IDENTIFICATION OF THE MONETARY POLICY SHOCK UNDER THE EFFECTIVE LOWER BOUND

Master's thesis

Author:
Samuel Rauhala

Supervisors:
Professor Heikki Kauppi

May 8, 2023
Turku

UNIVERSITY OF TURKU

ABSTRACT

- ☐ Bachelor's thesis
- ☑ Master's thesis
- ☐ Licenciate's thesis
- ☐ Doctor's thesis

Subject	Economics	Date	May 8, 2023
Author(s)	B.Soc.Sc. Samuel Rauhala	Number of pages	53+appendices
Title	Identification of the monetary policy shock under the effective lower bound		
Supervisor(s)	Professor Heikki Kauppi		

Abstract

Monetary policy plays a key role in steering the economy through business cycles fluctuations. Because of this, it is important to be able to estimate the consequences of a surprising or discretionary change in the monetary policy. These are changes in monetary policy that cannot be attributed to changes in the economic environment. However, in the aftermath of the 2007-2009 Great Recession, the emergence of low, zero or even negative interests as well as the unconventional monetary policy measures implemented during that era have made the old methods insufficient.

This thesis provides an overview of both the old insufficient structural vector autoregressive (SVAR) models as well as two new models by Mavroeidis (2021). These two models, kinked and censored vector autoregression, offer possible solutions to the challenges posed by the effective lower bound (ELB) in the interest rates as well as a method for incorporating unconventional monetary policy measures into the model without changing, for example, the identification scheme.

These approaches are compared and applied to euro area data in a simple New Keynesian three-variable model of the economy. The standard and kinked SVAR models suggest that monetary policy is not effective at steering the economy, whereas the censored SVAR suggests the opposite. Furthermore, while the censored model has certain attractive features over the kinked model, the empirical results lead to anomalies, which makes these results less clear. Overall, more research is needed and the conventional wisdom may not provide us with the answers we seek.

Keywords	Effective lower bound, structural vector autoregression, monetary policy
Further information	

UNIVERSITY OF TURKU

TIIVISTELMÄ

- ☐ Kandidaatintutkielma
- ☑ Pro gradu -tutkielma
- ☐ Lisensiaatintutkielma
- ☐ Väitöskirja

Oppiaine	Taloustiede	Päivämäärä	8.5. 2023
Tekijä(t)	VTK Samuel Rauhala	Sivumäärä	53+liitteet
Otsikko	Identification of the monetary policy shock under the effective lower bound (Rahapolitiikkashokin identifikaatio nollakorkorajoitteen alaisuudessa)		
Ohjaaja(t)	Professori Heikki Kauppi		

Tiivistelmä

Rahapolitiikalla on keskeinen rooli suhdannepolitiikassa. Tämän takia on tärkeää, että voidaan arvioida yllättävien tai harkinnanvaraisten rahapolitiikkapäätösten seurauksia. Nämä ovat siis päätöksiä, jotka eivät ole seurauksia taloudellisen ympäristön muutoksista. Vuosien 2007–2009 suuren taantuman myötä useissa teollistuneissa maissa kuitenkin saavutettiin ennätyksellisen matalia ja jopa negatiivisia korkoja. Samaisissa maissa alettiin myös käyttämään epäkonventionaalista rahapolitiikkaa, ja nämä faktat ovat tehneet vanhoista analysointimenetelmistä riittämättömiä.

Tämä tutkielma tarjoaa katsauksen sekä vanhaan riittämättömään rakenteelliseen vektoriautoregressioon (*engl. Structural vector autoregression, SVAR*) että kahteen uuteen Mavroeidiksen (2021) kehittämään malliin. Nämä mallit, mutkistunut (*engl. kinked*) ja sensuroitu (*engl. censored*) vektoriautoregressio, tarjoavat mahdolliset keinot ratkaista nollakorkorajoitteen (*engl. effective lower bound, ELB*) tuomat haasteet ja sisällyttää epäkonventionaalisen rahapolitiikan mahdollisuuden malliin ilman, että esimerkiksi identifikaatio strategiaa tarvitsee muuttaa.

Näitä eri menetelmiä verrataan ja sovelletaan euroalueen aineistoon yksinkertaisesssa uusikeynesiläisessä kolmen muuttujan mallissa. Konventionaalinen ja mutkistunut SVAR-malli viittaavat siihen, ettei rahapolitiikalla ole juuri kykyä vaikuttaa talouden kehityksen, kun taas sensuroidulla SVAR-mallilla tulokset ovat päinvastaiset. Edelleen, vaikka sensuroidulla mallilla on muihin malleihin verratessa puoleensavetäviä ominaisuuksia, empiriassa se tuottaa anomalioita, mikä tekee tuloksista kyseenalaiset. Kokonaisuudessaan lisää tutkimustyötä tarvitaan eivätkä vanhat opit välttämättä ole enää toimivia.

Keywords	Nollakorkorajoite, rakenteellinen vektoriautoregressio, rahapolitiikka
Further information	

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

Contents

List of Figures

List of Tables

Abbreviations

APP	Asset Purchase Program
AR	Autoregression
bp	basis points
CKSVAR	Censored and Kinked Structural Vector Autoregression
CSVAR	Censored Structural Vector Autoregression
ECB	European Central Bank
ELB	Effective Lower Bound
ESS	Effective Sample Size
FG	Forward Guidance
GDP	Gross Domestic Product
GIRF	Generalized Impulse Response Function
HICP	Harmonized Index of Consumer Prices
IRF	Impulse Response Function
IS	Investment–Saving
IV	Instrumental Variable
KSVAR	Kinked Structural Vector Autoregression
ML	Maximum Likelihood
MRO	Main Refinancing Operations
NIRP	Negative Interest Rate policy
NK	New Keynesian
QE	Quantitative easing
SIS	Sequence Importance Sampling
SRTSM	Shadow Rate Term Structure Model
SVAR	Structural Vector Autoregression
TLTRO	Targeted Long-Term Refinancing Operations
VAR	Vector Autoregression
VMA	Vector Moving Average

1 Introduction

Interest rates are the main tools used by central banks to steer the economy. However, the consequences of discretionary interest rate policy changes are by no means obvious to the naked eye. This is because the actual path of an interest rate is determined jointly by systematic reactions to an ever-changing economic environment, reactions to past interest rate policies as well as the actual discretionary changes in the interest rate policy. To estimate the effects of such changes, econometricians have long used structural vector autoregressive models (SVARs), which are flexible and allow for various different types of analysis (see, for example, Christiano et al. 1999; Stock and Watson 2001; Christiano et al. 2010).

Since low interest rates became widespread in industrialized countries in the aftermath of the 2007–2009 Great Recession, the assumptions that these models generally may no longer be valid. Two reasons for this stand out. Firstly, common sense and economic theory both suggest that there exists an effective lower bound (ELB) under which the interest rate cannot fall without it losing much of its meaning. Secondly, monetary policy is no longer carried out solely or even primarily by means of interest rate policy. Instead, various unconventional policy measures have become common in countries that have experienced low interest rates. (See, for example, the survey of Dell'Ariccia et al. 2018.)

The emergence of the ELB and unconventional monetary policy has lead to the development of more advanced methods that take these into account (see the survey of Rossi 2021). On one hand, some authors have constructed shadow rate time series that are, firstly, not bound by the ELB and, secondly, reflect also unconventional monetary policy measures (see, for example, Krippner 2013; Wu and Xia 2016, 2020). On the other hand, some authors have included censored interest rate variables into their models. These accomplish the same goal as shadow rates without relying on any external model of the shadow rates. (See, for example, Iwata and Wu 2006; Mavroeidis 2021; Aruoba et al. 2022.; Ikeda et al. 2022.) In addition, there are some more unique approaches (see, for example, Wright 2012; Getler and Karadi 2015).

The main focus in this thesis is on the two approaches by Mavroeidis (2021) that aim to solve the issue of the ELB but are built on very different assumptions. These represent the simpler and more straightforward end of the advanced approaches. The kinked SVAR (KSVAR) model merely introduces an ELB as a lower bound under which the interest rate may not fall. It does not however, take unconventional monetary policy into account. This is built around the idea that the central bank faces an asymmetric problem: It can raise the interest rate as much as it wants but it cannot drop it indefinitely. Meanwhile, the censored SVAR (CSVAR) model also takes into account unconventional monetary policy measures by estimating a shadow rate, i.e. an interest rate that is not generally observable but may fall under the ELB and reflects the true state of monetary policy. Unconventional means of conducting monetary policy make the central bank's problem symmetric. These models are used to analyze the monetary policy shock in euro area.

Before the advanced KSVAR and CSVAR can be analyzed, the standard SVAR setup is presented together with a simple New Keynesian (NK) model of the economy, which motivates the model structure. The identification scheme used in the SVAR models is traditional lower-triangular or short-term identification. This means that a discretionary monetary policy shock is identified as a shock that cannot be explained by the developments of key macroeconomic variables.

This thesis has two novel features. Firstly, all empirical considerations are carried out with euro area data, which has received far less attention in this type of research than the USA or Japan. This thesis also offers the reader context for understanding unconventional monetary policy in euro area. Secondly, this thesis uses the KSVAR and the CSVAR models to come up with concrete shadow rate time series together with confidence intervals that can be compared to other shadow rate time series, such as the one by Wu and Xia (2016).

The empirical results show that not only does the inclusion of the ELB change the results but also the way in which it is included also matters. Overall, the standard SVAR and the KSVAR imply that the monetary policy shock plays a very minor role in the economy. Meanwhile, the CSVAR reserves an important role for the monetary policy shock and also results in the price puzzle, a well-documented anomaly in the analysis of monetary policy. Indeed, CSVAR appears as the more attractive of the two advanced methods, but the anomalous results are troubling. The overall results are also reflected in the shadow rate series; unconventional monetary policy does not appear to have had much effect on the KSVAR shadow rate but clearly moves the CSVAR shadow rate. The policy implications of these results are that the existence of the ELB increases the uncertainty related to the outcomes of monetary policy and that more work is required. Indeed, one should exercise caution when applying the conventional wisdom in today's world.

The rest of this thesis is constructed as follows. Section 2 presents the standard SVAR framework that forms the backbone of this thesis but also largely of modern macroeconometrics. This is followed by Section 3 that introduces the conventional set up that ignores the ELB. This section takes a more practical stance by looking at the concrete problem of monetary policy shock identification and structural analysis. Section 4 takes a closer look at what the ELB means and also what type of unconventional monetary policy measures there are and how these have been used in euro area. Sections 5 and 6 present the KSVAR and the CSVAR models of Mavroeidis (2021) respectively. Sections 3, 5 and 6 end with empirical analyses of different models. Finally, Section 7 concludes.

2 The vector autoregression and structural vector autoregression toolkits

2.1 Vector autoregression

The workhorse of this thesis is the vector autoregression (VAR) and structural vector autoregression (SVAR) framework. This section provides a brief introduction to this vast topic.

I denote the vector of endogenous variables by $\mathbf{x}_t = (x_{1t}, x_{2t}, ..., x_{Kt})' \in \mathbb{R}^K$, where $t \in \mathbb{Z}$ is the period of observation. This process is said to follow a VAR(p) process, if

$$\mathbf{x}_t = \nu + A_1\mathbf{x}_{t-1} + A_2\mathbf{x}_{t-2} + ... + A_p\mathbf{x}_{t-p} + \varepsilon_t, \tag{1}$$

where $\nu \in \mathbb{R}^K$ and $A_j \in \mathbb{R}^{K\times K}$ for $\forall j = 1,2,...,p$ are parameters. The error term $\varepsilon_t \in \mathbb{R}^K$ is both independent and identically distributed and it has a zero mean. Within this thesis I'll further assume that ε_t follows a multivariate normal distribution $N_K(\mathbf{0},\Sigma)$, which I denote by $\varepsilon_t \sim n.i.d_K(\mathbf{0},\Sigma)$. The matrix Σ is a symmetric and positive definite matrix of dimension $K \times K$. Under these assumptions, ε_t is also called Gaussian white noise or innovation process. (Lütkepohl 2006, 13, 16; Kilian and Lütkepohl 2017, 24–25.)

In layman's terms, VAR(p) process of Equation (1) states that future realizations of a vector of endogenous variables $\mathbf{x}_t$ are linear combinations of the last p realizations of the same process with an intercept ν and an error term ε_t added. The number of lags p is called the order of the process.

It is desirable that the VAR(p) process is stable. This means that the expected value and autocovariance are constant over all values of t:

$$\mathbb{E}[\mathbf{x}_t] = \mu \text{ and } \mathbb{E}[(\mathbf{x}_t - \mu)(\mathbf{x}_{t+h} - \mu)'] = \Gamma(\pm h) \text{ for } \forall t \in \mathbb{Z}.$$

In other words, the expected value and the autocovariance function are constant over all values of t. A necessary and sufficient condition for this is

$$\det(I_K - A_1 z - A_2 z^2 - ... - A_p z^p) = 0 \quad \text{for } |z| \leq 1.$$

(Lütkepohl 2006, 14-16, 25; Kilian and Lütkepohl 2017, 25–26.)

An important consequence of the stability is that the VAR(p) process can be presented as a vector moving average process. More specifically it can be expressed as a so called VMA(∞) process such that

$$\mathbf{x}_t = \mu + \sum_{i=0}^{\infty} \Phi_i \varepsilon_{t-i} \tag{2}$$

For practical purposes, Φ_i can be constructed by

$$\Phi_0 = I_K \quad \text{and} \quad \Phi_i = \sum_{j=1}^{i} \Phi_{i-j} A_j, \quad i = 1, 2, ...$$

(Lütkepohl 2006, 18-23; Kilian and Lütkepohl 2017, 26.)

VAR(p) process can be predicted recursively. I denote by $\mathbb{E}_t[\mathbf{x}_{t+h}]$ the expected value of $\mathbf{x}_{t+h}$ where $h > 0$, conditional on the information available at period t, i.e. $\{\mathbf{x}_t, \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, ...\}$. The unbiased and minimum squared error prediction for $\mathbf{x}_{t+h}$ at period t is

$$\mathbb{E}_t[\mathbf{x}_{t+h}] = \nu + A_1 \mathbb{E}_t[\mathbf{x}_{t+h-1}] + A_2 \mathbb{E}_t[\mathbf{x}_{t+h-2}] + ... + A_p \mathbb{E}_t[\mathbf{x}_{t+h-p}],$$

where $\mathbb{E}_t[\mathbf{x}_{t+i}] = \mathbf{x}_{t+i}$ for $i \leq 0$. (Lütkepohl 2006, 33–35; Kilian and Lütkepohl 2017, 47–48.)

Naturally, the true model is not known to the econometrician and it has to be estimated. Let there be a sample $\mathbf{x}_{1-p}, \mathbf{x}_{1-p+1}, \mathbf{x}_{1-p+2}, ..., \mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_T$ to which the econometrician wants to fit a VAR(p) model. The first p observations are needed as initial values to fit the VAR(p) model. One characteristic feature of these models is that the number of parameters is high, namely $K + pK^2 + K(K+1)/2$.

A conventional VAR model can be fitted with relative ease using multivariate least squares (see for example Lütkepohl 2006, 69–77; Kilian and Lütkepohl 2017, 31–32). However, maximum likelihood (ML) estimation will prove more useful in later sections. This method hinges on the assumption that $\varepsilon_t \sim n.i.d_K(\mathbf{0}, \Sigma)$. In other words, the probability density of the error term in each period is given by multivariate normal distribution

$$f_\varepsilon(\varepsilon_t) = \frac{1}{(2\pi)^{(K/2)}} \det(\Sigma)^{-1} \exp\left(\frac{-1}{2}(\varepsilon_t' \Sigma^{-1} \varepsilon_t)\right).$$

Since $\mathbf{x}_t$ is given by the last p observations, the parameters $A_1, A_2, ..., A_p$ and ν and the error term ε_t as is shown in Equation (1), the probability density of x_t conditional on $\mathbf{X}_{t-1}$ and these parameters can be written as

$$f_x(\mathbf{x}_t | \mathbf{X}_{t-1}) = f_u(\mathbf{x}_t - \nu - A_1 \mathbf{x}_{t-1} - A_2 \mathbf{x}_{t-2} - ... - A_p \mathbf{x}_{t-p}),$$

where θ contains all parameters. This can then be turned to the probability density function of the whole sample given the initial values:

$$f_X(\mathbf{x}_1, ..., \mathbf{x}_T | \theta, \mathbf{x}_{1-p}, ..., \mathbf{x}_0) = \prod_{t=1}^{T} f_x(\mathbf{x}_t | \mathbf{X}_{t-1}).$$

The log-likelihood function is then

$$l(\theta|\mathbf{x}_{1-p},...,\mathbf{x}_T) = \sum_{t=1}^{T} \log(f_x(\mathbf{x}_t|\mathbf{X}_{t-1})) = -\frac{KT}{2} log(2\pi) - \frac{T}{2}\log(\det(\Sigma)) \tag{3}$$
$$-\frac{1}{2}\sum_{t=1}^{T}(\mathbf{x}_t - \nu - A_1\mathbf{x}_{t-1} - ... - A_p\mathbf{x}_{t-p})'\Sigma^{-1}(\mathbf{x}_t - \nu - A_1\mathbf{x}_{t-1} - ... - A_p\mathbf{x}_{t-p}).$$

The vector $\hat{\theta}$ that maximizes the log-likelihood function is the ML estimator. (Kilian and Lütkepohl 2017, 38–40.)

The ML and least squares estimates are equivalent. They can be expressed in analytical form. Start by defining the following objects:

$$B = (\nu, A_1, A_2, ..., A_p), \quad \beta = \text{vec}(B), \quad Z_t = \begin{bmatrix} 1 \\ \mathbf{X}_t \end{bmatrix}, \quad Z = (Z_0, Z_1, ..., Z_{T-1}),$$

$$X = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_T \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \text{vec}(X),$$

where vec(X) operator stacks the columns of X on top of each other. Now the estimator for β is

$$\hat{\beta} = ((ZZ')^{-1}Z \otimes I_K)\mathbf{x},$$

where $\otimes$ is the Kronecker product. Furthermore, the estimator for the covariance matrix Σ adjusted for degrees of freedom is

$$\hat{\Sigma} = \frac{1}{T - Kp - 1}\sum_{t=1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_t',$$

where $\hat{\varepsilon}_t$ are residuals $\mathbf{x}_t - \hat{\nu} - \hat{A}_1\mathbf{x}_{t-1} - ... - \hat{A}_p\mathbf{x}_{t-p}$. (Lütkepohl 2006, 70–75, 90.) These estimators are unbiased and consistent. Furthermore, they are normally distributed allowing for conventional hypothesis testing. (Kilian and Lütkepohl 2017, 38–40.)

Besides the parameters, the order p is usually unknown. It can be estimated for example by sequential testing, by using information criteria or by comparing forecast performances. (Kilian and Lütkepohl 2017, 51–66.) However, in the econometric literature it is not uncommon to choose the order a priori. For example, Stock and Watson (2001), Mavroeidis (2021) and Aruoba et al. (2022) choose $p = 4$. They work with quarterly data, which means that the oldest relevant observations are from a year ago. This can also be seen as a way to deal with the problem of seasonality. Because my goal is to replicate their studies, I will proceed with this assumption.

