DSGE Models and VARMA Representation*

Alessia Paccagnini†

University of Milano-Bicocca, Department of Economics

Raffaele Rossi‡

Lancaster University, Department of Economics

Abstract

We consider a small scale DSGE with trend inflation, where the price dispersion is a non-observable state variable. Hence, the model lacks of a finite VAR representation and the VAR analysis based on this model may suffer of the truncation bias problem. First, we use the DSGE model as data-generation process to create artificial pseudo-data. Second, using these pseudo-data, we employ a sign restrictions VAR to evaluate the effects of a monetary policy shock. The true generation process implies a strong response of output after a monetary shock, instead the VAR shows monetary neutrality. We conjecture that this discrepancy is due to a truncation bias which affects the statistical representation of our true data generation process.

JEL CODES: C11, C15, C32

KEYWORDS: Bayesian Analysis, DSGE Models, Vector Autoregressions, MonteCarlo experiments

---

*Preliminary version: July 2012

†E-mail : alessia.paccagnini@unimib.it, Università degli Studi - Milano Bicocca, Department of Economics

‡E-mail : raffaele.rossi1@gmail.com, Lancaster University, Department of Economics

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models are the main tool of macroeconomic policy analysis (Smets and Wouters (2003, 2004, 2007), Del Negro and Schorfheide (2004), Adolfson *et al.* (2008) and Christiano *et al.* (2005)).

In recent years there have been major advances in estimation methodology provided with various specifications of these models that are able to compete, in terms of data fit and predictability with more standard time-series models, such as vector autoregressions. For example, in the empirical macroeconomics literature, the DSGE models are used to explain the impact of an economic shock on macroeconomic variables (such as a monetary policy shock on output and inflation, or a technology shock on labor hours).

On one side, this empirical evidence is obtained from estimating a structural Vector Autoregressive model (VAR) as in Sims (1980, 1986). On the other side, the structural parameters of a DSGE model are estimated by minimizing the distance between the model's and the estimated VAR impulse response functions, as in Rotemberg and Woodford (1998) and Christiano, Eichenbaum and Evans (2005).

To test the performance of a DSGE model using a VAR, we need that the data-generating process which is consistent with the theoretical model, has a finite-order VAR representation. Indeed, this has been discussed in several papers (Canova and Pina , 2005, Chari, Kehoe, and McGrattan, 2005, Galì and Rabanal, 2005, Christiano, Eichenbaum and Vigfusson, 2006, Fernandez-Villaverde, Rubio-Ramirez, Sargent and Watson, 2006, and Ravenna, 2007).

Fernandez-Villaverde et al. (2006) show that a VAR representation mapping economic shocks to a vector of observable variables is admitted if and only if the vector moving representation (VMA) is invertible. On the contrary, if the VMA is not invertible, the VAR representation of a DSGE model may require an infinite number of lags. Hence whenever a researches uses a finite order VAR to approximate a DSGE that requires an infinite VAR, incurs in a truncation bias problem.

Ravenna (2007) studies the truncation bias problem in a DSGE model that lacks of a finite VAR representation. He shows that the truncated VAR(p) may return incorrect estimates of the impulse response function (IRF). In particular, IRF is affected by truncation bias through two different channels. First, the VAR(p) erroneously constraints to zero some coefficients in the true VAR representation. Second, the VAR(p) coefficients can lead to mistaken identification of the structural shocks. Moreover, he shows that the truncation bias can cause identification problems even when the identification strategy is consistent with the theoretical model. Similarly, Chari et al. (2008) find that the impulse response of labor hours to a technology shock, which are identified using long run restrictions, in a finite-order VAR is poor approximation to the true magnitude unless nontechnology shocks play a minor role.

In this paper, we build on Ravenna (2007) and Chari et al. (2008) by studying the truncation bias problem (TBP) in models that lack of a finite VAR representation. In particular, we address how TBP affects the IRF's of a monetary policy shock in a DSGE model with trend inflation (Ascari et al. (2011)). We choose this model for two main reasons. First, we consider this model as a good standard NK model by assuming a small level of trend inflation along the line of Ascari (2004) and Ascari and Ropele (2009). With this assumption, price dispersion becomes a relevant endogenous state variable even in the log-linear equilibrium. Second, the price dispersion is a non-observable state variable, consequently, the model lacks of a finite VAR representation and the VAR analysis based on this model may suffer of TBP.

The exercise consists in the following steps. First, we use the DSGE model with trend inflation as data-generation process (DGP) in order to create artificial pseudo-data. Second, using these pseudo-data, we employ different estimation strategies (essentially the Cholesky identification and the sign restrictions) to evaluate the effects of a monetary policy shock.

In further steps, we will produce a theoretical generalization of our result and we will implement different identification strategies (such the DSGE-VAR à la Del Negro and Schorfheide (2004)).

2 Small-scale DSGE model with trend inflation

The simple DSGE model with forward-looking features is usually referred to as a benchmark in the literature. For instance, Del Negro and Schorfheide (2004) used this model to introduce the DSGE-VAR, and investigate its predictive ability. In a DSGE setup, the economy is made up of three components. The first component is the representative household that maximizes a usual preference over leisure and consumption. The second component is a monopolistic competitive sector affected by nominal rigidities à la Calvo (1983). We depart from the standard NK model by assuming a small level of trend inflation along the line of Ascari (2004) and Ascari and Ropele (2009). With this assumption, price dispersion becomes a relevant endogenous state variable even in the log-linear equilibrium. For the sake of our exercise, we assume that this state is unobservable. Finally we close the model by characterizing the monetary policy as an automatic Taylor rule: it changes the nominal rate along with inflation and output. There are three economic shocks: an exogenous monetary policy shock (in the monetary policy rule), and two autoregressive processes, AR(1), which model government spending and technology shocks.

To solve the model, optimality conditions are derived for the maximization problems. After linearization around the steady-state, the economy is described, as in Ascari et al. (2012) by the following system of equations:

A New Keynesian Phillips Curve augmented with trend inflation. Note that price dispersion has a positive impact on current inflation

$$\hat{\pi}_t = [\beta\bar{\pi} + \eta\,(\varepsilon - 1)]E_t\hat{\pi}_{t+1} + \kappa\hat{y}_t - \lambda\varphi\hat{a}_t + \lambda\varphi\hat{s}_t + \hat{\phi}_{t+1} \tag{1}$$

The forward looking variable $\hat{\phi}_t$ does not have a precise economic interpretation. It can be considered as auxiliary flow variable

$$\hat{\phi}_t = (1-\sigma)\left(1-\theta\beta\bar{\pi}^{\varepsilon-1}\right)\hat{y}_t + \theta\beta\bar{\pi}^{\varepsilon-1}\left[(\varepsilon-1)\,\hat{\pi}_{t+1} + \hat{\phi}_{t+1}\right] \tag{2}$$

The price dispersion is positively related to current inflation, i.e. the higher the inflation, the higher the price dispersion

$$\hat{s}_t = \zeta\hat{\pi}_t + \theta\pi^{\varepsilon}\hat{s}_{t-1} \tag{3}$$

The dynamic IS curve simply states that the ex-ante interest rate is the opportunity cost of equating marginal utilities across time

$$\hat{y}_t = \hat{y}_{t+1} - \sigma^{-1}\left(\hat{R}_t - E_t\hat{\pi}_{t+1}\right) + \hat{g}_t$$

the Taylor rule links the current interest rate inflation, output and past interest rate.

$$\hat{R}_t = \rho_R\hat{R}_{t-1} + (1-\rho_R)(\psi_1\hat{\pi}_t + \psi_2\hat{y}_t) + \hat{r}_t \tag{4}$$

$$\hat{g}_t = \rho_g\hat{g}_{t-1} + \epsilon_{g,t} \tag{5}$$

$$\hat{a}_t = \rho_z\hat{a}_{t-1} + \epsilon_{z,t}, \tag{6}$$

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + \epsilon_{r,t} \tag{7}$$

Hatted variables indicate percentage deviations with respect to steady state values. $\hat{\pi}_t$ is the inflation rate, $\bar{\pi}$ is the steady state inflation rate, and $\hat{R}_t$ is the gross nominal interest rate, $\hat{y}_t$ stands for output, $\hat{a}_t$ is the technological shock, $\hat{g}_t$ is the demand shock and $\hat{r}_t$ stands for the monetary shock. Price dispersion is identified by $\hat{s}_t$, while $\hat{\phi}_t$ is an auxiliary variable. σ is the usual CRRA parameter, ε is the Dixit-Stiglitz aggregator parameter and θ is the Calvo parameter while

$$\lambda = \frac{\left(1 - \theta\bar{\pi}^{\varepsilon-1}\right)\left(1 - \theta\beta\bar{\pi}^{\varepsilon}\right)}{\theta\bar{\pi}^{\varepsilon-1}} \tag{8}$$

$$\eta = \beta\left(\bar{\pi} - 1\right)\left[1 - \theta\bar{\pi}^{\varepsilon-1}\right] \tag{9}$$

$$\kappa = \left(\lambda\left(\sigma + \varphi\right) - \eta\left(1 - \sigma\right)\right) \tag{10}$$

$$\zeta = \frac{\varepsilon\theta\bar{\pi}^{\varepsilon-1}\left(\bar{\pi} - 1\right)}{1 - \theta\bar{\pi}^{\varepsilon-1}} \tag{11}$$

or, in the case of output, from a trend path (King 2000; Woodford 2003). The model can be solved by applying the algorithm proposed by Sims (2002). Define the vector of variables $\tilde{\mathbf{Z}}_t = \left(\hat{y}_t, \hat{\pi}_t, \hat{R}_t, \hat{g}_t, \hat{a}_t, E_t\hat{y}_{t+1}, E_t\hat{\pi}_{t+1}, E_t\hat{\phi}_{t+1}\right)$ and the vector of shocks as $\epsilon_t = \left(\epsilon_{r,t}, \epsilon_{g,t}, \epsilon_{a,t}\right)$. Therefore the previous set of equations, (1) - (7), can be recasted into a set of matrices $(\mathbf{\Gamma}_0, \mathbf{\Gamma}_1, \mathbf{C}, \mathbf{\Psi}, \mathbf{\Pi})$ accordingly to the definition of the vectors $\tilde{\mathbf{Z}}_t$ and ϵ_t

$$\mathbf{\Gamma}_0\tilde{\mathbf{Z}}_t = \mathbf{C} + \mathbf{\Gamma}_1\tilde{\mathbf{Z}}_{t-1} + \mathbf{\Psi}\epsilon_t + \mathbf{\Pi}\eta_t \tag{12}$$

