barcelona gse graduate school of economics

Estimating Overidentified, Nonrecursive Time-Varying Coefficients Structural VARs

Fabio Canova
Fernando J. Pérez Forero

May 2012

Barcelona GSE Working Paper Series

Working Paper nº 637

Estimating overidentified, nonrecursive, time-varying coefficients structural VARs

Fabio Canova Fernando J. Pérez Forero*

May 8, 2012

Abstract

This paper provides a method to estimate time varying coefficients structural VARs which are non-recursive and potentially overidentified. The procedure allows for linear and non-linear restrictions on the parameters, maintains the multi-move structure of standard algorithms and can be used to estimate structural models with different identification restrictions. We study the transmission of monetary policy shocks and compare the results with those obtained with traditional methods.

JEL Classification: C11, E51, E52

Key words: Non-recursive overidentified SVARs, Time-varying coefficient models, Bayesian methods, Monetary transmission mechanism.

*ICREA-UPF, BSGE, CREMRED, CREI and CEPR; and UPF. Corresponding Address:Department of Economics, Universitat Pompeu Fabra, Ramon Trias Fargas 25–27, 08005 Barcelona, Spain.

1 Introduction

Vector autoregressive (VAR) models are routinely employed to summarize the properties of the data and new approaches to the identification of structural shocks have been suggested in the last 10 years (see Canova and De Nicoló, 2002, Uhlig, 2005, and Lanne and Lütkepohl, 2008). Constant coefficient structural VAR models are useful tools to understand how the economy responds to shocks but they may provide misleading information when the structure is changing over time. Cogley and Sargent (2005) and Primiceri (2005) were among the first to estimate time varying coefficient (TVC) VAR models and Primiceri also provides a structural interpretation of the dynamics using recursive restrictions on the impact response of shocks. Following Canova et al. (2008), Canova and Gambetti (2009), the literature nowadays mainly employes sign restrictions to identify structural shocks in TVC-VARs and the constraints used are, generally, theory based and robust to variations in the parameters of the DGP, see Canova and Paustian (2011).

While sign restrictions offer a simple and intuitive way to impose theoretical constraints on the data they are weak and identify a region of the parameter space, rather than a point. Furthermore, several implementation details are left to the researcher making comparison exercises difficult to perform. Because of these features, some investigators still prefer to use "hard" non-recursive identification restrictions, using the terminology of Waggoner and Zha (1999), even though these constraints are not theoretically abundant. However, when TVC models are used, existing estimation approaches can not deal with this type of identification restrictions.

This paper proposes a unified framework to estimate structural VARs. The framework can handle time varying coefficient or time invariant models, with hard recursive or non-recursive identification restrictions, and can be used in systems which are just-identified or overidentified. Non-recursive structures have been extensively used to accommodate structural models which are more complex than those permitted by recursive schemes. As shown, e. g., by Gordon and Leeper (1994), inference may crucially depend on whether a recursive or a non-recursive scheme is used. In addition, although just-identified systems are easier to construct and estimate, over-identified models have a long history in the literature (see e.g. Leeper et al., 1996, or Sims and Zha, 1998), and provide a natural framework to test interesting theoretical hypotheses.

TVC-VAR models are typically estimated using a Bayesian multi-move Gibbs sampling routine. In this routine, a state space system is specified (see Carter and Kohn,1994, and Kim and Nelson,1999) and the parameter vector is partitioned into blocks. When stochastic volatility is allowed for, an extended state space representation is used. If a recursive contemporaneous structure is assumed, one can sample the contemporaneous coefficients matrix equation by equation, taking as predetermined draws for the parameters belonging to previous equations. However, when the system is non-recursive, the sampling must be done differently.

To perform standard calculations, one also needs to assume that the covariance matrix of structural contemporaneous parameters is block-diagonal. When the structural model is overidentified, such an assumption may be implausible. However, relaxing the diagonality assumption complicates the computations since the blocks of the conditional distributions used in the Gibbs sampling do not necessarily have a known format. Primiceri (2005) suggests to use a Metropolis-step to deal with this problem. We follow his lead but nest this step into Geweke and Tanizaki (2001)'s approach to estimate general nonlinear state space models, modified to follow the multi-move Gibbs sampling logic. We employ Geweke and Tanizaki's setup because it allows for non-linear restrictions on the parameters, and thus can accommodate the structural systems which are the object of interest of this paper and others.