2.2 Structural vector autoregression

2.2.1 Identification

The VAR models by themselves can be used for prediction as well as some analysis of the multivariate time series (see for example the discussion on Granger causality in Kilian and Lütkepohl 2017, 48–50). However, in econometrics as well as in this thesis the goal is often to analyze so called structural shocks. These shocks are mutually uncorrelated and they have an interpretation in economic theory (Kilian and Lütkepohl 2017, 109). In particular, the interest of this thesis is in the monetary policy shock, which is an unexpected change in interest rate. In essence, the goal is to identify systematic and foreseeable interest rate changes and unexpected discretionary ones. In this section, I omit the intercept ν as it wont play a role in the subsequent methods. One can assume that the data is demeaned in this section.

The goal is to analyze the following modification of the standard reduced-form VAR model of Equation (1):

$$B_0\mathbf{x}_t = B_1\mathbf{x}_{t-1} + B_2\mathbf{x}_{t-2} + \dots + B_p\mathbf{x}_{t-p} + \mathbf{w}_t, \tag{4}$$

where $\mathbf{w}_t$ are structural shocks. Most importantly, $\mathbb{E}[\mathbf{w}_t\mathbf{w}_t'] = I_K$, which means that the components of $\mathbf{w}_t$ are mutually uncorrelated and have a unit variance. Note that the number of structural shocks matches the dimension of the error term. By rearranging the Equation (4), the reduced form Equation (1) can be derived:

$$\mathbf{x}_t = \underbrace{B_0^{-1}B_1}_{=A_1}\mathbf{x}_{t-1} + \underbrace{B_0^{-1}B_2}_{=A_2}\mathbf{x}_{t-2} + \dots + \underbrace{B_0^{-1}B_p}_{=A_p}\mathbf{x}_{t-p} + \underbrace{B_0^{-1}\mathbf{w}_t}_{=\varepsilon_t}.$$

In order for $\mathbf{w}_t$ to be identifiable from ε_t, B_0 has to be non-singular. This is because of $\varepsilon_t = B_0^{-1}\mathbf{w}_t$. Furthermore, $\Sigma = \mathrm{Cov}(\varepsilon_t) = \mathrm{Cov}(B_0^{-1}\mathbf{w}_t) = B_0^{-1}B_0^{-1\prime}$. (Kilian and Lütkepohl 2017, 109.)

The choosing of B_0 forms a large part of SVAR analysis. Here I'll restrict the discussion to so called B-models (see Lütkepohl 2006, 358–362, 364–367 for discussion on A- and AB-models). $\Sigma = B_0^{-1}B_0^{-1\prime}$ forms a system of equations consisting of $K(K+1)/2$ unique equations because Σ and $B_0^{-1}B_0^{-1\prime}$ are symmetric. However, B_0 has total of K^2 free variables meaning that there isn't a unique solution. In order for such to exist, $K(K-1)/2$ restrictions have to be put on B_0^{-1}. More can be added, but this results in an over-identified model, which puts restrictions on Σ, i.e. on the reduced form model. These restrictions could then further be tested. However, there cannot be fewer than $K(K-1)/2$ and each block of B_0^{-1} has to be identifiable as well. For example, block diagonal B_0^{-1} wouldn't work even if there were enough restrictions since the blocks themselves couldn't be identified. An example of a possible identification strategy, which will be studied in further detail below, is to restrict B_0^{-1} to a lower triangular matrix. This means that B_0^{-1} can be derived from a Choleski decomposition of Σ. (Lütkepohl 2006, 362–363.)

2.2.2 Impulse responses

Once structural shocks have been identified, methods of analyzing these are needed. A common approach is to study the impulse response functions (IRF). These are the responses of different components of $\mathbf{x}_{t+h}$ at different horizons h to different one-time impulses in $\mathbf{w}_t$:

$$\Theta_h = \frac{\partial \mathbf{x}_{t+h}}{\partial \mathbf{w}_t'},\ i = 0, 1, 2, ..., \tag{5}$$

where $\Theta_h \in \mathbb{R}^{K \times K}$. The elements of Θ_h are

$$\theta_{jk,h} = \frac{\partial x_{j,t+h}}{\partial w_{kt}},$$

so that $\Theta_h = [\theta_{jk,h}]$. Generally, the point of interest is the curve $\theta_{jk,h}$ over different values of h and not any single $\theta_{jk,h}$. (Kilian and Lütkepohl 2017, 110–111.)

For the sake of this thesis, the IRF is a change in the development of macroeconomic time series induced by a monetary policy shock or simply put the consequences of a discretionary change in monetary policy.

To construct impulse responses in the standard SVAR model, take the VMA(∞) presentation of Equation (2) and substitute $B_0^{-1}\mathbf{w}_t$ for ε_t:

$$\mathbf{x}_t = \mu + \sum_{i=0}^{\infty} \Phi_h B_0^{-1} \mathbf{w}_{t-h}, \tag{6}$$

which implies that

$$\Theta_h = \frac{\partial \mathbf{x}_{t+h}}{\partial \mathbf{w}_t'} = \Phi_h B_0^{-1}. \tag{7}$$

Note that $\Theta_0 = B_0^{-1}$. (Kilian and Lütkepohl 2017, 111–112.)

The Θ_h of Equation (7) is independent of history. That is, it does not matter under which initial conditions the shock $\mathbf{w}_t$ impacts, it will always lead to the same impulse responses. Furthermore, the actual impulse response is linearly dependent on the shock so that only the scale and not the shape of IRF depends on size of the $\mathbf{x}_t$. As stressed by Koop et al. (1996), however, impulse responses of nonlinear models often depend on the past values of $\mathbf{x}_t$. In addition, the shape of the impulse responses may also depend on the size of the shock $\mathbf{w}_t$. Instead, they offer generalized impulse response functions (GIRF) to be used with more complex models:

$$\text{GIRF}(h, \mathbf{w}_t, \mathbf{X}_t) = \mathbb{E}_t[\mathbf{x}_{t+h}|\mathbf{x}_t + \mathbf{w}_t] - \mathbb{E}_t[\mathbf{x}_{t+h}|\mathbf{x}_t], \tag{8}$$

where $\mathbf{X}_t$ contains the past and current values of $\mathbf{x}_t$ and $\mathbb{E}_t$ is the expected value operator conditional on the information available at period t, i.e. $\mathbf{x}_t$. Under the standard SVAR model considered in this section, $\text{GIRF}(h, \mathbf{w}_t, \mathbf{X}_t) = \Theta_h \mathbf{w}_t$.

In practice, Θ_h is unknown and has to be estimated using the estimators described in Section 2.1. It can then be shown that $\sqrt{T}\text{vec}(\hat{\Theta}_h - \Theta_h) \xrightarrow{d} N_{K^2}(\mathbf{0}, \Sigma_{\hat{\Theta}_h})$, where $\xrightarrow{d}$ denotes convergence in distribution. In other words, $\hat{\Theta}_h$ is consistent, unbiased and asymptotically normal. Although normality could be used to construct the confidence intervals for $\hat{\Theta}_h$, in practice, bootstrap is the most common method for this. (Lütkepohl 2006, 377.) The standard non-parametric bootstrap algorithm for constructing the bootstrap sample for confidence intervals is:

1) Fit a VAR model to the whole sample and store the parameter estimates $\hat{\nu}, \hat{A}_1, ..., \hat{A}_p$ and the residuals $\hat{\varepsilon}_t$ for $t = 1, 2, ..., T$. Center the residuals if necessary.
2) Randomly draw with replacement T bootstrap residuals ε_t^b from $\hat{\varepsilon}_t$.
3) Generate bootstrap time series recursively as $\mathbf{x}_t^b = \hat{\nu} + \hat{A}_1 \mathbf{x}_{t-1}^b + ... + \hat{A}_p \mathbf{x}_{t-p}^b + \varepsilon_t^b$. This can be initialized by setting $\mathbf{x}_{1-p}^b = \mathbf{x}_{1-p}$, $\mathbf{x}_{1-p+1}^b = \mathbf{x}_{1-p+1}$,..., $\mathbf{x}_0^b = \mathbf{x}_0$.
4) Estimate parameters $\hat{\nu}^b, \hat{A}_1^b, ..., \hat{A}_p^b$ based on the bootstrap time series $\mathbf{x}_t^b$.
5) Construct bootstrap impulse responses $\hat{\Theta}_h^b$ using $\hat{\nu}^b, \hat{A}_1^b, ..., \hat{A}_p^b$.
6) Repeat steps 2–5 many times storing $\hat{\Theta}_h^b$.

There is a number of ways that confidence intervals can be derived from the bootstrap sample of $\hat{\Theta}_h^b$. The most straightforward method is to directly construct the standard percentile intervals. Simply put, the $100\% - \gamma$ interval for $\theta_{jk,h}$ is approximately $[\theta_{jk,h,\gamma/2}^b, \theta_{jk,h,1-\gamma/2}^b]$, where $\theta_{jk,h,\gamma}^b$ is the γ-quantile of $\theta_{jk,h}^b$. (Lütkepohl 2006, 709–711.) The same approach can be used for $\text{GIRF}(h, \mathbf{w}_t, \mathbf{X}_t)$. Occasionally, it is necessary to use parametric bootstrap. This is accomplished by sampling ε_t^b from a centered multivariate normal distribution using the estimate for $\hat{\Sigma}$ instead of sampling them straight from the residuals.

3 Monetary policy shock in the absence of the ELB

In the broadest possible sense, the monetary policy shock is an exogenous shock to a monetary policy instrument or to a proxy of the central bank's monetary policy stance Ins_t. If $\mathbf{X}_t$ is the information set available to the central bank at time t and $f(\cdot)$ is the reaction function or feedback rule of the central bank to this information, the monetary policy shock m_t is

$$m_t = Ins_t - f(\mathbf{X}_t). \tag{9}$$

(Christiano et al. 1999.) For the purposes of this thesis, Ins_t is an interest rate that the central bank is able to control.

In this section, I will propose a linear candidate function $f(\cdot)$ and a model which describes how the monetary policy shocks, m_t, of Equation (9) propagate through the economy and analyze the effects of m_t on the economy, when the ELB is ignored.

3.1 A Simple New Keynesian model

In a conventional non-zero interest rate environment, the monetary policy shock can be identified in a simple New Keynesian (NK) three–equation model. The three building blocks of this model are the forward–looking Phillips curve, forward–looking Investment–Saving (IS) relation and a Taylor rule. While this model is extremely simple, it has proven popular in the literature. The model is easy to modify and it is also simple enough for the econometric analysis in the later sections.

To start off, the forward–looking Phillips curve simply relates inflation of the current period, π_t, to the output gap y_t and the expected future inflation $\mathbb{E}_t[\pi_{t+1}]$:

$$\pi_t = \kappa y_t + \beta \mathbb{E}_t[\pi_{t+1}] + u_t, \tag{10}$$

where $\kappa > 0$ and $1 > \beta > 0$ are parameters and u_t is an exogenous disturbance that is sometimes called a "cost-push shock". Unlike the old Keynesian non-forward looking Phillips curve, the curve of Equation (10) can in fact be derived from a lower level model with firms that try to select optimal prices under price frictions. In other words, it has micro-foundations. (Woodford 2010.)

The second element of the NK model is the forward–looking or inter–temporal IS relation. It relates the current output gap to expected future output gap and the current real interest rate, r_t. Using the Fisher equation, the real interest rate is the nominal interest rate, i_t, minus the expected future inflation. The relation is

$$y_t = \mathbb{E}_t[y_{t+1}] - \sigma(i_t - \mathbb{E}_t[\pi_{t+1}] - \rho_t), \tag{11}$$

where $\sigma > 0$ is a parameter and ρ_t is an exogenous disturbance that can be given the role of a natural rate of interest (Jung et al. 2005) or a demand preference shock (Ikeda et al. 2022). Much like the forward looking Phillips curve of equation (10), the forward–looking IS relation too has micro foundations. The IS relation can be derived from the Euler equation: Households try to even out their consumption over time and the real interest rate is the cost of doing so. (Woodford 2010.)

The shocks u_t and ρ_t can be given autoregressive AR(1) dynamics so that

$$u_t = \phi_u u_{t-1} + w_{ut} \quad \text{and} \quad \rho_t = \phi_\rho \rho_{t-1} + w_{\rho t}, \tag{12}$$

where $w_{ut} \sim n.i.d(0, \sigma_u)$ and $w_{\rho t} \sim n.i.d(0, \sigma_\rho)$. Stationarity also requires that $|\phi_u|, |\phi_\rho| < 1$.

The third equation is the Taylor rule. First proposed by Taylor (1993, see also 1999), the Taylor rule proposes that under optimal monetary policy, interest rate is set according to the rule

$$i_t = \pi_t + 0.5y_t + 0.5(\pi_t - 2\%) + 2\%.$$

However, for the econometrician, there is no reason to fixate on these exact numerical values for the coefficients. Furthermore, the Taylor rule can be derived using the forward looking Phillips curve and the forward looking IS relation together with some form of welfare function for households. The simplest choice is to consider a quadratic loss function:

$$\mathbb{E}_t[\Sigma_{j=0}^{\infty} \beta^j [(\pi_{t+j} - \pi^*)^2 + \omega y_{t+j}^2]], \tag{13}$$

where β is the discount factor, π^* is the inflation target and $\omega \geq 0$ is the relative weight of the output gap target. Note that there may also be a non-zero output gap target, and quite often the inflation rate target is zero. The appearance of β in Equation (10) is also worth noting. This results in more general versions of the Taylor rule. (Clarida et al. 1999; Woodford 2002.) The one I choose bears a close resemblance to the original rule:

$$i_t^{Taylor} = a_1 \pi_t + a_2 y_t, \tag{14}$$

where i_t^{Taylor} is the optimal or the Taylor rule implied interest rate.

The loss function of Equation (13) can be justified on one-hand by the mandates of central banks to maintain price stability, low unemployment and sturdy economic growth. However, it can also be justified using the household's utility function. In the latter case, the idea in a nutshell is that a higher level of production leads to a higher level of consumption, which increases utility, but also working hours, which reduces utility. This trade–off can be seen as making the closing of the output gap desirable. Furthermore, under sticky prices, inflation leads to non-optimal pricing, which leads to inefficiencies and thus costs.

Whilst these mechanics are very complex, under a very simple model, the utility function can be reasonably well approximated by a second–order Taylor approximation, where the first–order terms disappear. This allows the expression of the problem in the form of the quadratic loss function of Equation (13). (Woodford 2002.) Nevertheless, one should note that this is not the end of the story as more complex models can lead to more complex loss functions with additional terms such as "net worth gap" or "credit spread gap" (see for example Cúrdia and Woodford 2016; Hansen 2018). Additionally, using third–order approximation instead of second–order leads to asymmetric loss function (see Benigno and Rossi 2021). Finally, the ECB is officially only interested in price stability and not output stability. This goal also was not symmetric until recently (Lane 2022).

Of course, it is clear that the Taylor rule of Equation (14) is a nowhere near sophisticated enough policy rule that a central bank in the real world could follow it (Clarida et al. 1999). Nevertheless, it can be seen as a simple tool to identify discretionary changes and responses to more complex phenomena in monetary policy (Christiano et al. 2010). One of the simplest extensions to the model is to allow smoothing of interest rates by the central banks and allowing for discretionary monetary policy decisions. The smoothing may be motivated by the goal of keeping interest rates steady or by taking into account the uncertainty under which central banks make their decisions. Another addition is discretionary monetary policy shock m_t, which accounts for deviations from the rule. The smoothed Taylor rule with discretionary policy is

$$i_t = \lambda i_t^{Taylor} + (1-\lambda) i_{t-1} + m_t, \tag{15}$$

where $\lambda \in [0,1]$ (Ikeda et al. 2022). This expression can easily be expanded to include more lags of i_t or even i_t^{Taylor} to capture the frictions on monetary policy decision making.

3.2 The NK model in SVAR framework

The goal in this section is to construct a SVAR model to identify the monetary policy shock m_t and see how it propagates through the economy. Note that adding m_t to the model outlined in Section 3.1 allows for the identification of the shocks since this leads to three shocks in three equations.

The SVAR model can be derived by establishing a state space presentation of the three equation model using the method of unknown coefficients. Start by rewriting Equations (10) and (11) in terms of i_{t-1}, u_t, ρ_t and m_t:

$$\pi_t = b_{\pi i} i_{t-1} + b_{\pi u} u_t + b_{\pi \rho} \rho_t + b_{\pi m} m_t, \tag{16}$$

$$y_t = b_{yi} i_{t-1} + b_{yu} u_t + b_{y\rho} \rho_t + b_{ym} m_t. \tag{17}$$

Now the Taylor rule of Equation (15) can be rewritten:

$$\begin{aligned} i_t =& \lambda(a_1\pi_t + a_2 y_t) + (1-\lambda)i_{t-1} + m_t \\ =& (1-\lambda+\lambda a_1 b_{\pi i} + \lambda a_2 b_{yi})i_{t-1} + \lambda(a_1 b_{\pi u} + a_2 b_{yu})u_t + \lambda(a_1 b_{\pi\rho} + a_2 b_{y\rho})\rho_t \\ &+ (1+\lambda(a_1 b_{\pi m} + a_2 b_{ym}))m_t \\ =& b_{ii} i_{t-1} + b_{iu} u_t + b_{i\rho}\rho_t + b_{im} m_t. \end{aligned} \tag{18}$$

(Ikeda et al. 2022.)

Using matrix notation, the same can be presented as

$$\begin{aligned} \mathbf{x}_t = \begin{bmatrix} \pi_t \\ y_t \\ i_t \end{bmatrix} &= \begin{bmatrix} b_{\pi i} & b_{\pi u} & b_{\pi\rho} & b_{\pi m} \\ b_{yi} & b_{yu} & b_{y\rho} & b_{ym} \\ b_{ii} & b_{iu} & b_{i\rho} & b_{im} \end{bmatrix} \begin{bmatrix} i_{t-1} \\ u_t \\ \rho_t \\ m_t \end{bmatrix} = \begin{bmatrix} b_{\pi i} & b_{\pi u} & b_{\pi\rho} & b_{\pi m} \\ b_{yi} & b_{yu} & b_{y\rho} & b_{ym} \\ b_{ii} & b_{iu} & b_{i\rho} & b_{im} \end{bmatrix} \begin{bmatrix} i_{t-1} \\ \phi_u u_{t-1} + w_{ut} \\ \phi_\rho \rho_{t-1} + w_{\rho t} \\ m_t \end{bmatrix} \\ &= \begin{bmatrix} b_{\pi i} & \phi_u b_{\pi u} & \phi_\rho b_{\pi\rho} \\ b_{yi} & \phi_u b_{yu} & \phi_\rho b_{y\rho} \\ b_{ii} & \phi_u b_{iu} & \phi_\rho b_{i\rho} \end{bmatrix} \begin{bmatrix} i_{t-1} \\ u_{t-1} \\ \rho_{t-1} \end{bmatrix} + \begin{bmatrix} b_{\pi u} & b_{\pi\rho} & b_{\pi m} \\ b_{yu} & b_{y\rho} & b_{ym} \\ b_{iu} & b_{i\rho} & b_{im} \end{bmatrix} \begin{bmatrix} w_{ut} \\ w_{\rho t} \\ m_t \end{bmatrix} \\ &= A\mathbf{z}_{t-1} + B_0^{-1}\mathbf{w}_t. \end{aligned} \tag{19}$$

The vector $\mathbf{z}_t$ is not observed except for i_t, but it can be given the following law of motion, which is based on the AR(1) structure of u_t and ρ_t and the Equation (18):

$$\begin{aligned} \mathbf{z}_t &= \begin{bmatrix} b_{ii} & \phi_u b_{iu} & \phi_\rho b_{i\rho} \\ 0 & \phi_u & 0 \\ 0 & 0 & \phi_\rho \end{bmatrix} \mathbf{z}_{t-1} + \begin{bmatrix} b_{im} & b_{iu} & b_{i\rho} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{w}_t \\ &= C\mathbf{z}_{t-1} + D\mathbf{w}_t. \end{aligned} \tag{20}$$

Solving Equation (19) for $\mathbf{w}_t$ and substituting it for Equation (20) yields

$$\mathbf{z}_t = (C - DB_0A)\mathbf{z}_{t-1} + DB_0\mathbf{x}_t.$$

If $(C - DB_0A) = 0$, $\mathbf{z}_t$ can be written in terms of $\mathbf{x}_t$. This condition holds and this is proven in Appendix A. This fact can be used to turn Equation (19) into a VAR(1) representation:

$$\mathbf{x}_t = ADB_0\mathbf{x}_{t-1} + B_0^{-1}\mathbf{w}_t. \tag{21}$$

(Ikeda et al. 2022.)