where C is a vector of constants, ϵ_t is an exogenously evolving random disturbance and η_t is a vector of expectations errors, $\left(E_t\left(\eta_{t+1}\right) = \mathbf{0}\right)$, not given exogenously but to be treated as part of the model solution. In order to provide the mapping between the observable data and those computed as deviations from the steady state of the model we set the following measurement equations as in Del Negro and Schorfheide (2004)

$$\begin{aligned} \Delta \ln x_t &= \ln \gamma + \Delta \tilde{x}_t + \tilde{z}_t \\ \Delta \ln P_t &= \ln \pi^* + \tilde{\pi}_t \\ \ln R_t^a &= 4 \left[(\ln r^* + \ln \pi^*) + \tilde{R}_t \right] \end{aligned} \tag{13}$$

which can be also casted into matrices as

$$\mathbf{Y}_t = \mathbf{\Lambda}_0 (\theta) + \mathbf{\Lambda}_1 (\theta) \tilde{\mathbf{Z}}_t + v_t \tag{14}$$

where $\mathbf{Y}_t = (\Delta \ln x_t, \Delta \ln P_t, \ln R_t)'$, $v_t = 0$ and $\mathbf{\Lambda}_0$ and $\mathbf{\Lambda}_1$ are defined accordingly. For completeness, we write the matrices $\mathbf{T},$ $\mathbf{R},$ $\mathbf{\Lambda}_0$ and $\mathbf{\Lambda}_1$ as a function of the structural parameters in the model, θ. Such a formulation derives from the rational expectations solution. The evolution of the variables of interest, $\mathbf{Y}_t$, is therefore determined by (12) and (14) which impose a set of restrictions across the parameters on the moving average (MA) representation. Given that the MA representation can be very closely approximated by a finite order VAR representation, Del Negro and Schorfheide (2004) propose to evaluate the DSGE model by assessing the validity of the restrictions imposed by such a model with respect to an unrestricted VAR representation. The choice of the variables to be included in the VAR is however completely driven by those entering in the DSGE model regardless of the statistical goodness of the unrestricted VAR. Policy variables set by optimization - typically included s_t - are naturally endogenous as optimal policy requires some response to current and expected developments of the economy. Expectations at time t for some of the variables of the systems at time $t+1$ are also included in the vector $\mathbf{s}_t$,whenever the model is forward-looking. Models like (12) can be solved using standard numerical techniques as in Sims (2002) and the solution can be expressed as follows

$$s_t = A_0 + As_{t-1} + B\epsilon_t \tag{15}$$

where the matrices $\mathbf{A}_0, \mathbf{A}$,and $\mathbf{B}$ contain convolutions of the underlying model structural parameters. Consider the simple case in which all variables in the DSGE are observable and the number of structural shocks in ϵ_t is exactly equal to the number of variables in $\widetilde{\mathbf{Z}}_t$. In this case VAR are natural specifications for the data, therefore the estimated reduced form is

$$s_t = A_0 + As_{t-1} + u_t \tag{16}$$

3 The Mapping Between a DSGE Model and a VAR representation

Recent model evaluation of DSGE models exploits the fact that a solved RBC model is a statistical model. In fact, a solved DSGE model often generates a restricted MA representation for the vector of observable variables of interest, that can be approximated by a VAR of finite order (Fernandez-Villaverde *et al.,* 2007; Ravenna, 2007). Interestingly, this recent approach to model evaluation does not require identification of structural shocks but it is still potentially affected by lack of statistical identification. To make it clear, consider the general case of system (15) in which only a subset n of the m variables included in s_t is observable and define such a subset as x_t. Now, x_t has a VAR(∞) representation. This is usually approximated by a finite VAR representation at the cost of a truncation that can be relevant for purposes such as the identification of structural shocks (Ravenna 2007). Note that if the RBC model features a number of shocks smaller than the number of variables included in the VAR, some of the VAR shocks are interpreted as measurement error.

A log-linearized DSGE model yield a state-space representation of the following form:

$$s_t = As_{t-1} + B\epsilon_t \tag{17}$$

$$x_t = C\, s_{t-1} + D\epsilon_t \tag{18}$$

where s_t is a $k \times 1$ vector of state variables, x_t is an $n \times 1$ vector of observed variables, and ϵ_t is an $m \times 1$ vector of structural shocks. The variance-covariance matrix of these shocks is diagonal and given by Σ_ϵ. A, B, C, and D are matrices of comformable size whose elements are functions of the deep parameters of the model. We consider the same number of observed variables as shocks, $n = m$, consequently $\mathbf{D}$ is square and invertible.

We can solve for ϵ_t from eq. (18) as:

$$\epsilon_t = D^{-1}(x_t - Cs_{t-1})$$

Plugging this into (17) yields:

$$s_t = (A - BD^{-1}C)s_{t-1} + BD^{-1}x_t$$

Solving backwards, one obtains:

$$s_t = (A - BD^{-1}C)^{t-1}\mathbf{s}_0 + \sum_{j=0}^{t-1}(A - BD^{-1}C)^{j-1}BD^{-1}x_{t-j} \tag{19}$$

If $\lim_{t\to\infty}(A - BD^{-1}C)^{t-1} = 0$, the history of observables perfectly reveals the current state. This requires that the eigenvalues of $(A - BD^{-1}C)$ all be strictly less than one in modulus. If this condition is satisfied, eq. (19) can be plugged into (18) to yield a VAR in observables in which the VAR innovations correspond to the structural shocks:

$$x_t = C \sum_{j=0}^{t-1} (A - BD^{-1}C)^{j-1} BD^{-1} x_{t-1-j} + D\epsilon_t \tag{20}$$

The condition that the eigenvalues of $(A - BD^{-1}C)$ all be strictly less than unity is the "poor man's invertibility" condition given in Fernandez-Villaverde, Rubio-Ramirez, Sargent, and Watson (2007). It is a sufficient condition for a VAR on observables to have innovations that map directly back into structural shocks population. When satisfied, a finite order VAR(p) on s_t will yield a good approximation to eq. (20), and conventional estimation and identification strategies will allow one to uncover the model's impulse responses to structural shocks. When this condition for invertibility is not satisfied the state space system nevertheless yields a VAR representation in the observables, though the VAR innovations no longer correspond to the structural shocks.

The main issue when the invertibility condition is not met is that the observables do not perfectly reveal the state vector. To this aspect, use the Kalman filter to form a forecast of the current state, $\widehat{s}_t$ given observables and a lagged forecast:

$$\widehat{s}_t = (A - KC)\widehat{s}_{t-1} + Kx_t. \tag{21}$$

Here K is the Kalman gain. It is the matrix that minimizes the forecast error variance of the filter, i.e. $\Sigma_s = E(s_t - \widehat{s}_t)(s_t - \widehat{s}_t)'$. K and Σ_s are the joint solutions to the Riccati equations:

$$\Sigma_s = (A - KC)\Sigma_s(A - KC)' + B\Sigma_s B' + KD\Sigma_\epsilon D'K' - B\Sigma_\epsilon D'K' - KD\Sigma_\epsilon B'$$

$$K = (A\Sigma_s C' + B\Sigma_\epsilon D')(C\Sigma_s C' + D\Sigma_\epsilon D')^{-1}$$

Given values of K and Σ_s,add and subtract $C\widehat{s}_{t-1}$ from the right hand side of eq. (18) to obtain:

$$x_t = C\widehat{s}_{t-1} + u_t \quad (22)$$

$$u_t = c(s_{t-1} - \widehat{s}_{t-1}) + D\epsilon_t \quad (23)$$

Lagging (21) one period and recursively substituting into (22), one obtains an infinite order VAR representation in the observables:

$$x_t = (A - KC)^{t-1}\widehat{s}_0 + C\sum_{j=0}^{t-1}(A - KC)^j K x_{t-1-j} + u_t \quad (24)$$

Under weak conditions, Hansen and Sargent (2007) show that $(A - KC)$ is a stable matrix, so that the $(A - KC)^{t-1}\widehat{s}_0$ term disappears in the limit and the infinite sum on the lagged observables converges in mean square. The innovations in this VAR representation are comprised of two orthogonal components: the true structural shocks and the error in forecasting the state. The innovation variance is given by:

$$\Sigma_u = C\Sigma_s C' + D\Sigma_\epsilon D' \quad (25)$$

Fernandez-Villaverde, Rubio-Ramirez, Sargent, and Watson (2007) show that the eigenvalues of $(A - BD^{-1}C)$ being less than unity in modulus implies that $\Sigma_s = 0$.

When $\Sigma_s = 0$, then $\Sigma_u = D\Sigma_\epsilon D'$, and it is straightforward to show that (24) reduces to (20). If the "poor man's invertibility" condition is not satisfied, then $\Sigma_s \neq 0$, and the innovation variance from the VAR is strictly larger than the innovation variance in the structural model. The failure of invertibility is part and parcel of the observables to reveal the state vector. Non-invertibility is fundamentally an issue of missing information. Sims (2012) evidences (25) makes clear that the extent to which a failure of invertibility might "matter" quantitatively is how large Σ_s, i.e. how hidden the state is. We can find a lot of implications. First of all, even if the condition for invertibility fails, Σ_s may nevertheless be "small", meaning that $\Sigma_u \approx D\Sigma_\epsilon D'$. Put differently, the VAR innovations may very closely map into structural shocks even if a given system is technically non-invertible. Second, what observable variables are included in a VAR might matter - some observables may do a better job of forecasting the missing states, hence leading to smaller Σ_s and a closer mapping between VAR innovations and structural shocks. Finally, adding more observable variables should always work to lower Σ_s, and thus ameliorate problems due to non-invertibility. This means that estimating larger dimensional VARs may generally be advantageous relative to the small systems that are frequently estimated in the literature. Sims (2012) evidences the benefits of estimating factor augmented models, which can efficiently condition on large information sets.

4 Sign Restrictions

One main aspect of the impulse responses analysis is the identification of the restrictions. In the tradition of the macroeconometrics, there are essentially two important sets of restrictions. In the first set of restrictions (the linear restrictions on the structural parameters), we include the triangular identification (Christiano, Eichenbaum, and Evans (1996)) and the non-triangular identification (Sims (1986), King, Plosser, Stock, and Watson (1991), Gordon and Leeper (1994), Bernanke and Mihov (1998), Zha (1999), and Sims and Zha (2006)). In the second set of restrictions (the nonlinear restrictions on the structural parameters), we include ones directly imposed on the impulse responses, such as short-run and long-run restrictions (Blanchard and Quah (1993) and Galì (1992)). Rubio-Ramirez, Waggoner, and Zha (2007) find a way to transform nonlinear restrictions on the original parameter space to linear restrictions on the transformed parameter space, since working on the linear restrictions on the transformed parameter space is easier.

