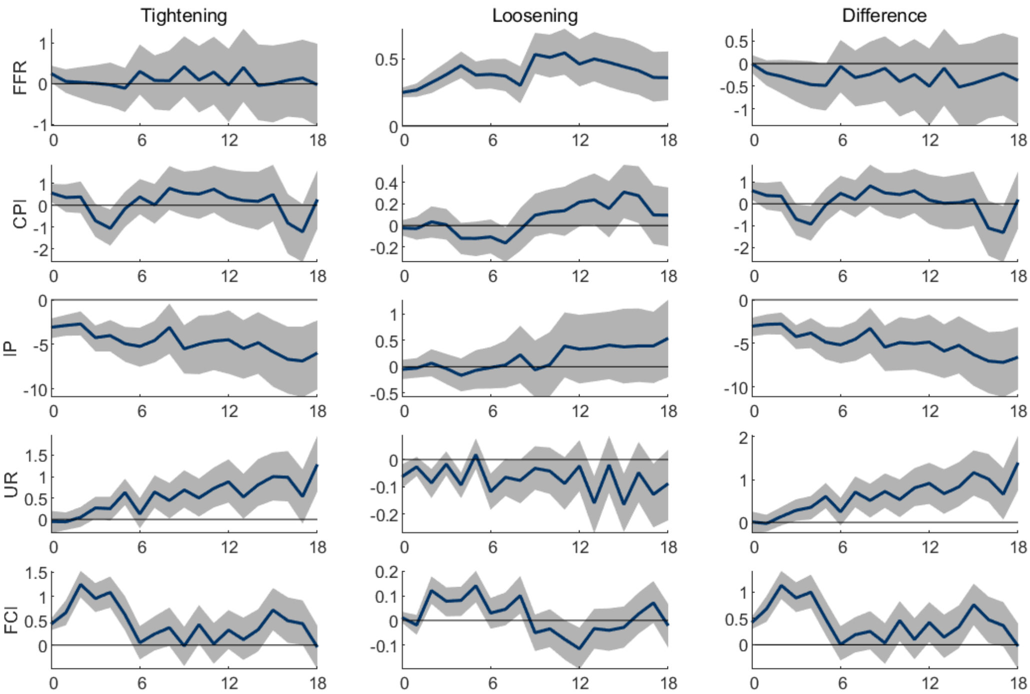

Figure C.4: Non-Linear Impulse Responses: Swanson's factors

Notes: The shaded area represents the 68 percent HPD credible sets, and the blue line represents the median. The left panel presents impulse responses to positive shocks. The central panel shows the impulse responses to negative shocks. They are flipped to facilitate comparison. The right panel shows their difference. The monetary policy shock has been normalised to have an impact of 25 basis points on the federal funds rate.

iv. BAA spread

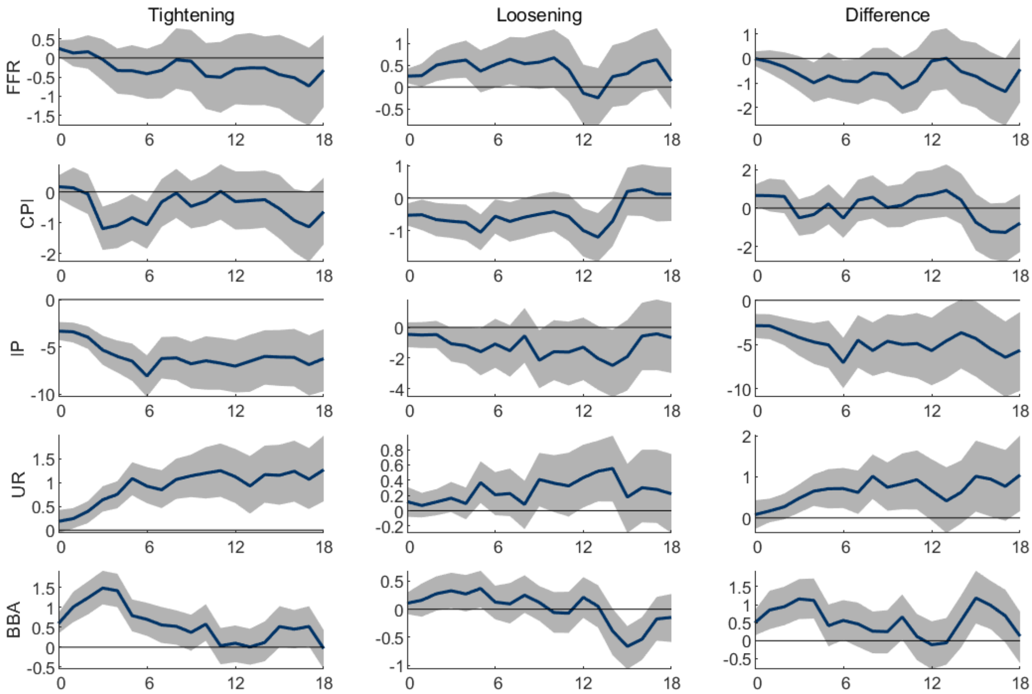

Figure C.5: Non-Linear Impulse Responses: BAA spread

Notes: The shaded area represents the 68 percent HPD credible sets, and the blue line represents the median. The left panel presents impulse responses to positive shocks. The central panel shows the impulse responses to negative shocks. They are flipped to facilitate comparison. The right panel shows their difference. The monetary policy shock has been normalised to have an impact of 25 basis points on the federal funds rate.