Equation (21) is a VAR(1) model instead of a more general VAR(p) model where the order $p \in \mathbb{N}_+$. This is not surprising since only a single lag was used all throughout Section 3.1. To account for AR(p) processes where $p > 1$ in Equation (12) and for a central bank that smooths interest rates over more than just one period in Equation (15), the p can be increased when fitting the VAR model. Furthermore, the VAR(1) model is nested within a VAR(p) model, which means that increasing p does not impose further restrictions on the model. Indeed, following Mavroeidis (2021) and Ikeda et al. (2022), I choose $p = 4$ in the empirical analysis of this thesis.

Once a VAR(p) model is fitted to data, the monetary policy shock can be identified by imposing restrictions on B_0^{-1} of Equation (21). I make the common assumption that the monetary policy shock does not have an immediate impact on π_t or y_t. The idea is that if the monetary policy shock would have an immediate impact on π_t and y_t, the central bank could react to these changes in a systematic non-stochastic manner. By Equation (9), such a shock would cease to be an exogenous shock. This is by far the most popular identification scheme (Examples include but are not limited to Christiano et al. 1999, 2010; Stock and Watson 2001; Iwata and Wu 2006; Wu and Xia 2016; Mavroeidis 2021). In practice, this restriction means that $b_{\pi m} = b_{ym} = 0$ in Equation (19). This results in

$$B_0^{-1} = \begin{bmatrix} b_{\pi u} & b_{\pi \rho} & 0 \\ b_{yu} & b_{y\rho} & 0 \\ b_{iu} & b_{i\rho} & b_{im} \end{bmatrix}. \tag{22}$$

Unfortunately, this isn't sufficient as it is still impossible to identify w_{ut} or $w_{\rho t}$. Fortunately, however, these are not the focuses of interest here and the exact identification can be accomplished for example by setting $b_{\pi\rho} = 0$. This does not have an effect on the identification of m_t, which is the point of interest. If there are more monetary policy instruments, they can be added below the interest rate so that the monetary policy shock has an immediate impact on them. If there are more endogenous variables that are not monetary policy instruments, they can be added above the interest rate, so that the monetary policy shock doesn't have an immediate impact on them. (Christiano et al. 1999.)

A slight modification of the identification strategy of Equation (22) is offered by Stock and Watson (2001). They consider a forward-looking Taylor rule to identify the monetary policy shock. They assume that Federal Reserve sets interest rates so that a 1%-point increase in inflation forecast leads to a 150bp increase in the rate of interest and a 1%-point increase in unemployment forecast leads to a 125bp increase in the rate of interest. They construct forecasts using the reduced-form VAR model. Any deviation of the rate of interest from this rule is identified as the monetary policy shock. This identification, of course, relies heavily on the assumption that the central bank follows a very specific form of the Taylor rule.

3.3 Alternative identification strategies

The short-term restriction method of Equation (22) is far from the only strategy that can be used to identify a monetary policy shock. This subsection provides a brief overview of different methods. (For a more detailed overview and more exhaustive list, see Rossi 2021 and references therein.)

A very common identification strategy in structural modelling is to use heteroskedasticity. This method was developed by Rigobon (2003). Under this scheme, Σ is made conditional on its regime j at period t, $\Sigma_{j(t)}$. Then $\Sigma_{j(t)}$ is estimated within different regimes. Usually, there are only two regimes with Σ_1 and Σ_2, but more can be added. These can be derived from the residuals in periods $\{t \in \{1,...,T\}|j(t) = 1\}$ and $\{t \in \{1,...,T\}|j(t) = 2\}$ respectively. The individual shocks can then be estimated from decompositions $\Sigma_1 = B_0^{-1}B_0^{-1\prime}$ and $\Sigma_2 = B_0^{-1}\Lambda B_0^{-1\prime}$, where B_0^{-1} is a $K \times K$ matrix and Λ is a diagonal $K \times K$ matrix with only positive components on its diagonal. This results in $K \times K + K$ unknowns and $K \times (K+1)/2 \times 2 = K \times K + K$. In other words, the model is perfectly identified with just two regimes. Increasing the number of regimes leads to over-identification. This identification leads to $\Theta_0 = B_0^{-1}$ in regime 1 but $\Theta_0 = \Lambda B_0^{-1}$ in regime 2. However, since Λ is diagonal, IRFs only differ from each other in scale but not in shape. (Kilian and Lütkepohl 2017, 491–501.)

Identification by heteroskedasticity can be used to identify the monetary policy shock in two ways. Firstly, it can be seen as a method to identify orthogonal shocks, which are then afterwards given a structural interpretation by the econometrician. For example, Brunnmeier et al. (2021) construct a ten variable VAR model for the US economy. They divide the economic history into six periods or regimes, and use these regimes to identify shocks by heteroskedasticity. Then they conduct impulse response analysis and label one of the shocks the monetary policy shock due to the impulse responses corresponding to their preconception of the effects of the monetary policy shock.

The second approach by Wright (2012) is similar to so–called high frequency identification. This strategy relies on two ideas. Firstly, only one shock is allowed to vary across regimes. The procedure starts with estimating $\hat{\Sigma}_1$ and $\hat{\Sigma}_2$ and then finding a vector $\mathbf{b}$ so that $\mathbf{b}\mathbf{b}'$ is as close as possible to $\hat{\Sigma}_1 - \hat{\Sigma}_2$. This leads to only partial identification as only the shock associated with $\mathbf{b}$ is actually identified. Secondly, daily data is used. The assumption is that the variance of the monetary policy shock is larger on the days on which there is a monetary policy announcement. In other words, one regime is the policy announcement days and the other is every other day.

It is also possible to test other identification schemes using heteroskedasticity to impose additional over-identifying restrictions. Lanne and Lütkepohl (2008) examine a group of models where a monetary policy shock is identified by short-term restrictions much like I have done in Section 3.2. However, these models include a variety of variables that can be expected to react to a monetary policy shock immediately, such as bank reserves. The models vary in the way in which these variables respond to a monetary policy shock. To test which one is the most plausible identification scheme, they argue much like Brunnmeier et al. (2021) that there have been structural breaks in the way in which monetary policy has been carried out and use these structural breaks to impose further over-identifying restrictions. They then proceed to test the over-identifying restrictions and reject some but not all models they consider. This way heteroskedasticity can be used to validate other schemes.

The heteroskedasticity approaches have their downsides. The most apparent problem is that different regimes may not actually exist, although this can be easily tested. Additionally, the approach by Brunnmeier et al. (2021) relies on their subjective interpretation about the monetary policy shock. On the other hand, the approach by Wright (2012) uses high frequency data that isn't available for most variables, such as GDP or inflation. Furthermore, the announcement days can be difficult to separate from other days as central bankers make speeches, data is released and such also on days on which there are not policy announcements. What is worse, since the approach uses the monetary policy announcement as a dummy that identifies the monetary policy shock, it would have to be exogenous. Naturally, however, monetary policy decisions are often made in response to the state of the economy so that they are endogenous. Regardless, there are significant benefits to these approaches as well. The identification by Wright (2012) does not rely on any monetary policy instrument making it more robust to monetary policy being carried out by different methods compared to, for example, the short-term identification. This flexibility can of course also be a problem, since the shocks do not correspond to any specific change in policy, such as a surprise 25bp rate hike. In addition, the uncertain announcement days can sometimes be identified from the volatility so that this problem can be overcome. (Rossi 2021.) The shocks identified by heteroskedasticity can also have an immediate impact on variables that are not monetary policy instruments but are assumed to respond immediately, such as asset prices or interest on bonds.

Besides heteroskedasticity, instrumental variables (IV) can be used to identify the monetary policy shock. Let $\mathbf{z}_t$ be a vector of instrumental variables at time t. Assume that $\mathbf{E}[\mathbf{z}_t m_t] = \omega \neq \mathbf{0}$ and $\mathbf{E}[\mathbf{z}_t w_t] = \mathbf{0}$, where m_t is the monetary policy shock and w_t is any other structural shock. With these assumptions, IV can be used to identify a monetary policy shock in two ways.

Stock and Watson (2012) identify various shocks by relying heavily on the assumptions made. Let $B_0 = [\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_p]$, where $\mathbf{b}_i \in \mathbb{R}^p$ and $\mathbf{b}_m$ is the column that corresponds to the monetary policy shock. Based on the assumptions, $\mathbf{b}_p$ and ω can be identified up to a sign and scale from

$$\mathbb{E}[\varepsilon_t \mathbf{z}_t] = \mathbb{E}[B_0 \mathbf{w}_t \mathbf{z}_t'] = [\mathbf{b}_1 \cdots \mathbf{b}_p] \begin{bmatrix} \mathbb{E}[w_{1t}\mathbf{z}_t'] \\ \vdots \\ \mathbb{E}[w_{pt}\mathbf{z}_t'] \end{bmatrix} = \mathbf{b}_m \omega.$$

Stock and Watson (2012) identify all structural shocks in their model with instruments. This means that they can identify the whole B_0 and thus B_0^{-1} up to scale and sign. Scale and sign can be chosen so that the impulse responses are easy to interpret.

Gertler and Karadi (2015) identify only the monetary policy shock by choosing an explicit monetary policy instrument and using two-stage least squares. The goal is the partial identification of the SVAR model. More specifically, the identification of the column $\mathbf{b}_m$ of B_0^{-1}, where $\Sigma = B_0^{-1} B_0^{-1\prime}$, so that the immediate impact of the monetary policy shock is $\mathbf{b}_m m_t$. Let ε_t^m be the reduced form error that corresponds to the monetary policy instrument and ε_t^w be any other reduced form error. Also let b^m be the component of $\mathbf{b}_m$ that corresponds to the monetary policy instrument and b^w any other component. In the first stage, regress ε_t^m on $\mathbf{z}_t$ to derive fitted values $\hat{\varepsilon}_t^m$. Then regress all other residuals on $\hat{\varepsilon}_t^m$:

$$\varepsilon_t^w = \frac{b^w}{b^m} \hat{\varepsilon}_t^m + u_t.$$

This results in $K-1$ equations and K unknowns. The K:th equation is, however, formed by the fact that $\mathbf{b}_m' \mathbf{b}_m$ must be equal to the diagonal component of Σ that corresponds to the monetary policy instrument. Thus the shock is identified.

Gertler and Karadi (2015) use surprise changes in interest rate futures as instruments in their analysis of the monetary policy shock in the US. They define the surprise as change in the rates that occurs around monetary policy announcement. They also use futures for different interest rates and for different maturities. However, it isn't clear which monetary policy instrument one should choose. They try the federal funds rate and 1- and 2-year government bond yields. Meanwhile, Stock and Watson (2012) use multiple shock series developed by different authors simultaneously. These shock series are derived from economic models, which naturally makes them sensitive to model specifications. A common challenge that both Gertler and Karadi (2015) and Stock and Watson (2012) face is that they find their instruments rather weak. This means that their instruments explain very little of the variation in ε_t^m. This is rather concerning as it is difficult to actually test whether or not instruments are weak, as autocorrelation can invalidate many commonly used tests, including the F-test (Rossi 2021).

It is also possible to identify monetary policy shocks by restricting the sign of the impact. These are called sign restrictions and they were developed by Faust (1998), Canova and Nicoló (2002) and Uhlig (2005). Sign restrictions are identified by simulating series of uncorrelated shocks with a unit variance η_t, converting these to candidate series for structural shocks $\mathbf{w}_t^* = Q'\eta_t$, where Q is an orthogonal matrix, and choosing a lower triangular matrix P so that the reduced form errors $\varepsilon_t = P\eta_t = PQ\mathbf{w}_t^*$. If PQ satisfies the sign restrictions, it is retained. Otherwise, it is rejected. The goal is to generate many PQ, which can be used to construct candidate IRFs. The mean or median IRF can then be interpreted as the true point estimate for IRF. (Kilian and Lütkepohl 2017, 421–425.)

Canova and Nicoló (2002) identify a model where a negative monetary policy shock has a positive impact on output, employment, consumption, inflation and real cash balances but which causes short-term interest rates to decline in relation to longer ones. They justify this by constructing what they call a limited participation model of the economy. Similarly Debortoli et al. (2019) assume that a positive monetary policy shock has a positive impact on prices and output but a negative impact on the long term interest rate. However, they further impose that these signs persist at least for a year, whereas Canova and Nicoló (2002) only impose these restriction on the immediate impact. The former is called a dynamic sign restriction and the latter a static sign restriction (Kilian and Lütkepohl 2017, 424–425, 432).

The sign restriction approach relies on the econometrician's subjective interpretations about the monetary policy shock. This means that one needs a credible macroeconomic model to justify these choices. Additionally, under sign restriction, the restrictions put on other shocks also matter for the monetary policy shock. This is not true for short-term restriction identification. There is no consensus on how this matter should be handled. For example, both Canova and Nicoló (2002) and Debortoli et al. (2019) identify all structural shocks. This is not necessary when using short-term restrictions. Sign restriction identification also does not result in unique point estimates for IRFs, which can make inference challenging. Finally, sign restriction identification requires a method for generating lots of random orthogonal Q matrices. This can be computationally challenging especially when K is large. (Kilian and Lütkepohl 2017, 425–437.)

There are also other methods than SVARs such as event studies and local projections. These, however, lie outside the scope of this thesis. (See Rossi 2021.)

3.4 The effects of the monetary policy shock

3.4.1 Data

I identify the effects of a monetary policy shock in the euro area using quarterly data from the first quarter of 1999 up to the first quarter of 2020. This choice of time range is made due to the availability of data and to avoid the Covid-19 recession. Covid-19 is avoided because the shock appears as a clear outlier and appears to violate the normality assumption inherent in the VAR models discussed here. I use the annualized log-change in harmonized index of consumer prices (HICP) for inflation and the end-of-period Main Refinancing Operations rate (MRO) for interest rate. I use the end-of-period interest rate as the assumption that monetary policy shock doesn't have immediate impact on other endogenous variables is more justified with the end-of-period interest rates than with the average interest rates. The MRO rate is the interest rate that ECB charges from banks for collateralized loans with one week maturity. HICP is taken from Eurostat and MRO is from the ECB.[1]

I construct the output gap by filtering out the cyclical component of logarithmic Gross Domestic Product (GDP) using the modified Hamilton filter of Quast and Wolters (2022). In contrast to the more popular Hodrick-Prescott filter, modified Hamilton filter is as real time as the publication of GDP figures. This means that these output gap estimates reflect the decision making environment of central banks, firms and households better. Furthermore, in contrast to the original Hamilton filter of Hamilton (2018), the modified filter is robust to varying business cycle frequencies. The filter is described in further detail in the Appendix B.

A popular alternative to output gap is to use unemployment instead. This can be attractive as unemployment is available in almost real time, has a higher publication frequency and does not require any filtering. However, unemployment does not contain all the information of the output gap. (See Christiano et al. 2010 for a discussion and, for example, Stock and Watson 2001; Mavroeidis 2021 for examples of similar studies using unemployment.) GDP data is also taken from Eurostat.

The used time series are depicted in figure 1. A couple of observations can be made. Firstly, the interest rate seems to have generally followed inflation and the output gap especially before early 2010s just as the Taylor rule would imply. Secondly, the interest rates have been very low starting from the early 2010s. In fact, they were zero from 2016:Q1 all the way to 2022:Q2. The subsequent post-coronavirus rate hikes are left out of the sample. Under the standard SVAR framework, this ELB period is very dubious, since 0 is treated like any other real number.

[1] https://ec.europa.eu/eurostat/en/ and https://sdw.ecb.europa.eu/ respectively. The MRO rate is spliced together from variable and fixed tender rates.

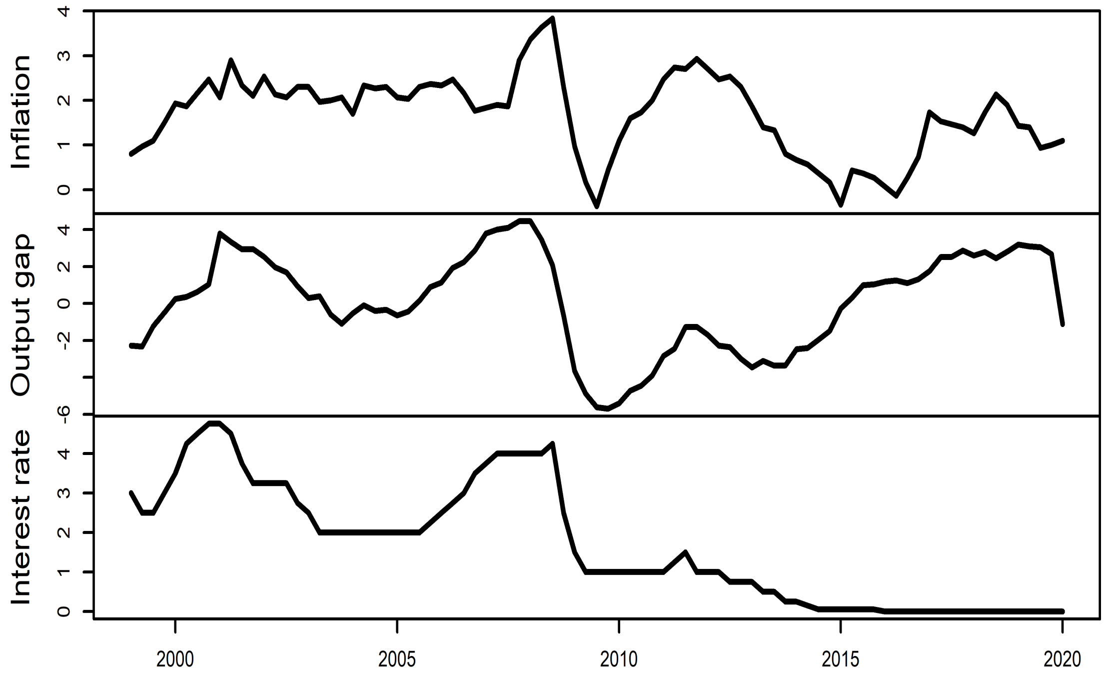

Figure 1: Euro area inflation, output gap and interest rate rate between 1999:Q1–2020:Q2.