Employing both sets of restrictions, the main objective is to identify structural shocks.

For example, according to many DSGE models (and essentially to the "conventional wisdom"), after a monetary policy contraction the short term interest rate should increase, the prices and the output should not increase, and the nonborrowed reserves should decrease. Hence, a successful identification strategy should produce impulse responses coherent to the theory. Unfortunately, it happens that some restrictions, such as the triangular identification, may not generate impulse responses that have the desired signs, generating a "puzzle". Considering these identification issues, Faust (1998), Canova and De Nicolò (2002), Peersman (2005) and Uhlig (2005) propose an alternative approach: sign restrictions. The idea is to impose ex post sign restrictions on a set of moments generated with the VAR, e.g. a set of impulse responses to a given shock. Considering, for example, the case of a contractionary monetary shock, we can impose a restriction such that the interest rate rises while money, output, and prices fall in response to the shock. In other words, the usual identification scheme based on the Cholesky decomposition of the variance-covariance matrix imposes and inertial restriction only on the quarter in which the shock occurs. Instead, sign restrictions assume something about the future of the system and, thus, the future should be modeled appropriately.

Considering the VAR representation:

$$\mathbf{y}_t = A(L)\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$$

$\mathbf{y}_t$ is a $(T \times n)$ matrix with rows y_t'. $\boldsymbol{\varepsilon}_t$ is a $(T \times n)$ matrix with rows ε_t'

We estimate the reduced-form VAR coefficients $A(L)$ and covariance matrix (Σ) from the data via OLS. After, we orthogonalize the VAR residuals considering the eigenvalue-eigenvector decomposition such that:

$$\Sigma = PDP^{'}$$

where P is the matrix of eigenvectors and D is the diagonal matrix of eigenvalues. The non-uniqueness of the MA representation of the VAR is exploited to provide a set of alternative proposals for the shocks of interest using three Givens rotation matrices.

In the context of a three variable VAR, a 3×3 Givens matrix (called a Givens rotation matrix) Q_{12} has the form:

$$Q_{12} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

i.e., the matrix is the identity matrix in which the block consisting of the first and the second columns and rows has been replaced by cosine and sine terms and θ are drawn randomly from a uniform distribution $U\,[0,\pi]$ for each θ_k. as in Peersman (2005). Then $Q_{12}^{'}Q_{12} = I_3$ (considering that $\cos^2\theta + \sin^2\theta = 1$). Essentially, there are three possible Givens rotations for a three variable system; the others being Q_{13} and Q_{23}. Each Q_{ij} depends on a different parameter θ_k.

More specifically, $Q = \prod_{m,n} Q_{m,n}(\theta) = Q_{12}(\theta_1) \times Q_{13}(\theta_2) \times Q_{23}(\theta_3)$. Q is orthogonal and $\widetilde{B}(\theta) = PD^{\frac{1}{2}}Q(\theta)$ is the impulse matrix, where Q depends upon three different θ_k. For each vector $\Theta = [\theta_1, \theta_2, \theta_3]$, we obtain a set of impulse responses. To compute an impulse response, we take the estimated parameters of a reduced-form model from the relevant quarter. More specifically, for the computation of the impulse response for quarter t, we use the estimated standard deviations and the matrix of contemporaneous effects from quarter t and the coefficients at lags of variables from quarters $t+1, t+2, ..., t+T_{HOR}$, where T_{HOR} denotes the number of quarters for which the impulse responses are computed.

Under a computational point of view, if the impulse responses to the "candidate" shock satisfy all the required restrictions, then the draw of the orthonormal vector θ and the corresponding responses are retained. Otherwise, the responses are discarded. As specified in Castelnuovo (2012), an equal and strictly positive weight is assigned to the draws which meet the restrictions (retained draws). Instead, a zero prior weight is assigned to responses violating the constrains. In Uhlig (2005), there is a discussion about the possibility to set up a penalty function to penalize violations and reward large and correct responses. In Canova and De Nicolò (2002), a grid of M values for each of the values of θ_k is suggested between 0 and π, and then compute all the possible Q. All these models which are distinguished by different numerical values for θ_k are able to produce an exact fit to the variance of the data. Only those Q producing shocks that agree with maintained sign restrictions would be retained. Rubio-Ramirez, Waggoner, and Zha (2010) propose an algorithm to compute the rotation matrix Q efficiently, called in Fry and Pagan (2011), Householder Transformation.

This alternative method of forming an orthogonal matrix Q implies to generate some 3×3 random variables W from an $N(0, I_3)$ density (in case of a three variable VAR). Then we decompose $W = Q_R R$, where Q_R is an orthogonal matrix and R is a triangular matrix. The algorithm implemented to produce Q_R is the QR decomposition. Since many draws of W can be made, one can find many Q_R. According to Rubio-Ramirez, Waggoner, and Zha (2010), as the size of the VAR grows, this is a computationally efficient strategy relative to the Givens approach. Fry and Pagan (2007) show that the methods (Givens matrices and Householder method) are equivalent, except for the computational speed in favor to the second method. Canova and Paustian (2012) propose an algorithm which derives a set of robust restrictions from a class of structural DSGE models that one may exploit to identify the shock(s) of interest with VAR.

5 Empirical Analysis

In this paper, we use the model with trend inflation described in Section 2 as Data Generating Process. We generate 100000 observations and we take the last 1000 observations, avoiding the influence of the initial conditions[1].

We aim at comparing the true (DSGE-consistent) impulse responses with those produced with a VAR with the Cholesky identification and with a VAR whose monetary policy shock is identified with sign restrictions. We identify the monetary policy shock by imposing constraints on the impulse responses of inflation and the policy rate to a monetary policy shock, as reported in Table 1. The constrains are imposed for the first K=2 pseudo-quarters, i.e., the one in which the shock occurs and the following quarter as proposed in Castelnuovo (2012). This choice is in line with Uhlig's (2005), which sets K=5 but deals with monthly data, instead in our paper we consider quarterly frequency. The VAR coefficients ad the variance-covariance matrix Σ have been fixed at the Maximum Likelihood Estimation (MLE) point estimate. We implement the Rubio-Ramirez, Waggoner, and Zha (2010) algorithm.

[1]We use Hodrick-Prescott filter to detrended our pseudo-data: inflation, output, and short-term interest rate.

Table 1 Sign restrictions for our DSGE model Shocks

$Variable/Shock$	y	π	R
Monetary Policy Shock		$\leq$	$\geq$

The VAR is implemented considering different number of lags. According to Schwarz criterion, we should implement a VAR with 2 lags. We compare the true impulse responses of the DSGE model with the Cholesky and sign restrictions identifications, considering as number of lags: 1,2, and 4.

5.1 Results

Many authors, e.g. Uhlig, 2005, have argued that an agnostic approach to sign restrictions can lead to a mute response (called as monetary neutrality) of output when the economy faces a contractionary monetary policy shock. In our exercise, we show that this result may well be driven by the fact that some endogenous state variables are unobservable. Our experiment shows that even if the true generation process implies a strong response of output after a monetary shock, the corresponding VAR analysis, characterized by an unobservable state variable shows a mute response of output to monetary policy shock. We conjecture that this discrepancy is due to a truncation bias which affects the statistical representation of our true data generation process as in Ravenna (2007) Chari et al. (2008).

From Figure 1 to 6, we report the comparison between the impulse responses of the DSGE (called true IRF) and the impulse responses with sign-restrictions, considering 1, 2, and 4 lags in the VAR representation. Specifically, in Figure 2, 4 and 6, we report the confidence interval (considering the 5th and the 95th percentiles of the distribution) of the VAR with sign-restrictions.

The inflation have been restricted not to be positive for two quarters, and plots show that the response is negative up the second quarter, after it becomes close to zero. We notice that the response from the DSGE model is included in the intervals.

Focusing on the impact of a monetary policy shock on the output, we can notice that the response is close to zero across lags (from 1 to 4). Considering the confidence intervals, we evidence that the response of the model is not included in the intervals in the first periods (around 6 quarters) for VAR(1) and VAR(4). Instead in case of the VAR(2), the response of the model is included in the confidence interval after the first quarter, and the intervals show that the response can be significantly different from zero.

The responses of the short-term interest rate given by the VAR with sign restrictions is close to the response produced by the DSGE and confidence intervals show that the responses are significantly different from zero.

6 A Theoretical Generalization

[TBA]

References

[1] Adolfson Malin, Stefan Laséen, Jesper Lindé and Mattias Villani (2008): "Evaluating an Estimated New Keynesian Small Open Economy Model", *Journal of Economic Dynamics and Control* Elsevier, vol. 32(8), pages 2690-2721.

[2] Ascari, Guido (2004): "Staggered Prices and Trend Inflation: Some Nuisances," *Review of Economic Dynamics*, Elsevier for the Society for Economic Dynamics,

vol. 7(3), pages 642-667, July.

[3] Ascari, Guido and Tiziano Ropele (2009): "Trend Inflation, Taylor Principle, and Indeterminacy," *Journal of Money, Credit and Banking*, Blackwell Publishing, vol. 41(8), pages 1557-1584, December.

[4] Ascari, Guido, Efrem Castelnuovo, and Lorenza Rossi (2011): "Calvo vs. Rotemberg in a Trend Inflation World: An Empirical Investigation", *Journal of Economic Dynamics and Control*, 35(11), 1852-1867.

[5] Bernanke, Ben S. and Ilian Mihov (1998), "Measuring Monetary Policy", *Quarterly Journal of Economics*, 113 (3),

[6] 869–902.

[7] Blanchard, Olivier J. and Danny Quah (1993): "The Dynamic Effects of Aggregate Demand and Supply Disturbances", *American Economic Review*, 83 (3), 655–673.

[8] Calvo, Guillermo A. (1983): "Staggered prices in a utility-maximizing framework," *Journal of Monetary Economics*, Elsevier, vol. 12(3), pages 383-398, September.

[9] Canova, Fabio and Gianni Denicolò (2002): "Monetary Disturbances Matter for Business Fluctuations in the G-7", *Journal of Monetary Economics*, 49 (6), 1121–1159.

