The Macroeconomic Effects of Conventional Monetary Policy Shocks in the Euro Area pre- and post-1999

Uriel Braham

March 31 2018

Contents

1 Introduction

In January 1999, eleven European countries agreed to form the European Economic and Monetary Union (EMU) with the Euro as its single currency, thereby delegating national monetary autonomy into the hands of the newly formed European Central Bank (ECB). By the Maastricht Treaty of 1992 aggregate price stability in form of an inflation rate of below, but close to, 2% over the medium-run was set as the ultimate objective. To achieve this, the ECB faces the challenge of tailoring an optimal monetary policy for a body of heterogenous economies as Euro members naturally differ in the stability of the transmission mechanism and are most likely in different stages of the business- and inflation cycle. It is therefore of utmost importance for the ECB to have a sound understanding of the monetary transmission mechanism within individual countries and on aggregate in the Euro Area. As more than 19 years have passed since the EMU was formed a sufficient amount of data is now available to study if and how the EMU and the corresponding economic integration changed the monetary transmission.

The focus of this paper is to empirically model and characterize the Euro Area aggregate macroeconomic dynamics in response to a conventional monetary tightening using a (synthetic) Euro Area dataset from 1980 to 2016.[1] One obvious difficulty in this investigation is the monetary regime switch as well as the corresponding integration process that occurred when delegating monetary policy to the ECB in 1999. Several studies (e.g. Peersman and Smets 2001; Monticelli and Tristani 1999) analyze the monetary transmission in the Euro Area relying mostly only on synthetic pre-EMU samples where aggregate monetary policy shocks were fictitiously identified as if an ECB had existed, or use only a very short sample with EMU data. Weber *et al.* (2009) estimate an area-wide VAR for 1980–2006 and conclude that there has been no considerable change in the transmission of monetary policy with the creation of the EMU. Furthermore, the authors identify data-driven and statistically a structural break in the transmission mechanism between 1996 and 1999. In contrast, Boivin *et al.* (2008) use a Factor Augmented VAR with a sample ranging from 1980–2007 and conclude that the EMU has indeed changed the transmission with an overall reduction in the impact on real output, prices and long-term interest rates.

The contribution of this paper is to update the inference on the monetary transmission in the Euro Area by adding an investigation of a larger sample after the creation of the EMU to document possible changes that occurred. The methodological approach followed is the estimation of a vector autoregression (VAR) model for two (sub-)sample periods: pre-1999 and post-1999. In consequence, estimation of the model is constrained by the relatively short

[1]The paper abstracts from the unconventional monetary policy measures (e.g. the Asset Purchase Programme) undertaken by the ECB since September 2015.

time series length on key macroeconomic variables. This paper uses a recursive identification scheme to identify an unanticipated monetary policy shock, and estimates impulse response functions that provide a useful description of dynamic effects on macroeconomic variables. To obtain reliable estimates, techniques of Bayesian inference will be applied to account for general parameter uncertainty in VAR models, the short sample length and to derive probabilistic statements about impulse responses.

The estimated impulse responses of conventional monetary policy shocks in the Euro Area before and after the adoption of the Euro deliver plausible dynamics, are consistent with stylized facts about monetary policy and are remarkably similar to related studies for the US (e.g. Christiano *et al.* 1999). Furthermore, the qualitative transmission to real activity and prices has not considerably changed. For both sample periods, an unexpected monetary contraction temporarily increases short-term interest rates, leads to a hump-shaped contraction of real output that peaks after 6-8 quarters and a reduction in the aggregate price level in the long-run. The interest rate passthrough in the term structure is sticky, but consistent with the expectations hypothesis. As documented by other studies (e.g. Musso *et al.* 2010; Jarocinski and Smets 2008) the monetary contraction leads to temporary decline in the real property prices. However, the quantitative impact and timing of significant responses changed considerably since 1999. The decline in real GDP roughly doubles and becomes more persistent. Prices are more flexible, respond faster and stronger in magnitude. From this perspective the introduction of the Euro had its desired effects in increasing real efficiency and price flexibility for monetary policy. Possible explanations for this can be derived when considering more specific transmission channels. Co-movement of real property prices with real output has increased considerably in magnitude and persistence with EMU. This provides evidence that house prices constitute an increasingly important channel in the transmission of monetary policy. Furthermore, the passthrough in the term structure to long-term interest rates has only gained slight more persistence since 1999.

The remainder of this paper is organized as follows. Section 2 reviews the econometric methodology of Bayesian Vector Autoregression (BVAR) models and provides an identification scheme for the structural model and the monetary policy shock. Section 3 describes the data employed and discusses related issues for the Euro Area. Section 4 studies the effects of monetary policy shocks in the Euro Area and documents changes in the transmission that occurred since 1999. Finally, section 5 concludes.

2 Econometrics of Bayesian VARs

This section reviews the purpose and econometric methodology of Bayesian methods for Vector Autoregression models. Vector Autoregression models are flexible, linear multivariate time series models, designed to capture the complex dynamics of multiple time series variables without imposing *a-priori* too rigid identifying assumptions. They are large, parameter-rich and by no means parsimonious. As time series data of typical macroeconomic variables, e.g. GDP growth, CPI inflation among others, involve mostly monthly, quarterly or annual observations the sample will only be of moderate size relative to the number of model parameters. For example, a model with 10 variables and 4 lags contains at least 400 distinct coefficients while a quarterly time series over 20 years would only offer 800 data points. In these environments, classical estimation methods such as Maximum Likelihood (MLE) and Ordinary Least Squares (OLS) fail to deliver precise estimates or are even infeasible when parameters can not be identified. Moreover, with noisy macroeconomic time series data issues of in-sample overfitting arising from over-parameterization constitute serious concerns that significantly reduces forecasting performance and precision in structural inference (Koop 2017).

These practical issues call for the desire to impose restrictions on model parameters. Shrinkage of parameters towards specific values by imposing prior restrictions greatly reduces the dimensionality and over-parameterization problems of VAR models, allows for more efficient parameter estimation and thus to greatly improved forecasting performance (Koop 2013). Bayesian methods provide a coherent and formally consistent framework to incorporate prior beliefs into estimation and inference. It also allows to use extraneous economic information on model parameters that would be impossible in a frequentist setting. Importantly, with Bayesian methods it is possible to make probabilistic statements about parameters of interest, e.g. coefficients, forecasts or impulse responses, given the sample and the stochastic model under consideration. Policy makers increasingly monitor the uncertainty measures of these estimates. These practical advantages lead to a widespread use of Bayesian methods in macroeconomic forecasting and structural modeling for policy analysis (Koop and Korobilis 2009).

2.1 *Econometric Methodology*

This section provides the theoretical background necessary for the estimation of a generic Vector Autoregression model. The evolution of M jointly modeled time series variables in $\mathbb{R}^M$ in its reduced-form is described by a p^{th}-order linear difference equation

$$\boldsymbol{y}_t = \sum_{k=1}^{p} \mathcal{A}_k \, \boldsymbol{y}_{t-k} + \boldsymbol{c} + \boldsymbol{\epsilon}_t, \quad t = 1, \ldots, T \tag{1}$$

with $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{Mt})'$ being the a $M \times 1$ vector of endogenous variables of interest; $\mathcal{A}_k$ $k = 1, \ldots, p$ denoting $M \times M$ autoregressive coefficient matrix corresponding to the k^{th} lag of the vector of endogenous variables, and $\boldsymbol{c} = (c_1, \ldots, c_M)'$ is the $M \times 1$ vector of intercepts. The $M \times 1$ vector of stochastic disturbances $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \ldots, \epsilon_{Mt})'$ is assumed to be white noise with expectation $\mathbb{E}[\boldsymbol{\epsilon}_t] = \mathbf{0}_M$ and variance-covariance matrix $\mathbb{E}[\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_s'] = \boldsymbol{\Sigma}_\epsilon$ for $s = t$ and $\mathbb{E}[\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_s'] = \mathbf{0}_{M \times M}$ for $s \neq t$. To fully characterize the distribution by which $\boldsymbol{y}_t$ is generated conditional on its past realizations $\boldsymbol{y}_{1-p:t-1} = \{\boldsymbol{y}_{1-p}, \ldots, \boldsymbol{y}_0, \ldots, \boldsymbol{y}_{t-2}, \boldsymbol{y}_{t-1}\}$ it is assumed that the stochastic disturbance is Gaussian distributed with $\boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}_M, \boldsymbol{\Sigma}_\epsilon)$. Hence by linearity of the VAR(p) model $\boldsymbol{y}_t$ is conditionally on its past realizations Normally distributed. Furthermore, it is assumed that all roots of the characteristic polynomial of Eq. (1) lie outside the unit circle ensuring stationarity of the stochastic process.

The reduced-form VAR model can also be concisely represented in terms of the matrix-variate Normal distribution. Therefore, define $\boldsymbol{Y} = (\boldsymbol{y}_{p+1}, \ldots, \boldsymbol{y}_T)'$ as a $T \times M$ matrix of the T observations on each of the M variables stacked in columns. The regressor vector for all endogenous variables in period t, $\boldsymbol{x}_t$, containing the p lags is then defined as $\boldsymbol{x}_t = (\boldsymbol{y}_{t-1}', \ldots, \boldsymbol{y}_{t-p}', 1)'$. Stacking all T observations on top of each other defines the $T \times K$ regressor matrix $\boldsymbol{X} = (\boldsymbol{x}_{p+1}', \ldots, \boldsymbol{x}_T')'$ where $K = 1 + Mp$ is the number of coefficients in each equation. Hence, with M endogenous variables the total number of coefficients in the VAR(p) is $k = MK = M(1 + Mp)$. To be consistent with the dimensions of $\boldsymbol{Y}$ the Gaussian white noise $\boldsymbol{\epsilon}_t$ is stacked such that $\mathcal{E} = (\boldsymbol{\epsilon}_{p+1}, \ldots, \boldsymbol{\epsilon}_T)'$. Finally, the $K \times M$ coefficient matrix $\mathcal{A}$ is defined as $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_p, \boldsymbol{c})'$. The VAR($p$) model in Eq. (1) is then represented by

$$\boldsymbol{Y} = \boldsymbol{X}\mathcal{A} + \mathcal{E}. \tag{2}$$

Since each equation in the VAR(p) contains identical covariates the model can also be reformulated in its vectorized form as

$$\boldsymbol{y} = (\boldsymbol{I}_M \otimes \boldsymbol{X})\boldsymbol{\alpha} + \boldsymbol{\epsilon}, \tag{3}$$

where $\boldsymbol{y} \equiv vec(\boldsymbol{Y})$, $\boldsymbol{\alpha} \equiv vec(\mathcal{A})$ and $\boldsymbol{\epsilon} \equiv vec(\mathcal{E})$ with $vec(\cdot)$ denoting the column stacking operator and $\boldsymbol{I}_M$ is the $M-$ dimensional identity matrix. The stochastic disturbance in vectorized form follows a multivariate Normal distribution such that $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}_M, \boldsymbol{\Sigma}_\epsilon \otimes \boldsymbol{I}_M)$.