We use the methodology to identify monetary policy shocks in a non-recursive, overidentifed TVC system similar to the one employed by Robertson and Tallman (2001), Waggoner and Zha (2003) and Sims and Zha (2006). We compare the results with those obtained with a recursive, just-identified TVC models and with an overidentified, but fixed coefficient model. We show that there are important time variations in the variance of the monetary policy shock and in the estimated contemporaneous coefficients. These time variations translate in important changes in the transmission of monetary policy shocks which are consistent with the idea that the ability of monetary policy to influence the real economy has waned, especially in the 2000s. We also show that a recursive identification scheme or a fixed coefficient model produce a different characterization of the liquidity effect. Furthermore, the properties of the money demand function and its time variations would have been considerably different.

The paper is organized as follows: Section 2 presents the methodology employed to estimate non-recursive, overidentified TVC-VAR models. Section 3 applies it to study the transmission of monetary policy shocks. Section 4 summarizes the conclusions and discusses potential applications of the approach.

2 The methodology

Consider a $M \times 1$ vector of non-stationary variables y_t, $t = 1, \ldots, T$ and assume that it can be represented with a finite order autoregression:

$$y_t = B_{0,t}D_t + B_{1,t}y_{t-1} + ... + B_{p,t}y_{t-p} + u_t \tag{1}$$

where $B_{0,t}$ is a matrix of coefficients on a $\bar{M} \times 1$ vector of deterministic variables D_t; $B_{i,t}; i = 1, \ldots, p$ are square matrices containing the coefficients of the lags of the endogenous variables and $u_t \sim N(0, \Omega_t)$, where Ω_t is symmetric, positive definite, and full rank for every t. For the sake of presentation, we do not allow for exogenous variables, but the setup can be easily extended to account for them, if they exist. Equation (1) is a reduced form and the error u_t does not have an economic interpretation. Denote the structural shocks by $\varepsilon_t \sim N(0, I)$. Let the mapping between structural and reduced form shocks be

$$u_t = A_t^{-1} \Sigma_t \varepsilon_t \tag{2}$$

where A_t denotes the contemporaneous coefficients matrix and Σ_t is a diagonal matrix containing the standard deviations of the structural shocks. The structural $VAR(p)$ model is:

$$y_t = X_t' B_t + A_t^{-1} \Sigma_t \varepsilon_t \tag{3}$$

where $X_t' = I_M \otimes \left[D_t', y_{t-1}', \ldots, y_{t-k}'\right]$ and $B_t = \left[vec(B_{0,t})', vec(B_{1,t})', \ldots, vec(B_{p,t})'\right]'$ are a $M \times K$ matrix and a $K \times 1$ vector, where $K = \bar{M} \times M + pM^2$. As it is standard in the literature, assume that the parameter blocks (B_t, A_t, Σ_t) evolve as independent random-walks:

$$\begin{aligned} B_t &= B_{t-1} + \upsilon_t && (4) \\ \alpha_t &= \alpha_{t-1} + \zeta_t && (5) \\ \log(\sigma_t) &= \log(\sigma_{t-1}) + \eta_t && (6) \end{aligned}$$

where α_t denotes the vector of free parameters of A_t, $\sigma_t = diag(\Sigma_t)$ and let:

$$V = Var\left(\begin{bmatrix} \varepsilon_t \\ \upsilon_t \\ \zeta_t \\ \eta_t \end{bmatrix}\right) = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & Q & 0 & 0 \\ 0 & 0 & S & 0 \\ 0 & 0 & 0 & W \end{bmatrix} \tag{7}$$

Thus, the setup is able to capture time variations in i) the lag structure (see 4), ii) the contemporaneous reaction parameters (see 5) and iii) the structural variances (see 6). As shown in Canova et al. (2012), models with breaks at a specific date can be accommodated by adding restrictions on the law of motion (4) – (6).