[10] Canova, Fabio and Joaquim P. Pina (2005): "What VAR tells about DSGE models? ", in Deibolt, C. and Krytsou. C. (eds.) New Trends and Macroeconomics, Springer Verlag.

[11] Canova, Fabio and Sala, Luca (2009): "Back to Square One: Identification issues in DSGE models", *Journal of Monetary Economics*, vol. 56(4), pages 431-449.

[12] Canova, Fabio and Matthias Paustian (2012): "Business cycle measurement with

some theory", *Journal of Monetary Economics*, 58 (4), 345-361, 2011.

[13] Castelnuovo, Efrem (2012): "Monetary Policy Neutrality? Sign Restrictions Go to Monte Carlo", mimeo.

[14] Chari, V.V., Patrick Kehoe and Ellen R. McGrattan (2005): "A Critique of Structural VARs Using Real Business Cycle Theory", Federal Reserve Bank of Minneapolis, Working Paper 631.

[15] Chari, V.V., Patrick Kehoe and Ellen R. McGrattan (2008): "Are structural VARs with long-run restrictions useful in developing business cycle theory?," *Journal of Monetary Economics*, Elsevier, vol. 55(8), pages 1337-1352, November.

[16] Christiano, Lawrence J., Martin Eichenbaum and Charles L. Evans (1996): "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds", *Review of Economics and Statistics*, 78 (1), 16–34.

[17] Christiano, Lawrence J., Martin Eichenbaum and Charles L. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, University of Chicago Press, vol. 113(1), pages 1-45, February.

[18] Christiano, Lawrence J., Martin Eichenbaum and Robert Vigfusson (2006): "Assessing Structural VARs", in: Acemoglu, D., Rogoff, K. and Woodford, M. (Eds.), NBER Macroeconomics Annual 2006, Cambridge: The MIT Press.

[19] Cooley, Thomas F and Mark Dwyer (1998): "Business Cycle Analysis Without Much Theory. A Look at Structural VARs", *Journal of Econometrics* 83(1-2), 57-88.

[20] Del Negro, Marco and Frank Schorfheide (2004): "Priors from General equilibrium Models for VARs", *International Economic Review*, 45, pp 643-673.

[21] Del Negro, Marco and Frank Schorfheide (2006): "How Good Is What You've Got? DSGE-VAR as a Toolkit for Evaluating DSGE Models", *Federal Reserve Bank of Atlanta Economic Review* 91, Second Quarter 2006.

[22] Del Negro, Marco, Frank Schorfheide, Frank Smets and Raf Wouters (2007a): "On the Fit of New-Keynesian Models", *Journal of Business and Economic Statistics*, Volume 25, Number 2 April 2007, pp123-143 .

[23] Del Negro, Marco, Frank Schorfheide, Frank Smets and Raf Wouters (2007b): "On the Fit of New-Keynesian Models: Rejoinder", *Journal of Business and Economic Statistics*, Volume 25, Number 2 April 2007, pp159-162 .

[24] Faust, Jon (1998): "On the Robustness of Identified VAR Conclusions about Money", *Carnegie-Rochester Conference Series on Public Policy 49, 207-244.*

[25] Fernandez-Villaverde, Jesus, Juan Francisco Rubio-Ramirez , Thomas Sargent and Mark Watson (2007): "A, B, C's (and D's) for understanding VARs", *the American Economic Review,* 97, 3, pp 1021–1026.

[26] Fry, Renee and Adrian Pagan (2007): "Some Issues in Using Sign Restrictions for Identifying Structural VARs," NCER Working Paper Series 14, National Centre for Econometric Research.

[27] Fry, Renee and Adrian Pagan (2011): "Sign Restrictions in Structural Vector Autoregressions: A Critical Review," *Journal of Economic Literature*, American Economic Association, vol. 49(4), pages 938-60, December.

[28] Gali, Jordì (1992): "How Well Does the IS-LM Model Fit Postwar U.S. Data?", *Quarterly Journal of Economics*, 107 (2), 709–738.

[29] Gali, Jordì and Rabanal, Pau, (2005): "Technology shocks and aggregate fluctuations: how well does the RBC model fit postwar US data?", in Gertler, M. and Rogoff, K., eds., NBER Macroeconomics Annual, MIT Press.ù

[30] Gordon, David B. and Eric Leeper, (1994): "The Dynamic Impacts of Monetary Policy: An Exercise in Tentative

[31] Identification", *Journal of Political Economy*, 102 (6), 1228–1247.

[32] Hansen, Lars and Thomas Sargent (2008): "Recursive Linear Models of Dynamic Economies", Princeton University Press.

[33] Ingram, Beth and Charles Whiteman (1994): "Supplanting the Minnesota Prior - Forecasting Macroeconomics Time Series using Real Business Cycle Model Priors", *Journal of Monetary Economics,* 34,pp 497-510.

[34] Iskrev, Nikolay (2010): "Local identification in DSGE models", *Journal of Monetary Economics*, vol. 57(2), pages 189-202.

[35] King, Robert G., Plosser, Charles I., Stock, James H. and Watson, Mark W. (1991): "Stochastic Trends and Economic

[36] Fluctuations", *American Economic Review*, 81 (4), 819–840.

[37] King, Robert G. (2000): "The New IS-LM Model: Language, Logic, and Limits", *Federal Reserve Bank of Richmond Economic Quarterly*, 86, pp 45-103.

[38] Peersman, Gert (2005): "What caused the early millennium slowdown? Evidence based on vector autoregressions," *Journal of Applied Econometrics*, John Wiley & Sons, Ltd., vol. 20(2), pages 185-207.

[39] Ravenna, Federico (2007): "Vector Autoregressions and Reduced Form Representations of DSGE Models", *Journal of Monetary Economics,* 54, pp. 2048-2064.

[40] Rotemberg, Julio J. and Michael Woodford (1998): "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy: Expanded Version," NBER Technical Working Papers 0233, National Bureau of Economic Research, Inc.

[41] Rubio-Ramirez, Juan, Daniel F. Waggoner and Tao Zha (2008): "Structural Vector Autoregressions: Theory of

[42] Identification and Algorithms for Inference" (Federal Reserve Bank of Atlanta Working Paper 2008-18, September 2008).

[43] Rubio-Ramirez, Juan, Daniel F. Waggoner and Tao Zha (2010): "Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference," *Review of Economic Studies*, Wiley Blackwell, vol. 77(2), pages 665-696, 04.

[44] Sims, Christopher A. (1980): "Macroeconomics and Reality", *Econometrica*, 48: pp 1-48.

[45] Sims, Christopher A. (1986): "Are forecasting models usable for policy analysis?," Quarterly Review, Federal Reserve Bank of Minneapolis, issue Win, pages 2-16.

[46] Sims, Chistopher. A. (2002): "Solving Linear Rational Expectations Models", *Computational Economics*, 20 (1-2), pp 1-20.

[47] Sims, Chistopher. A. and Tao Zha (2006): "Were There Regime Switches in US Monetary Policy?", *American Economic Review*, 96, 54–81.

[48] Sims, Eric R. (2012): "News, Non-Invertibility, and Structural VARs", mimeo.

[49] Smets, Frank and Raf Wouters (2003): "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area", *Journal of the European Economic Association*, 1, 1123-75.

[50] Smets, Frank and Raf Wouters (2004): "Forecasting with a Bayesian DSGE Model: An Application to the Euro Area," *Journal of Common Market Studies*, Wiley Blackwell, vol. 42(4), pages 841-867, November.

[51] Smets, Frank and Raf Wouters (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach", *American Economic Review*, American

Economic Association, vol. 97(3), pages 586-606, June.

[52] Uhlig, Harald (2005): "What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure." *Journal of Monetary Economics* 52(2), 381-419.

[53] Woodford, Michael (2003): "*Interest and Prices*", Princeton University Press.

[54] Zha, Tao (1999): "Block Recursion and Structural Vector Autoregressions", *Journal of Econometrics*, 90 (2), 291–316.

7 Calibration

Parameter	Description	Value
β	Discount factor	0.99025
π	Trend Inflation	$1.02^{1/4}$
θ	Calvo Parameter	0.75
φ	Inverse Frisch	1
σ	CRRA	2
ε	Market Power	6
μ	Indexation Parameter	0.8
χ	Indexation Parameter	0
ω	Backward Looking Euler	0.7
ρ	Backward Looking Mon.Rule	0.6
ϕ_π	Inflation policy parameter	2
ϕ_y	Output Policy Parameter	0.05
ρ^π	Persistence Cost Push shock	0.8
ρ^y	Persistence Demand shock	0.7
ρ^r	Persistence Monetary Shock	0.6
λ	Convolution of Parameters	
η	Convolution of Parameters	
ψ	Convolution of Parameters	