The conditional density of the t^{th} observation is

$$p(\boldsymbol{y}_t | \boldsymbol{y}_{1-p:t-1}; \mathcal{A}, \boldsymbol{\Sigma}_\epsilon) = (2\pi)^{-M/2} |\boldsymbol{\Sigma}_\epsilon|^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{y}_t - \mathcal{A}'\boldsymbol{x}_t)'\boldsymbol{\Sigma}_\epsilon^{-1}(\boldsymbol{y}_t - \mathcal{A}'\boldsymbol{x}_t)\right\} \tag{4}$$

The conditional joint density, i.e. conditional likelihood function, for observations 1 through t, $\boldsymbol{y}_{1:T}$ conditional on pre-sample realizations $\boldsymbol{y}_{0:1-p}$ when assuming stochastically independent Gaussian errors is

$$p(\boldsymbol{y}_{1:T}|\boldsymbol{y}_{1-p:0};\mathcal{A},\boldsymbol{\Sigma}_\epsilon) = \prod_{t=1}^{T} p(\boldsymbol{y}_t|\boldsymbol{y}_{1-p:t-1};\mathcal{A},\boldsymbol{\Sigma}_\epsilon) \tag{5}$$

$$= \prod_{t=1}^{T}(2\pi)^{-M/2}|\boldsymbol{\Sigma}_\epsilon|^{-1/2}\exp\left\{-\frac{1}{2}(\boldsymbol{y}_t - \mathcal{A}'\boldsymbol{x}_t)'\boldsymbol{\Sigma}_\epsilon^{-1}(\boldsymbol{y}_t - \mathcal{A}'\boldsymbol{x}_t)\right\} \tag{6}$$

$$\propto (2\pi)^{-TM/2}|\boldsymbol{\Sigma}_\epsilon|^{-T/2}\exp\left\{-\frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{y}_t - \mathcal{A}'\boldsymbol{x}_t)'\boldsymbol{\Sigma}_\epsilon^{-1}(\boldsymbol{y}_t - \mathcal{A}'\boldsymbol{x}_t)\right\}. \tag{7}$$

In terms of the matrix-variate reformulation of the model the conditional likelihood function can be re-expressed as

$$p(\boldsymbol{y}_{1:T}|\boldsymbol{y}_{1-p:0};\mathcal{A},\boldsymbol{\Sigma}_\epsilon) = (2\pi)^{-MT/2}|\boldsymbol{\Sigma}_\epsilon|^{-T/2}\exp\left\{-\frac{1}{2}tr[\boldsymbol{\Sigma}_\epsilon^{-1}(\boldsymbol{Y} - \boldsymbol{X}\mathcal{A})'(\boldsymbol{Y} - \boldsymbol{X}\mathcal{A})]\right\} \tag{8}$$

$$\propto |\boldsymbol{\Sigma}_\epsilon|^{-T/2}\exp\left\{-\frac{1}{2}tr[\boldsymbol{\Sigma}_\epsilon^{-1}(\boldsymbol{Y} - \boldsymbol{X}\mathcal{A})'(\boldsymbol{Y} - \boldsymbol{X}\mathcal{A})]\right\}. \tag{9}$$

Kadiyala and Karlsson (1997) show that the (conditional) likelihood function in Eq. (9) in its vectorized form can be decomposed into two parts: one distribution for α conditional $\boldsymbol{\Sigma}_\epsilon$ and y and another for $\boldsymbol{\Sigma}_\epsilon$ conditional y. This implies a Normal distribution for α given $\boldsymbol{\Sigma}_\epsilon$ and $\boldsymbol{Y}$

$$\boldsymbol{\alpha}|\boldsymbol{\Sigma}_\epsilon, \boldsymbol{Y} \sim \mathcal{N}(\hat{\boldsymbol{\alpha}}, \boldsymbol{\Sigma}_\epsilon \otimes (\boldsymbol{X}'\boldsymbol{X})^{-1}), \tag{10}$$

and an inverse Wishart distribution for $\boldsymbol{\Sigma}_\epsilon$ given $\boldsymbol{Y}$

$$\boldsymbol{\Sigma}_\epsilon|\boldsymbol{Y} \sim \mathcal{IW}(\boldsymbol{S}, \nu), \tag{11}$$

where $\hat{\boldsymbol{\alpha}} = vec(\hat{\mathcal{A}})$ with $\hat{\mathcal{A}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$ is the Maximum Likelihood estimate for $\boldsymbol{\alpha}$ and $\boldsymbol{S} = (\boldsymbol{X} - \boldsymbol{X}\hat{\mathcal{A}})'(\boldsymbol{X} - \boldsymbol{X}\hat{\mathcal{A}})$ and $\nu = T - K - M - 1$. A classical (frequentist) approach using unrestricted MLE or OLS would obtain these or similar quantities depending on the degrees of freedom adjustments.

From inspection of the estimated quantities in Eq. (10) and Eq. (11) it can be inferred that when the number of parameters (autoregressive coefficients and error variances and covariances) is large relative to the number of data points in the sample, statistical inference on the basis of confidence intervals and statistical tests will be imprecise due to low degrees of freedom. If the number of parameters in the VAR model exceeds the number of usable data points – which is possible in large VAR models – parameters in the likelihood function $p(\boldsymbol{y}_{1:T}|\boldsymbol{y}_{1-p:0};\mathcal{A},\boldsymbol{\Sigma}_\epsilon)$

can not be identified by classical methods. Hence, estimation and inference will be infeasible (Cameron and Trivedi 2005).

Bayesian methods offer solutions to these problems by incorporating of (subjective) prior information in the estimation. As Bayesian methods combine the likelihood function (sample density) with the prior density it is possible to obtain valid posterior densities in large models even if some parameters in the likelihood function are not identified. However, the prior information becomes increasingly important as the number of model parameters increases relative to the data points (Koop 2017). Therefore, as with any Bayesian regression model, sensible priors for VAR coefficients have to be elicited.

2.2 *Minnesota prior*

The first approach to incorporate prior information into VAR models was done by Doan *et al.* (1984) and Litterman (1986) and is known as the "Minnesota" or "Litterman prior". The Minnesota prior contains a set of data centric beliefs that shrink the VAR parameters toward a stylized representation of macroeconomic time series data that is based on long-run properties. The Minnesota prior tackles the problem of overfitting from over-parameterization by introducing shrinkage in on autoregressive coefficients. Doan *et al.* (1984) argue that most macroeconomic time series variables seem to be characterized by unit roots such that their changes behave like a random walk (with drift) when considered in levels. In addition, the authors believe that the own lags of a variable are a more important source to explain variation over time than the lagged values of foreign variables, and that more recent lags are more informative than more distant lags. The prior beliefs amount to different prior moments for the autoregressive coefficients. These Minnesota beliefs are summarized in the prior Normal distribution for the parameters of the model $\boldsymbol{\alpha} \sim \mathcal{N}(\underline{\boldsymbol{\alpha}}, \underline{\boldsymbol{\Omega}})$. How these priors are set will be elaborated in the following.

For the Minnesota prior, the prior mean has the reduced-form VAR representation of a random walk

$$\boldsymbol{y}_t = \boldsymbol{I}_{M\times M}\boldsymbol{y}_{t-1} + \boldsymbol{\epsilon}_t, \tag{12}$$

where the coefficient on the first own lag of each variable is assumed to have a mean of unity and the prior mean of all other coefficients is set to zero. A modification of the prior that allows for more flexibility with different measurements of variables is not to impose prior unit roots but rather to allow some time series variables to be characterized by mean reversion instead of high persistence. For example, with stationary (first) differenced data a white noise process

($\delta_i = 0$) for the respective equation is the appropriate prior to use. The identity matrix in (8) is replaced by $diag\{\delta_1, \ldots, \delta_M\}$. The first moment of the prior distribution for the autoregressive coefficients can then be summarized as follows

$$\mathbb{E}[(\mathcal{A}_k)_{ij}|\boldsymbol{\Sigma}_\epsilon] = \begin{cases} \delta_i, & j = i\,,\, k = 1 \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

where $(\mathcal{A}_k)_{ij}$ is the i, j^{th} element of $\mathcal{A}_k$, i.e. the coefficient on variable j in equation i at lag k, and δ_i denotes the prior mean such that $\mathbb{E}[(\mathcal{A}_1, \ldots, \mathcal{A}_p)|\boldsymbol{\Sigma}_\epsilon] = (diag\{\delta_1, \ldots, \delta_M\}, \mathbf{0}_M, \ldots, \mathbf{0}_M)$. The prior mean on the intercepts is $\mathbb{E}[(c_1, \ldots, c_M)|\boldsymbol{\Sigma}_\epsilon] = \mathbf{0}'_M$. It follows that the prior mean for $\boldsymbol{\alpha}$ is then $\mathbb{E}[\boldsymbol{\alpha}|\boldsymbol{\Sigma}_\epsilon] = \underline{\boldsymbol{\alpha}} = vec(\underline{\mathcal{A}}) = ((diag\{\delta_1, \ldots, \delta_M\}, \mathbf{0}_M, \ldots, \mathbf{0}_M)', (\mathbf{0}_M)')$.

Doan *et al.* (1984) assume that the importance of own and foreign variables in explaining a variables time series variation falls off in lag length. Hence, the prior variance on the autoregressive coefficients falls with lag distance of autoregressive coefficients. Assuming that own relative to foreign lags are more important in explaining variation it is possible to discriminate in the relative prior tightness. The second moment of the priors on the autoregressive coefficients takes the form of

$$\mathbb{V}[(\mathcal{A}_k)_{ij}|\boldsymbol{\Sigma}_\epsilon] = \begin{cases} \frac{\lambda_1^2}{k^2}, & j = i, k = 1, \ldots, p \\ \frac{\lambda_2^2 \sigma_{ii}}{k^2 \sigma_{jj}}, & j \neq i, k = 1, \ldots, p \end{cases} \tag{14}$$

The prior variance on the intercept is set to $\mathbb{V}[c_i|\boldsymbol{\Sigma}_\epsilon] = \lambda_3^2 \sigma_{ii}$, $i = 1, \ldots, M$. The term $1/k^2$ controls the rate at which the prior variance decreases with lag length k, $k = 1, \ldots, p$ and σ_{ii}/σ_{jj} controls for the possibility of differing variability and scale across variables in the model. The (relative) tightness of the prior on the parameters on own and foreign lags is determined by the hyperparameters λ_1 and λ_2, while λ_3 controls the tightness of the prior on the constant (and exogenous variables) and therefore govern the relative importance of the prior belief to the information in the sample data. Smaller values of these hyperparameters imply smaller prior variance and thus stronger shrinkage towards the prior means. If $\lambda_1 > \lambda_2$, foreign lags are shrunk to zero more strongly relative to own lags. The specification of a high-dimensional prior distribution is thus broken down into selecting only three hyperparameters.

Further, σ_{ii} is set to $\hat{\sigma}_i^2 = s_i^2$, the residual mean squared error (MSE) of a univariate autoregression of order p for $\boldsymbol{y}_{it}$ $i = 1, \ldots, p$. Furthermore, all coefficients $\boldsymbol{c}, \mathcal{A}_1, \ldots, \mathcal{A}_p$ are *a-priori* independent. Stacking all the prior variances for the coefficients in a way that is consistent with $\underline{\boldsymbol{\alpha}}$ yields the prior variance covariance matrix $\underline{\boldsymbol{\Omega}}$. The original Minnesota prior treats the error variance-covariance matrix $\boldsymbol{\Sigma}_\epsilon$ as known and to be of diagonal structure $\boldsymbol{\Sigma}_\epsilon = diag\{\sigma_1^2, \ldots, \sigma_M^2\}$. $\boldsymbol{\Sigma}_\epsilon$ is then replaced by an estimate $\hat{\boldsymbol{\Sigma}}_\epsilon$ based on the full sample where the diagonal elements in $\hat{\mathbf{\Sigma}}_\epsilon$ are set to the estimates of the mean squared error (MSE) of an univariate AR(p) regression of variable i. This restrictive assumption of uncorrelated errors is relaxed by using the estimate $\hat{\mathbf{\Sigma}}_\epsilon = (\boldsymbol{Y} - \boldsymbol{X}\hat{\mathcal{A}})'(\boldsymbol{Y} - \boldsymbol{X}\hat{\mathcal{A}})/(T-K)$. The posterior distribution for the parameter vector α is then obtained by combining the (conditional) prior distribution $p(\mathcal{A}, \mathbf{\Sigma}_\epsilon)$ with the sample likelihood $p(\boldsymbol{y}_{1:T}|\boldsymbol{y}_{1-p:0}; \mathcal{A}, \mathbf{\Sigma}_\epsilon)$ using Bayes' theorem (Koop 2003). With the assumption of a known (estimated) error variance-covariance matrix the Minnesota prior allows for a simple, tractable analytical solution for the posterior distribution of $\boldsymbol{\alpha}$ given the data $\boldsymbol{Y}$

$$\boldsymbol{\alpha}|\boldsymbol{Y} \sim \mathcal{N}(\overline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\Omega}}), \tag{15}$$