3.4.2 Impulse responses

Typically, when using the standard SVAR model with the same identification as I have used, econometricians have found that a monetary policy shock that increases the interest rate reduces inflation and output gap. If unemployment has been used instead, then unemployment increases. (See, for example, Christiano et al. 1999, 2010; Jarociński 2010.) However, this result is by no means universal. Indeed, many have observed a so–called price puzzle. This means that inflation has increased, not decreased, in reaction to the monetary policy shock. For example, Sims (1992) and Stock and Watson (2001) observe that in short term inflation increases in response to a monetary policy shock whereas Hanson (2004) and Estrella (2015) finds this increase to be persistent. Meanwhile, Castelnuovo and Surico (2010) find the price puzzle to exist in the US only in some sub-samples. Nevertheless, Rusnak et al. (2013) find in their meta-analysis that the majority of authors have not reported a significant price puzzle, although they partly attribute this to publication bias.

The price puzzle is usually attributed to unobserved variable bias, and a common solution is to include a separate commodity price inflation. The idea is that commodity price inflation induces consumer price inflation with a lag. To counter this, the central bank increases interest rates. This makes it appear as if it is the central bank increasing the interest rates that is causing inflation. (Sims 1992; Christiano et al. 1999.) An alternative candidate for the omitted variable is inflation expectations (Castelnuovo and Surico 2010).

Note that the omitted variable bias is not the only explanation. For example, the price puzzle could also be a result of increased borrowing costs for producers that are then transferred to consumers. This is called working capital channel. (Christiano et al. 2010. For additional alternative explanations and solutions, see Barth and Ramey 2001; Leeper and Zha 2001; Estrella 2015.) Alternative identification schemes can as well solve this problem. For example, if one explicitly specifies the monetary policy shock to be such that an increase in the interest rate reduces inflation and output gap, the price puzzle is made impossible. Similarly, using heteroskedasticity and then picking the monetary policy shock afterwards solves this issue.

Additionally, there exists a rarer anomaly where the output gap also increases in the short term after a monetary policy shock. This is observed, for example, by Gertler and Karadi (2015) when using short-term restriction strategy.

Figure 2 reports the impulse responses of a SVAR-model to a monetary policy shock identified by Equation (22). The underlying VAR model has 4 lags and inflation, output gap and interest rate as endogenous variables. The shock itself is scaled to correspond to a 25bp immediate shock to interest rate. Figure 2 also has 90% confidence intervals that have been constructed using a non-parametric bootstrap with 5000 realizations.

An obvious takeaway from figure 2 is that the monetary policy shock does not appear to have a particularly large effect on the other variables. Indeed, looking at the 90% confidence intervals, the monetary policy shock has a statistically significant effect only on the output gap with a horizon of 1 quarter as well as on the interest rate itself. This raises doubts about the ability of the central bank to exercise discretionary monetary policy to achieve its goals. This is, of course, partly due to the rather wide choice of confidence intervals. It is not uncommon to encounter 66% or 67% confidence intervals in the literature (see, for example, Stock and Watson 2001; Ikeda et al. 2022). It is also worth noting that inflation seems to increase in response to the monetary policy shock, although this is insignificant. This could be seen as weak evidence for the price puzzle.

The weakness of the monetary policy shock used to be exceptional for standard SVAR models with the current identification. For example, Christiano et al. (1999; 2010) and Stock and Watson (2001) found that in the US the economy responded strongly to a monetary policy shock. Sims (1992) found similar results in many different countries. However, some more recent studies have found the monetary policy shock weaker (see, for example, Jarociński 2010 for a survey of European countries and the EU; Estrella 2015 for the US). This does, however, depend on the identification strategy. For example, when comparing their IV identification strategy to the short-term restriction strategy, Gertler and Karadi (2015) observe very similar results to those of figure 2 when using the short-term identification. With their IV strategy, however, the monetary policy shock has the expected effects, i.e. prices and output fall. Furthermore, when using sign restrictions,

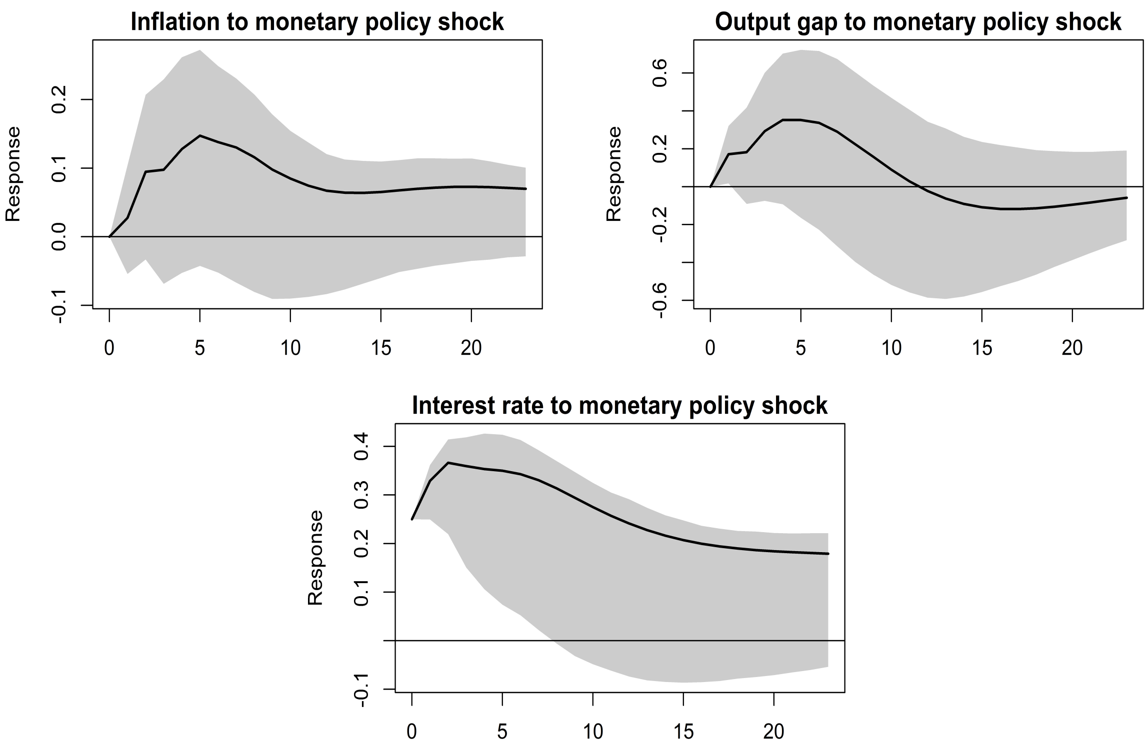

Figure 2: Impulse response functions of the monetary policy shock in linear SVAR framework.
Notes: Solid lines are the point estimates of the impulse response functions to a 25bp monetary policy shock. The shaded areas are their 90% bootstrap confidence intervals with a non-parametric bootstrap sample of 5000 simulations. The quarterly horizon is up to 24 quarters (6 years).

Debortoli et al. (2019) find no evidence for a change in the IRF of the monetary policy shock in the US. Instead, they observe the expected effects. Although since they use sign restrictions, this is by design.

4 Views on the ELB

4.1 What is the ELB?

In the Sections 3, the interest rate was assumed to adjust without being bounded. This implies a rather odd assumption that the central bank has chosen to keep the interest rate at 0bp even though they could have chosen any other rate, including a negative one. In other words, in their quest to fulfill their mandate, the ECB found 0bp to be the optimal rate since 2016:Q1 to the end of the sample. Instead of this, interest rate can be assumed to be bounded by some level called the ELB. This level need not be one that the rates cannot fall under but rather a rate under which the changes in the interest rate either no longer reflect the true state of monetary policy or are inconsequential.

From the central banks point of view, the ELB introduces a certain asymmetric aspect to monetary policy. Taking the Taylor rule of Equation (**??**) as an example, the ELB limits the central bank's ability to answer low inflation or output gap with lower interest rates. Meanwhile, nothing limits the central bank's ability to increase interest rates when inflation or output gap is high. This means that over heated economy with high inflation and output gap is easier for the central bank to deal with than a stagnant economy with low inflation and output gap.

From statistical point of view, the ELB can be accounted for by a censoring rule. Such a rule can take the form of:

$$i_t = \max(i_t^*, b), \tag{23}$$

where i_t is the observed and bounded nominal interest rate, i_t^* is the unbounded nominal shadow rate and b is the censoring threshold or the ELB. The shadow rate i_t^* is not observed when it falls below b. However, it is still known to be under that level. This means that i_t^* is a latent variable. (Iwata and Wu 2006.)

The ELB may be at zero but can also take different values. For example, Wu and Xia (2016), Mavroeidis (2021) and Aruoba et al. (2022) use 25bp and Ikeda et al. (2022) 20bp as the ELB in the U.S. whilst Iwata and Wu (2006) use 50bp as the ELB in Japan. This choice can be based on visually observing the behaviour of the interest rates. Iwata and Wu (2006), for example, note a qualitative shift in their sample as interest rates fell below 50bp. The ELB can also be based on the behaviour of other rates. For example, Wu and Xia (2016) base their choice of 25bp on the interest rate on reserves that the Federal Reserve had paid. It is also possible that the ELB changes over time. For example, Hayashi and Koeda (2019) and Ikeda et al. (2022) set the ELB as the interest on bank reserves in Japan, and Wu and Xia (2020) set it as the lowest observed deposit rate up to date in the euro area.

In my sample period for the euro area and for the MRO rate, the deposit facility rate is not an option for the ELB. The problem is that the deposit facility rate is by design lower than the MRO rate, and it has occasionally been negative whereas the MRO rate has only been non-negative. However, looking at the figure 1, two possible choices seem natural. Firstly, 0bp is both the lowest value that the MRO rate has taken, and the rate was stuck at 0bp since 2016:Q1 to the end of the sample in 2020:Q1. This implies that the movement of the MRO rate has clearly been restricted to some extent by the level 0. Additionally, 0 as a minimum interest rate is an intuitive choice. A second option would be 5bp. The MRO rate fell to 5bp in 2014:Q3 rather smoothly without getting stuck at other levels. Then the MRO rate stayed at 5bp from 2014:Q3 all the way to 2015:Q4. It could thus be argued that since this extremely low interest rate regime was new to the euro area and to the ECB decision makers, they hesitated to lower the rate to 0 until 2016:Q1. Of course, these can be spliced together by setting the ELB to 5bp up to 2015:Q4 and to 0bp since 2016:Q1. Similarly to the arguments of Wu and Xia (2020), it can be argued that prior to 2016:Q1 it was perceived that 5bp was the ELB. However, as the MRO rate fell beneath 5bp, the new ELB was set at 0bp. If the MRO rate is increased afterwards, as has happened after the sample studied here, there is no reason to increase the ELB as the public and the policy makers now know from experience that the ELB is at 0bp. This would, however, make the construction of bootstrap confidence intervals more challenging and because of this I opt to use $b = 5$bp.

4.2 Shadow rates

Generally speaking, there are two schools of thought when it comes to the ELB. One set of researchers have focused on the way in which the ELB restricts the ability of the central bank to carry out monetary policy. The ELB is seen as something that the central bank must adjust to and which leads to the inability of the central bank to adopt optimal policies by setting the interest rate. (See Eggertsson and Woodford 2003; Jung et al. 2005; Gust et al. 2017.) The second set of researchers have instead tried to assess the policy stance of the central bank without relying on the observed interest rates that are bound to zero. The justification is that in such circumstances, the central bank exercises monetary policy by other unconventional means. In practice, this implies estimating the shadow rate i_t^* of Equation (23) or estimating the monetary policy shock without actually including a monetary policy instrument into the model. The latter can be accomplished, for example, by the heteroskedasticity identification scheme of Wright (2012). (See also Rossi 2021 for an overview of the topic.)

The possibility of ignoring the ELB allows for an additional research: Does the ELB matter? In other words, can the monetary policy makers circumvent the limitations set by the ELB by exercising unconventional monetary policy? This question has lead to two competing hypotheses: The ELB irrelevance and the ELB relevance hypotheses.

The first approach for estimating i_t^* uses an auxiliary model of the term structure and other variables. Krippner (2013) links the term structure of bonds to the expected short rates, $\mathbb{E}_t[i_{t+h}]$. Since investors can always hold cash, these rates can be expected to be non-negative, though in the absence of cash, they could turn negative. These hypothetical rates are expected shadow short rates, $\mathbb{E}_t[i^*_{t+h}]$. Krippner (2013) constructs a model to estimate these rates. Choosing $h = 0$ leads to estimates of i_t^*. A related idea is that of the "lift-off" from zero of Swanson and Williams (2014). It is possible to derive the market expectations for the period when interest rates lift-off from zero by looking at interest rates of different maturities since according to the expectations theory, longer interest rates are averages of future interest rates.

Wu and Xia (2016, 2020) suggest an alternative approach to deriving shadow rates by using a shadow rate term structure model (SRTSM). The SRTSM explains the shadow rates by a linear model of factors, whereas the observed rate is a non-linear transformation of the shadow rate:

$$
\begin{aligned}
i_t^* &= \beta_0 + \beta_1' Fac_t, \\
i_t &= b_t + \sigma^* g(\frac{i_t^* - b_t}{\sigma^*}) + \varepsilon_t, \quad \varepsilon_t \sim n.i.d(0, \sigma),
\end{aligned}
$$

where Fac_t is a vector of factors that summarise the state of the economy and $g(x) = x\Phi(x) + \phi(x)$, where $\Phi(x)$ and $\phi(x)$ are the cumulative distribution function and the density function of a standardized normal distribution, respectively. Regardless of the method used, these shadow rates can then be used as i_t^* in a conventional linear SVAR model described in Section 2 and used in Section 3.4.

The problem associated with these models is that the estimates for shadow rates are as artificial constructs sensitive to the model specification (Krippner 2020). If the specification is not correct, any measurement errors can be large and persistent (Mavroeidis 2021). Fortunately, the estimation of shadow rates is not necessary for the identification of the effects of the monetary policy shock. The shadow rate can be included as an unobserved latent variable directly in the SVAR model.

The easiest way to include the shadow rate in the SVAR framework is to keep the observed rate as is as an explanatory variable but to interpret the interest rate on the response side as censored, i.e. unobserved if it falls to or below the ELB. In the terminology of Mavroeidis (2021), this is the kinked SVAR model (KSVAR). The consequence of this model structure is that the shadow rate itself does not have a causal effect on the economy, only the observed rate has. The shadow rate thus becomes more-so the rate that central bank would prefer in the absence of the ELB rather than a proxy for the monetary policy stance. Because of this, other policy instruments, such as the money supply or excess reserves, must be added to the model to account for the unconventional monetary policy (see, for example, Iwata and Wu 2006; Hayashi and Koeda 2019). If other monetary policy instruments are not included, this would imply an assumption that unconventional monetary policy is irrelevant. This fact can be seen either as a claim that these unconventional measures are ineffective or as a limitation of the KSVAR model.

Overall, the KSVAR changes the dynamics of the NK-model of Section 3 a little if one does not introduce alternative instruments. The Equations (10) and (11) can be taken as is since neither the inflation nor the output gap is constrained directly. However, Eggertsson and Woodford (2003) and Jung et al. (2005) find that the optimal interest rate is not perfectly captured by the conventional Taylor rule of Equation (15) even in the simplistic NK-model if the central bank can credibly commit to a future policy. Instead, interest rates are kept at the ELB longer than the Taylor rule would suggest. Jung et al. (2005) go as far as to suggest that interest rates stay below the ELB until the cumulative deviation from the observed rate, $\sum_{j=0}^{t} i_j^* - i_j$, is zero. This stimulates the economy as the central bank commits to the ELB for a longer period of time. In other words, the central bank exercises forward guidance.

In practice, the models by Eggertsson and Woodford (2003) and Jung et al. (2005) depict an extreme case since a central bank cannot be expected to fully commit to future monetary policy. Instead, the late exit can be approximated in the SVAR model by introducing different regimes, one for a non-ELB regime and another for an ELB regime. These regimes have different reduced form parameters $\nu, A_1, ..., A_p$ and perhaps also Σ, which can account for the changing persistence of the monetary policy. In the simplest case, the process is assumed to switch between these two regimes as the interest rate falls below to or to the ELB, or surpasses it. (See Mavroeidis 2021; Aruoba et al. 2022 for implementation.) In more complex cases, a more sophisticated exit rule can be formulated. For example, Hayashi and Koeda (2019) require both the shadow rate to exceed the ELB and inflation to exceed a stochastic inflation target. In practice, however, such formulations require a more complex setup or a large number of switches between the ELB regime and non-ELB regime in order for the models to be estimated. The latter is not the case with the pre-covid euro area sample period studied in this thesis and the former leads to models that are no longer the KSVAR.

A somewhat more challenging approach is to include the shadow rate both as a response and a covariate. In the terminology of Mavroeidis (2021), this is the censored SVAR model (CSVAR). The challenge is that the covariate is unobservable and thus its distribution has to be estimated. Again, the Equations (10) and (11) can be taken as is except that they depend on the shadow rate i_t^* and not the observed rate. This is because i_t^* captures the whole range of unconventional monetary policies that are implicitly assumed to work just like conventional policy. Of course, this makes i_t^* a proxy for the monetary policy stance and not an actual interest rate that banks or other institutions pay to the central bank. Because of this, there is no need to introduce additional monetary policy instruments into the model unless they specifically are expected to work somehow differently. Furthermore, the late exit or forward guidance is now taken care of since such policy is associated with a change in the shadow rate today, which leads to changes in the expected future shadow rates. (Ikeda et al. 2022.)

The most difficult approach is to use both the observed and the shadow rates as covariates. This is likely necessary if one wishes to introduce regime switching as the estimation of such regimes may be impossible without the introduction of a shadow rate as a covariate. This formulation – censored and kinked SVAR (CKSVAR) in Mavroeidis's (2021) terminology – allows the shadow rate and all the unconventional monetary policy measures that influence it to have an effect on the economy but differently from the observed rate and the conventional monetary policy instruments. However, the CKSVAR is by design identified as is and thus it relies on a wholly different identification scheme from the standard SVAR, the KSVAR or the CSVAR considered in this thesis. Because of this a more rigorous overview of the CKSVAR is left outside this thesis. (Mavroeidis 2021.)