8 Figures

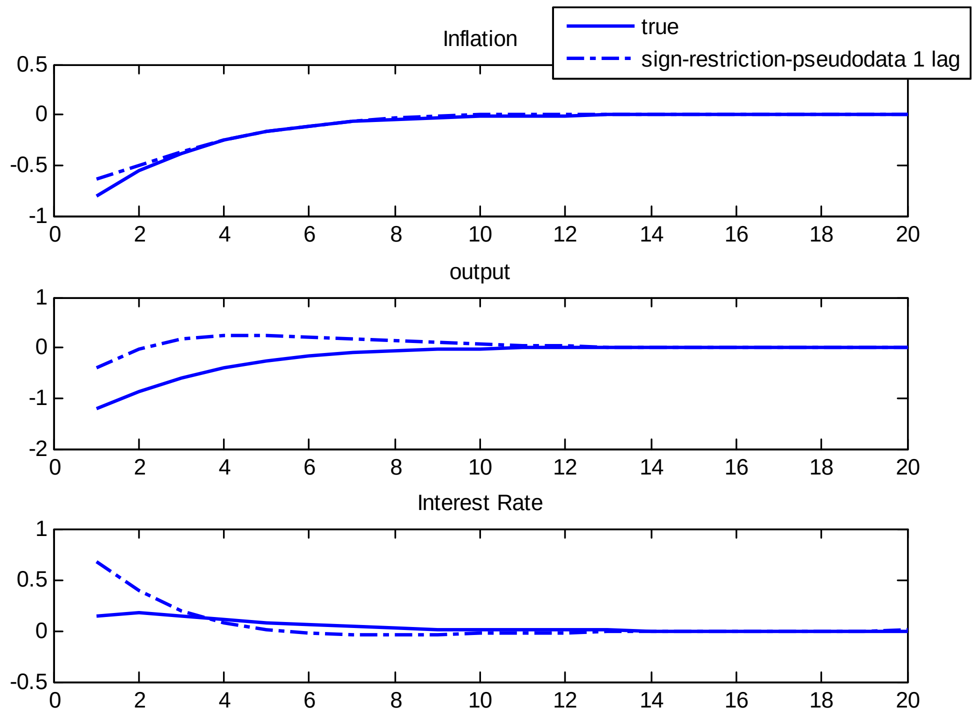


Figure 1

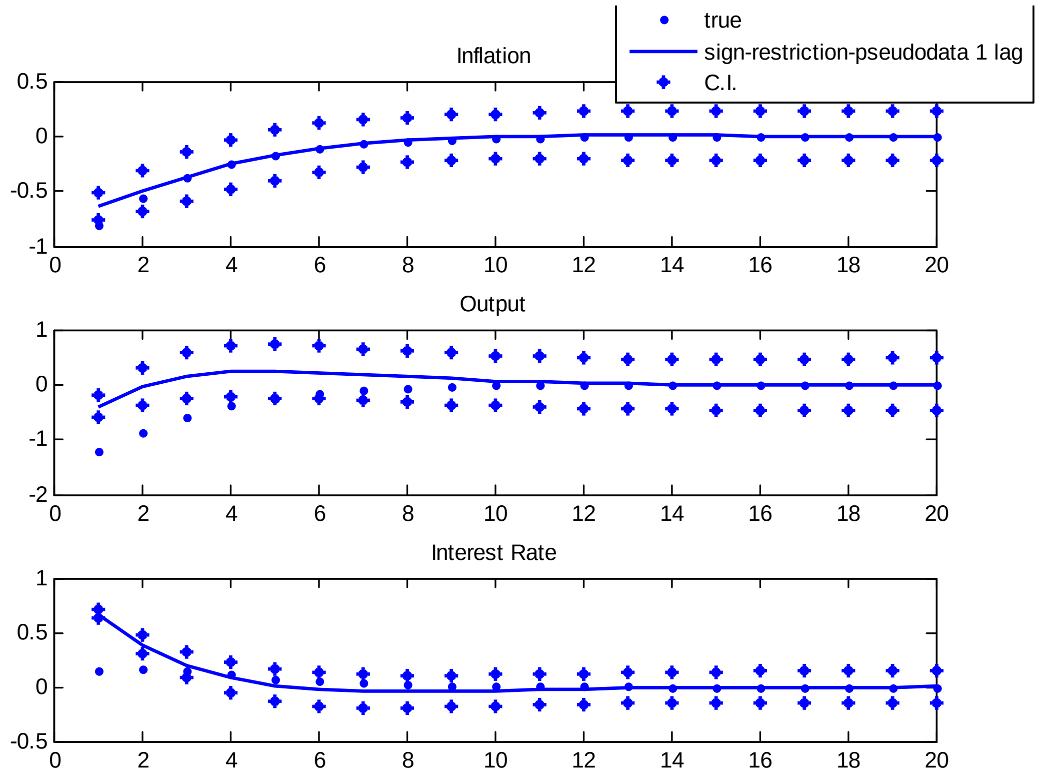
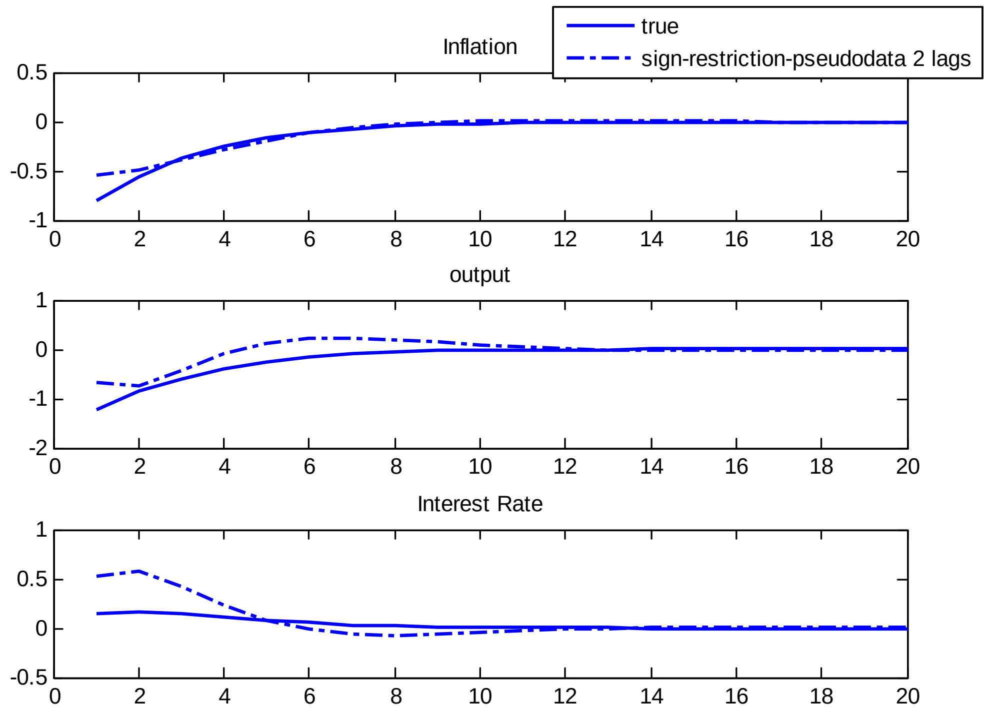

true
sign-restriction-pseudodata 1 lag
C.I.
Inflation
0.5
0
-0.5
-1
0
2
4
6
8
10
12
14
16
18
20
Output
1
0
-1
-2
Interest Rate
1
0.5
0
-0.5

Figure 2



true
sign-restriction-pseudodata 2 lags
Inflation
0.5
0
-0.5
-1
0
2
4
6
8
10
12
14
16
18
20
output
1
0
-1
-2
Interest Rate
1
0.5
0
-0.5

Figure 3

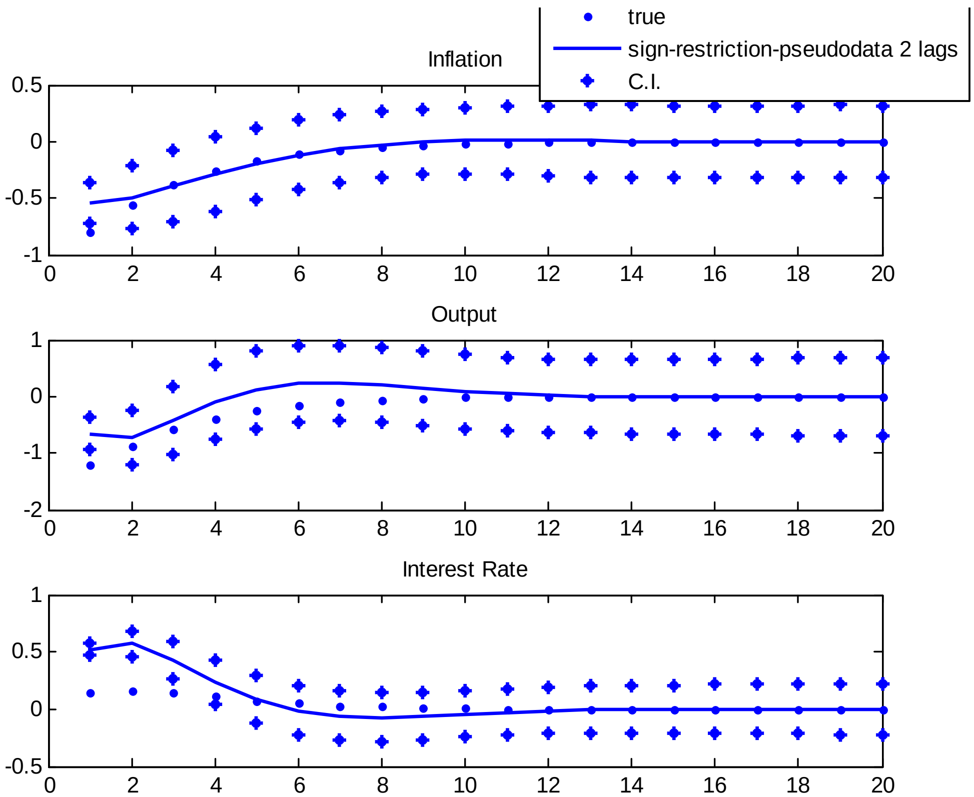

true
sign-restriction-pseudodata 2 lags
C.I.
Inflation
0.5
0
-0.5
-1
0 2 4 6 8 10 12 14 16 18 20
Output
1
0
-1
-2
0 2 4 6 8 10 12 14 16 18 20
Interest Rate
1
0.5
0
-0.5
0 2 4 6 8 10 12 14 16 18 20

Figure 4

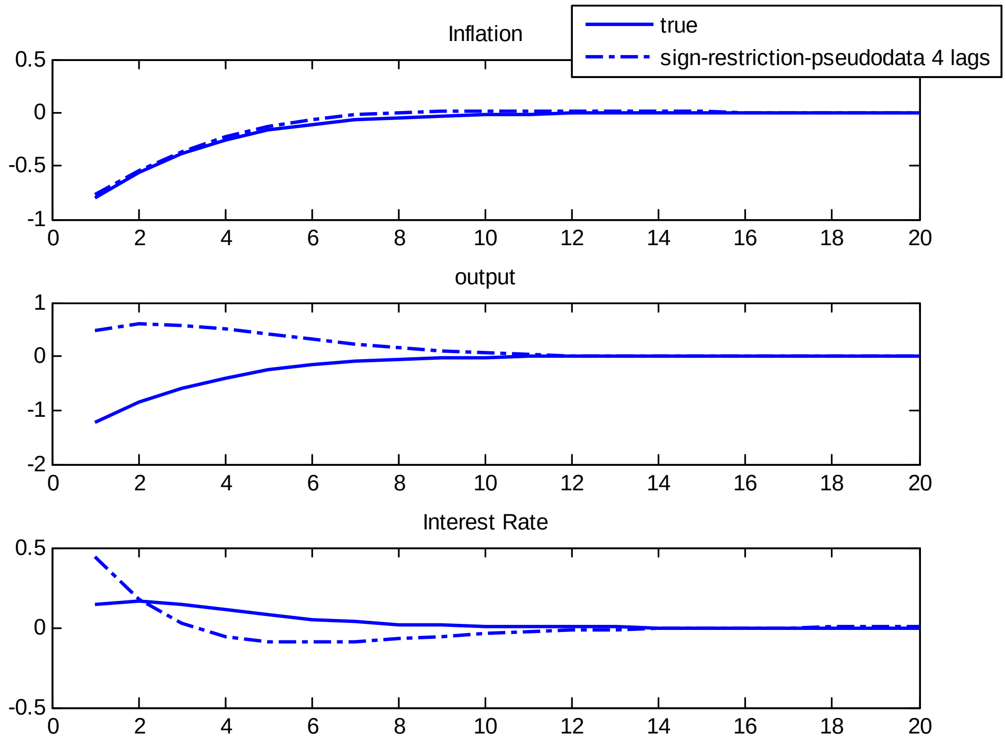

Inflation
true
sign-restriction-pseudodata 4 lags
0.5
0
-0.5
-1
0
2
4
6
8
10
12
14
16
18
20
output
1
0
-1
-2
0
2
4
6
8
10
12
14
16
18
20
Interest Rate
0.5
0
-0.5
0
2
4
6
8
10
12
14
16
18
20