v. Excess bond premium

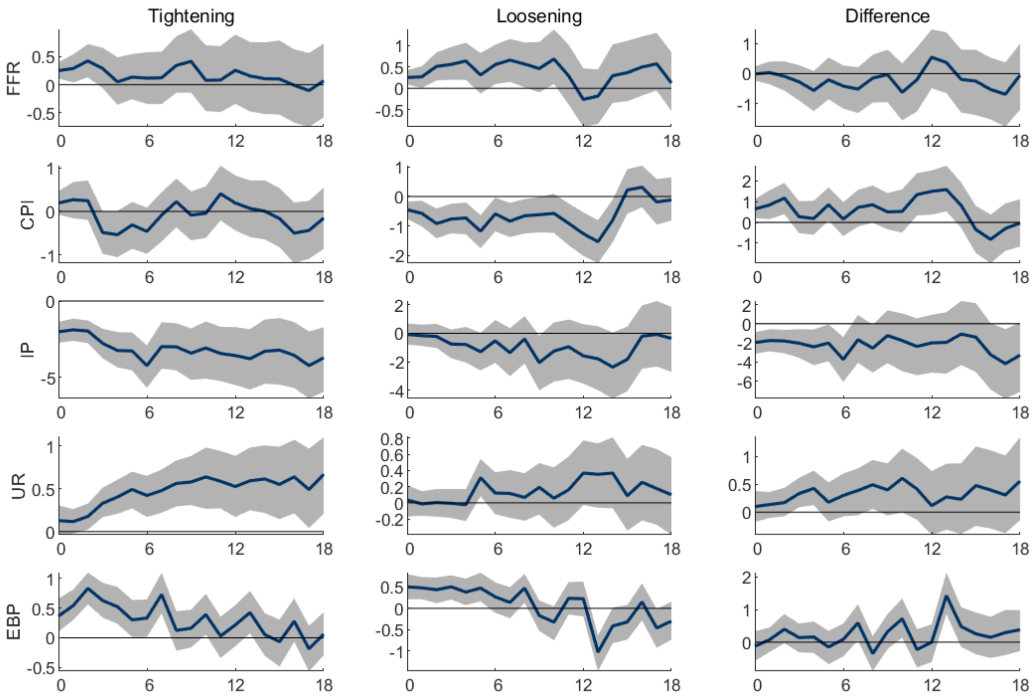

Figure C.6: Non-Linear Impulse Responses: EBP

Notes: The shaded area represents the 68 percent HPD credible sets, and the blue line represents the median. The left panel presents impulse responses to positive shocks. The central panel shows the impulse responses to negative shocks. They are flipped to facilitate comparison. The right panel shows their difference. The monetary policy shock has been normalised to have an impact of 25 basis points on the federal funds rate.

D. Euro area

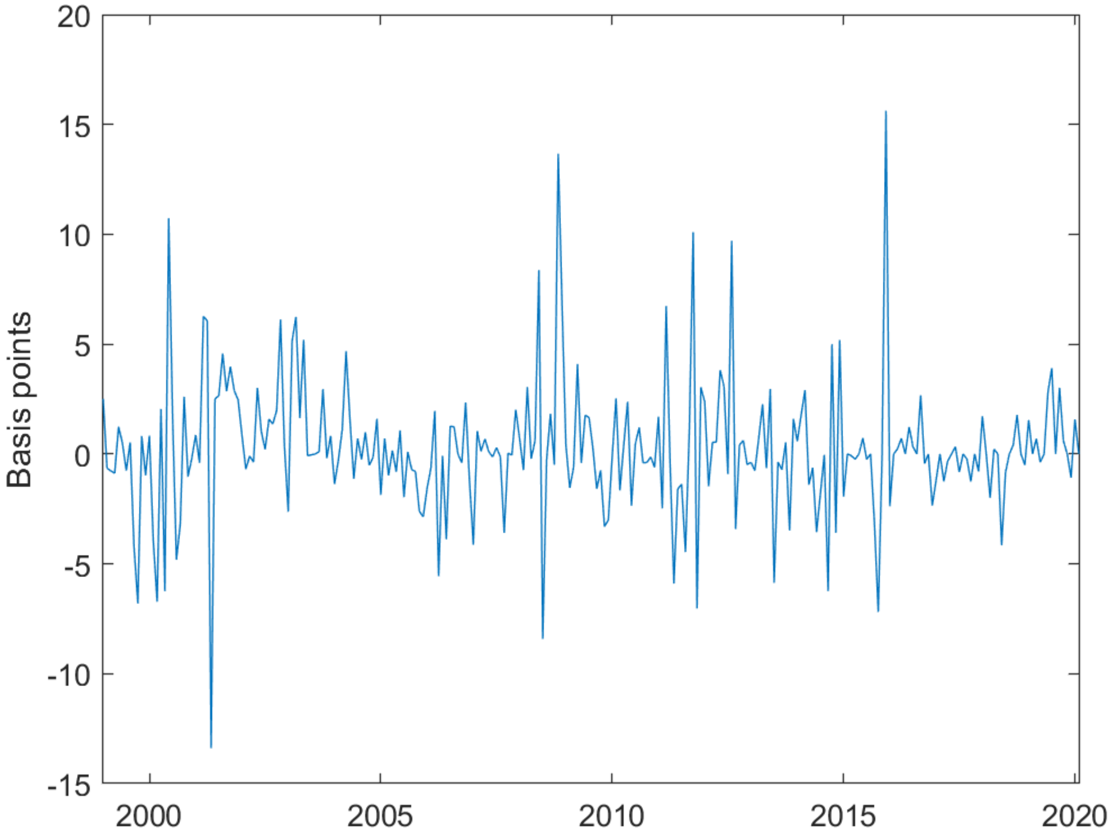

Figure C.7: Updated Jarociński and Karadi (2020)'s MP Shocks