The advantage of Mavroeidis's (2021) setup is that a model where the ELB is irrelevant, i.e. the CSVAR, and a model where the ELB makes monetary policy impossible, i.e. the KSVAR, are nested within a single model, i.e. the CKSVAR. Since these models are Gaussian, the ELB (ir)relevance hypothesis can be tested using a likelihood ratio test. Of course, within the euro area this does not necessarily perfectly identify the ELB (ir)relevance since it could confuse time dependence and the ELB relevance. More precisely, since non-ELB and ELB periods are separated by a single period, it is not clear that Mavroeidis's (2021) test comparing the CKSVAR and the CSVAR truly tests ELB relevance since it could just as well be a structural break in time, comparable to Chow's test (see Lütkepohl 2006, 182–184 for a description of Chow's test). On the other hand, it is also likely that any structural break that coincides with the interest rate reaching the ELB is caused by the ELB. Indeed, some researchers have used time varying VAR models to see if the parameters have changed since the beginning of the ELB era (see, for example, Swanson and Williams 2014; Debortoli et al. 2019). Alternatively, Ikeda et al. (2022) compare the impulse responses of monetary policy shocks that impact at different dates. This, naturally, requires a model that has time dependent impulse responses, and is merely focused on the ELB's (ir)relevanse to the monetary policy shock. Also, the challenge of identifying the change in impulse responses that are specifically induced by the ELB persists.

Wu and Xia (2016) investigate the effects of unconventional monetary policy by comparing the observed development of the economy and a counterfactual time series where there were no unconventional monetary policy shocks. This can be done using the VMA(∞) representation of Equation (6) and setting the monetary policy shocks to zero if the ELB is binding. However, since the sample is finite, any VMA decomposition, also called historical decomposition, is only approximate (see Kilian and Lütkepohl 2017, 116–123 for an overview of historical decomposition and 131–139 for counterfactuals). Furthermore, this approach requires that the shocks are observed, which isn't always the case. Lastly, because their model assumes monetary policy shocks to be equally as effective in the ELB as outside it, they don't actually test the ELB (ir)relevance hypothesis.

4.3 Unconventional monetary policy

4.3.1 The extended toolkit

Dell'Ariccia et al. (2018) list three unconventional monetary policy tools that a central bank may use. They are forward guidance (FG), quantitive easing (QE) and negative interest rates policies (NIRP). The main goal of these policies is to flatten the yield curve, that is to make borrowing money cheaper for a longer period of time.

Campbell et al. (2012) make a distinction between two types of FG: Delphic and Odyssean. Engaging in Delphic forward guidance means that a central bank clarifies what principles it will follow in future and what its expectations are whereas engaging in Odyssean forward guidance refers to the central bank committing to some interest rate policy regardless of how the economy develops. The Odyssean variety of FG is similiar to that studied by Eggertsson and Woodford (2003) and Jung et al. (2005).

FG can take the form of a conditional promise to maintain certain interest rate policy. Examples include the statement by the Bank of Japan in October 2010 to keep interest rates low until "price stability is in sight" or the statement by the Bank of England in August 2013 to keep interest rates low until unemployment has fallen below 7%. (Dell'Ariccia et al. 2018.) This can be very attractive for central bankers who have to operate in an uncertain environment (Panetta 2023).

FG faces many challenges that may hinder its effectiveness. First of all, in order for FG to have any effect, central bank has to commit to some policy which differs from market expectations. On the other hand, if the statement differs too much from these expectations, it may not be credible. Secondly, a promise to keep interest rates low also informs the markets that the central bank is pessimistic about the future. This may make FG counter-productive. (Dell'Ariccia et al. 2018.) On the other hand, FG can also be the most straightforward way to lower interest rates on loans with higher maturities, since it targets these directly.

QE is a large scale asset purchase operation by a central bank. While these usually target government bonds, they can also target corporate bonds in general, and banks and mortgage lenders in particular. Such operations usually occur at a time when the reserves banks hold at the central bank increase, meaning that they can be financed by these reserves. Furthermore, as bonds mature, QE wears off naturally without any sale program. (Dell'Ariccia et al. 2018.)

QE is usually enacted through a schedule that specifies the timeline of purchases and their amount and asset classes. This can be seen as a strong complement to FG. The channel by which QE has an effect on the economy through such signalling is called the signalling channel. Furthermore, by choosing what maturities and types of bonds the central bank purchases, they can easily target different maturities. (Dell'Ariccia et al. 2018.)

NIRP has been enacted, for example, by the ECB and Bank of Japan by charging banks' interest on the deposits that they make at the central bank. In theory, such policies should be impossible as banks would prefer to hold cash instead of paying interest. However, in practice such a course of action would involve considerable transaction costs. This means that as long as these transaction costs are larger than the negative interest rate that banks have to pay to the central bank, NIRP is possible. (Dell'Ariccia et al. 2018.)

NIRP can be a double edged sword. On one hand, it encourages lending at a lower interest rate by banks. On the other hand, it reduces bank profitability, which can reduce lending and hurt the banking sector that is probably already weakened by a recession. Usually NIRP also targets only a part of the reserves. Commonly, mandatory reserves are exempted. (Dell'Ariccia et al. 2018.)

Unconventional policies have received a good deal of criticism as well. Firstly, unconventional policies can threaten bank profitability, as they lower interest rates beyond what conventional policy would do. This fall in interest rates is not caused by banks' access to cheaper credit but rather by directly lowering the interest rates that they can charge on the market. Secondly, unconventional policies can also encourage riskier behaviour since they typically lower the yields of safe assets and may also make investors and banks more complacent and confident that they will have easy access to cheap credit in the future as well. Finally, unconventional policy holds political risks as central banks venture onto uncharted waters. (Dell'Ariccia et al. 2018.)

4.3.2 Unconventional monetary policy in the euro area

At the onset of the great recession in 2008 and 2009, the ECB was particularly cautious in enacting unconventional policies. It slashed interest rates but still mainly functioned as the lender-of-last-resort. Most notable somewhat unconventional measures were the Longer-Term Refinancing Operations offered to banks in need of liquidity and Covered Bond Purchase Program. The former were essentially Main Refinancing Operations with longer maturities and the latter was a comparably small program where the ECB purchased some bonds from banks. (Dell'Ariccia et al. 2018.)

It was only with the sovereign debt crisis of 2010–2012 that the ECB started a bond purchase program to restore liquidity to the financial market and lower the borrowing costs of the governments of some euro area countries. These measures were initially a part of a larger joint program of EU and International Monetary Fund to save Greece from default and to avoid forcing it to abandon euro all together. Later this Securities Market Program was expanded to include also Spanish, Irish, Italian and Portuguese government bonds. Regardless, although the euro was preserved, the countries that received help from Securities Market Program still experienced high borrowing costs and a recession in 2011–2012. (Dell'Ariccia et al. 2018.)

The unconventional monetary policy of the ECB shifted to a higher gear in 2012. First came the ECB's president Mario Draghi's "whatever it takes" speech. In it he stated that ECB will do whatever it takes to preserve the euro. Second came the Outright Monetary Transactions (OMT). This was a program offered to euro area countries by ECB where ECB could purchase a large amount of country's bonds in exchange for being allowed to monitor the governments finances. By 2013 the sovereign debt crisis had been averted even though OMT was not activated even once. Indeed, OMT operated wholly through the signalling channel. (Dell'Ariccia et al. 2018.)

Up until the end of the sovereign debt crisis in 2012, the ECB had mostly enacted unconventional monetary policy measures to save euro area countries from default and the euro area from collapsing. Even though the worst was averted, the economy was still under-performing, as growth was slow and inflation was well below the target of 2%. To help the economy, ECB started to engage in FG in 2013 by telling that they expected interest rates to remain low for an extended period of time. NIRP was first started in June 2014 by cutting the deposit rate to -10bp and further cutting it to -40bp in 2016. The beginning of NIRP coincides with the fall of the MRO rate to 5bp, which further justifies the choice of $b = 0.05$. Targeted longer-term refinancing operations (TLTRO) were also started where banks could receive cheap loans with long maturities if they offered more favourable loans to households and firms. The goal of this policy was to transmit NIRP to actual loan contracts. Nevertheless, most notable policy was perhaps ECB's own Asset Purchase Program (APP) that was aimed at boosting the economy as a whole and not to save specific governments from default. APP reached its peak in late 2016 to early 2017 when the average monthly target for purchases was 80 billion euros. However, by 2019 net purchases had practically reached 0. (Driffill 2016; Dell'Ariccia et al. 2018; Asset purchase programs.)

In response to the Covid-19 pandemic, the ECB announced Pandemic Emergency Purchase Program in March 2020. This was a large QE program that was aimed at euro area government bonds and private sector securities that essentially restarted APP. This was soon followed by an update to TLTRO in April, which expanded the amount that banks could borrow. Also an additional Pandemic Emergency Longer-Term Refinancing Operations were introduced as a backup if TLTROs proved insufficient. (Hutchinson and Mee 2020; Asset purchase programs.)

Post-Covid, the ECB has abandoned unconditional FG. This means that any FG that the ECB exercises is conditional on the economic outlook. For example, the ECB's commitment to 2% inflation can be seen as a form of FG that the ECB exercises currently. (Panetta 2023.) Also as interest rates have increased since Covid, this has not meant an end to unconventional policy measures. Instead, the ECB will complement conventional interest rate policy with unconventional policies, such as asset purchase programs. (Lane 2022.) Nevertheless, APP effectively ended in 2022 (Asset purchase programs).

All the unconventional monetary policy measures are collected in the table 1 together with their start date and type.

Compared to other central banks, such as the Bank of England, Federal Reserve and the Bank of Japan, unconventional monetary policy of the ECB had some notable differences. Firstly, ECB was slow to engage in unconventional monetary policy. The Bank of England, for instance, engaged in QE as early as 2009 and the Bank of Japan started to exercise FG already in 2010. This can partly be attributed to the European treaties that limited the ability of the ECB to carry out such unconventional policies, but also to the fact that the ECB did by and large succeed in its goal of 2% inflation without such measures. Secondly, the ECB did not only face an environment of slow economic development in the aftermath of the Great Recession but it also faced a very real possibility that the currency area may not survive. (Driffill 2016; Dell'Ariccia et al. 2018.)

Table 1: Notable unconventional monetary policies in the euro area.

Start date	Policy	Type
July 2012	"What ever it takes"–speech	*FG*
August 2012	Outright Monetary Transactions (OMT)	*QE*
July 2013	Forward Guidance (FG)	FG
June 2014	Negative Interest Rate Policy (NIRP)	NIRP
September 2014	Asset Purchase Program (APP)	QE
March 2016	Targeted Longer-Term Refinancing Operations (TLTRO)	Other
March 2020	Pandemic Emergency Purchase Program	QE
March 2020	Pandemic Emergency Longer-Term Refinancing Operations	Other

Notes: Types are abbreviated; FG for forward guidance, QE for quantitative easing and NIRP for negative interest rate policies. Types are in italics if the type is somewhat ambiguous. See text for references.

5 Kinked structural vector autoregression

5.1 Kinked vector autoregression

Linear models with censored responses are called tobit models:

$$y_i = \max\{\beta' \mathbf{x}_i + \varepsilon_i, b\},$$

where y_i is the response, x_i are the covariates, β is a parameter vector, $\varepsilon_i \sim N(0,\sigma)$ are i.i.d. error terms and b censoring is threshold under which the response simply cannot fall or under which it is unobservable. The same model can be formulated if there exists a maximum value for the response or both a minimum and a maximum. When fitting the model, the basic idea is to use the conventional maximum likelihood estimation when ever the response is observed. However, when the response is censored, the conditional probability that the response is censored is used instead. (Tobin 1958; Cameron and Trivedi 2009, 536–544.)

Let $f_i(y_i;\theta)$ and $F_i(b;\theta)$ be the conditional density and the conditional probability distribution function for observation $y_i \in \mathbb{R}$ respectively of any linear model with parameter vector θ. Furthermore, let there be N observations and S be the group of censored observations, formally $S = \{i | y_i \leq b, i = 1,2,...,N\}$. Now the log-likelihood can be written as

$$l(\theta) = \sum_{i \in S} \log[F_i(b;\theta)] + \sum_{i \notin S} \log[f_i(y_i;\theta)].$$

If correctly specified, ML estimation leads to unbiased and consistent estimator for θ. (Tobin 1958; Cameron and Trivedi 2009, 536–544.)

In order to construct a VAR model for $\mathbf{x}_t = (\pi_t, y_t, i_t)'$ with censoring, the most obvious approach is to simply regard i_t as censored whenever $i_t \leq b$. This leads to so called kinked VAR model in the terminology of Mavroeidis (2021):

$$\mathbf{x}_t^* = \nu + A_1\mathbf{x}_{t-1} + A_2\mathbf{x}_{t-2} + ... + A_p\mathbf{x}_{t-p} + \varepsilon_t, \tag{24}$$

where $\mathbf{x}_t^* = (\pi_t, y_t, i_t^*)'$. Note the subtle assumption made here; While the uncensored and unobservable interest rate or shadow rate i_t^* appears on the left-hand side, it does not appear on the right. This means that the shadow rate itself does not have an effect on the other variables except immediately through the covariance structure of the error term ε_t.

The economic interpretation of the model of Equation (24) is that if the interest rate hits the lower bound and unconventional monetary policy instruments are not explicitly included in $\mathbf{x}_t$, the central bank becomes limited in its ability to carry out monetary policy. In other words, the interest rate is the only significant tool available to the central bank and if it cannot lower the interest rate any further due to the lower bound, it is out of options to influence the economy. The exception to this is the immediate effect from the covariance of reduced form errors ε. However, the structural form of the model specified in Section 5.2 makes this also irrelevant.

The actual fitting of kinked VAR is based on the same principles as conventional tobit model. If the response is observed, the conditional density, $f_t(\mathbf{x}_t;\theta)$, can be used. However, when dealing with a censored response, i.e. the possibly censored variable within $\mathbf{x}_t$ is b, the likelihood of that observation is $f_t(\mathbf{x}_{1t} \wedge x_t^* \leq b;\theta)$, where $\mathbf{x}_{1t} \in \mathbb{R}^{K-1}$ are the components that cannot be censored and $x_t^* \in \mathbb{R}$ is the component that can be censored. When constructing the likelihood function, it proves useful to use the relation $P(A \wedge B) = P(A)P(B|A)$:

$$l(\theta) = \sum_{k \in S} \log[f_{1k}(\mathbf{x}_{1k};\theta)F_k^*(b|\mathbf{x}_{1k};\theta)] + \sum_{k \notin S} \log[f_k(\mathbf{x}_k;\theta)], \tag{25}$$

where $f_{1k}(\mathbf{x}_{1k};\theta)$ is the marginal density of x_{1k}, $F_k^*(b|\mathbf{x}_{1k};\theta)$ is the cumulative distribution function of x_t^* at b conditional on $\mathbf{x}_{1k}$. (Mavroeidis 2021.)

When $\varepsilon_t \sim n.i.d._K(0,\Sigma)$, the kinked VAR model can be given an analytic likelihood function. The function $f(\cdot)$ of Equation (25) is identical to the one in Section 2.1 if there is no censoring. However, in the presence of censoring, likelihood is best split to two, as is done in Equation (25). The marginal density of non-censored variables is

$$f_{1t}(\mathbf{x}_{1t};\theta) = \int_{-\infty}^{\infty} f_t(\mathbf{x}_t;\theta)dx_t^* = \frac{1}{(2\pi)^{(K-1)/2}} \det(\Sigma_{11})^{-1} \exp\Big(\frac{-1}{2}(\varepsilon_{1t}'\Sigma_{11}^{-1}\varepsilon_{1t})\Big), \tag{26}$$

where Σ_{11} is Σ without the row and column corresponding to x_t^* and ε_{1t} is the error term of $\mathbf{x}_{1t}$. ε_{1t} can be written $\varepsilon_{1t} = \mathbf{x}_{1t} - \nu_1 - A_{1,1}\mathbf{x}_{t-1} - A_{1,2}\mathbf{x}_{t-2} - \ldots - A_{1,2}\mathbf{x}_{t-p}$, where $A_{1,i}$ is A_i with the row corresponding to x_t^* removed and ν_1 is ν with the component corresponding to the intercept ν^* of x_t^* removed. (Mavroeidis 2021.) In the case of the three variable kinked VAR for (π_t, y_t, i_t), $x_t^* = i_t^*$ and $\mathbf{x}_{1t} = (\pi_t, y_t)'$. Furthermore, Σ_{11} is Σ without the third column and row, ε_{1t} is ε_t without the third component, $A_{1,i}$ is A_i without the third row and ν_1 is ν without the third component.

The conditional cumulative distribution function is somewhat more difficult. Fortunately, $\varepsilon_t^*|\varepsilon_{1t}$, where ε_t^* is the error term of x_t^*, is normal if ε_t is normal. Its expected value and variance can then be derived from conditional moments of the multivariate normal distribution. Its expected value or fitted value is

$$\hat{x}_t^* = \nu^* + A_{2,1}\mathbf{x}_{t-1} + A_{2,2}\mathbf{x}_{t-2} + \ldots + A_{2,p}\mathbf{x}_{t-p} + \Sigma_{12}'\Sigma_{11}^{-1}\varepsilon_{1t}, \tag{27}$$

where $A_{2,i}$ is the row of A_i that corresponds to x_t^* and Σ_{12} is a vector of the covariances of

ε_{1t} and ε_t^*. In the three variable case considered here, $A_{2,i}$ is the bottom row of A_i, Σ_{12} is the third column of Σ without the third component and Σ_{11} is the 2×2 top left corner of Σ. Likewise the conditional variance is

$$\sigma^{*2} = \Sigma_{22} - \Sigma_{12}' \Sigma_{11}^{-1} \Sigma_{12}. \tag{28}$$

This leads to cumulative distribution function

$$F_t(b; \mathbf{x}_{1t}, \theta) = \Phi\left(\frac{b - \hat{x}_t^*}{\sigma^*}\right). \tag{29}$$

(Mavroeidis 2021.)

5.2 Shock identification and impulse responses

The structural model – KSVAR in the terminology of Mavroeidis (2021) – can be derived easily:

$$\mathbf{x}_t^* = B_0^{-1} B_1 \mathbf{x}_{t-1} + B_0^{-1} B_2 \mathbf{x}_{t-2} + \ldots + B_0^{-1} B_p \mathbf{x}_{t-p} + B_0^{-1} \mathbf{w}_t, \tag{30}$$

where B_0^{-1} is a lower triangle matrix just as before. Note that since the monetary policy shock has an immediate impact only on i_t^* and i_t^* itself has an effect on the other variables π_t and y_t only if $i_t^* > b$, the monetary policy shock has no effect if $i_t^* \leq b$ after the shock has hit. This means that in the KSVAR expansionary monetary policy is ineffective when the ELB is binding. Unconventional monetary policy cannot change this. This is a consequence of the model structure and not of the parameter values or data. (Iwata and Wu 2006; Mavroeidis 2021.)

Iwata and Wu (2006) consider a slight modification of the model of Equation (30), where the growth of money supply is included as an additional monetary policy instrument. They define the monetary policy shock as a shock that doesn't have an immediate effect on y_t or π_t. In fact, since the number of shocks has to match the number of variables, this identification scheme leads to two monetary policy shocks. The first one has an immediate effect on both the interest rate and money supply and the second has an immediate effect only on money supply. They call the former money supply shock and the latter money demand shock. Since they focus on the supply shock, under their scheme the monetary policy shock is effective even when the interest rate is firmly stuck at the ELB.