with $\overline{\boldsymbol{\Omega}} = [\underline{\boldsymbol{\Omega}}^{-1} + \hat{\mathbf{\Sigma}}_\epsilon^{-1} \otimes \boldsymbol{X}'\boldsymbol{X}]^{-1}$ and $\overline{\boldsymbol{\alpha}} = \overline{\boldsymbol{\Omega}}[\underline{\boldsymbol{\Omega}}^{-1}\underline{\boldsymbol{\alpha}} + (\hat{\mathbf{\Sigma}}_\epsilon^{-1} \otimes \boldsymbol{X}'\boldsymbol{X})\hat{\boldsymbol{\alpha}}]$ (Koop and Korobilis 2009).[2] The original specification of the Minnesota prior suffers heavily from the drawback that the error variance-covariance matrix $\mathbf{\Sigma}_\epsilon$ is treated as a know quantity. Uncertainty surrounding this parameter is therefore suppressed by assumption. Extensions of the Minnesota prior relax the assumptions on $\mathbf{\Sigma}_\epsilon$ and prior specifications. Kadiyala and Karlsson (1997) use the property of the likelihood function in Eq. (5) that it can be decomposed into a conditional Normal distribution for $\boldsymbol{\alpha}$ given $\boldsymbol{Y}$ and $\mathbf{\Sigma}_\epsilon$, and an inverse Wishart distribution for $\mathbf{\Sigma}_\epsilon$ given $\boldsymbol{Y}$ to generalize the Minnesota prior to treat $\mathbf{\Sigma}_\epsilon$ as an unknown and non-diagonal quantity. Thus VAR(p) model is treated in a truly Bayesian fashion where all model parameters are seen as random variables. They propose the modified Minnesota prior as a natural conjugate prior.[3] To obtain an analytical solution for the posterior densities of $p(\boldsymbol{\alpha}|\boldsymbol{Y})$ and $p(\mathbf{\Sigma}_\epsilon|\boldsymbol{Y})$ the prior on the model coefficients has to follow a conditional Normal distribution, and for the prior on the error variance-covariance matrix an inverted Wishart distribution has to be assumed

$$\boldsymbol{\alpha}|\mathbf{\Sigma}_\epsilon \sim \mathcal{N}(\underline{\boldsymbol{\alpha}}, \mathbf{\Sigma}_\epsilon \otimes \underline{\boldsymbol{\Phi}}) \quad \text{and} \quad \mathbf{\Sigma}_\epsilon \sim \mathcal{IW}(\underline{\boldsymbol{S}}, \underline{\nu}), \tag{16}$$

where elements $\underline{\boldsymbol{\alpha}}$, $\underline{\boldsymbol{\Phi}}$, $\underline{\boldsymbol{S}}$ and $\underline{\nu}$ will be functions of hyperparameters that have to be chosen. These can be set in such a way to match Minnesota moments discussed before. The variance-covariance matrix of $\boldsymbol{\alpha}$ is a special case of the original Minnesota variance-covariance matrix where $\mathbf{\Sigma}_\epsilon$ is diagonal and $\lambda_1 = \lambda_2 = \lambda$.

[2] If the prior precision matrix $\underline{\boldsymbol{\Omega}}^{-1}$ were set to a zero matrix as with uninformative prior (Jeffry's prior) the Bayesian approach would yield the same quantities as the unrestricted frequentist approach. Still, the interpretation is fundamentally different (Koop 2003).

[3] Natural conjugate priors are those where the prior, likelihood and posterior come from the same family of distributions (Koop 2003).

One desirable property of the natural conjugate prior is that it allows for analytical solutions of the conditional posterior distributions. As shown by Koop (2003) the natural conjugate prior can be interpreted as a sample from a fictitious population with the prior moments. Banbura *et al.* (2010) show that the Minnesota prior can thus be implemented by appending a set of artificial observations $\boldsymbol{Y}_d$ and $\boldsymbol{X}_d$ of length T_d to the matrix-variate representation of the reduced-form VAR in Eq. (2). This equivalent to imposing the Normal inverse Wishart prior with $\underline{\mathcal{A}} = (\boldsymbol{X}_d'\boldsymbol{X}_d)^{-1}\boldsymbol{X}_d'\boldsymbol{Y}_d$, $\underline{\boldsymbol{\alpha}} = vec(\underline{\mathcal{A}})$, $\underline{\boldsymbol{\Phi}} = (\boldsymbol{X}_d'\boldsymbol{X}_d)^{-1}$, $\underline{\boldsymbol{S}} = (\boldsymbol{Y}_d - \boldsymbol{X}_d\underline{\mathcal{A}})'(\boldsymbol{Y}_d - \boldsymbol{X}_d\underline{\mathcal{A}})$ and $\underline{\nu} = T_d - K$. To match the Minnesota moments the set of artificial observations is

$$\boldsymbol{Y}_d = \begin{pmatrix} diag\{\delta_1 s_1, \ldots, \delta_M s_M\}/\lambda \\ \mathbf{0}_{(Mp-M+1)\times M} \\ diag(s_1, \ldots, s_M) \end{pmatrix} \quad \boldsymbol{X}_d = \begin{pmatrix} J_p \otimes diag\{s_1, \ldots, s_M\}/\lambda & \mathbf{0}_{Mp\times 1} \\ \mathbf{0}_{1\times Mp} & 1/\lambda_3 \\ \mathbf{0}_{M\times Mp} & \mathbf{0}_{M\times 1} \end{pmatrix}, \tag{17}$$

where $\boldsymbol{J}_p = diag\{1, 2, \ldots, p\}$ and δ_i, $i = 1, \ldots, M$ denotes the prior mean on the first own lag of each variable. As with the original Minnesota prior s_i accounts for different scaling and variability of variables and is set to the root mean squared error (RMSE) from univariate AR(p) regressions for variable y_{it}, $i = 1, \ldots, p$. The first block of dummies represents the priors in the autoregressive coefficients, the second block on the intercepts (exogenous variables) and the last block corresponds to the prior on the variance-covariance matrix. Appending the artificial data to the model in Eq. (2) with $\boldsymbol{Y}_* = (\boldsymbol{Y}', \boldsymbol{Y}_d')'$, $\boldsymbol{X}_* = (\boldsymbol{X}', \boldsymbol{X}_d')'$ and $\mathcal{E}_* = (\mathcal{E}', \mathcal{E}_d')'$ then reads as $\boldsymbol{Y}_* = \boldsymbol{X}_*\mathcal{A} + \mathcal{E}_*$. To insure the existence of the prior expectation of $\boldsymbol{\Sigma}_\epsilon$ it is necessary to add an improper prior on $\boldsymbol{\Sigma}_\epsilon$, i.e. $\boldsymbol{\Sigma}_\epsilon \sim |\boldsymbol{\Sigma}_\epsilon|^{-(M+3)/2}$. This allows to obtain the conditional posterior distributions

$$\boldsymbol{\alpha}|\boldsymbol{\Sigma}_\epsilon, \boldsymbol{Y} \sim \mathcal{N}(\overline{\boldsymbol{\alpha}}, \boldsymbol{\Sigma}_\epsilon \otimes (\boldsymbol{X}_*'\boldsymbol{X}_*)^{-1}) \quad \boldsymbol{\Sigma}_\epsilon|\boldsymbol{Y} \sim \mathcal{IW}(\overline{\boldsymbol{S}}, \overline{\nu}), \tag{18}$$

with $\overline{\boldsymbol{\alpha}} = vec(\overline{\mathcal{A}})$ where $\overline{\mathcal{A}} = (\boldsymbol{X}_*'\boldsymbol{X}_*)^{-1}\boldsymbol{X}_*'\boldsymbol{Y}_*$, $\overline{\boldsymbol{S}} = (\boldsymbol{Y}_* - \boldsymbol{X}_*\overline{\mathcal{A}})'(\boldsymbol{Y}_* - \boldsymbol{X}_*\overline{\mathcal{A}})$ and $\overline{\nu} = T + T_d - K + 2$.[4]

With the (natural conjugate) Normal inverted Wishart prior only one hyperparameter λ can exist that determines shrinkage on all autoregressive coefficients and hence the tightness of the prior, compared to the original Minnesota prior where it was possible to discriminate between own and foreign lags. As $\lambda \to \infty$ the prior becomes uninformative such that posterior quantities coincide with MLE estimates, whereas when $\lambda \to 0$ posterior will equal the prior. The hyperparameter λ_3 that controls the shrinkage on the coefficients for the constant (and

[4]Appending the artificial observations resolves the matrix inversion problem arising in medium- and large-scale VAR models that arises with original Minnesota prior (Banbura *et al.* 2010)

exogenous variables) is commonly set to a large number (e.g. 10^4) reflecting a diffuse prior. The Minnesota prior as a natural conjugate is used in the estimation.

2.3 *Structural Analysis*

Without a structural representation of the reduced-form VAR in Eq. (1) any inference on the dynamics of the system to innovations (impulse response functions (IRF) or forecast error variance decompositions (FEVD)) is meaningless as the reduced-form error $\boldsymbol{\epsilon}_t$ in Eq. (1) is the one-step ahead forecast error. In general, each element in ϵ_t will reflect all fundamental economic shocks. Hence, there is no reason to presume that any element of $\boldsymbol{\epsilon}_t$ corresponds to a particular orthogonal economic shock, e.g. a monetary policy shock, and thus it cannot be interpreted structurally (Kilian and Lütkepohl 2017). The goal is to build a bridge between the reduced-form model and its structural representation. The identification of the structural VAR, the econometric approximation of the economy, permits to obtain the impulse responses of $\boldsymbol{y}_t$ to orthogonal economic disturbances that can be given meaningful economic interpretations. The parameters of the structural VAR are estimated from the reduced-form representation and with sensible identifying assumptions it is possible to recover the structural model.

The structural VAR corresponding to the reduced-form representation in Eq. (1) is provided by

$$\mathcal{D}_0 \boldsymbol{y}_t = \mathcal{D}_1 \boldsymbol{y}_{t-1} + \cdots + \mathcal{D}_p \boldsymbol{y}_{p-1} + \boldsymbol{\nu} + \boldsymbol{u}_t, \tag{19}$$

where $\mathcal{D}_i$, $i = 1, 2, \ldots, p$ are $M \times M$ autoregressive coefficient matrices and $\boldsymbol{\nu}$ is a $M \times 1$ vector of intercepts. $\boldsymbol{u}_t$ is the zero mean serially uncorrelated $M \times 1$ vector of structural innovations with the variance-covariance matrix $\mathbb{E}[\boldsymbol{u}_t \boldsymbol{u}_t'] = \boldsymbol{\Sigma}_u = \boldsymbol{I}_{M \times M}$. Hence, the number of structural shocks is equal to the number of endogenous variables in the model. They are mutually orthogonal which is necessary to give impulse response functions meaningful interpretations. The structural matrix $\mathcal{D}_0$ governs the contemporaneous relations among the variables in the structural VAR, and premultiplying Eq. (19) by $\mathcal{D}_0^{-1}$ relates the structural VAR to the the reduced-form VAR in (1) with $\boldsymbol{c} = \mathcal{D}_0^{-1} \boldsymbol{\nu}$, $\mathcal{A}_i = \mathcal{D}_0^{-1} \mathcal{D}_i$, $i = 1, \ldots, p$. Moreover, it establishes a necessary relationship between the reduced-form disturbances $\boldsymbol{\epsilon}_t$ and fundamental shocks $\boldsymbol{u}_t$ as $\boldsymbol{\epsilon}_t = \mathcal{D}_0^{-1} \boldsymbol{u}_t$. The structural model shows that the reduced-form one-step ahead prediction errors are a weighted average of the structural shocks, confirming the intuition that reduced-form error can not be given any structural interpretation and would provide a misleading picture of the actual dynamics of the economy (Kilian and Lütkepohl 2017).