Figure 5

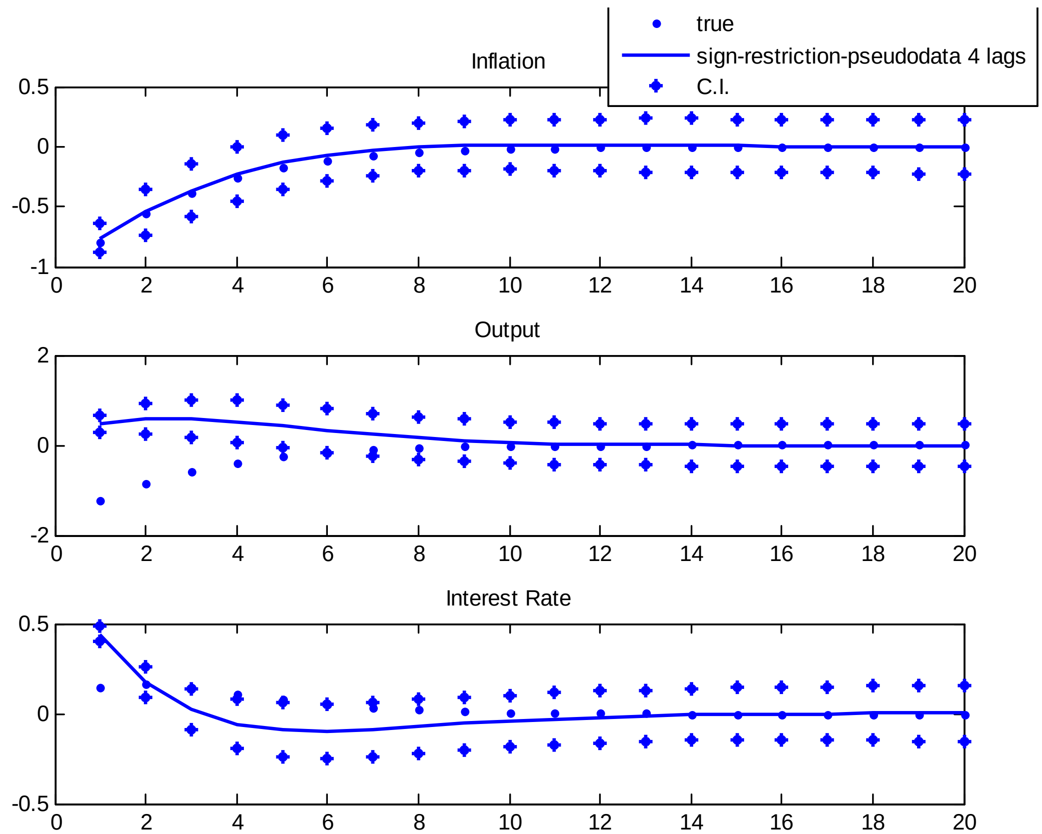
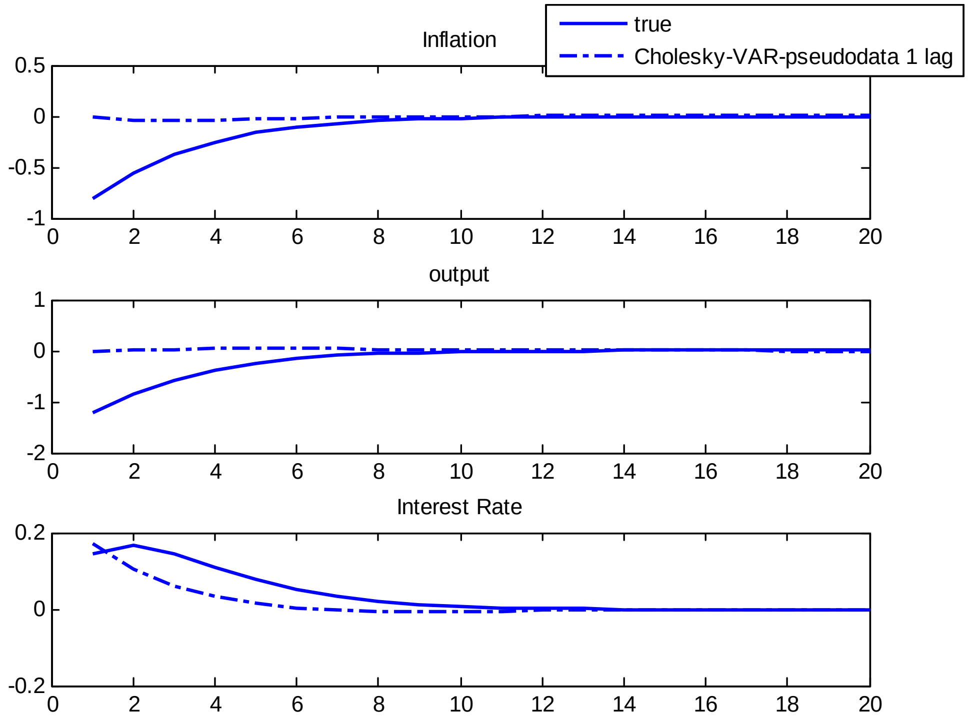
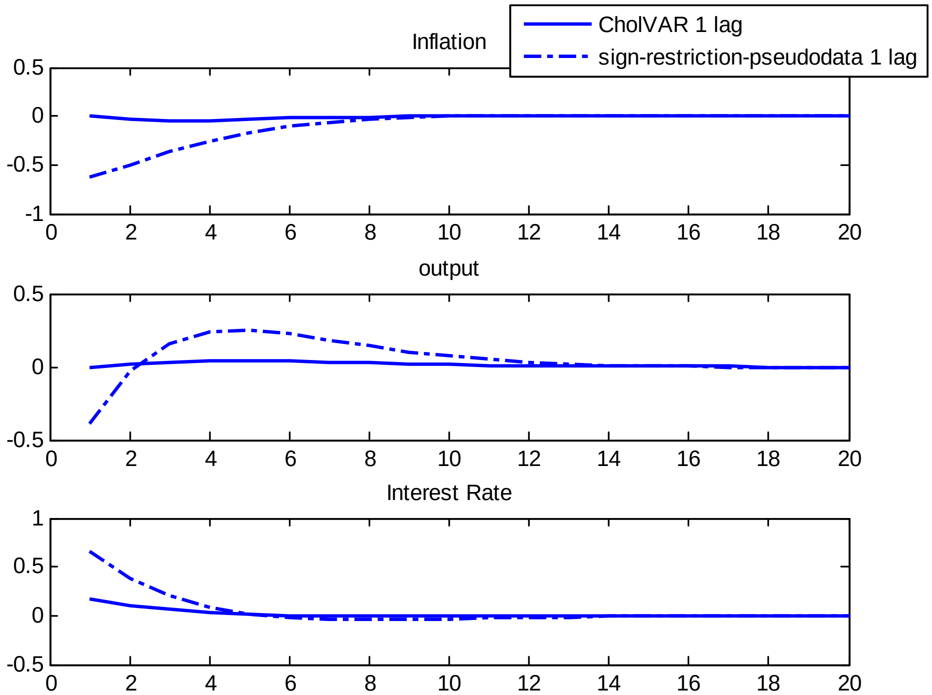
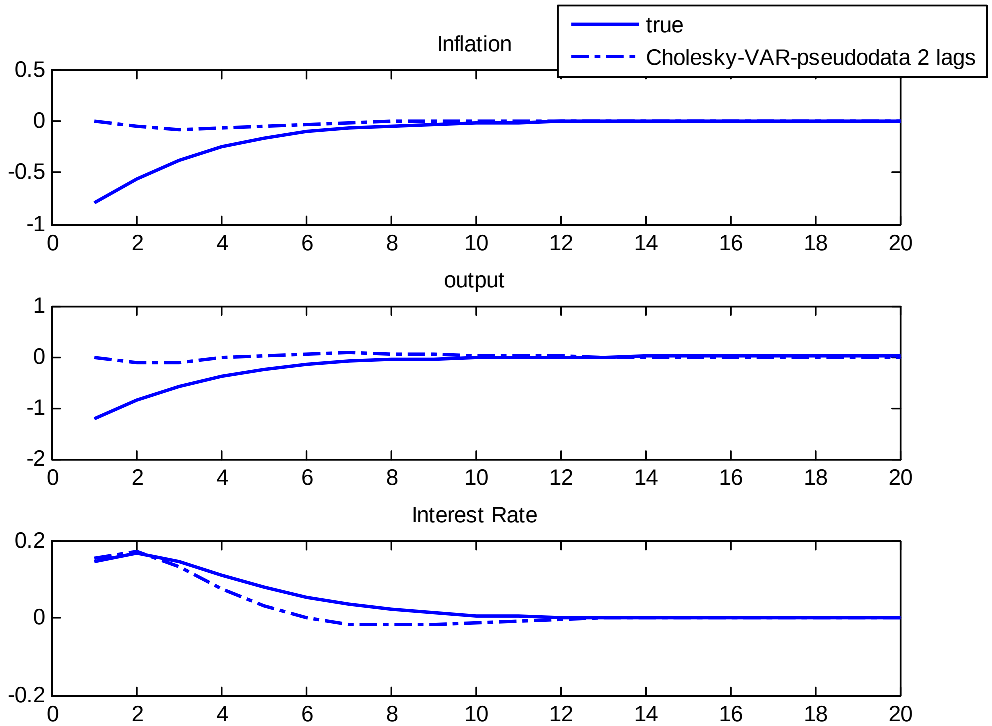

true
sign-restriction-pseudodata 4 lags
C.I.
Inflation
0.5
0
-0.5
-1
0
2
4
6
8
10
12
14
16
18
20
Output
2
0
-2
0
2
4
6
8
10
12
14
16
18
20
Interest Rate
0.5
0
-0.5
0
2
4
6
8
10
12
14
16
18
20