The construction of GIRFs requires predictions. However, predicting with a kinked VAR is complicated by the nonlinear kinked structure of the model. More specifically, these difficulties arise from $\mathbb{E}_t[i_{t+1}^*] \neq \mathbb{E}_t[i_{t+1}]$ even if $i_t > b$ and $\mathbb{E}_t[i_{t+1}^*] > b$. To derive predictions for $\mathbf{x}_t$, start by formulating forecasts for $\mathbf{x}_t^*$:

$$\mathbb{E}_{t-1}[\mathbf{x}_t^*] = \nu + A_1\mathbf{x}_{t-1} + A_2\mathbf{x}_{t-2} + ... + A_p\mathbf{x}_{t-p}. \tag{31}$$

Then derive prediction for $\mathbf{x}_t$ using $\mathbb{E}_{t-1}[\mathbf{x}_t^*]$. The $\mathbf{x}_{1t}$ are obviously the same for both $\mathbb{E}_{t-1}[\mathbf{x}_t^*]$ and $\mathbb{E}_{t-1}[\mathbf{x}_t]$. However, $\mathbb{E}_{t-1}[i_t]$ is more challenging but an analytic solution can be derived from a truncated normal distribution and the relation $\mathbb{E}[A] = \mathbb{E}[A|B=b]P(B=b) + \mathbb{E}[A|B \neq b]P(B \neq b)$. Let $\hat{i}_t^* = \mathbb{E}_{t-1}[i_t^*]$:

$$\begin{aligned}\mathbb{E}_{t-1}[i_t] =& \left(\hat{i}_t^* + \frac{\phi(\frac{\hat{i}_t^*-b}{\sqrt{\Sigma_{22}}})}{\Phi(\frac{\hat{i}_t^*-b}{\sqrt{\Sigma_{22}}})}\sqrt{\Sigma_{22}}\right)\Phi\left(\frac{\hat{i}_t^*-b}{\sqrt{\Sigma_{22}}}\right) + b\Phi\left(\frac{b-\hat{i}_t^*}{\sqrt{\Sigma_{22}}}\right) \\ =& b + (\hat{i}_t^* - b)\Phi\left(\frac{\hat{i}_t^*-b}{\sqrt{\Sigma_{22}}}\right) + \phi\left(\frac{\hat{i}_t^*-b}{\sqrt{\Sigma_{22}}}\right)\sqrt{\Sigma_{22}}. \end{aligned} \tag{32}$$

Note that this is exactly the same equation as in the SRTSM of Wu and Xia (2016). Unfortunately, after deriving $\mathbb{E}_{t-1}[i_t]$, the Equation (32) cannot be used anymore as i_{t+1}^* doesn't have a normal distribution. This is rather common in non-linear time series models and the solution is to simulate many realizations of $\varepsilon_t, \varepsilon_{t+1}, ..., \varepsilon_{t+h}$, use these to construct possible future paths of $\mathbf{x}_t$ with the reduced form Equation (24) and take the average of these paths (Iwata and Wu 2006). Nevertheless, these functions can be used recursively to derive approximate forecasts with less computational power.

Once predictions for the kinked VAR have been laid out, constructing GIRFs becomes a straight forward exercise. The relevant point here is to apply the shock to $\mathbf{x}_t^*$. This is done the same way for the KSVAR as for regular SVAR:

$$\mathbb{E}_{t-1}[\mathbf{x}_t^*|\mathbf{w}_t] = \nu + A_1\mathbf{x}_{t-1} + A_2\mathbf{x}_{t-2} + ... + A_p\mathbf{x}_{t-p} + B_0^{-1}\mathbf{w}_t. \tag{33}$$

Since the shock $\mathbf{w}_t$ is the only stochastic term of the model, this fixes $\mathbf{x}_t|\mathbf{w}_t = \max\{b, \mathbf{x}_t^*|\mathbf{w}_t\}$. Because of this, $\mathbf{x}_{t+1}^*$ is normal and the Equation (32) can be used for the second step. However, after this simulation has to be used to derive $\mathbb{E}_{t-1}[\mathbf{x}_{t+h}|\mathbf{w}_t]$ when $h > 1$. Once the forecasts $\mathbb{E}_{t-1}[\mathbf{x}_{t+h}|\mathbf{w}_t]$ and $\mathbb{E}_{t-1}[\mathbf{x}_{t+h}]$ are constructed, their difference forms the GIRF of Equation (8). Confidence intervals can also be derived using the bootstrap algorithm described in Section 2.2.2. The only difference is that, since the residuals aren't observed except when $i_t > b$, it is safer to use parametric bootstrap and to sample ε_t^b from a centered multivariate normal distribution using the estimate for the covariance matrix $\hat{\Sigma}$ (Mavroeidis 2021). Since both the construction of $\mathbb{E}_{t-1}[\mathbf{x}_{t+h}|\mathbf{w}_t]$ and $\mathbb{E}_{t-1}[\mathbf{x}_{t+h}]$ and the bootstrap require simulation, this part of the analysis is rather computationally intensive.

5.3 Shadow target rate

Using the KSVAR framework, the shadow rate i_t^* can be estimated in a few different ways. Let $\hat{i}_t^*$ be the shadow rate estimate. The naive approach would be to use the one-step forecast of i_t^*, i.e. $\hat{i}_t^* = \mathbb{E}_{t-1}[i_t^*]$. This can be directly derived from the fitted values of the model. However, such a definition for $\hat{i}_t^*$ would not be consistent with the censoring rule of Equation (23), since as a forecast it does not include the error term and can rise above the ELB even when the observed rate does not. A more logically sound approach is to use $\hat{i}_t^* = \mathbb{E}_t[i_t^*]$ as the definition. With this, when the ELB isn't binding, $\hat{i}_t^* = i_t$ and when it is, $\hat{i}_t^*$ is the expected value of a truncated normal distribution with maximum of b. The expected value of the normal distribution is the same as in Equation (27) and variance is the same as in Equation (28). This leads to the conditional estimator for shadow rate when the ELB is binding:

$$\hat{i}^*_{t|i_t=b} = \mathbb{E}_{t-1}[i_t^*] + \Sigma_{12}'\Sigma_{11}^{-1}\hat{\varepsilon}_{1t} - \frac{\phi\left(\frac{b-\mathbb{E}_{t-1}[i_t^*]-\Sigma_{12}'\Sigma_{11}^{-1}\hat{\varepsilon}_{1t}}{\sqrt{\Sigma_{22}-\Sigma_{12}'\Sigma_{11}^{-1}\Sigma_{12}}}\right)}{\Phi\left(\frac{b-\mathbb{E}_{t-1}[i_t^*]-\Sigma_{12}'\Sigma_{11}^{-1}\hat{\varepsilon}_{1t}}{\sqrt{\Sigma_{22}-\Sigma_{12}'\Sigma_{11}^{-1}\Sigma_{12}}}\right)}\sqrt{\Sigma_{22} - \Sigma_{12}'\Sigma_{11}^{-1}\Sigma_{12}} \tag{34}$$

It is also possible to construct confidence intervals using a bootstrap similar to that of the GIRFs.

Unlike the series constructed by for example Wu and Xia (2016), the shadow rate estimates constructed using Equation (34) are best seen as the desired, target or Taylor rule implied rates that the central bank would have set if it was suddenly allowed to set the interest rate without being bound by the ELB. They are made with the assumption that the interest rate in the past periods were at or above the ELB. This means that they do not reflect the monetary policy stance nor are they counterfactuals reflecting a reality where interest rates aren't bound by the ELB.

5.4 Empirical results

Figure 3 reports the generalized impulse responses of a KSVAR model to a monetary policy shock. The shock is a 25bp shock to shadow rate that impacts in 2020:Q2. The model has 4 lags, an ELB of 5bp and inflation, output gap and interest rate as endogenous variables. Figure 3 also reports the 90% confidence intervals constructed by parametric bootstrap with a bootstrap sample of 500. In order to construct the forecasts necessary for GIRFs, 50 000 realizations were made of the point estimate and 10 000 of each bootstrap sample. In addition, the comparative linear SVAR IRFs are shown. These are the same as in figure 2.

Figure 3 shows that the KSVAR results in considerably different impulse responses to

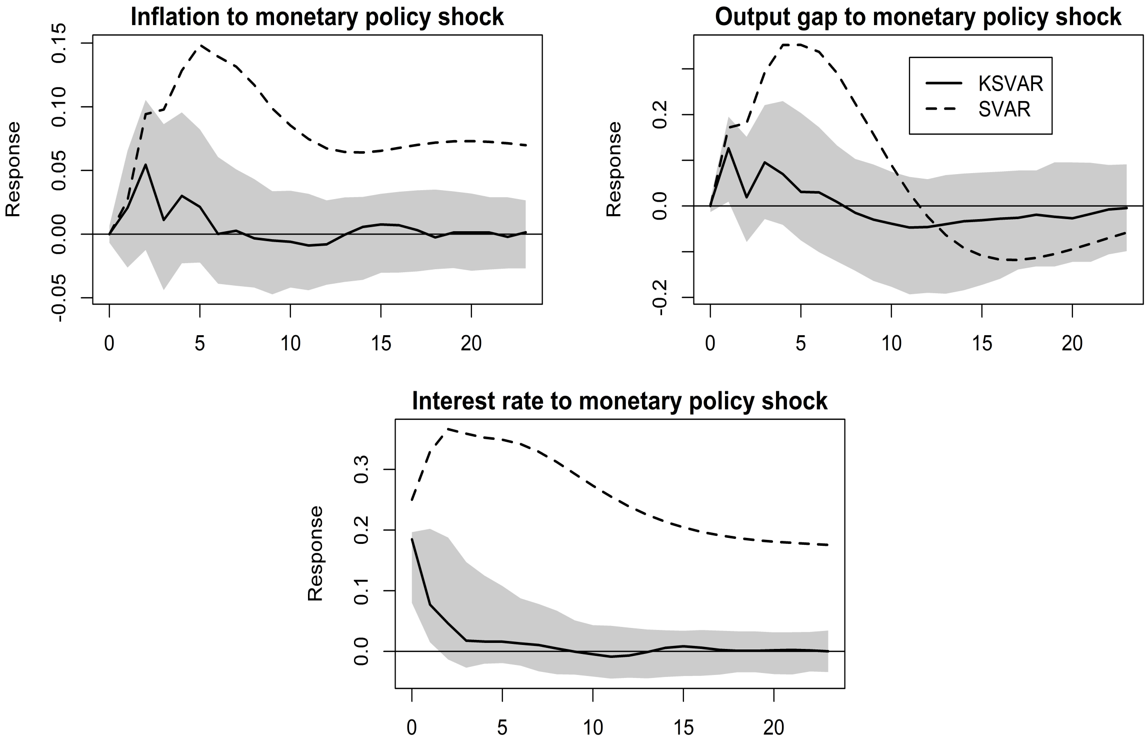

Figure 3: Impulse response functions of the monetary policy shock in the KSVAR framework.
Notes: Thick solid lines are the point estimates of the generalized impulse response functions to a 25bp monetary policy shock to the shadow rate in 2020:Q2 in the KSVAR model. The shaded areas are their 90% bootstrap confidence intervals with a parametric bootstrap sample of 500 simulations. Forecasting was done with 50 000 realization of the point esimate and 10 000 of each bootstrap sample. The dashed lines are the point estimates of the impulse responses in the linear model. The quarterly horizon is up to 24 quarters (6 years) and the thin horizontal solid line is at 0.

the standard linear SVAR. Firstly, since in 2020:Q2 the interest rate was at the ELB and thus the shadow rate was below it, the shock is considerably muted. The observed interest rate would have only experienced a shock of just below 20bp. In other words, a shock of about 5bp would have been necessary just to get out of the ELB. Secondly, since the effect of the monetary policy shock is muted, its effects on other variables are also somewhat muted. Thirdly, the effects of the shock are far less persistent. In fact, the effects of the monetary policy shock on the interest rate fade to statistical insignificance within a year. The direction and statistical significance of the impulse responses stay mostly the same. The prize puzzle is non-existent.

The shock is identified using the same Equation (22) that is used in Section 3.4. This identification means that the shock impacts i_t^*. If, as was the case in 2020:Q2, $i_t = b$ this shock can be inconsequential if it doesn't lift the interest rate out of the ELB. Furthermore, as i_t^* is unobserved and thus has to be derived using the model and each bootstrap sample estimates its own model used to estimate i_t^*, there is uncertainty as to what the immediate impact of the shock is. The shock is here chosen to match a 25bp shock to the shadow rate. Alternatively, one could scale the shock to correspond to a 25bp shock in the observed rate. This can be done analytically by solving $\mathbb{E}_{t-1}[i_t|\mathbf{w}_t] - \mathbb{E}_{t-1}[i_t] = 25\text{bp}$ for $\mathbf{w}_t$ using the Equation (32).

Iwata and Wu (2006) examine a very similar model using Japanese data. They find that the monetary policy shock increases inflation and output gap as one would expect. However, their analysis suffers from the fact that they do not consider confidence intervals and thus the significance of their results is unclear.

Figure 4 shows the KSVAR estimates for the shadow rate constructed using Equation (34), where the censoring threshold is 5bp shown with the thin horizontal line. Note that prior to 2014:Q3 the shadow rate and observed rate match. The greyed out area is the 90% confidence interval constructed using a 500 sample parametric bootstrap. From the figure, it seems that the target rate never ventured particularly far below the ELB. Of course, here the target rate is that which the central bank would have immediately adopted. Had the ELB been taken away, over time the rate could have ventured much further below the ELB. Also, the shadow rate appears to have been at its lowest in 2014, almost immediately after the MRO rate hit the chosen 5bp censoring threshold. This also coincides with the beginning of NIRP and APP. Instead in 2016, when NIRP reached its lowest level and APP peaked, the shadow rate experienced only a very moderate dip.

Figure 4: The KSVAR shadowrate estimates.
Notes: The thick line prior to 2014:Q3 is the observed MRO rate and the rate afterwards is the KSVAR estimate for the shadow rate constructed using the Equation (34). The thin horizontal line is at the censoring threshold of 5bp. The gray area is 90% confidence interval for the shadow rate constructed from 500 sample parametric bootstrap.

6 Censored structural vector autoregression

6.1 Censored vector autoregression

In contrast to the model of Equation (24) the unbounded shadow rate i_t^* appears on both sides of the censored vector autoregression model:

$$\mathbf{x}_t^* = \nu + A_1\mathbf{x}_{t-1}^* + A_2\mathbf{x}_{t-2}^* + ... + A_p\mathbf{x}_{t-p}^* + \varepsilon_t. \tag{35}$$

In other words, we are dealing with a rather conventional VAR model, like that of Equation (1) with the only difference that in practice $\mathbf{x}_t^*$ isn't always observed. In some sense, the process is unchanged and the only challenges come about in fitting the model. In the terminology of Mavroeidis (2021), this is a censored VAR model.

Start with the assumption that every $\mathbf{x}_{t-1}^*, ..., \mathbf{x}_{t-p}^*$ is observed. It is evident that the likelihood function of such an observation is the same as that of kinked VAR. The difficulty, however, lies in the observations where at least one $\mathbf{x}_{t-1}^*, ..., \mathbf{x}_{t-p}^*$ isn't observed. Ideally, one would derive the likelihood of $\mathbf{X}$ using marginal distributions:

$$L(\mathbf{X};\theta) = \int_{-\infty}^{b} ... \int_{-\infty}^{b} f(\mathbf{X}^*)dx^*,$$

where $\mathbf{X}$ contains all observations $\mathbf{x}_t$, f is the joint density function of all observations given the parameter vector θ, x^* is all censored and thus unobserved x_t^* and $\mathbf{X}^*$ is $\mathbf{X}$ with these unobserved x_t^*. Obviously, this integral is far too unwieldy to use as is. Instead, following Mavroeidis (2021), I use the Sequential Importance Sampling (SIS) algorithm for dynamic tobit models of Lee (1999) adjusted for multivariate data.

Start by recalling notation $S = \{i|x_i \leq b, i = 1,2,...,N\}$ and rewriting the joint density function f in terms of individual observations:

$$\begin{aligned} f(\mathbf{X}^*) &= \prod_{i\in S} P(\mathbf{x}_{1i} \wedge x_i^* \wedge i \in S|\mathbf{X}_{i-1}^*) \times \prod_{i\notin S} f(\mathbf{x}_i|\mathbf{X}_{i-1}^*) \\ &= \prod_{i\in S} f(\mathbf{x}_{1i}|\mathbf{X}_{i-1}^*)F^*(b|\mathbf{x}_{1i},\mathbf{X}_{i-1}^*,x_i^*)g(x_i^*|\mathbf{x}_{1i},\mathbf{X}_{i-1}^*) \times \prod_{i\notin S} f(\mathbf{x}_i|\mathbf{X}_{i-1}^*) \\ &= \prod_{i\in S} f(\mathbf{x}_{1i}|\mathbf{X}_{i-1}^*)F^*(b|\mathbf{x}_{1i},\mathbf{X}_{i-1}^*)g(x_i^*|x_i^* \leq b,\mathbf{x}_{1i},\mathbf{X}_{i-1}^*) \times \prod_{i\notin S} f(\mathbf{x}_{1i}|\mathbf{X}_{i-1}^*), \end{aligned} \tag{36}$$

where $\mathbf{X}_i^*$ consists of $\mathbf{x}_t^*$, where $t = 1,2,...,i$. $g(x_i^*|\mathbf{x}_{1i},\mathbf{X}_{i-1}^*)$ is the conditional density of x_i^* given $\mathbf{x}_{1i}$ and $\mathbf{X}_{i-1}^*$. The idea is that g can be used as a kernel to generate x_i^*.

The top row of Equation (36) presents a conditional mixed joint density, $P(\cdot)$. It consists of the joint densities of the observed $\mathbf{x}_{1i}$ and unobserved x_i^* as well as the probability that the unobserved x_i^* is indeed unobserved. This is then broken into pieces two different ways. On the second row, x_i^* is generated from the non-truncated distribution and the probability that the generated x_i^* is unobserved is then made conditional on that generated x_i^*. However, since the density of x_i^* is of course 0 if $i \in S$, the probability that $x_i^* \leq b$ is also 0 and its complement is 1. In other words, $F(\cdot)$ degenerates into an indicator function. This leads to a non-smooth density, which may be computationally more intensive. On the bottom row, this order is flipped. First, the probability that x_i^* is unobserved is constructed and then x_i^* is generated from a truncated $g(\cdot)$ that ensures $x_i^* \leq b$. This way not only does the density become smooth but the estimator is also more efficient. (Lee 1999.) This is the approach I and Mavroeidis (2021) use.

The idea behind the likelihood simulation executed here, is to generate a sample of x^*, calculate the likelihood of $\mathbf{X}^*$ with each x^* and to sum them together. For this purpose, a method of sampling x^* must be established. Assuming that the error terms are normal, $g(x_i^*|x_i^* \leq b, \mathbf{x}_{1i}, \mathbf{X}_{i-1}^*)$ is the density function of a conditional truncated normal distribution with a maximum of b. The relevant moments of the untruncated conditional normal distribution have already been derived in the context of the KSVAR model in Section 5.1. The mean, however, requires a minor revision to Equation (27):

$$\hat{x}_t^* = \nu^* + A_{2,1}\mathbf{x}_{t-1}^* + A_{2,2}\mathbf{x}_{t-2}^* + ... + A_{2,p}\mathbf{x}_{t-p}^* + \Sigma_{12}'\Sigma_{11}^{-1}\hat{\varepsilon}_{1t}.$$

The conditional variance σ^{*2} of Equation (28) can be taken as is. Now let $u_t^{(j)}$, $j = 1,...,M$, be independent draws from a uniform distribution $U_t^{(j)} \sim U(0,1)$. Now the j:th draw of x_t^j from $g(x_t^*|x_t^* \leq b, \mathbf{x}_{1t}, \mathbf{X}_{t-1}^*)$ is given by

$$x_t^{*(j)} = \hat{x}_t^* + \sigma^*\Phi^{-1}\left(u_t^{(j)}\Phi\left(\frac{b - \hat{x}_t^*}{\sigma^*}\right)\right). \tag{37}$$

(Lee 1999; Mavroeidis 2021.)