The central question therefore is how to obtain an estimate for $\mathcal{D}_0^{-1}$ that permits to recover the structural shocks from the reduced-form errors using $\mathbb{E}[\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t'] = \boldsymbol{\Sigma}_\epsilon = \mathcal{D}_0^{-1}\mathcal{D}_0^{-1'}$ with $\boldsymbol{\Sigma}_u = \boldsymbol{I}_M$. Given that $\boldsymbol{\Sigma}_\epsilon$ is known (estimated or drawn) all the information about $\mathcal{D}_0^{-1}$ is contained in the relationship $\boldsymbol{\Sigma}_\epsilon = \mathcal{D}_0^{-1}\mathcal{D}_0^{-1'}$. However, as $\mathcal{D}_0^{-1}$ has M^2 distinct elements and the variance-covariance matrix $\boldsymbol{\Sigma}_\epsilon$ provides only $M(M+1)/2$ free parameters and there is an identification problem of the structural model.[5] At least (additional) $M(M-1)/2$ restrictions on $\mathcal{D}_0^{-1}$ have to be imposed to uniquely identify $\mathcal{D}_0^{-1}$ and thus the structural model from the information provided by reduced-form VAR. The most common approach to identify $\mathcal{D}_0^{-1}$ is to impose exclusion (zero) restrictions on selected elements of $\mathcal{D}_0^{-1}$.[6]

The identification of the structural model then allows the computation of the impulse response functions when fundamental economic innovations disturb the economy. Questions such as, what is the impact of a monetary policy shock in the form of a 100 basis point increase in the key short-term interest rate on the real economic activity, inflation and other macroeconomics variables of interest, can be given meaningful answers.

2.4 *Identification of the Monetary Policy Shock*

Identification of the structural model is obtained by using a recursive identification scheme based on a Cholesky decomposition of the reduced-form error variance-covariance matrix $\boldsymbol{\Sigma}_\epsilon$.[7] The focus of this paper lies on the identification of a monetary policy shock. Christiano *et al.* (1999) point out that it is important not to exclusively focus on the actions of monetary policy makers, i.e. observed changes in the monetary policy instrument, to identify monetary policy shocks. Actions might only reflect endogenous reactions of the monetary authority to developments in the economy like changes in the output gap or inflationary pressures. Exogenous movements in the monetary policy instrument have thus to be identified.

Therefore, the $M \times 1$ vector $\boldsymbol{y}_t$ is partitioned into three subsets: m_1 slow $(\mathcal{S}_t)$, m_2 fast $(\mathcal{F}_t)$ moving variables and the monetary policy instrument $(\mathcal{I}_t)$ with $M = m_1 + m_2 + 1$. The first group of slow moving variables corresponds generally to a set of real and price variables while the fast moving variables contain mostly financial variables. The endogenous variables are then ordered as $\boldsymbol{y}_t = (\mathcal{S}_t', \mathcal{I}_t, \mathcal{F}_t')'$. The recursive Wold causality chain that is implicitly postulated with this ordering states that the monetary authority's reaction function contains the current and lagged values of the slow moving variables $\mathcal{S}_t$ and only the lagged but not the

[5]See Kilian and Lütkepohl (2017) for a more thorough discussion of the identification problem in structural VARs.

[6]The normalizing assumption on the variance-covariance matrix $\boldsymbol{\Sigma}_u = \boldsymbol{I}_M$ is with out loss of generality when the diagonal elements in $\mathcal{D}_0$ remain unrestricted.

[7]The recursive identification scheme was first proposed by Sims (1980).

contemporaneous values of fast moving variables $\mathcal{F}_t$ in the information set in t. The central banks reaction function may then be summarized as

$$\mathcal{I}_t = f(\Omega_t) + u_{\mathcal{I},t}, \tag{20}$$

where Ω_t is the information set in t, $f(\cdot)$ is some linear function that relates Ω_t to the instrument $\mathcal{I}_t$ and $u_{\mathcal{I},t}$ is the monetary policy shock. Based on its information set Ω_t the monetary authority will respond to contemporaneous structural innovations in the slow moving variables but will respond only with a period delay to shocks of the fast moving variables. Any movement in the monetary policy instrument not explained by the information set in t is therefore identified as an exogenous monetary policy shock corresponding to $u_t^{\mathcal{I}}$.

Moreover, the recursive identification scheme implies that monetary policy shocks do not have a contemporaneous impact on the slow moving variables $\mathcal{S}_t$, e.g. real GDP and CPI inflation, but affect the fast moving variables $\mathcal{F}_t$, e.g. asset prices and interest rates, instantaneously.[8] Since this paper is only concerned with identifying a monetary policy shock the order of the variables within the group of slow and fast moving variables does not matter (Christiano *et al.* 1999).

3 Data Description

The variables employed mainly come from the 17th (most recent) update of the ECB's Area-wide Model (AWM) database maintained by the Euro Area Business Cycle Network (EABCN). The AWM database is a unique source on macroeconomic time series for the Euro Area at a quarterly frequency covering the period from 1970Q1 to 2016Q4. Since the Euro was only introduced as a common currency in January 1999 the historical time series pre-1999 are backdated using individual country information in a coherent aggregation scheme.[9] Financial variables for the Euro Area are retrieved from the ECB Statistical Warehouse and the Bank of International Settlements (BIS) database. Since for most of the financial variables observations before 1980 do not exist the sample period considered is restricted to 1980Q1–2016Q4.[10]

Examining monetary policy in the Euro Area over the entire sample period remains a

---

[8]Given the ordering of the variables the (lower triangular) Cholesky decomposition of $\mathbf{\Sigma}_\epsilon$ exactly identifies D_0^{-1} through $\mathbf{\Sigma}_\epsilon = \mathcal{D}_0^{-1'}\mathcal{D}_0^{-1} = \mathcal{P}'\mathcal{P}$ since $\mathcal{D}_0^{-1} = \mathcal{P}$ is a valid solution where $\mathcal{P}$ is the lower triangular Cholesky factor and thereby places $M(M-1)/2$ zero restrictions on $\mathcal{D}_0^{-1}$.

[9]Fagan *et al.* (2005) provide more details on the construction of the historical series.

[10]A more detailed description of the dataset and sources, including the information on the transformations, is provided in Appendix A.

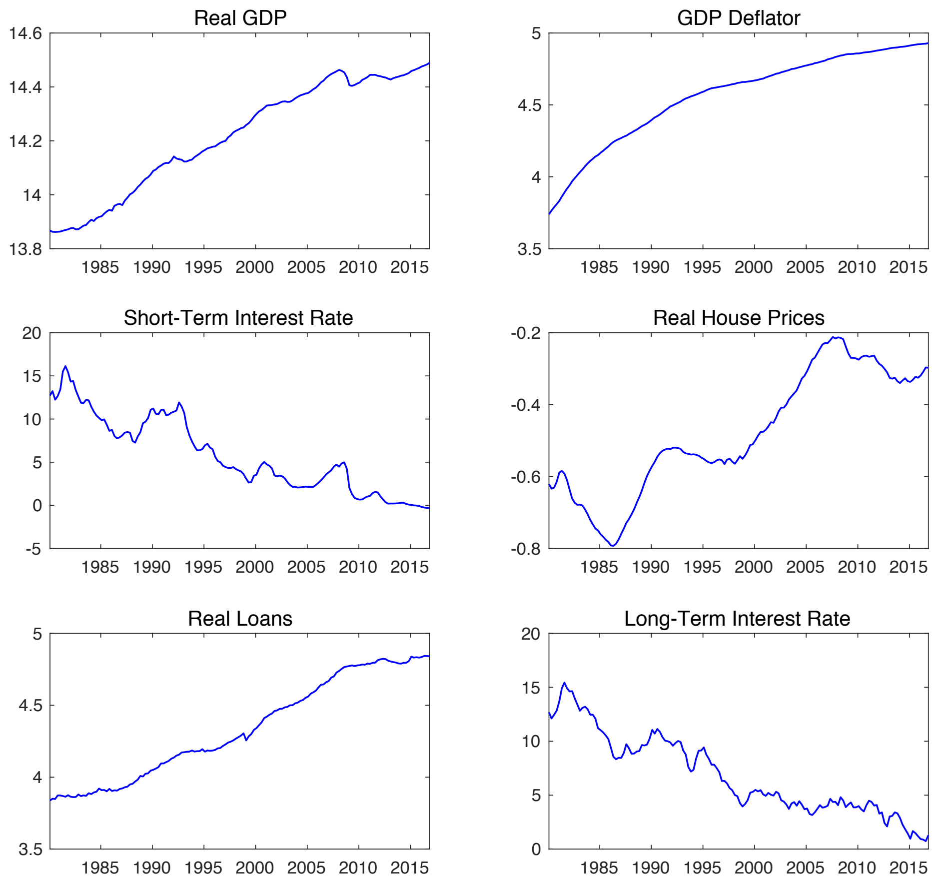
delicate issue. It has to be stressed that given EMU did not exist before 1999 no unified monetary policy for the entire sample period exists. Identification of the monetary policy shock using for example the nominal short-term interest rate is appropriate for the post-1999 sample period where the ECB is in place. The monetary policy shock pre-1999 is identified as a fictitious shock that would have been generated by the ECB if it had existed, and assumes an area wide aggregate (representative) monetary reaction function. This identification, however, may be inappropriate and mislead inference since each national central bank differed in their reaction function, i.e. preferences toward output and inflation gaps, and in principle had full autonomy over the short-term interest rates. Pre-EMU the German Bundesbank played a central role in setting nominal short-term interest rates for all countries participating in the European Monetary System (EMS) that pegged their currencies to the German Deutsch Mark. Therefore, most of the national central banks were constrained the Bundesbank's policy decisions to defend their parities. Assuming that the Bundesbank, the fictitious ECB and the ECB have similar reaction functions allows to meaningfully compare the pre-EMU and post-EMU impulse responses to monetary policy shocks.

Variables of special interest in the transmission of monetary policy are a measure of real economic activity, a measure of overall prices, a suitable monetary policy instrument and additional financial variables that potentially play a role in the propagation mechanism. For the benchmark model with six variables real GDP ($RGDP_t$) is used as an indicator of real economic activity, the price level is proxied by the GDP Deflator ($DGDP_t$), and following standard practice in the VAR literature the Euribor 3-month interest rate (I_t) serves as the monetary instrument. From theoretical work on the transmission mechanisms of monetary policy (e.g. Mishkin 1996, 2007) a real house price index (RHP_t) is used to proxy the asset price channel and the real loan volume to the private sector ($RLOANS_t$) adds the dimension of financial intermediation. Both variables combined possibly give rise to the workings of the balance-sheet channel as well as the credit channel and thus the financial accelerator that is perceived as a key feature of the transmission following Bernanke and Gertler (1995). The 10–year benchmark government bond yield ($GB10Y_t$) captures further aspects of long-term financing conditions of the traditional interest rate channel. With section 2.4 in mind the variables are grouped as follows: $\mathcal{S}_t$=$(RGDP_t, DGDP_t)'$, $\mathcal{I}_t = (I_t)$ and $\mathcal{F}_t$=$(RHP_t, RLOANS_t, GB10Y_t)'$. This identification structure allows output and prices only to respond with an one quarter delay to the monetary policy shock, whereas real house prices, real loans and the long-term interest rate are free to respond instantaneously. The model includes an intercept and a linear trend. In light of the quarterly frequency of the sample the lag order is set to $p = 4$.[11]

[11] The lag oder of $p = 4$ is also selected by the Akaike and Bayesian Information Criteria.

Variables except those already expressed in rates are in logarithms. There are several reasons to estimate the VAR in log-levels instead of first-differences or annualized growth rates that ensure stationarity for most variables. Using log-levels with co-integrating relationships between endogenous variables allows to consistently estimate the systems dynamics while differencing could discard information contained in the log-levels and may lead to misspecification. Further, non-stationarities of some variables does not pose a problem in a Bayesian framework, as the presence of unit roots does not alter the likelihood function and thus the posterior distribution (Sims *et al.* 1990). For non-stationary variables (in log-levels) the random walk prior ($\delta_i = 1$) and for stationary variables the white noise prior ($\delta_i = 0$) is assumed.

FIGURE 1: Benchmark model variables 1980–2016



NOTES: The figure depicts the time series of the variables used in the benchmark model for the period 1980Q1–2016Q4.

To mitigate the subjectivity of the prior elicitation for the overall shrinkage parameter λ Giannone *et al.* (2015) develop a hierarchical prior for the hyperparameters. λ is set according to the posterior mean of the medium-scale model in their analysis. Hence, $\lambda = 0.175$ and the shrinkage parameter λ_3 is set to a large number 10^4, reflecting a diffuse prior on the intercept and exogenous variables. How the choice of λ affects the results will be considered in a robustness analysis.

4 Results

This section analyzes the impulse response functions to a conventional monetary policy tightening in the Euro Area. To account for the issues presented in the last section various sample periods are considered: the entire sample period (1980Q1–2016Q4), the sample preceding the introduction of the Euro (1980Q1–1998Q4), and the sample period with the EMU in place (1999Q1–2016Q4).