2.1 Identification

From the restrictions we have imposed, we have

$$\Omega_t = A_t^{-1} \Sigma_t \Sigma_t' \left(A_t^{-1}\right)' \tag{8}$$

Thus, a TVC-SVAR system is locally identified if there exists a parameter vector $\Theta_t = (\alpha_t', \sigma_t')'$ which is a solution to the system of equations (8) for each t. The system is globally-identified if there exists a parameter vector $\Theta_t = (\alpha_t', \sigma_t')'$, which is a unique solution to the system of equations (8) at each t. As shown by Rubio-Ramírez et al. (2010), a globally-identified SVAR can be, at the same time, non-recursive and/or overidentified.

2.2 Relaxing standard assumptions

In a constant coefficient SVAR one typically has the option to identify shocks imposing short run, long run, or heteroschedasticity restrictions. In TVC-VARs only restrictions on A_t are employed, but there is no special reason for this choice, apart from estimation tractability. Furthermore, it is common to assume a lower triangular format for A_t even though this choice is restrictive in terms of the structural models it can accommodate. We next show how one can embed a general identification scheme into a TVC-VAR setup and perform proper Bayesian inference.

Consider the concentrated model obtained with estimates of the reduced-form coefficients $\widehat{B}_t$:

$$A_t\left(y_t - X_t'\widehat{B}_t\right) = A_t\widehat{y}_t = \Sigma_t\varepsilon_t \tag{9}$$

Noticing that $vec(A_t\widehat{y}_t) = vec(I_M A_t\widehat{y}_t) = (\widehat{y}_t' \otimes I_M)vec(A_t)$; and that $vec(\Sigma_t\varepsilon_t) = \Sigma_t\varepsilon_t$; and since we can decompose A_t as

$$vec\left(A_t\right) = S_A\alpha_t + s_A \tag{10}$$

where S_A and s_A are matrices with ones and zeros of dimensions $M^2 \times \dim(\alpha)$ and $M^2 \times 1$, respectively (see Amisano and Giannini, 1997, and Hamilton, 1994), the concentrated model is

$$\left(\widehat{y}_t' \otimes I_M\right)\left(S_A\alpha_t + s_A\right) = \Sigma_t\varepsilon_t$$

Thus, adding the law of motion of the α_t,the state space is

$$\widetilde{y}_t = Z_t\alpha_t + \Sigma_t\varepsilon_t \tag{11}$$

$$\alpha_t = \alpha_{t-1} + \zeta_t \tag{12}$$

where $\widetilde{y}_t \equiv \left(\widehat{y}_t' \otimes I_M\right) s_A$, $Z_t \equiv -\left(\widehat{y}_t' \otimes I_M\right) S_A$, and $Var\left(\zeta_t\right) = S$. The question is how to draw $\alpha^T \equiv \{\alpha_t\}_{t=1}^T$, from its smoothed posterior distribution $p\left(\alpha^T \mid \widetilde{y}^T, \Sigma^T, S, \widehat{B^T}\right)$.

The standard approach is to partition α_t into blocks associated with each equation, say $\alpha_t = \left[\alpha_t^{1\prime}, \alpha_t^{2\prime}, \ldots, \alpha_t^{M\prime}\right]'$, and assume that these blocks are independent, so that $S = diag\left(S_1, \ldots, S_M\right)$. Under these assumptions, the posterior is

$$p\left(\alpha^T \mid \widetilde{y}^T, \Sigma^T, S, \widehat{B^T}\right) = \prod_{m=2}^{M} p\left(\alpha^{m,T} \mid \alpha^{m-1,T}, \widetilde{y}^T, \Sigma^T, S, \widehat{B^T}\right) \times p\left(\alpha^{1,T} \mid \widetilde{y}^T, \Sigma^T, S, \widehat{B^T}\right) \tag{13}$$

Thus, for each equation m, the coefficients in equation $m-j, j \geq 1$ are treated as predetermined and changes in coefficients across equations are uncorrelated. The setup is convenient because equation by equation estimation is possible. Since the factorization does not necessarily have an economic interpretation, it may make sense to assume that the innovations in the α_t blocks are uncorrelated. However, if we insist that each element of α_t has some economic interpretation, the diagonality assumption of S is no longer plausible. For example, if in the SVAR there are policy and non-policy parameters, it will be hard to assume that non-policy parameters are strictly invariant to changes in the policy parameters (see e.g. Lakdawala, 2011).