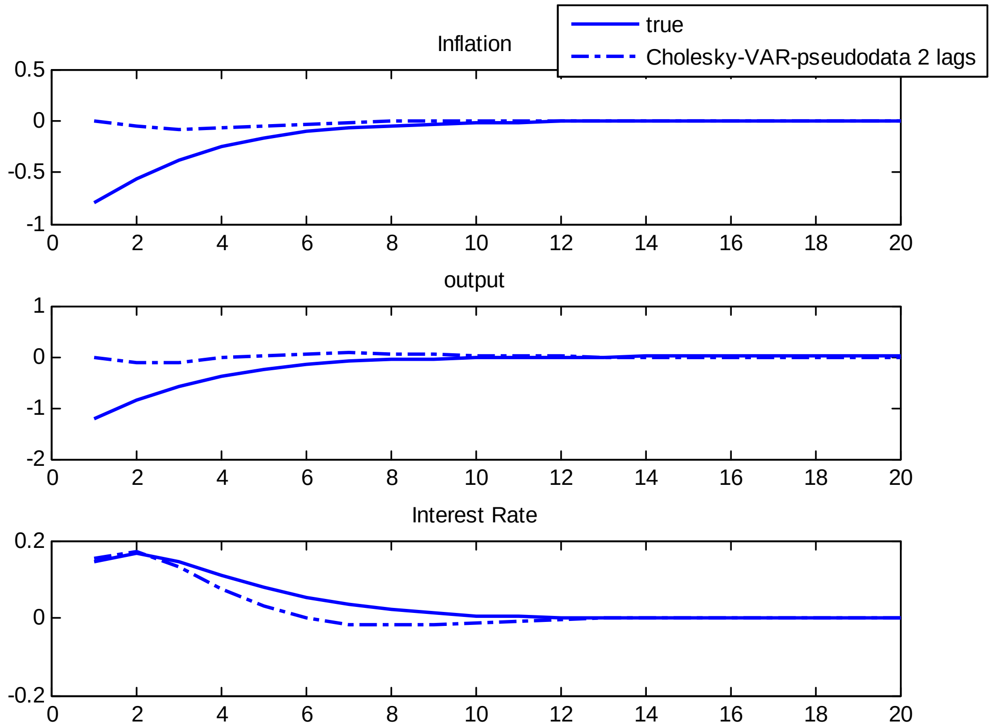
Figure 6

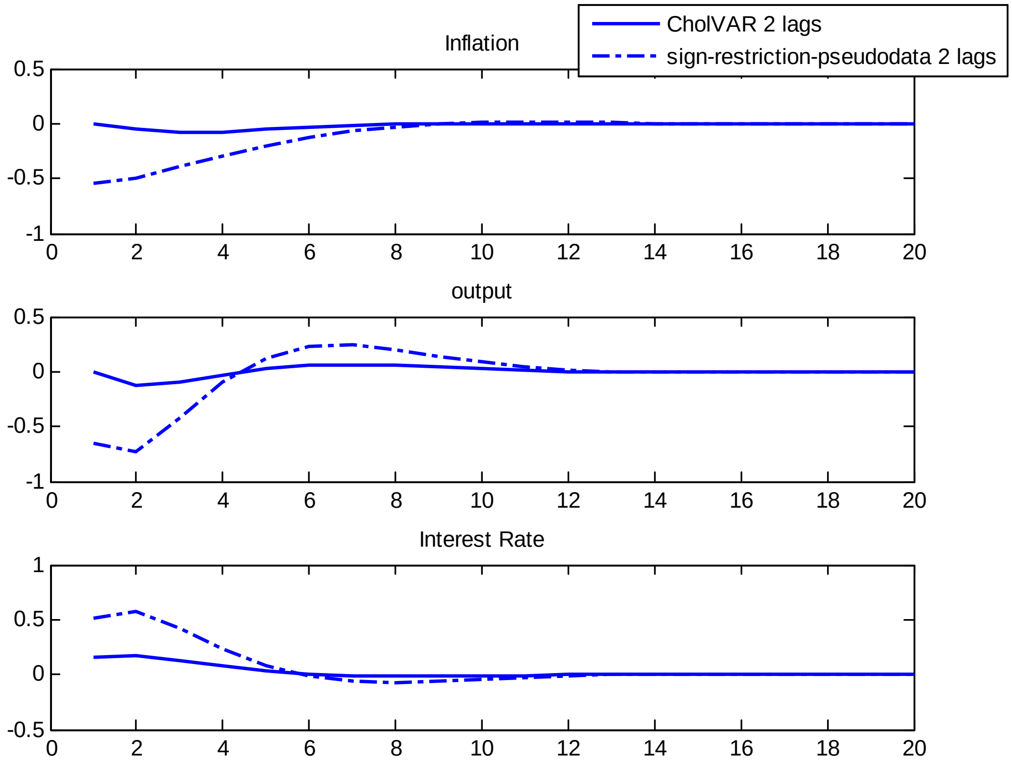
8.1 Figures: Robustness with Cholesky Decomposition

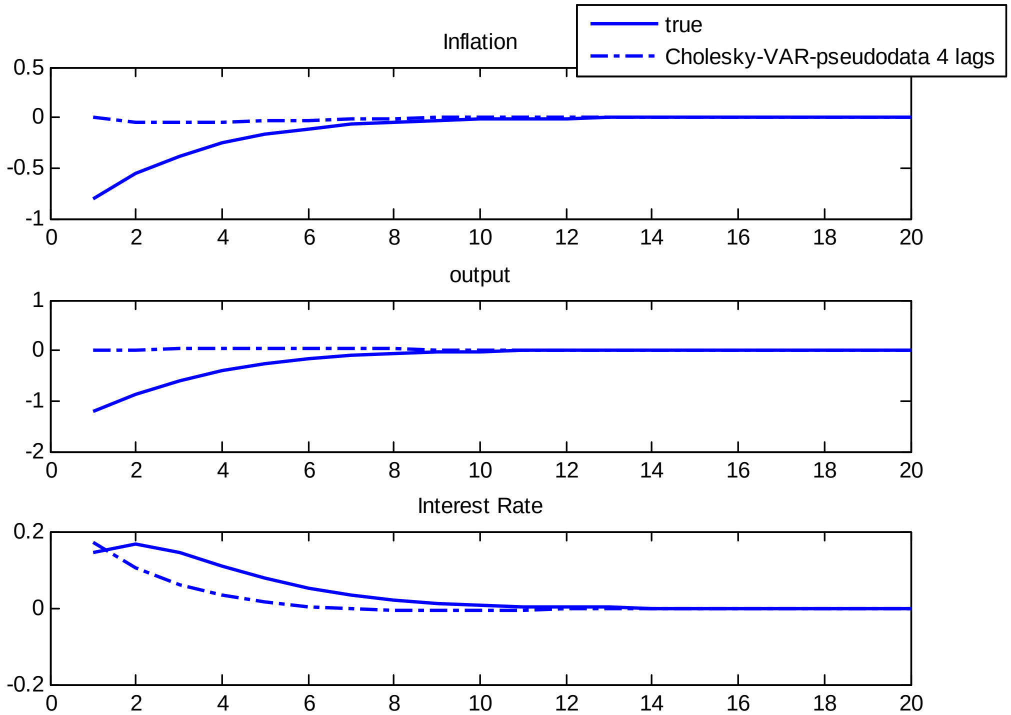



CholVAR 1 lag
sign-restriction-pseudodata 1 lag
Inflation
0.5
0
-0.5
-1
0
2
4
6
8
10
12
14
16
18
20
output
0.5
0
-0.5
0
2
4
6
8
10
12
14
16
18
20
Interest Rate
1
0.5
0
-0.5
0
2
4
6
8
10
12
14
16
18
20

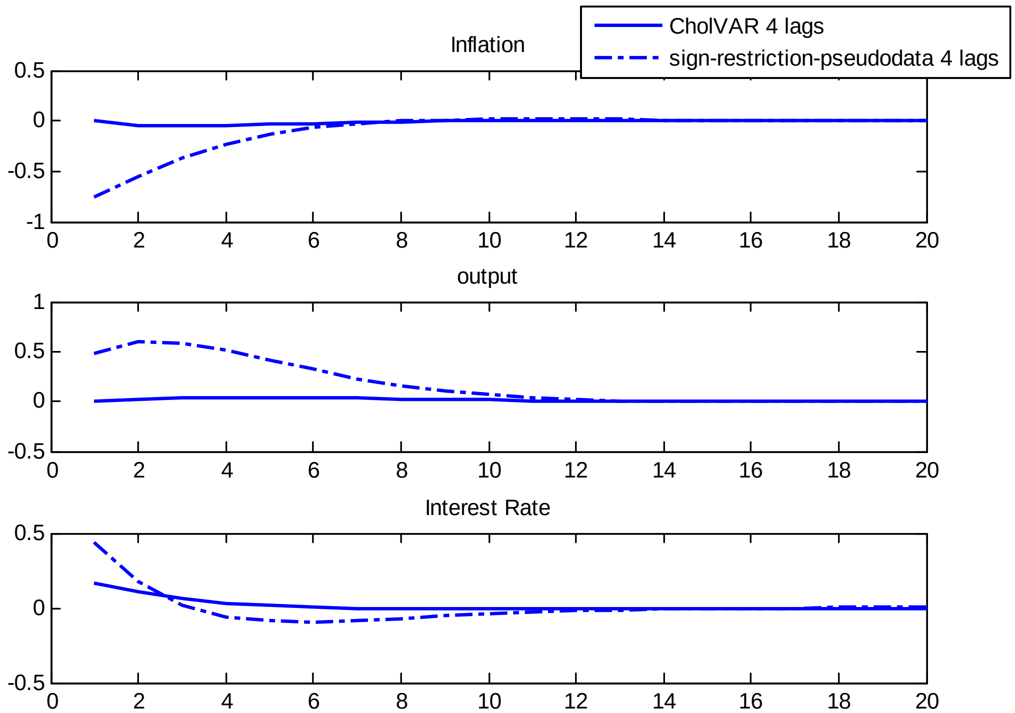

true
Cholesky-VAR-pseudodata 2 lags
Inflation
0.5
0
-0.5
-1
0 2 4 6 8 10 12 14 16 18 20
output
1
0
-1
-2
0 2 4 6 8 10 12 14 16 18 20
Interest Rate
0.2
0
-0.2
0 2 4 6 8 10 12 14 16 18 20


true
Cholesky-VAR-pseudodata 2 lags
Inflation
0.5
0
-0.5
-1
0
2
4
6
8
10
12
14
16
18
20
output
1
0
-1
-2
0
2
4
6
8
10
12
14
16
18
20
Interest Rate
0.2
0
-0.2
0
2
4
6
8
10
12
14
16
18
20