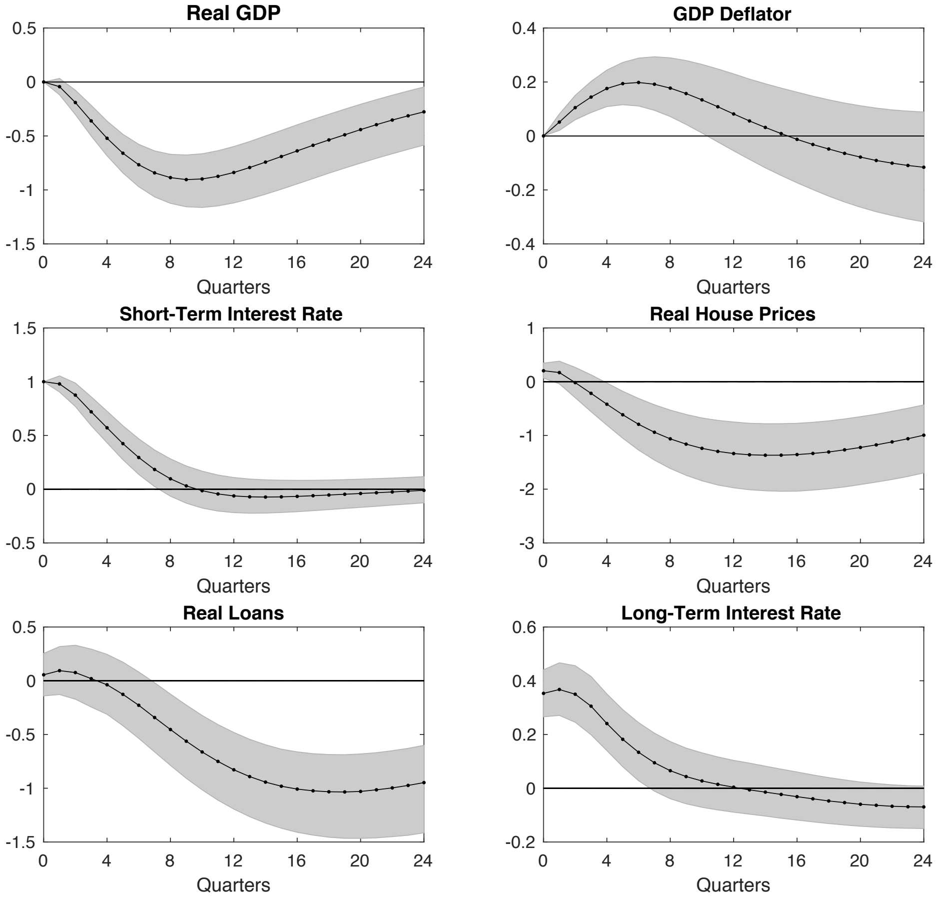
The impulse responses correspond to an unexpected increase to the Euribor 3-month interest rate which is normalized to unity, i.e. 100 basis points, to make the responses comparable over different sample periods. Impulse responses are traced out for 24 quarters following the shock period. The median (dotted line) with a 68% credible set enclosed by the 16^{th} and 84^{th} percentiles (grey shaded region) of the posterior distribution of the impulse responses are reported. For the interpretation, "significance" is referred to as the 68% credible set not containing a zero response. The responses of log-level variables are in percent deviations from baseline (trend), while the response of the rates are expressed in percentage points (at an annual rate).

4.1 *Period 1980–2016*

Figure 2 displays the impulse responses for the benchmark model with real GDP, the GDP Deflator, the short-term interest rate, real house prices, real loans to the private sector, and the long-term interest rate estimated over the entire sample period from 1980Q1 to 2016Q4.

The results are not fully in line with economic theory and stylized facts on the effects of monetary policy. Specifically, a monetary tightening significantly reduces real output in the medium-run with a median impact of -0.50% even after 24 quarters, contradicting in part the medium-run neutrality of money. Furthermore, a significant increase in the price level for the first 12 quarters following the shock period is observed. The literature refers to the positive response of prices to an increase of short-term interest rates as the "price puzzle". The short-term interest mimics its own positive shock, shows strong persistence and returns to its baseline level after 8 quarters. The passthrough in the term-structure in line with the expectation hypothesis as the long-term interest rates response is positive but half the impact magnitude of the short-term rate. It returns to baseline after 8 quarters and show very strong persistence. Real house prices and real loans to the private sector fall significantly after 8 and 4 quarters and decline in the long-run by about 1% (at the median). Similar responses are found by Weber *et al.* (2009) for the sample period 1980–2006.

FIGURE 2: Impulse Responses Benchmark 1980–2016

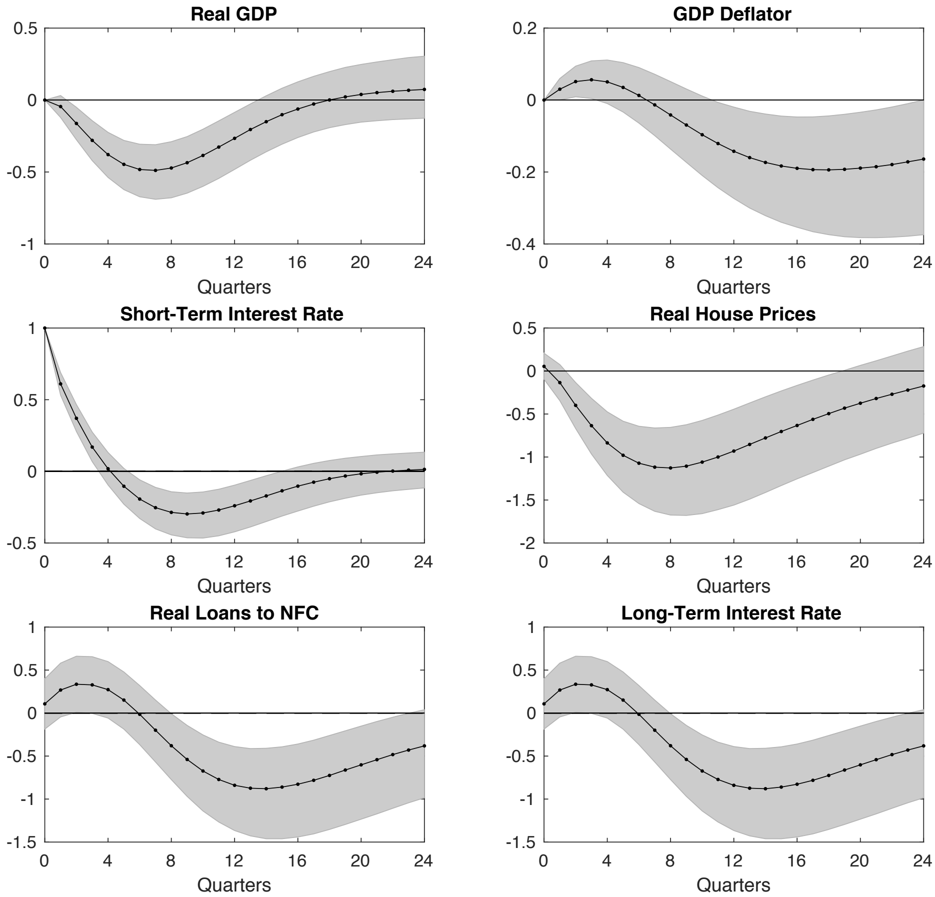


NOTES: The figure depicts the 50^{th} (dotted line), 16^{th} and 84^{th} percentiles (grey lines) of the posterior distribution of the impulse response functions to a 100 basis point shock to the short-term interest rate.

The first explanation for the non-neutrality of money and the "price puzzle" goes in the direction that the benchmark VAR model for the Euro Area is misspecified and thus inadequate in capturing macroeconomic dynamics. Specifically concerning the "price puzzle", Sims (1992) argues that simple low-dimensional VAR models that are only backward-looking may correspond to an insufficient description of the monetary authority's expectation about future inflation, thus misspecifying the reaction function that is used to identify the monetary policy shock. Sims (1992) proposes to include a commodity price index that should in part proxy expectations on future inflationary pressures. However, including a commodity price index as an endogenous or exogenous variable in the benchmark model does not resolve the "price puzzle". Since the sample period 1980Q1 to 2016Q4 is sufficiently long to estimate a larger VAR the benchmark model is augmented by various financial and price variables which typically appear in monetary VAR studies.[12]

[12]The variables are grouped as follows: $\mathcal{S}_t = (RGDP_t, DGDP_t, PCOM_t)'$, $\mathcal{I}_t = (I_t)$ and $\mathcal{F}_t = (RHP_t, M2_t, M3_t, EEN_t, RLNFC_t, RLHH_t, GB2Y_t, GB10Y_t)'$. Note that in this specification credit to

Figure 3a: Impulse Responses Augmented Benchmark 1980–2016



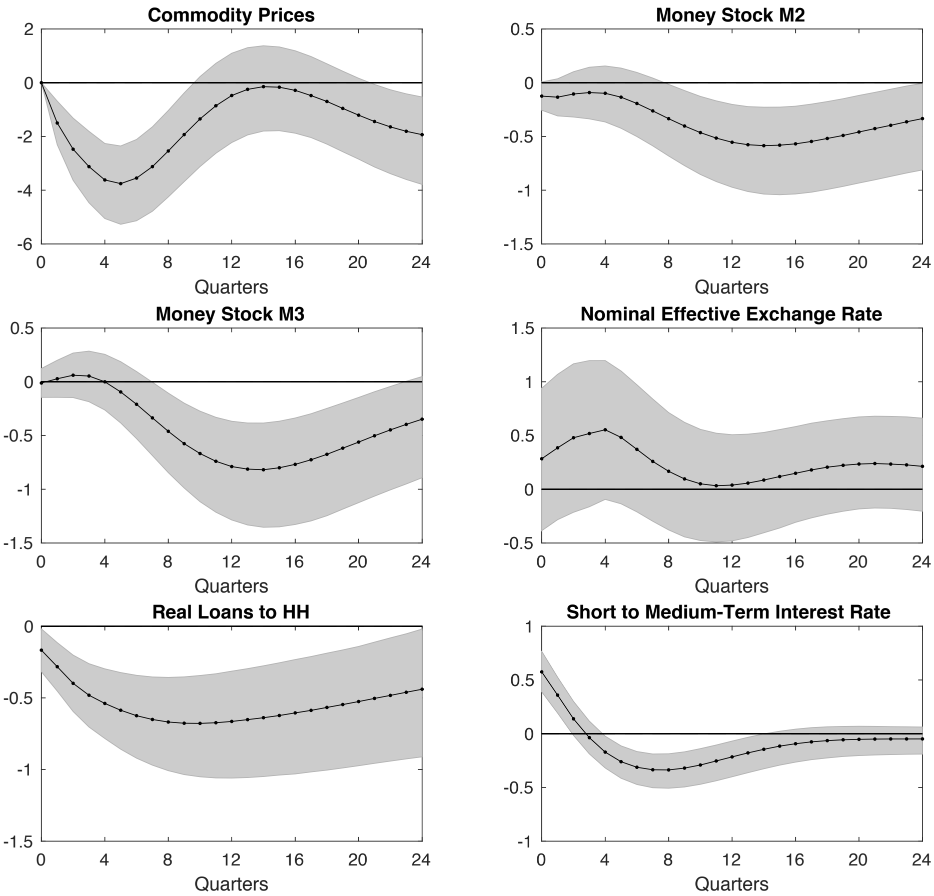
Notes: The figure depicts the 50th (dotted line), 16th and 84th percentiles (grey lines) of the posterior distribution of the impulse response functions to a 100 basis point shock to the short-term interest rate. The hyperparameter λ is set in accordance with Giannone *et al.* (2015) to $\lambda = 0.1$

Figure 3A and 3B depict the respective impulse responses for this specification. Increasing the set of variables in the monetary authority's reactions functions yields plausible responses. Specifically, the response of real output is now humped-shape with a maximum decline after 8 quarters that reverts back to the baseline level after 3 years. Augmenting the benchmark model can mitigate the "prize puzzle" as the response of prices falls significantly below zero after 12 quarters. The responses of the remaining variables are largely in line with other studies on the transmission of monetary policy in the Euro Area (e.g. Peersman and Smets 2001; McCallum and Smets 2007; Musso *et al.* 2010). However, for reasons discussed below inference based on this specification has to be taken with great caution.