The SIS algorithm itself can now be introduced. It returns the log-likelihood of parameters θ (Mavroeidis 2021):

1) Initialize the algorithm by setting $W_0^j = 1$ and $\mathbf{X}_0^{*j} = (\mathbf{x}_0^j, ..., \mathbf{x}_{-p+1}^j)'$ for all $j = 1,...,M$.

2) Simulate likelihood by repeating the following for each $t = 1,...,T$:

   a) For $j = 1,...,M$ construct incremental weights

$$\begin{aligned} w_{t-1|t}^j =& P(\mathbf{x}_t|\mathbf{X}_{t-1}^{*j}, \theta) \\ =& \left[f(\mathbf{x}_{1t}|\mathbf{X}_{t-1}^{*j})F^*(b|\mathbf{x}_{1t}, \mathbf{X}_{t-1}^{*j})\right]^{1(t\in S)} f(\mathbf{x}_t|\mathbf{X}_{t-1}^{*j})^{1(t\notin S)}, \end{aligned}$$

   where $1(a)$ is an indicator function that returns 1 if the statement a is true and

0 otherwise.

b) Create an approximation for the likelihood of the observation t:

$$S_t = \frac{1}{M}\sum_{j=1}^{M} w^j_{t-1|t} W^j_{t-1}.$$

c) For $j = 1,...,M$ set $\mathbf{X}^{*j}_t = (\mathbf{x}^{*j}_t, \mathbf{X}^{*j'}_{t-1})'$, where $\mathbf{x}^{*j}_t = \mathbf{x}_t$ if $t \notin S$ and otherwise $\mathbf{x}^{*j}_t = (\mathbf{x}'_{1t}, x^{*j}_t)'$, where x^{*j}_t is drawn from $g(x^*_t | x^*_t \leq b, \mathbf{x}_{1t}, \mathbf{X}^{*j}_{t-1})$ using Equation (37).

d) For $j = 1,...,M$ update the weights:

$$W^j_t = \frac{w^j_{t-1|t} W^j_{t-1}}{S_t}$$

3) Return the log-likelihood approximation:

$$\hat{l}(\mathbf{X};\theta) = \sum_{t=1}^{T} \log(S_t).$$

The first step of the SIS algorithm sets the initial x^*_t for $t \leq 0$ to their observed values and makes the weight of every simulation equal ($W^j_0 = 1$ for all j). The second step loops through the data and does the real work. The step a) first constructs the likelihoods $w^j_{t-1|t}$ of observation t with all the simulated paths j. Then in step b) the observation itself is given a likelihood S_t by taking a weighted average of all $w^j_{t-1|t}$. These weights W^j_{t-1} add up to M and are themselves the weighted likelihood of path j of the previous period $t-1$: $w^j_{t-2|t-1} W^j_{t-2}$. This weighing has to be done since SIS doesn't only simulate single observations, x^*_t, but rather paths of simulations, $..., x^*_{t-1}, x^*_t, x^*_{t+1}, ...$. Then in step c) if necessary x^*_t is simulated and in step d) the weights are updated. Finally, the last step returns the log-likelihood of the whole sample. If u^j_t are kept constant all through out optimization process, $\hat{l}(\theta)$ is a smooth function, making the computations easier. (Mavroeidis 2021.)

A possible challenge with SIS is that the sample may degenerate. This means that weights W^j_t become such that few of them are large and most are very close to 0. This is obviously undesirable as it would be equivalent to using very few simulations in the likelihood simulation. This is explored briefly in Appendix C.

6.2 Shock identification and impulse responses

Once the censored VAR model has been estimated, it is possible to do inference on the model by turning it into a CSVAR model. Note that since the dynamics of a censored VAR are completely captured by the Equation (35), which is linear, the methods for analyzing i_t^* are the same as for standard linear SVAR. This means that unlike for the KSVAR the impulse responses are global and linear for i_t^*.

Few problems do, however, arise when the point of interest is the observed rate i_t. One only has to construct the forecasts for i_t^* with and without the shock of interest, use the Equation (32) to move from i_t^* to i_t piece-wise at each horizon and then take the difference between the two. This is made even simpler by the linear structure of the underlying model, which means that, unlike for the KSVAR, simulation of forecast errors is unnecessary. In practice, however, the i_{t-i}^* for $i = 1, ..., p$ may not be observed. Mavroeidis (2021) avoids this issue by constructing the GIRFs at a period when the US interest rates were sufficiently high for him to consider them observed. In the euro area there is no such luck since 2014 or 2015 as the interest rates have only very recently and after the sample period surpassed 5bp. The easy solution would be to use observed rates and accept that the GIRFs are only approximations. A somewhat better approach is to use the generated shadow rate series i_t^*. Of course, when deriving bootstrap confidence intervals, one has to choose whether to keep these shadow rates constant or also bootstrap them. I opt for the latter.

6.3 Shadow rates

For the CSVAR, the shadow rate i_t^* is somewhat more difficult than for the KSVAR as $\mathbb{E}_{t-1}[i_t^*]$ relies on covariates that are unobservable, namely lagged i_t^* itself. A naive approach that may result in decent approximations would be to use the Equation (34) recursively; Start from $t = 1$, estimate $\hat{i}_{1+1}^*$ and move on to the next period. Whenever $i_t \leq b$ set $i_t^* = \hat{i}_{t|i_t=b}^*$. The problem with this approach is that once an ELB spell has lasted for more than one period, the distribution of i_t^* is no longer normal since it depends on $\hat{i}_{t-1|i_{t-1}=b}^*$ which follows a truncated normal distribution. Fortunately however, the SIS algorithm already returns many shadow rate estimates, x_t^{*j}. An inefficient approach would be to take a simple unweighted average of these. However, a more efficient approach is to use the SIS algorithm and weigh these by W_t^j: $i_t^* = \frac{1}{M}\sum_{j=1}^{M} W_t^j i_t^{*j}$. After all, W_t^j is an approximation for the relative probability density of the j:th simulated path of ε up to period t. This does, however, mean that if the sample degenerates these estimates can rely on just a few realizations. This increases the variance of this estimator. Again bootstrap confidence intervals can be derived just as they are derived for GIRFs.

Unlike for the KSVAR, the CSVAR shadow rates are more similar to those of Wu and

Xia (2016) and others since it does reflect the monetary policy stance. However, unlike in studies that start by estimating shadow rate and only then construct SVAR models, here the shadow rate series is only a side product of the analysis even though the shadow rate is relevant for the model structure.

6.4 Empirical results

Figure 5 reports the generalized impulse responses of a CSVAR model to a monetary policy shock together with the impulse responses of other models considered. I consider a shock of 25bp to shadow rate that impacts in 2020:Q2. The model has 4 lags, an ELB of 0.05 and inflation, the output gap and the shadow interest rate as endogenous variables. Figure 5 also reports the 90% confidence intervals constructed by parametric bootstrap with a bootstrap sample of 500. For estimation purposes, SIS used $M = 100$ for the point estimate and $M = 60$ for the bootstrap. Since interest rates were at the ELB in 2020:Q2, the true shadow rate must be estimated as well. These were simulated using a sample of 50 replications.

Figure 5 shows that the CSVAR results in considerably different impulse responses to the KSVAR model. Compared with it, the CSVAR's impulse responses are not muted by the ELB as only the unbounded shadow rate is relevant to the responses. In fact, a monetary policy shock leads to a strong, somewhat persistent and definitely significant increase in both inflation and output gap, whereas in the KSVAR the shock was mostly irrelevant. On the other hand the impulse responses aren't particularly far from those of the linear SVAR except in regards to statistical significance. Indeed, the monetary policy shock of 25bp increases inflation by about 0.12% at most, with this effect still remaining significant after 20 quarters at the 90% confidence level. This implies the presence of the price puzzle. Output gap too is increased by 0.4% at its peak, however the effect is no longer significant after 10 quarters at 90% confidence level. Note that since the underlying model of $\mathbf{x}_t^*$ is linear, the impulse responses of inflation, the output gap and the shadow rate are not state dependent. Only those of the observed rate in the bottom right corner are.

Figure 5 reports both the impulse responses to the observed rate and to the shadow rate. For convenience's sake, the interest rate responses in other models are reported in the same panel as the shadow rate impulse responses of the CSVAR model. The shadow rate experiences a very persistent increase, which seems almost permanently stuck at around 25bp. This is very similiar to the linear SVAR. Meanwhile, the responses of the observed rates are quite a bit different. Since at the time of impact in 2020:Q2 the shadow interest rate was well below the ELB, as can be seen in figure 6 below, the shock would have had to be exceptionally strong to lift the observed rate off the ELB. Indeed, a 25bp

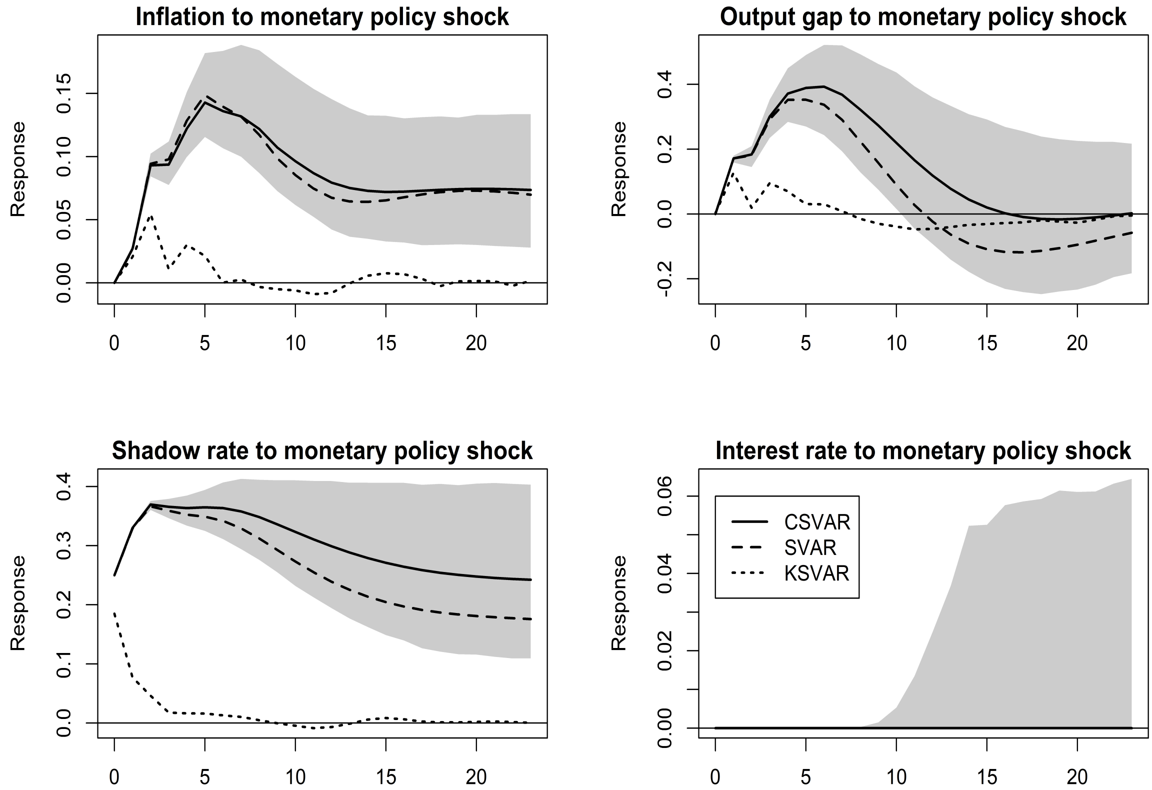

Figure 5: Impulse response functions of the monetary policy shock in the CSVAR framework.
Notes: Thick solid lines are the point estimates of the generalized impulse response functions to a 25bp monetary policy shock to the shadow rate in 2020:Q2 in the CSVAR model with an ELB at 5bp, $p = 4$ and $M = 100$. The shaded areas are their 90% bootstrap confidence intervals with a parametric bootstrap sample of 500 simulations, using $M = 60$ in SIS. The shadow rate is estimated using 50 simulations. The dashed lines are the point estimates of the impulse responses in the linear model and dotted lines are the point estimates of the generalised impulse responses in the KSVAR model. These are the same as in figures 2 and 2 respectively. For convenience, the interest rate responses of these models are in the same panel as the shadow rate responses of the CSVAR. The quarterly horizon is up to 24 quarters (6 years) and the horizontal thin black line is at 0.

shock seems to have been inadequate to do this and thus the observed interest rate would have stayed at the ELB regardless. Since both the forecast conditional on the shock and the unconditional forecast are the same, the GIRF is 0. Nevertheless, since the shock is very persistent, it does increase the probability of a lift off later on. This is why the upper bound of 90% confidence interval does raise above 0.

The results shown in figure 5 are quite different from those of Mavroeidis (2021). He considers both the CSVAR and the CKSVAR model with US data. With the CSVAR, while he does find that the monetary policy shock increases inflation in the short term this effect dissipates quickly. Similarly, unemployment increases in response to the monetary policy shock although with great delay, which would be equivalent to output gap decreasing. He broadly speaking repeats the same results with the CKSVAR(4) model. The only difference is that inflation does not increase at any horizon. Similar results have been achieved by others using models similar to the CKSVAR (see, for example, Hayashi and Koeda 2019; Aruoba et al. 2022; Ikeda et al. 2022).

Figure 6 shows the CSVAR estimates for the shadow rate, where the censoring threshold is 5bp shown with the thin horizontal line. The SIS algorithm used $M = 100$. Note that prior to 2014:Q3 the shadow rate and observed rate match. The greyed out area is the 90% confidence interval constructed using a 500 sample parametric bootstrap with $M = 60$. Also shown in the figure are the KSVAR shadow rate estimate of figure 4 in dashed line and Wu's and Xia's (2016) estimate for the shadow rate in dotted line[2]. The latter only starts in 2004:Q4 and the original monthly data is converted to quarterly by averaging. Vertical lines show the dates when various unconventional monetary policies were started. Note that this generally differs from the date when these policies reached their zenith.

From figure 6 it is clear that the CSVAR shadow rate, unlike the KSVAR shadow rate, has fallen well below the ELB. However, the shadow rate estimate of Wu and Xia (2016) isn't even included in the 90% confidence interval. It is clear that the CSVAR shadow rates have been significantly higher than those of Wu and Xia (2016). Indeed, whereas Wu and Xia (2016) shadow rate steadily fell from 2013 until 2019, the CSVAR shadow rate lacks a clear trend. This is of course partly due to the fact that the CSVAR assumes the series to be stationary.

In the figure 6, there is a similar and in fact larger dip in the CSVAR than in KSVAR shadow rate right after the MRO rate hit the censoring threshold in 2014. However, this dip is much smaller than the dip in 2015 and 2016 when the NIRP hit its lowest level and the APP started to peak. This is dip is almost non-existent in KSVAR. In general it appears that the launching of an unconventional monetary policy was followed by a fall in the shadow rate. There is also a large dip in 2020:Q2. Based on figure 1, this is probably due to the fall in output. Interestingly, there is no similar drop in either of the two other shadow rates.

---

[2]The series is available for the whole ELB period and is available at https://sites.google.com/view/jingcynthiawu/shadow-rates. Since the end of the ELB Wu recommends using conventional interest rates.

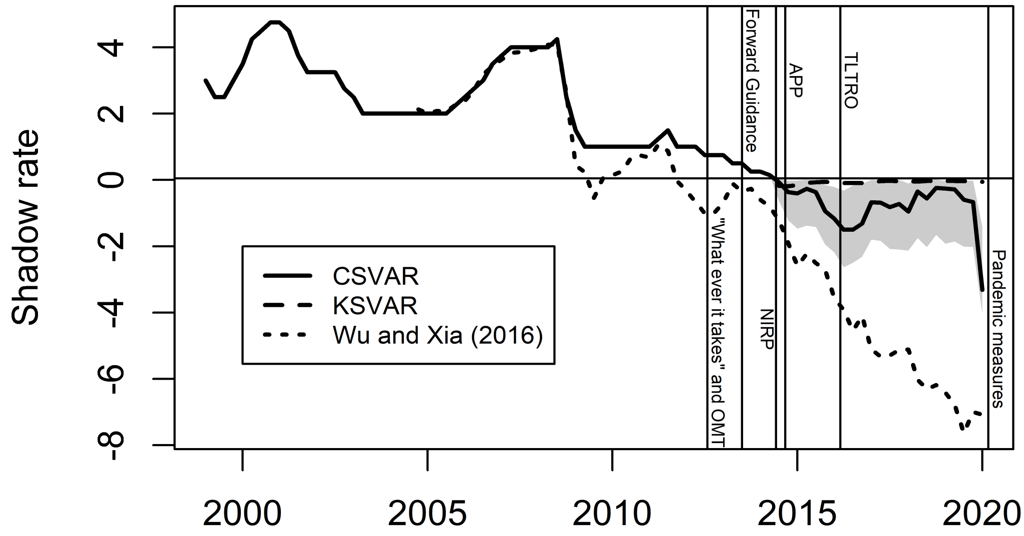

Figure 6: The CSVAR shadowrate estimates.
Notes: Thick line prior to 2014:Q3 is the observed MRO rate and the rate afterwards is the CSVAR estimate for the shadow rate constructed using SIS algorithm with $M = 100$. Thin horizontal line is at the censoring threshold of 5bp, the thick dashed line is the KSVAR shadow rate shown in figure 4 and the thick dotted line is the shadow rate estimate of Wu and Xia (2016). The gray area is 90% confidence interval for the shadow rate constructed from 500 sample parametric bootstrap using $M = 60$. Thin vertical lines indicate the starting dates of notable unconventional monetary policy measures (see table 1).

7 Conclusions

This thesis provided a general overview of structural vector autoregressive (SVAR) models and the identification and analysis of the monetary policy shock in the conventional framework that was popular prior to the low interest rate era that followed the Great Recession of 2007–2009. This analysis of the monetary policy shock is important since it means in essence the analysis of discretionary monetary policy. The modelling choices were motivated by a simple three-equation NK model. It was, however, found that the emergence of the ELB and unconventional monetary policy made this framework generally insufficient.

Two solutions developed by Mavroeidis (2021) were offered in this thesis: the KSVAR and CSVAR. The KSVAR is built on the assumption that interest rate simply cannot fall under the ELB. This introduces non-linearity to the model. Since no unconventional monetary policy instruments were included in the model, this means in practise that unconventional monetary policy is ineffective. Meanwhile, the CSVAR is built on the assumption that while the interest rate is unobservable below the ELB, it can effectively still go below it because of unconventional monetary policy measures. In practise, this means that the unconventional monetary policy is as effective as conventional and that the model is still linear. The differences of the three models are summarized in the table 2.