CholVAR 2 lags
sign-restriction-pseudodata 2 lags
Inflation
0.5
0
-0.5
-1
0 2 4 6 8 10 12 14 16 18 20
output
0.5
0
-0.5
-1
0 2 4 6 8 10 12 14 16 18 20
Interest Rate
1
0.5
0
-0.5
0 2 4 6 8 10 12 14 16 18 20



Inflation
true
Cholesky-VAR-pseudodata 4 lags
0.5
0
-0.5
-1
0
2
4
6
8
10
12
14
16
18
20
output
1
0
-1
-2
0
2
4
6
8
10
12
14
16
18
20
Interest Rate
0.2
0
-0.2
0
2
4
6
8
10
12
14
16
18
20



CholVAR 4 lags
sign-restriction-pseudodata 4 lags
Inflation
0.5
0
-0.5
-1
0
2
4
6
8
10
12
14
16
18
20
output
1
0.5
0
-0.5
0
2
4
6
8
10
12
14
16
18
20
Interest Rate
0.5
0
-0.5
0
2
4
6
8
10
12
14
16
18
20

9 Appendix

9.1 DSGE-VAR

The basic idea of the Del Negro-Schorfheide (2004) approach is to use the DSGE model to build prior distributions for the VAR.

The starting point for the estimation is an unrestricted VAR of order p

$$\mathbf{Y}_t = \mathbf{\Phi}_0 + \mathbf{\Phi}_1 \mathbf{Y}_{t-1} + ... + \mathbf{\Phi}_p \mathbf{Y}_{t-p} + \mathbf{u}_t \tag{26}$$

In compact format:

$$\mathbf{Y} = \mathbf{X\Phi} + \mathbf{U} \tag{27}$$

$\mathbf{Y}$ is a $(T \times n)$ matrix with rows Y_t', $\mathbf{X}$ is a $(T \times k)$ matrix $(k = 1 + np, p =$number of lags) with rows $X_t' = [1, Y_{t-1}', ..., Y_{t-p}']$, $\mathbf{U}$ is a $(T \times n)$ matrix with rows u_t' and $\mathbf{\Phi}$ is a $(k \times n) = [\Phi_{0,} \Phi_{1,} ..., \Phi_p]'$.The one-step-ahead forecast errors u_t have a multivariate normal distribution $N(0, \mathbf{\Sigma}_u)$ conditional on past observations of Y.The log-likelihood function of the data is a function of $\mathbf{\Phi}$ and $\mathbf{\Sigma}_u$

$$L(\mathbf{Y}|\mathbf{\Phi}, \mathbf{\Sigma}_u) \propto |\mathbf{\Sigma}_u|^{-\frac{T}{2}} \exp\left\{ -\frac{1}{2} tr \left[\Sigma_u^{-1} \left(\mathbf{Y'Y} - \mathbf{\Phi' X' Y} - \mathbf{Y' X \Phi} + \mathbf{\Phi' X' X \Phi} \right) \right] \right\} \tag{28}$$

The prior distribution for the VAR parameters proposed by Del Negro and Schorfheide (2004) is based on the statistical representation of the DSGE model given by a VAR approximation. Let Γ_{xx}^*, Γ_{yy}^*, Γ_{xy}^* and Γ_{yx}^* be the theoretical second-order moments of the variables $\mathbf{Y}$ and $\mathbf{X}$ implied by the DSGE model, where

$$\begin{aligned} \mathbf{\Phi}^* (\theta) &= \mathbf{\Gamma}^{*-1}_{xx} (\theta) \, \mathbf{\Gamma}^*_{xy} (\theta) \\ \mathbf{\Sigma}^* (\theta) &= \mathbf{\Gamma}^*_{yy} (\theta) - \mathbf{\Gamma}^*_{yx} (\theta) \, \mathbf{\Gamma}^{*-1}_{xx} (\theta) \, \mathbf{\Gamma}^*_{xy} (\theta) \end{aligned} \tag{29}$$

The moments are the dummy observation priors used in the mixture model. These vectors can be interpreted as the probability limits of the coefficients in a VAR estimated on the artificial observations generated by the DSGE model. Conditional on the vector of structural parameters in the DSGE model θ, the prior distributions for the VAR parameters $p(\mathbf{\Phi}, \mathbf{\Sigma}_u|\theta)$ are of the Inverted-Wishart (IW) and Normal forms

$$\begin{aligned} \mathbf{\Sigma}_u \, |\theta &\sim IW \left((\lambda T \mathbf{\Sigma}^*_u (\theta) \, , \lambda T - k, n \right) \\ \mathbf{\Phi} \, |\mathbf{\Sigma}_u, \theta &\sim N \left(\mathbf{\Phi}^* (\theta) \, , \mathbf{\Sigma}_u \otimes (\lambda T \mathbf{\Gamma}_{XX} (\theta))^{-1} \right) \end{aligned} \tag{30}$$

where the parameter λ controls the degree of model misspecification with respect to the VAR; for small values of λ the discrepancy between the VAR and the DSGE-VAR is large and a sizeable distance is generated between the unrestricted VAR and DSGE estimators. Large values of λ correspond to small model misspecification and for $\lambda = \infty$ beliefs about DSGE misspecification degenerate to a point mass at zero. Bayesian estimation could be interpreted as estimation based on a sample in which data are augmented by a hypothetical sample where observations are generated by the DSGE model, the so-called dummy prior observations (Theil and Goldberg 1961; Ingram and Whiteman 1994). Within this framework λ determines the length of the hypothetical sample.

The posterior distributions of the VAR parameters are also of the Inverted-Wishart and Normal forms. Given the prior distribution, posterior distributions are derived by the Bayes theorem

$$\mathbf{\Sigma}_u \, |\theta, \mathbf{Y} \sim IW \left((\lambda + 1) \, T \hat{\mathbf{\Sigma}}_{u,b} (\theta) \, , (\lambda + 1) \, T - k, n \right) \tag{31}$$

$$\mathbf{\Phi}\left|\mathbf{\Sigma}_u, \theta, \mathbf{Y} \sim N\left(\hat{\mathbf{\Phi}}_b(\theta), \mathbf{\Sigma}_u \otimes \left[\lambda T \mathbf{\Gamma}_{XX}(\theta) + \mathbf{X}'\mathbf{X}\right]^{-1}\right)\right. \tag{32}$$

$$\hat{\mathbf{\Phi}}_b(\theta) = \left(\lambda T \mathbf{\Gamma}_{XX}(\theta) + \mathbf{X}'\mathbf{X}\right)^{-1}\left(\lambda T \mathbf{\Gamma}_{XY}(\theta) + \mathbf{X}'\mathbf{Y}\right) \tag{33}$$

$$\hat{\mathbf{\Sigma}}_{u,b}(\theta) = \frac{1}{(\lambda+1)T}\left[\left(\lambda T \mathbf{\Gamma}_{YY}(\theta) + \mathbf{Y}'\mathbf{Y}\right) - \left(\lambda T \mathbf{\Gamma}_{XY}(\theta) + \mathbf{X}'\mathbf{Y}\right)\hat{\mathbf{\Phi}}_b(\theta)\right] \tag{34}$$

where the matrices $\hat{\mathbf{\Phi}}_b(\theta)$ and $\hat{\mathbf{\Sigma}}_{u,b}(\theta)$ have the interpretation of maximum likelihood estimates of the VAR parameters based on the combined sample of actual observations and artificial observations generated by the DSGE. Equations (31) and (32) show that the smaller λ is, the closer the estimates are to the OLS estimates of an unrestricted VAR. Instead, the higher λ is, the closer the VAR estimates will be tilted towards the parameters in the VAR approximation of the DSGE model ($\hat{\mathbf{\Phi}}_b(\theta)$ and $\hat{\mathbf{\Sigma}}_{u,b}(\theta)$). In order to obtain a non-degenerate prior density (30), which is a necessary condition for the existence of a well-defined Inverse-Wishart distribution, and for computing meaningful marginal likelihoods λ has to be greater than λ_{MIN}

$$\begin{aligned}
\lambda_{MIN} &\geq \frac{n+k}{T}; k = 1 + p \times n \\
p &= \text{lags} \\
n &= \text{endogenous variables.}
\end{aligned}$$

Hence, the optimal lambda must be greater than or equal to the minimum lambda $\left(\hat{\lambda} \geq \lambda_{MIN}\right)$.

Essentially, the DSGE-VAR tool allows the econometrician to draw posterior inferences about the DSGE model parameters θ. Del Negro and Schorfheide (2004) explain that the posterior estimate of θ has the interpretation of a minimum-distance estimator, where the discrepancy between the OLS estimates of the unrestricted VAR parameters and the VAR representation of the DSGE model is a sort of distance function. The estimated posterior of parameter vector θ depends on the hyperparameter λ. When $\lambda \to 0$, in the posterior the parameters are not informative, so the DSGE model is of no use in explaining the data. Unfortunately, the posteriors (32) and (31) do not have a closed form and we need a numerical method to solve the problem. The posterior simulator used by Del Negro and Schorfheide (2004) is the Markov Chain Monte Carlo Method and the algorithm used is the Metropolis-Hastings acceptance method. This procedure generates a Markov Chain from the posterior distribution of θ and this Markov Chain is used for Monte Carlo simulations. The optimal λ is given by maximizing the log of the marginal data density

$$\hat{\lambda} = \underset{\lambda \geqslant \lambda_{MIN}}{\arg\max} \ln p(\mathbf{Y}|\lambda)$$

According to the optimal lambda $\left(\hat{\lambda}\right)$, a corresponding optimal mixture model is chosen. This hybrid model is called DSGE-VAR$\left(\hat{\lambda}\right)$ and $\hat{\lambda}$ is the weight of the priors. It can also be interpreted as the restriction of the theoretical model on the actual data.