The procedure we suggest relaxes both assumptions, that is, the vector α_t is jointly drawn and S is not necessarily block diagonal. This modification allows us to deal with recursive, non-recursive, just-identified or overidentified structural models in a unified framework. The modification is not without costs. In fact, it is not immediate to sample the vector α^T using the system $(11)-(12)$, while preserving Carter and Kohn (1994)'s multi-move approach because a non-diagonal S makes the posterior for α^T non-standard.

To solve this problem, we follow a lead of Primiceri (2005, pp. 850) and use a Metropolis step to draw α^T but we embed the process into a modified version of Geweke and Tanizaki (2001)'s routine for estimating general state space models, which takes into account the fact that parameters are time varying. The setup is appealing because it allows us to consider SVAR with linear or non-linear parameter restrictions and Gaussian or non-Gaussian shocks within the same framework, and this greatly expands the type of structural models one may want to consider.

2.3 A non-recursive, overidentified SVAR

Next, we provide an example of a particular overidentified structural model.

The vector of endogenous variables is $y_t = (Pcom_t, M_t, R_t, GDP_t, P_t, U_t)'$, where $Pcom_t$ represents a commodity price index, M_t a money aggregate, R_t the nominal interest rate, GDP_t a measure of aggregate output, P_t a measure of aggregate prices and U_t the unemployment rate. Since researchers working with this set of variables are typically interested in their dynamic response to monetary policy shocks, see e.g. Sims and Zha (2006), the structure of A_t is as in table 1, where X indicates a non-zero coefficients.

Thus, the structural form is *identified* from the VAR as follows:

1. *Information equation*: Commodity prices ($Pcom_t$) convey information about recent developments in the economy. Therefore, it is assumed that they react contemporaneously to every structural shock.

Reduced form \ Structural	$Pcom_t$	M_t	R_t	GDP_t	P_t	U_t
Information	X	X	X	X	X	X
Money demand	0	X	X	X	X	0
Monetary policy	0	X	X	0	0	0
Non-policy 1	0	0	0	X	0	0
Non-policy 2	0	0	0	X	X	0
Non-policy 3	0	0	0	X	X	X

Table 1: Identification restrictions

2. *Money demand equation:* Within the period money demand $\left(M_t^d\right)$, is a function of core macroeconomic variables (R_t, GDP_t, P_t).

3. *Monetary policy equation*: The interest rate (R_t) is used as an instrument for controlling the money supply (M_t^s). No other variable contemporaneously affects this equation.

4. *Non-policy block*: Following, e.g., Bernanke and Blinder (1992), the non-policy variables (GDP_t, P_t, U_t) react to policy changes, money changes or informational changes only with delay. This setup can be formalized by assuming that the private sector only considers lagged values of these variables as states or that private decisions have to be taken before the current values of these variables are known. The relationship between the variables in the non-policy block is left unmodeled and, for simplicity, a recursive structure is assumed.

Once the intuition behind the identification assumptions is clear, it is easy to understand why independence in coefficients of different equations is unappealing: changes in policy and non-policy coefficients are likely to be correlated. Let the vector of structural innovations be $\varepsilon_t = \begin{bmatrix} \varepsilon_t^i & \varepsilon_t^{md} & \varepsilon_t^{mp} & \varepsilon_t^y & \varepsilon_t^p & \varepsilon_t^u \end{bmatrix}'$. Then, the structural model is