The second explanation for the non-neutrality of money and the "price puzzle" in the benchmark model is more fundamental than just an inadequate macroeconomic modeling of the Euro Area monetary policy passthrough dynamics. Monetary policy before 1999, where national cen-

the private sector is split into households and non-financial corporations. See Appendix A for a detailed description of the variables.

Figure 3b: Impulse Responses Augmented Benchmark 1980–2016



Notes: The figure depicts the 50^{th} (dotted line), 16^{th} and 84^{th} percentiles (grey lines) of the posterior distribution of the impulse response functions to a 100 basis point shock to the short-term interest rate. The hyperparameter λ is set in accordance with Giannone *et al.* (2015) to $\lambda = 0.1$

tral banks were characterized by heterogenous reactions functions (preferences toward inflation and output gaps), and after the adoption of a single currency and common central bank, differed substantially. Identification of the monetary policy shock on the basis of an aggregate monetary authority's reaction function for the entire sample period (1980Q1–2016Q4) can raise misspecification doubts as a single constant coefficient VAR can not capture multiple monetary regimes. Furthermore, the transition into the Monetary Union that was also anticipated years in advance might perturbed the structure of the Euro Area economy. This intuition is underlined by Weber *et al.* (2009) who statistically identify a structural break in the transmission mechanism between 1996 and 1999. Not accounting for this significantly biases the results in low-dimensional models and thus inference.

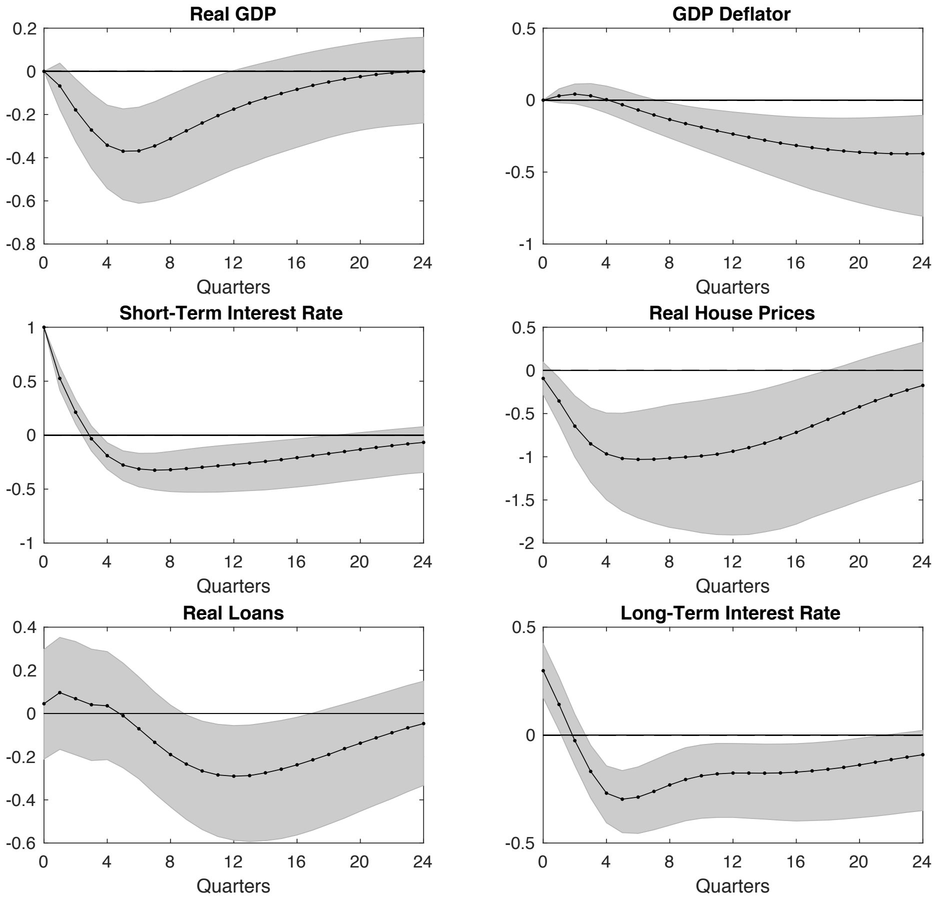
4.2 *Period 1980–1998 and 1999–2016*

To document if there have been changes in the monetary transmission mechanism before and after the creation of the EMU in 1999 the impulse responses for the benchmark model are re-estimated for the sample periods 1980Q1–1998Q4 and 1999Q1–2016Q4. These are depicted in the Figure 4 and Figure 5 respectively. In contrast to Figure 2, the impulse responses in the benchmark model for both sub-samples display plausible, well-behaved dynamics of the endogenous variables to a contractionary monetary policy shock. Specifically, neutrality of money on real GDP is restored, the significant "prize puzzle" vanishes and a significant long-run decline in the prices level is observed. Real GDP temporarily falls with the negative effect for both samples peaking roughly after 6 and 10 quarters. Prices respond more sluggish and start to significantly fall after 8 and 4 quarters. On impact, the short-term interest rate rises by 1 percentage point, but falls below baseline after 4 quarters and then converges to baseline thereafter. The response of long-term interest rates is in line with the expectations hypothesis of the term structure. However, the increase is sticky and lower in magnitude. While the response of long-term interest rates becomes significantly negative after 2 quarters in the pre-1999 sample before returning to baseline, the response in the post-1999 sample is more persistent and never becomes significantly negative. Real house prices respond immediately to the monetary policy shock, initially fall significantly, but return gradually to baseline after 5 years for both sub-samples. The maximum median decline in in the pre-EMU sample is after 6 quarters while for the post-1999 the maximum median is reached after 10 quarters. At first glance, the response of real loans to the private sector to the monetary policy shock opposes theoretical considerations that higher interest rates should reduce the equilibrium lending activities of financial intermediaries. For both samples the maximum median response takes place after 14 quarters. In the pre-EMU sample the response converges to zero in the long-run, whereas in the post-1999 sample a significant negative response can be observed after even 6 years. Given the relatively short samples size for both periods, it has to be stressed that the uncertainty around the median response is high.

Overall, key stylized facts on the transmission of monetary policy can be capture by a low-dimensional VAR models for the Euro Area when accounting for a structural break arising from the creation of the EMU. Furthermore, it is remarkable that the impulse response are very similar in shape and magnitude to those found in estimated VAR models for the United States (e.g. Christiano *et al.* 1999; Giannone *et al.* 2015; Stock and Watson 2001).

With the Bayesian approach the posterior distribution on which inference is based is derived from a recursive update of the prior with the data. It is therefore necessary to analyze if the results are mainly driven by the choice of the hyperparameter λ, governing the importance of the prior relative to the information contained in the data. Figure 6 and 7 display the impulse responses to a conventional monetary policy shock for different assumptions on λ in the pre- and post-1999 sample respectively. While a tighter prior (smaller λ) tends to smooth the

Figure 4: Impulse Responses Benchmark 1980–1998



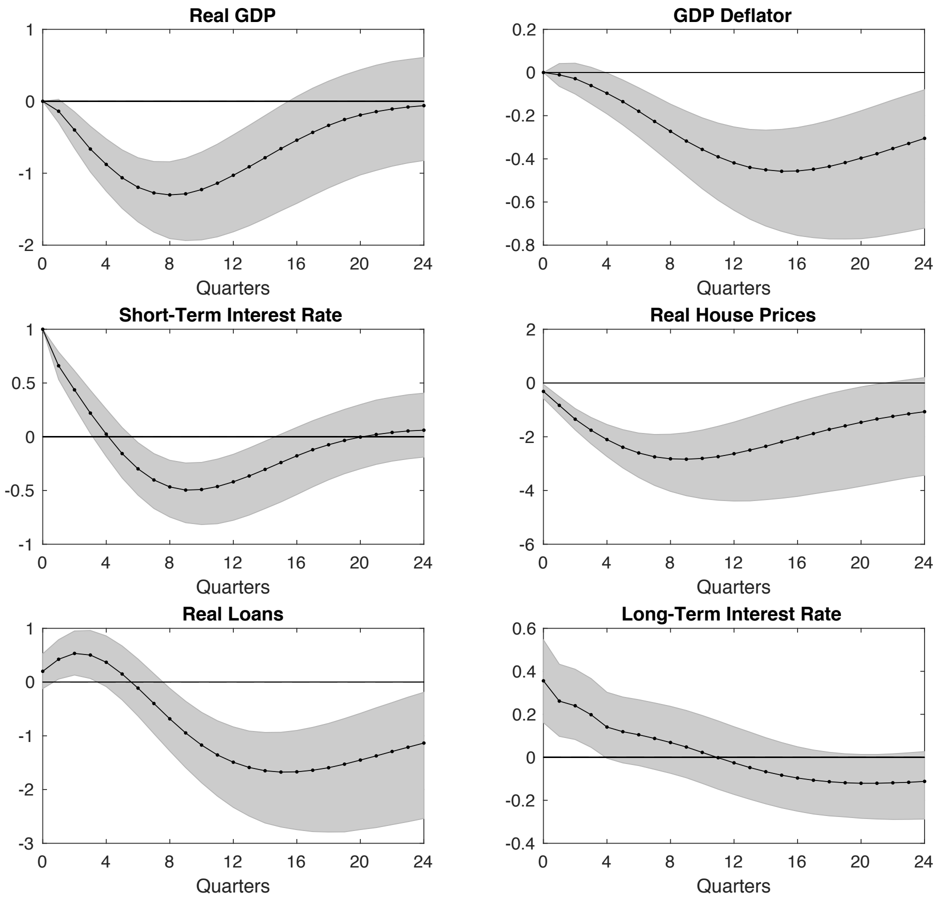
Notes: The figure depicts the 50th (dotted line), 16th and 84th percentiles (grey lines) of the posterior distribution of the impulse response functions to a 100 basis point shock to the short-term interest rate.

impulse responses, the overall the qualitative results are not driven by choice of the tightness of the prior information. However, the quantitative results differ and underlines the necessity for theoretically founded choice of hyperparameters (Giannone *et al.* 2015).[13]

Comparing Figures 6 and 7 and neglecting strong uncertainty surrounding the median responses allows to derive the main conclusion of the VAR analysis. With the creation of the EMU in 1999 qualitatively there have been only minor changes in the transmission of monetary policy. However, there are quantitative perturbations to the transmission. The maximum impact of a monetary policy shock on real output at the median has more than doubled from -0.7% to -2%, and the response of real GDP became more persistent such that the timing maximum response increased from 6–8 quarters to 10–12 quarters after the shock period. Furthermore, for the post-1999 sample the price level immediately falls (at the median) while for the pre-EMU sample prices are adjusting more sluggish and start to fall only after 4–8 quarters

[13]The response of $\lambda = 0.1$ stands out. Following Giannone *et al.* (2015) and Banbura *et al.* (2010) $\lambda = 0.1$ is possibly to small for the model size of six variables.