When it comes to the actual evaluation of the monetary policy shock in euro area, the choice of model matters substantially for the conclusions. The standard SVAR model leads one to believe that, while the point estimates for the impulse responses are large, the monetary policy shock is not a statistically significant driver of economic activity. The KSVAR leads to very similar conclusions but would suggest that even the point estimates are small and transitory. Meanwhile, the CSVAR leads to strikingly different conclusions; while the point estimates are very similar to the standard model, the impulse responses are very clearly statistically significant. Furthermore, the shocks appear very persistent.

Of the three models, CSVAR appears the most attractive as it accounts for both the ELB and unconventional monetary policy measures. However, the existence of the price puzzle, which is an anomaly, creates an aura of uncertainty around the exact effects of a

Table 2: Summary of the structural autoregressive models of this thesis.

Model	Is linear?	Accounts for... the ELB?	Accounts for... unconventional monetary policy?	Produces a shadow rate series?
SVAR	Yes	No	No	No
KSVAR	No	Yes	No	Yes
CSVAR	Yes	Yes	Yes	Yes

discretionary monetary policy decision. Overall, these results strongly suggest that further research into the effects of the monetary policy shock in the euro area is warranted.

Both the KSVAR and the CSVAR can be also be used to derive a shadow rate time series together with confidence intervals. These two are clearly different as the CSVAR shadow rate is consistently below the KSVAR one. Furthermore, both are clearly different from the shadow rate time series by Wu and Xia (2016). This would suggest that Wu and Xia's (2016) estimates may overestimate the expansionary effects of the unconventional monetary policy. Also the negative interest rate policy (NIRP) and the Asset purchase program (APP) seem to appear in the CSVAR shadow rate, which would suggest that these policies were indeed successful in easing the monetary policy.

This thesis did not consider the CKSVAR model of Mavroeidis (2021) that combines the two approaches largely because it relies on an identification strategy that is different from the one used in this thesis, and is thus not directly comparable. Future research could expand this study to include the CKSVAR. Additionally more variables could be added to the three model or the model could be built with non-Gaussian errors and regime-switching type dynamics, which could perhaps allow the investigation of the pandemic period. It is also possible that other macroeconomic shocks may have different effects in the ELB regime than outside it (see, for example, Bonam et al. 2022).

References

Asset purchase programs. <https://www.ecb.europa.eu/mopo/implement/app/html/index.en.html>, retrieved 21.3.2023.

Aruoba, S. B. – Mlikota, M. – Schorfheide, F. – Villalvazo, S. (2022) SVARs with occasionally-binding constraints. *Journal of Econometrics*, Vol. 231 (2), 477–499.

Barth III, M. J. – Ramey, V. A. (2001) The cost channel of monetary transmission. *NBER macroeconomics annual 2001* Vol. 16, eds. Ben S. Bernanke – Kenneth Rogoff, 199-256. University of Chicago Press, Chicago.

Benigno, P. – Rossi, L. (2021) Asymmetries in monetary policy. *European economic review*, Vol. 140, 103945.

Bonam, D. – De Haan, J. – Soederhuizen, B. (2022) The effects of fiscal policy at the effective lower bound. *Macroeconomic Dynamics*, Vol. 26 (1), 149–185.

Brunnermeier, M. – D. Palia – K. A. Sastry – C. Sims (2021) Feedbacks: financial markets and economic activity. *American Economic Review*, Vol. 111 (6), 1845–1879.

Cameron, A. Colin – Trivedi, Pravin K. (2009) Microeconometrics: method and application. 8th ed. *Cambridge University Press,* New York.

Canova, F. – G. De Nicoló (2002) Monetary disturbances matter for business fluctuations in the G-7. *Journal of Monetary Economics*, Vol. 49 (6), 1131—1159.

Castelnuovo, E. – Surico, P. (2010) Monetary policy, inflation expectations and the price puzzle. *The Economic Journal*, Vol. 120 (549), 1262–1283.

Christiano, L. J. – Eichenbaum, M. – Evans, C. L. (1999) Monetary policy shocks: What have we learned and to what end? In: *Handbook of macroeconomics* Vol. 1, eds. John B. Taylor – Michael Woodford, 65–148. Elsevier North Holland, Amsterdam.

Christiano, L. J. – Trabandt, M. – Walentin, K. (2010) DSGE models for monetary policy analysis. In: *Handbook of monetary economics* vol. 3, eds. Benjamin M. Friedman – Michael Woodford, 285–367. Elsevier North Holland, Amsterdam.

Clarida, R. – Galí, J. – Gertler, M. (1999) The science of monetary policy: a new Keynesian perspective. *Journal of economic literature*, Vol. 37 (4), 1661–1707.

Cúrdia, V. – Woodford, M. (2016) Credit frictions and optimal monetary policy. *Journal of Monetary Economics*, Vol. 84, 30–65.

Debortoli, D. – J. Galí – L. Gambetti (2019) On the empirical (ir)relevance of the zero lower bound constraint. In: *NBER Macroeconomics Annual 2019* vol. 34, eds. Martin S. Eichenbaum – Erik Hurst – Jonathan A. Parker, 141–170. University of Chicago Press, Chicago.

Dell'Ariccia, G. – Rabanal, P. – Sandri, D. (2018) Unconventional monetary policies in the euro area, Japan, and the United Kingdom. *Journal of Economic Perspectives*, Vol. 32 (4), 147–172.

Driffill, J. (2016) Unconventional monetary policy in the euro zone. *Open Economies Review*, Vol. 27, 387–404.

Eggertsson, Gauti B. – Woodford, Michael (2003) The zero bound on interest rates and optimal monetary policy. *Brookings Papers on Economic Activity*, Vol. 34 (1): 139–235.

Estrella, A. (2015) The price puzzle and var identification. *Macroeconomic Dynamics*, Vol. 19 (8), 1880–1887.

Faust, J. (1998) The robustness of identified VAR conclusions about money. *Carnegie-Rochester conference series on public policy*, Vol. 49, 207–244.

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

Gust, C. – E. Herbst – D. Lopez-Salido – M. E. Smith (2017) The empirical implications of the interest-rate lower bound. *American Economic Review*, Vol. 107 (7), 1971—2006.

Hamilton, J. D. (2018) Why You Should Never Use the Hodrick–Prescott Filter. *Review of Economics and Statistics*, Vol. 100 (5), 831-–843.

Hansen, J. (2018) Optimal monetary policy with capital and a financial accelerator. *Journal of Economic Dynamics and Control*, Vol. 92, 84–102.

Hanson, M. S. (2004) The "price puzzle" reconsidered. *Journal of Monetary Economics*, Vol. 51 (7), 1385–1413.

Hayashi, F. – Koeda, J. (2019) Exiting from quantitative easing. *Quantitative Economics*, Vol. 10 (3), 1069–1107.

Herbst, E. P. – F. Schorfheide (2015) Bayesian Estimation of DSGE Models. *Princeton University Press.*, Princeton.

Hutchinson, John – Mee, Simon (2020) The impact of the ECB's monetary policy measures taken in response to the COVID-19 crisis. In: *ECB Economic Bulletin*, Issue 5.

Ikeda, D. – Li, S. – Mavroeidis, S. – Zanetti, F. (2022). Testing the effectiveness of unconventional monetary policy in Japan and the United States. *arXiv preprint* arXiv:2012.15158.

Iwata, S. – Wu, S. (2006) Estimating monetary policy effects when interest rates are close to zero. *Journal of Monetary Economics*, Vol. 53 (7), 1395–1408.

Jarociński, M. (2010) Responses to monetary policy shocks in the east and the west of Europe: a comparison. *Journal of Applied Econometrics*, Vol. 25 (5), 833–868.

Jung, T. – Teranishi, Y. – Watanabe, T. (2005) Optimal monetary policy at the zero-interest-rate bound. *Journal of Money, credit, and Banking*, Vol. 37 (5), 813–835.

Kilian, L. – Lütkepohl, H. (2017) Structural Vector Autoregressive Analysis. *Cambridge University Press*, Cambridge.

Koop, G. – Pesaran, M. H. – Potter, S. M. (1996) Impulse response analysis in nonlinear multivariate models. *Journal of econometrics*, Vol. 74 (1), 119–147.

Krippner, L. (2013) Measuring the stance of monetary policy in zero lower bound environments. *Economics Letters*, Vol. 118 (1), 135–138.

Krippner, L. (2020) A note of caution on shadow rate estimates. *Journal of Money, Credit and Banking*, Vol. 52 (4), 951–962.

Lane, Philip R. (2022) The monetary policy strategy of the ECB: the playbook for monetary policy decisions. Speech at Hertie School, Berlin 2.3.2022. <https://www.ecb.europa.eu/press/key/date/2022/html/ecb.sp220302 8031458eab.en.html>, retrieved 21.3.2023.

Lanne, M. – Lütkepohl, H. (2008) Identifying monetary policy shocks via changes in volatility. *Journal of Money, Credit and Banking*, Vol. 40 (6), 1131–1149.

Lanne, M. – Lütkepohl, H. – Maciejowska, K. (2010) Structural vector autoregressions with Markov switching. *Journal of Economic Dynamics and Control*, Vol. 34 (2), 121–131.

Lee, L. F. (1999) Estimation of dynamic and ARCH Tobit models. *Journal of Econometrics*, Vol. 92 (2), 355–390.

Leeper, E. M. – Zha, T. (2001) Assessing simple policy rules: a view from a complete macroeconomic model. *Economic Review-Federal Reserve Bank of Atlanta*, Vol. 86 (4), 13–36.

Lütkepohl, Helmut (2006) New Introduction to Multiple Time Series Analysis. 1st ed. *Springer*, Berlin / Heidelberg.

Mavroeidis, S. (2021) Identification at the zero lower bound. *Econometrica*, Vol. 89 (6), 2855–2885.

Panetta, Fabio, Member of the Executive Board of the ECB. Interview 23.1.2023, conducted by Andreas Kröner, Jan Mallien and Frank Wiebe. <https://www.ecb.europa.eu/press/inter/date/2023/html/ecb.in230124 f50d72e488.en.html>, retrieved 21.3.2023.

Quast, J. – Wolters, M. H. (2022) Reliable real-time output gap estimates based on a modified Hamilton filter. *Journal of Business & Economic Statistics*, Vol. 40 (1), 152–168.

Rigobon, R. (2003) Identification through heteroskedasticity. *Review of Economics and Statistics*, Vol. 85 (4), 777–792.

Rossi, B. (2021) Identifying and estimating the effects of unconventional monetary policy: How to do it and what have we learned? *The Econometrics Journal*, Vol. 24 (1), C1–C32.

Rusnák, M. – Havranek, T. – Horváth, R. (2013) How to solve the price puzzle? A meta-analysis. *Journal of Money, Credit and Banking*, Vol. 45 (1), 37–70.

Sims, C. A. (1992) Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review*, Vol. 36 (5), 975–1000.

Stock, J. H. – Watson, M. W. (2001) Vector autoregressions. *Journal of Economic Perspectives*, Vol. 15 (4), 101–115.

Stock, J. H. – Watson, M. W. (2012) Disentangling the channels of the 2007-09 recession. *Brookings Papers on Economic Activity*, 120–157.

Swanson, E. T. – J. C. Williams (2014) Measuring the effect of the zero lower bound on medium- and longer-term interest rates. *American Economic Review* Vol. 104 (10), 3154–3185.

Taylor, J. B. (1993) Discretion versus policy rules in practice. *Carnegie-Rochester conference series on public policy*, Vol. 39, 195-–214.

Taylor, J. B. (1999) A historical analysis of monetary policy rules. In *Monetary policy rules* eds. John B. Taylor, 319–348. University of Chicago Press, Chicago.

Tobin, J. (1958) Estimation of relationships for limited dependent variables. *Econometrica*, Vol. 26 (1), 24–36.

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

Woodford, Michael (2002) Inflation Stabilization and Welfare. *The B.E. Journal of Macroeconomics*, Vol. 2 (1), 1–53.

Woodford, M. (2010) Optimal monetary stabilization policy. In: *Handbook of monetary economics* vol. 3, eds. Benjamin M. Friedman – Michael Woodford, 723–828. Elsevier North Holland, Amsterdam.

Wright, J. H. (2012) What does monetary policy do to long-term interest rates at the zero lower bound? *Economic Journal*, Vol. 122 (564), F447—F466.

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

Wu, J. C. – Xia, F. D. (2020) Negative interest rate policy and the yield curve. *Journal of Applied Econometrics*, Vol. 35 (6), 653–672.

Appendices

Appendix A Proof of Equation (22)

The goal is to prove that $C - DB_0A = 0$. The proof presented here is from Ikeda et al. (2022). Knowing the C and D matrices the condition is equivalent to

$$B_0A = \begin{bmatrix} b_{ii}/b_{im} & 0 & 0 \\ 0 & \phi_u & 0 \\ 0 & 0 & \phi_\rho \end{bmatrix}.$$

Furthermore knowing the matrices B_0 and A, the condition becomes $b_{yi} = b_{ym}\frac{b_{ii}}{b_{im}}$ and $b_{\pi i} = b_{\pi m}\frac{b_{ii}}{b_{im}}$. Substituting the Equation (19) into Equations (11) and (12) and focusing solely on the terms including i_{t-1} or m_t and containing the terms that include only u_t or ρ_t into $f_\pi(u_t, \rho_t)$ and $f_y(u_t, \rho_t)$ yields

$$\begin{aligned} \pi_t =& (\beta b_{\pi i} + \kappa b_{yi})i_{t-1} + (\beta b_{\pi i} b_{im} + \kappa b_{ym})m_t + f_\pi(u_t, \rho_t), \\ y_t =& (b_{yi} - \sigma + \sigma b_{\pi m})b_{ii}i_{t-1} + (b_{yi} - \sigma + \sigma b_{\pi i})b_{im}m_t + f_y(u_t, \rho_t). \end{aligned}$$

Since these expressions must match those of Equations (17) and (18), the conditions hold:

$$\begin{aligned} b_{\pi i} &= \beta b_{\pi i} + \kappa b_{yi}, \\ b_{\pi m} &= \beta b_{\pi i} b_{im} + \kappa b_{ym}, \\ b_{yi} &= (b_{yi} - \sigma + \sigma b_{\pi m})b_{ii}, \\ b_{ym} &= (b_{yi} - \sigma + \sigma b_{\pi i})b_{im}. \end{aligned}$$

The last two imply that $b_{yi} = b_{ym}\frac{b_{ii}}{b_{im}}$, and substituting this in to upper two yields $b_{\pi i} = b_{\pi m}\frac{b_{ii}}{b_{im}}$.

Appendix B Modified Hamilton filter of Quast and Wolters (2022)

The modified Hamilton filter of Quast and Wolters (2022) is

$$\begin{aligned} \hat{y}_t &= \frac{1}{9}\sum_{i=4}^{12} \hat{v}_{t,i}, \\ \hat{v}_{t,i} &= Y_t - \hat{\beta}_{0,i} - \hat{\beta}_{1,i}Y_{t-i} - \hat{\beta}_{2,i}Y_{t-i-1} - \hat{\beta}_{3,i}Y_{t-i-2} - \hat{\beta}_{4,i}Y_{t-i-3}, \end{aligned}$$

where $\hat{y}_t$ is the estimate for output gap or the cyclical component of any other time series, and Y_t is logarithmic GDP or any other I(1) time series that has a cyclical component. A time series is I(1) if $\Delta Y_t = Y_t - Y_{t-1}$ is stationary. The variable $\hat{v}_{t,i}$ is the forecast error of Y_t when it was predicted i periods ago with a simple AR(4) model.

The idea behind the standard Hamilton filter of Hamilton (2018) is that the logarithmic GDP is being forecasted i quarters ahead with a simple AR(p) model and the forecast errors are estimates for the output gap. For quarterly data, Hamilton (2018) suggests $i = 8$ and $p = 4$. The prior is based on the idea that forecast errors at this horizon forecast errors can mostly be attributed to cyclical variations, and it is a multiple of 4 to deal with any seasonality. The latter too is a natural choice for quarterly data to avoid problems associated with seasonality, but it also ensures a parsimonious model that can be fitted to small samples. Hamilton (2018) also provides a modification of the base model that can be used for any $I(d)$ times series, where $d \geq 1$. This means time series where $\Delta^d Y_t$ is stationary. In the modified Hamilton filter of Quast and Wolters (2022), the output gap is the average of forecast errors over different forecast horizons. This makes it more robust to business cycles of different frequencies, whilst retaining the main benefits of the original filter over its alternatives.

Appendix C Sample degeneracy in the CSVAR

Mavroeidis (2021) suggests using effective sample size (ESS) of Herbst and Schorfheide (2015) to gauge at the possibility of sample degeneracy. This is defined as:

$$\text{ESS}_t = \frac{M}{\frac{1}{M}\sum_{j=1}^{M}(W_t^j)^2}.$$

This equal to the inverted sum of squared relative weights $(W_t^j/M)^2$. Note that this results in a time series.

Figure 7 shows the effective sample size of Herbst and Schorfheide (2015) for the CSVAR with $M = 100$. This corresponds to the one used in section 6 to derive the point estimates. ESS was at its lowest in 2020:Q1 when it was about 5.43. Whilst this is low, it should be noted that this is the effective sample size only for 2020:Q1. The average ESS is about 80.20.

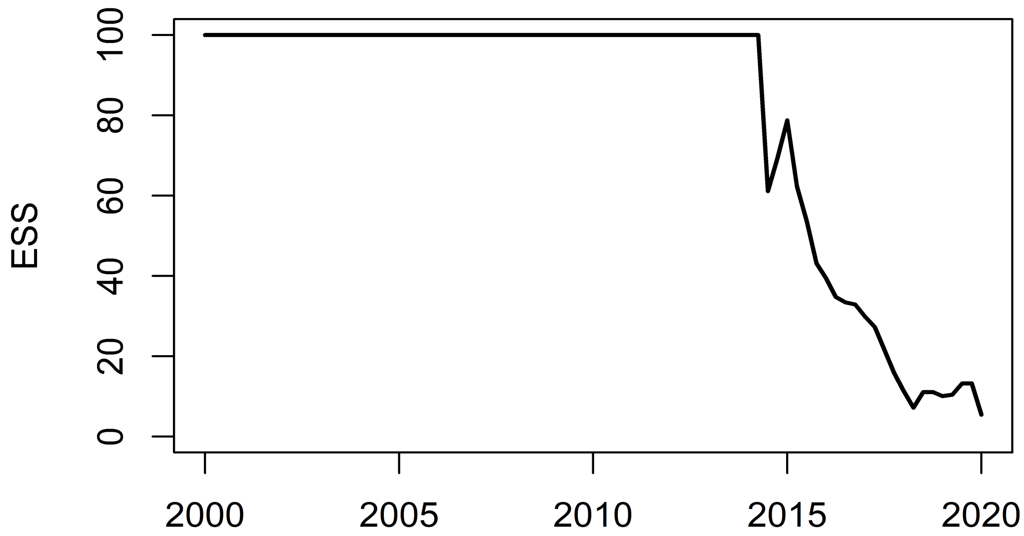

Figure 7: Effective sample size of Herbst and Schorfheide (2015) for the CSVAR with $M = 100$.