$$\underbrace{\begin{bmatrix} 1 & \alpha_{1,t} & \alpha_{3,t} & \alpha_{5,t} & \alpha_{9,t} & \alpha_{12,t} \\ 0 & 1 & \alpha_{4,t} & \alpha_{6,t} & \alpha_{10,t} & 0 \\ 0 & \alpha_{2,t} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \alpha_{7,t} & 1 & 0 \\ 0 & 0 & 0 & \alpha_{8,t} & \alpha_{11,t} & 1 \end{bmatrix}}_{A_t} \times \begin{bmatrix} Pcom_t \\ M_t \\ R_t \\ GDP_t \\ P_t \\ U_t \end{bmatrix} = A_t^+(L) \begin{bmatrix} Pcom_{t-1} \\ M_{t-1} \\ R_{t-1} \\ GDP_{t-1} \\ P_{t-1} \\ U_{t-1} \end{bmatrix} + \Sigma_t \begin{bmatrix} \varepsilon_t^i \\ \varepsilon_t^{md} \\ \varepsilon_t^{mp} \\ \varepsilon_t^y \\ \varepsilon_t^p \\ \varepsilon_t^u \end{bmatrix} \tag{14}$$

where we normalize the main diagonal of A_t so that the left-hand side of each equation corresponds to the dependent variable and $A_t^+(L)$ is a function of A_t and B_t. Moreover,

$$\Sigma_t = \begin{bmatrix} \sigma_t^i & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_t^{md} & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_t^{mp} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_t^y & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_t^p & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_t^u \end{bmatrix}$$

is the matrix of standard deviations of the structural shocks.

The structural model (14) is non-recursive and overidentified by 3 restrictions. Overidentification obtains because the policy equation is different from the Taylor rule generally employed in the literature. Nevertheless, the system is globally identified and therefore suitable for interesting policy experiments. In addition, while the structural model is conditionally linear, it is easy to conceive setups where the structural model has a non-linear state space representation. For example, Rubio-Ramírez et al. (2010) impose identification restrictions directly on the impulse response function and these imply non-linear constraints on A_t and $A_t^+(L)$. Long run restrictions of the type originally imposed by Blanchard and Quah (1989) also generate non-linear constraints on the structural parameters. Clearly, one could also relax the random walk assumption and consider a non-linear law of motion for the free parameters. Thus, it seems reasonable to have a general setup that can accommodate for all these possibilities.

2.4 Estimation

The structural model that needs to be estimated has a large number of stochastic parameters. Since $(3), (4), (5), (6)$ define a hierarchical structure, a Bayesian perspective is employed and posterior distributions for sequences of the parameters are derived. As it is standard in the literature, see e.g. Kim and Nelson (1999) and Koop (2003), the Gibbs sampler will be used to draw posterior sequences. For this purpose, consider the general non-linear state-space system

$$\widehat{y}_t = z_t(\alpha_t) + \varepsilon_t \tag{15}$$

$$\alpha_t = t_t(\alpha_{t-1}) + R_t(\alpha_{t-1})\,\eta_t \tag{16}$$

where $\widehat{y}_t$ and ε_t are $M \times 1$ vectors; α_t and η_t are $N \times 1$ vectors; $\varepsilon_t \sim N(0, Q_t^{\varepsilon})$ and $\eta_t \sim N(0, Q_t^{\eta})$. Assume that $z_t(.)$, $t_t(.)$ and $R_t(.)$ are vector-valued functions.

To estimate such a system, it is typical to linearize it around the previous forecast of the state vector, so that

$$z_t(\alpha_t) \simeq z_t\left(\widehat{a}_{t|t-1}\right) + \widehat{Z}_t\left(\alpha_t - \widehat{a}_{t|t-1}\right)$$

$$t_t(\alpha_{t-1}) \simeq t_t\left(\widehat{a}_{t-1|t-1}\right) + \widehat{T}_t\left(\alpha_{t-1} - \widehat{a}_{t-1|t-1}\right)$$

$$R_t(\alpha_{t-1}) \simeq \widehat{R}_t$$

where $\widehat{Z}_t$, $\widehat{T}_t$ and $\widehat{R}_t$ are $N \times N$ matrices corresponding to the Jacobians of $z_t(.)$, $t_t(.)$ and $R_t(.)$, respectively, evaluated at $\alpha_t = \widehat{a}_{t|t-1}$. Thus, the approximated linear state-space is