FIGURE 5: Impulse Responses Benchmark 1999–2016

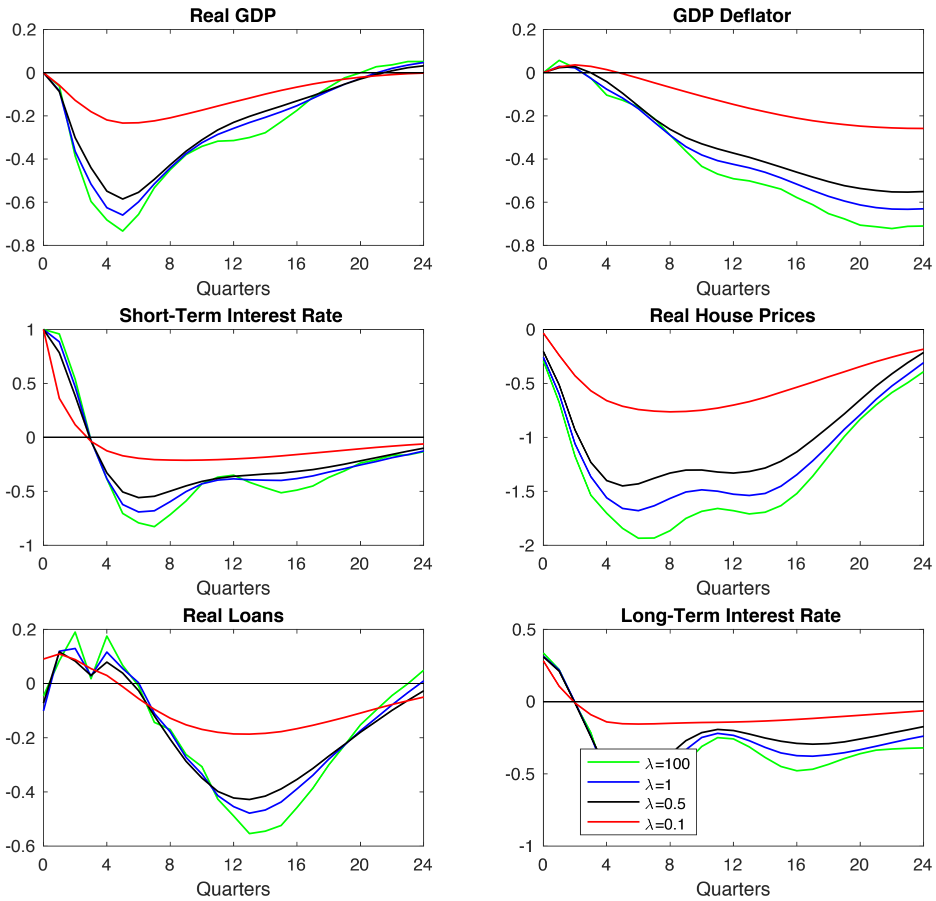


NOTES: The figure depicts the 50th (dotted line), 16th and 84th percentiles (grey lines) of the posterior distribution of the impulse response functions to a 100 basis point shock to the short-term interest rate.

after a slight increase. The maximum impact has slightly increased from −0.7% to −0.8%. The shape of the impulse response of the short-term interest rates has hardly changed. With the EMU the response of the short-term interest rate has gained slightly more persistence as the median response falls below zero after 6–8 quarters. Notable is also the very strong response of real house prices in the post-1999 sample where the maximum response has roughly tripled from −1.5% of up to −5%. The strong and lasting impact on real house prices corroborates the rising importance of this channel in the transmission of monetary policy to real economic activity. These results are even stronger than those obtained by Musso *et al.* (2010). The impact response of the long-term interest rate before the Euro and after the adoption of the Euro is roughly identical. However, an increase in persistence can be observed as the response falls below baseline after 12 compare to 4 quarters. Besides the response of real house prices, this provides a further possible explanation for the increased impact of monetary policy on real GDP. Additionally, the response of credit to the private sector has a gained a considerable amount in magnitude with an increase of the maximum impact from −0.6% to −3%.

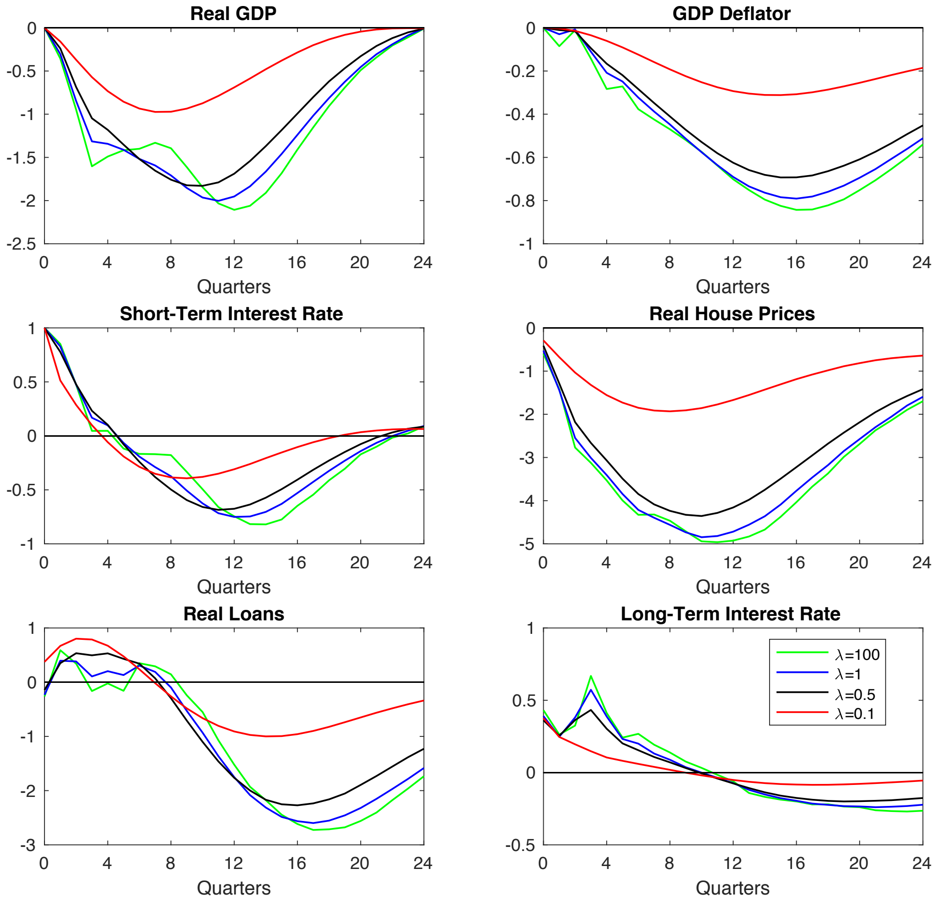
These observations stand in contrast to Boivin *et al.* (2008) who conclude that with the EMU the effect of monetary policy on real output, prices and long-term interest rates has declined. However, the results of this paper in general corroborate the conclusions drawn by Weber *et al.* (2009). The findings are consistent with the introduction of the Euro having achieved its objectives for inflation and real efficiency.

FIGURE 6: Prior Robustness Check 1980–1998



NOTES: The figure depicts the 50[th] percentile of the posterior distribution of the impulse response functions to a 100 basis point shock to the short-term interest rate for various values of the hyperparameters λ.

Figure 7: Prior Robustness Check 1998–2016



Notes: The figure depicts the 50[th] percentile of the posterior distribution of the impulse response functions to a 100 basis point shock to the short-term interest rate for various values of the hyperparameters λ.

5 Conclusion

The creation of the European Monetary Union in 1999, the transition period and the establishment of the ECB with its straightforward mandate in maintaining price stability might have changed the transmission mechanism of monetary policy in the Euro Area. This paper quantitatively investigates through the lens of a Bayesian Vector Autoregression model the macroeconomic dynamics to a conventional monetary policy shock based on three (sub-)samples periods from 1980–2016, 1980–1998 and 1999–2016 and documents possible changes since 1999. Using a recursive identification scheme proposed by Sims (1980) this paper uncovers plausible impulse responses in the Euro Area that provide a comprehensive picture of the transmission mechanism.

Even though a large-scale VAR model estimated over 1980–2016 can display plausible dynamics, the fact that a low-dimensional benchmark model delivers in part "puzzles" for the sample period 1980–2016 raises the question of misspecification when not controlling for a structural break in the transmission of monetary policy in the Euro Area. Accounting for this possibility, the impulse responses for the 1980–1998 and 1999–2016 suggest that qualitatively no significant changes in the transmission of monetary policy have taken place. Still, with the creation of the EMU quantitatively differences in impact and timing of significant responses emerged. The temporary depressing impact of a contractionary monetary policy on real GDP has roughly doubled and became more persistent. Prices start to fall faster, and slightly stronger in magnitude. Furthermore, monetary policy has a strong and lasting impact on real house prices, and these effects are roughly three times as large as its effect on real economic activity, underlining the importance of property prices in the transmission mechanism. Regarding the traditional interest rate channel, the increased persistence in the passthrough in the term structure to long-term interest rates provides some further indication for the causes in the increased effect of monetary policy on real activity.

However, it is important to recognize the limitations of this investigation. First, the VAR model is silent on cross-country differences and within country changes in the transmission of monetary policy that might occurred since 1999. Secondly, the descriptive VAR evidence of the impulse responses is not informative on the specific sources of changes in the transmission channels. Structural changes in the economy may have offsetting effects in a specific channel, thus making it hard to detect variations in the monetary transmission mechanism. Only fully-fledged structural DSGE models can shed light into this.

References

BANBURA, M., GIANNONE, D. and REICHLIN, L. (2010). Large bayesian vector autoregressions. *Journal of Applied Econometrics*, **25** (1), 71–92.

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

BOIVIN, J., GIANNONI, M. P. and MOJON, B. (2008). How has the euro changed the monetary transmission mechansism? In D. Acemoglu, K. Rogoff and M. Woodford (eds.), *NBER Macroeconomics Annual 2008*, *2*, University of Chicago Press, pp. 77–125.

CAMERON, A. C. and TRIVEDI, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

CHRISTIANO, L., EICHENBAUM, M. and EVANS, C. (1999). Monetary policy shocks: What have we learned and to what end? In J. B. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, vol. 1, Part A, *02*, 1st edn., Elsevier, pp. 65–148.

DOAN, T., LITTERMAN, R. and SIMS, C. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, **3** (1), 1–100.

FAGAN, G., HENRY, J. and MESTRE, R. (2005). An area wide model (awm) for the euro area. *Economic Modelling*, **22** (1), 39–59.

GIANNONE, D., LENZA, M. and PRIMICERI, G. (2015). Prior selection for vector autoregressions. *Review of Economics and Statistics*, **97** (2), 436–451.

JAROCINSKI, M. and SMETS, F. (2008). *House Prices and the Stance of Monetary Policy*. Working Paper Series 891, European Central Bank (ECB).

KADIYALA, R. and KARLSSON, S. (1997). Numerical methods for estimation and inference in bayesian var-models. *Journal of Applied Econometrics*, **12** (2), 99–132.

KILIAN, L. and LÜTKEPOHL, H. (2017). *Structural Vector Autoregressive Analysis*. Themes in Modern Econometrics, Cambridge: Cambridge University Press.

KOOP, G. (2003). *Bayesian Econometrics*. John Wiley & Sons Ltd.

— (2013). Forecasting with medium and large bayesian vars. *Journal of Applied Econometrics*, **28** (2), 177–203.

— (2017). Bayesian methods for empirical macroeconomics with big data. *Review of Economic Analysis*, **9** (1), 33–56.

— and KOROBILIS, D. (2009). Bayesian multivariate time series methods for empirical macroeconomics. In *Foundations and Trends in Econometrics*, vol. 3, now publishers, pp. 267–358.

LITTERMAN, R. (1986). Forecasting with bayesian vector autoregressions five years of experience. *Journal of Business and Economic Statistics*, **4** (1), 25–38.

MCCALLUM, A. H. and SMETS, F. (2007). *Real Wages and Monetary Policy Transmission in the Euro Area*. Working paper, European Central Bank (ECB).

MISHKIN, F. S. (1996). *The Channels of Monetary Transmission: Lessons for Monetary Policy*. Working Paper 5464, National Bureau of Economic Research.

— (2007). *Housing and the Monetary Transmission Mechanism*. Working Paper Series 13518, National Bureau of Economic Research.

MONTICELLI, C. and TRISTANI, O. (1999). *What does the single monetary policy do? A SVAR benchmark for the Europenan Central Bank*. Working Paper 2, European Central Bank (ECB).

MUSSO, A., NERI, S. and STRACCA, L. (2010). *Housing, Consumption and Monetary Policy: How different* *are the US and the Euro Area*. Working Paper Series 1161, European Central Bank (ECB).

Peersman, G. and Smets, F. (2001). *The Monetary Transmisson Mechanism in the Euro Zone: More Evidence from VAR Analysis*. Working Paper Series 91, European Central Bank (ECB).

Sims, C. (1980). Macroeconomics and reality. *Econometrica*, **48** (1), 1–48.

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

—, Stock, J. and Watson, M. (1990). Inference in linear time series model with some unit roots. *Econometrica*, **58** (1), 113–144.

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

Weber, A., Gerke, R. and Worms, A. (2009). *Has the monetary transmission process in the euro area changed? Evidence based on VAR estimates*. BIS Working Papers 276, Bank of International Settlements.

Appendix A – Data

The dataset is at a quarterly frequency covering the period 1980Q1–2016Q4. All variables are transformed to log-levels, except for the interest rates which remain in levels. Sources are indicated as AWM=Area-wide Model database (*https://eabcn.org/page/area-wide-model*), ECB=European Central Banks Statistical Warehouse (*http://sdw.ecb.europa.eu*) and BIS=Bank for International Settlements statistics (*https://www.bis.org/statistics/index.htm*) with the Mnemonic of the original series following in brackets.

Real GDP ($RGDP$). Gross Domestic Product (GDP) at market prices, Million Euro, Chain linked volume, Calendar and seasonally adjusted data, Reference year 1995.
Source: AWM (YER)

GDP Deflator ($DGDP$). GDP Deflator, Index, Index base year 1995 (1995=100). Defined as the ratio of nominal, and real gross domestic product (GDP).
Source: AWM (YED)

Short-Term Interest Rate (I). Nominal Short-Term Interest Rate, Euribor 3-month, Percent per annum, Last trade price.
Source: AWM (STN)

Real House Price Index (RHP). Residential property prices (2007=100), New and existing dwellings; Residential property in good and poor condition; Whole country; Neither seasonally nor working day, deflated by GDP Deflator (PGDP).
Source: ECB (RPP.Q.I8.N.TD.00.3.00)

Real Loans to Private Sector ($RLOANS$). Sum of "Credit to Non-financial corporations from All sectors at Market value" and "Credit to Households and NPISHs from All sectors", Euro (Billions), Domestic currency, Adjusted for breaks, deflated with the GDP Deflator (PGDP).

Long-Term Interest Rate ($GB10Y$). Benchmark bond - Euro area 10–year Government Benchmark bond yield - Yield - percent per annum.
Source: ECB (FM.M.U2.EUR.4F.BB.U2_10Y.YLD)

Money Stock M2 ($M2$). Euro area (changing composition), Outstanding amounts at the end of the period (stocks), MFIs, central government and post office giro institutions reporting sector - Monetary aggregate M2, All currencies combined - Euro area (changing composition)

counterpart, Non-MFIs excluding central government sector, denominated in Euro, data Working day and seasonally adjusted.
Source: ECB (BSI.M.U2.Y.V.M20.X.1.U2.2300.Z01.E)

Money Stock M3 ($M3$). Outstanding amounts at the end of the period (stocks), MFIs, central government and post office giro institutions reporting sector - Monetary aggregate M3, All currencies combined - Euro area (changing composition) counterpart, Non-MFIs excluding central government sector, denominated in Euro, data Working day and seasonally adjusted.
Source: ECB (BSI.M.U2.Y.V.M30.X.1.U2.2300.Z01.E)

Short to Medium-Term Interest Rate ($GB2Y$). Benchmark bond - Euro area 2–year Government Benchmark bond yield - Yield - Euro.
Source: ECB (FM.M.U2.EUR.4F.BB.U2_2Y.YLD)

Nominal Effective Exchange Rate (EEN). Nominal Effective Exchange Rate (NEER), Euro Euro area-19 countries vis-à-vis the NEER–38 group of main trading partners, Base year 1999 (1999Q1 = 100).
Source: AWM (EEN)

Real Loans to NFC ($RLNFC$). Credit to Non-financial corporations from All sectors at Market value, Euro (Billions), Domestic currency, Adjusted for breaks, deflated with the GDP Deflator (PGDP).
Source: BIS (Q:XM:H:A:M:XDC:A)

Real Loans to HH ($RLHH$). Credit to Households and NPISHs from All sectors, Euro (Billions), Domestic currency, Adjusted for breaks, deflated with the GDP Deflator.
Source: BIS (Q:XM:N:A:M:XDC:A)

Commodity Prices ($PCOM$). Non-oil Commodity Prices, ECB commodity price index US dollar denominated, Import weighted, Total non-energy commodity, Neither seasonally nor work ing day adjusted data.
Source: AWM (COMPR)

Appendix B – Gibbs Sampler Algorithm

This section provides a brief sketch of the algorithm employed to estimate the Bayesian VAR. All estimates in section 4 are based on Gibbs sampling for posterior numerical evaluation of the distribution of parameters of interest (coefficients and impulse response functions). These are based on $N = 25,000$ valid Gibbs draws from the (conditional) posterior distribution where the first $\overline{n} = 5,000$ draws are discarded for the burn-in phase of the sampler. A stylized representation of the Gibbs sampler algorithm is provided below.

Algorithm 1 Gibbs sampler

```
Choose starting value Σ_ϵ^(0).
for n = 1, ..., N do
    Randomly sample α^(n) | Σ_ϵ^(n−1), Y from N(ᾱ, Σ_ϵ^(n−1) ⊗ (X'_* X_*)^(−1)).
    Randomly sample Σ_ϵ^(n) | Y from IW(S̄, ν̄).
    if n ≥ n̄ then
        Store the pair {α^(n), Σ_ϵ^(n)}.
        Calculate the sequence {Ψ_i^(n)}_{i=0}^T = (Ψ_0^(n), Ψ_1^(n), Ψ_2^(n), ..., Ψ_T^(n)).
        Store {Ψ_i^(n)}_{i=0}^T.
    end if
end for
```

The starting value $\boldsymbol{\Sigma}_{\epsilon}^{(0)}$ is the OLS estimate of the variance-covariance matrix. The sequence of impulse response functions $\{\boldsymbol{\Psi}_i^{(n)}\}_{i=0}^{T}$ is then computed as described in the following. The starting point to derive the structural impulse response function $\{\boldsymbol{\Psi}_i\}_{i=0}^{T}$ are the responses of $\boldsymbol{y}_{t+h}$ to the reduced form errors $\boldsymbol{\epsilon}_t$. These are obtained by considering the VAR(1) representation of the reduced-form VAR(p) model in Eq. (1) which is

$$\boldsymbol{Y}_t = \boldsymbol{C} + \boldsymbol{F}\boldsymbol{Y}_{t-1} + \boldsymbol{\eta}_t \tag{B.1}$$

where

$$\underset{(Mp\times 1)}{\boldsymbol{Y}_t} \equiv \begin{pmatrix} \boldsymbol{y}_t \\ \vdots \\ \boldsymbol{y}_{t-p+1} \end{pmatrix}, \quad \underset{(Mp\times Mp)}{\boldsymbol{F}} \equiv \begin{pmatrix} \mathcal{A}_1 & \mathcal{A}_2 & \dots & \mathcal{A}_{p-1} & \mathcal{A}_p \\ \boldsymbol{I}_{M\times M} & \boldsymbol{0}_{M\times M} & \dots & \boldsymbol{0}_{M\times M} & \boldsymbol{0}_{M\times M} \\ \boldsymbol{0}_{M\times M} & \boldsymbol{I}_{M\times M} & \dots & \boldsymbol{0}_{M\times M} & \boldsymbol{0}_{M\times M} \\ \vdots & & \ddots & \vdots & \vdots \\ \boldsymbol{0}_{M\times M} & \boldsymbol{0}_{M\times M} & \dots & \boldsymbol{I}_{M\times M} & \boldsymbol{0}_{M\times M} \end{pmatrix},$$

$$\underset{(Mp\times 1)}{\boldsymbol{\eta}_t} \equiv \begin{pmatrix} \boldsymbol{\epsilon}_t \\ \mathbf{0}_{M\times 1} \\ \mathbf{0}_{M\times 1} \\ \vdots \\ \mathbf{0}_{M\times 1} \end{pmatrix}, \quad \underset{(Mp\times 1)}{\boldsymbol{C}} \equiv \begin{pmatrix} \boldsymbol{c} \\ \mathbf{0}_{M\times 1} \\ \mathbf{0}_{M\times 1} \\ \vdots \\ \mathbf{0}_{M\times 1} \end{pmatrix}.$$

From recursive substitution by k times we get

$$\boldsymbol{Y}_t = \boldsymbol{C}\sum_{i=0}^{k}\boldsymbol{F}^i + \boldsymbol{F}^{k-1}\boldsymbol{Y}_{t-k-1} + \sum_{i=0}^{k}\boldsymbol{F}^i\boldsymbol{\eta}_{t-i}. \tag{B.2}$$

Thus when k approaches ∞ the Vector Moving Average (VMA(∞)) follows

$$\begin{aligned}\boldsymbol{Y}_t =& \lim_{k\to\infty}\boldsymbol{C}\sum_{i=0}^{k}\boldsymbol{F}^i + \boldsymbol{F}^{k+1}\boldsymbol{Y}_{t-k-1} + \sum_{i=0}^{k}\boldsymbol{F}^i\boldsymbol{\eta}_{t-i} \\ =& \boldsymbol{M} + \sum_{i=0}^{\infty}\boldsymbol{F}^i\boldsymbol{\eta}_{t-i},\end{aligned} \tag{B.3}$$

under the standard covariance-stationarity conditions for multivariate stochastic processes that all eigenvalues of the companion-matrix $\boldsymbol{F}$ have modulus less than unity (Kilian and Lütkepohl 2017: Chapter 2) and where $\boldsymbol{M} \equiv \boldsymbol{C}(\boldsymbol{I}_{Mp\times Mp} - \boldsymbol{F})^{-1}$. Premultiplying Eq. (23) by $\boldsymbol{J} \equiv (\boldsymbol{I}_{M\times M}\ \mathbf{0}_{M\times M(p-1)})$ and using $\boldsymbol{\eta}_t = \boldsymbol{J}'\boldsymbol{\epsilon}_t$ and $\boldsymbol{\mu} = \boldsymbol{J}'\boldsymbol{M}$ gives the solution for the VAR(p) process

$$\boldsymbol{y}_t = \boldsymbol{\mu} + \sum_{i=0}^{\infty}\boldsymbol{J}\boldsymbol{F}^i\boldsymbol{J}'\boldsymbol{\epsilon}_{t-i}. \tag{B.4}$$

Given the estimates for $\mathcal{A}_j$, $j = 1,\ldots,p$ orthogonalization of the reduced from errors is then obtain by using relation between structural and reduced form errors $\boldsymbol{\epsilon}_t = \mathcal{D}_0^{-1}\boldsymbol{u}_t$ to give the structural VMA(∞) representation

$$\boldsymbol{y}_t = \boldsymbol{\mu} + \sum_{i=0}^{\infty}\boldsymbol{J}\boldsymbol{F}^i\boldsymbol{J}'\mathcal{D}_0^{-1}\boldsymbol{u}_{t-i}. \tag{B.5}$$

Now define $\boldsymbol{\Psi}_i \equiv \boldsymbol{J}\boldsymbol{F}^i\boldsymbol{J}'\mathcal{D}_0^{-1}$ for $i = 0, 1, \ldots$ and iterating Eq. (B.5) forward by h periods yields

$$\boldsymbol{y}_{t+h} = \boldsymbol{\mu} + \sum_{i=0}^{\infty}\boldsymbol{\Psi}_i\boldsymbol{u}_{t+h-i}. \tag{B.6}$$

Taking the partial derivative of $\boldsymbol{y}_{t+h}$ in Eq. (B.6) w.r.t. $\boldsymbol{u}_t'$

$$\frac{\partial \boldsymbol{y}_{t+h}}{\partial \boldsymbol{u}_t'} = \underset{(M\times M)}{\boldsymbol{\Psi}_h}, \tag{B.7}$$

provides us with the dynamic multiplier at horizon h, i.e. the impulse response (matrix) of the endogenous variables in $t+h$ to a structural (unit) shock in t. The $(i,j)^{\text{th}}$ element of $\boldsymbol{\Psi}_h$, denoted by $\boldsymbol{\psi}_{ij,h}$, is the impulse response of variable i to a unit innovation to variable j at time t at horizon $t+h$ holding all other innovations constant $(\partial y_{i,t+h}/\partial u_{j,t} = \boldsymbol{\psi}_{ij,h})$. The sequence for $h = 0,\ldots,T$ is then the impulse response function $\{\boldsymbol{\Psi}_i\}_{i=0}^{T}$ for all endogenous variables and all possible shocks for propagation horizon T periods.