$$\widehat{y}_t \simeq \widehat{Z}_t\alpha_t + \widehat{d}_t + \varepsilon_t \tag{17}$$

$$\alpha_t \simeq \widehat{T}_t\alpha_{t-1} + \widehat{c}_t + \widehat{R}_t\eta_t \tag{18}$$

where

$$\widehat{d}_t = z_t\left(\widehat{a}_{t|t-1}\right) - \widehat{Z}_t\widehat{a}_{t|t-1} \tag{19}$$

$$\widehat{c}_t = t_t\left(\widehat{a}_{t-1|t-1}\right) - \widehat{T}_t\widehat{a}_{t-1|t-1} \tag{20}$$

Equations (17) and (18) are similar to our original equations (11) and (12). When $z_t(.)$ and $t_t(.)$ are linear, $\widehat{d}_t = \mathbf{0}$ and $\widehat{c}_t = \mathbf{0}$. In the cases considered by Rubio-Ramírez et al. (2010) or Blanchard and Quah (1989) $\widehat{d}_t \neq \mathbf{0}$, while if the law of motion of the coefficient is non-linear $\widehat{c}_t \neq \mathbf{0}$. This kind of system can be estimated with the Extended Kalman Filter described below.

The posterior distribution of the parameters $\Theta \equiv (B^T, \alpha^T, \Sigma^T, s^T, V)$ is

$$p(\Theta \mid y^T) = \frac{p(y^T \mid \Theta)p(\Theta)}{p(y^T)} \propto p(y^T \mid \Theta)p(\Theta) \tag{21}$$

where $p(y^T \mid \Theta)$ is the likelihood, $p(\Theta)$ is the prior and s an indicator defined below. Let

$$p(\Theta) = p(B^T)p(\alpha^T)p(\Sigma^T)p(s^T)p(V) \tag{22}$$

and factor the likelihood function $L(\Theta \mid y^T) = p(y^T \mid \Theta)$ as

$$\begin{aligned} L(\Theta \mid y^T) &= L(V \mid y^T) \times L(s^T \mid V, y^T) \times L(\Sigma^T \mid s^T, V, y^T) \\ &\times L(\alpha^T \mid \Sigma^T, s^T, V, y^T) \times L(B^T \mid \alpha^T, \Sigma^T, s^T, V, y^T) \end{aligned} \tag{23}$$

To evaluate the posterior $p(\Theta \mid y^T)$ via the Gibbs sampler, we need to construct the conditionals $p\left(\Theta_i \mid \Theta_j, y^T\right), j \neq i$, and draw from them. The algorithm below describes how this can be done.

2.4.1 The algorithm

Let $\left\{\{s_{j,l,t}\}_{t=1}^{T}\right\}_{l=1}^{M}$ be a discrete indicator variable taking $j = 1, \ldots, k$ possible values. The procedure has 7 steps and 5 sampling blocks:

1. Set an initial value for $(B_0^T, \alpha_0^T, \Sigma_0^T, s_0^T, V_0)$.

2. Draw B_i^T from from $p\left(B_i^T \mid \alpha_{i-1}^T, \Sigma_{i-1}^T, s_{i-1}^T, V_{i-1}\right) \cdot I_B\left(B_i^T\right)$, using Kalman smoothed estimates $B_{t|T}$ obtained from the system (3), and compute $\widehat{y}_i^T$, where $I_B(.)$ truncates the posterior to insure stationarity of impulse responses.

3. Draw α_i^T from

$$\begin{aligned} p\left(\alpha_i^T \mid \widehat{y}_i^T, \Sigma_{i-1}^T, s_{i-1}^T, V_{i-1}\right) &= p\left(\alpha_{i,T} \mid \widehat{y}_i^T, \Sigma_{i-1,T}, s_{i-1,T}, V_{i-1}\right) \times \\ &\quad \prod_{t=1}^{T-1} p\left(\alpha_{i,t} \mid \alpha_{i,t+1}, \widehat{y}_i^t, \Sigma_{i-1,t}, s_{i-1,t}, V_{i-1}\right) \\ &\propto p\left(\alpha_{i,T} \mid \widehat{y}_i^T, \Sigma_{i-1,T}, s_{i-1,T}, V_{i-1}\right) \times \\ &\quad \prod_{t=1}^{T-1} p\left(\alpha_{i,t} \mid \widehat{y}_i^t, \Sigma_{i-1,t}, s_{i-1,t}, V_i\right) \times \\ &\quad f_{t+1}\left(\alpha_{i,t+1} \mid \alpha_{i,t}, \Sigma_{i-1,t}, s_{i-1,t}, V_{i-1}\right) \end{aligned} \tag{24}$$

   using the approach described below.

4. Draw Σ_i^T using a log-normal approximation to their distribution as in Kim et al. (1998). After sampling $\left(B_i^T, \alpha_i^T\right)$, the state space is linear but the error term is not normally distributed. In fact, given $\left(B_i^T, \alpha_i^T\right)$, the model is composed of

$$\widehat{A}_t \widetilde{y}_t = y_t^{**} = \Sigma_t \varepsilon_t$$

   and (6). Consider the $l-th$ equation $y_{l,t}^{**} = \sigma_{l,t}\varepsilon_{l,t}$, where $\sigma_{l,t}$ is the $l-$th element in the main diagonal of Σ_t and $\varepsilon_{l,t}$ is the $l-$th element of ε_t. Then

$$y_t^* = \log\left[\left(y_{l,t}^{**}\right)^2 + \overline{c}\right] \approx 2\log\left(\sigma_{l,t}\right) + \log\varepsilon_{l,t}^2 \tag{25}$$

   where $\overline{c}$ is a small constant. Since $\varepsilon_{l,t}$ is Gaussian, $\log\varepsilon_{l,t}^2$ is $\log\left(\chi^2\right)$ distributed, it can be approximated by a mixture of normals. Conditional on s_t, the indicator for the mixture of normals, the model is linear and Gaussian. Hence, standard Kalman Filter recursions can be used to draw $\{\Sigma_t\}_{t=1}^T$ from the system $(25)-(6)$. To ensure independence of the structural variances, each element of the sequence $\{\sigma_{l,t}\}_{l=1}^M$ is sampled assuming a diagonal W.

5. To draw s_i^T, conditional on Σ_i^T, y_t^*, given l and t, draw $u \sim U(0,1)$ and compare it to

$$\begin{aligned} P\left(s_{l,t} = j \mid y_{l,t}^*, \log(\sigma_{l,t})\right) &\propto q_j \times \phi\left(\frac{y_{l,t}^* - 2\log(\sigma_{l,t}) - m_j + 1.2704}{\upsilon_j}\right); \\ j &= 1, \ldots, k; l = 1, \ldots, M \end{aligned}$$

where $\phi(.)$ is the normal probability density function, and the term inside the function is the standardized error term $\log \varepsilon_{l,t}^2$.Then assign $s_{l,t} = j$ iff $P\left(s_{l,t} \leq j-1 \mid y_{l,t}^*, \log(\sigma_{l,t})\right) < u \leq P\left(s_{l,t} \leq j \mid y_{l,t}^*, \log(\sigma_{l,t})\right)$.

6. Draw V_i from $P\left(V_i \mid \alpha_i^T, \widehat{y}_i^T, \Sigma_{i-1}^T, s_{i-1}^T\right)$. The matrix V_i is sampled assuming that each block follows an independent Inverted-Wishart distribution.

7. Use $B_i^T, \alpha_i^T, \Sigma_i^T, s_i^T, V_i$ as initial values for the next iteration. Repeat 2 to 6 for $i = 1, \ldots, N$.

2.4.2 The details in step 3

For step 3 we use a Metropolis step to decide whether the draw from a proposal distribution is retained or not. The densities $p(\alpha_t \mid \widehat{y}^t, \Sigma_t, s, V)$ are obtained applying the Extended Kalman Smoother (EKS) to the original system of nonlinear equations. To draw α_i^T given $\widehat{y}_i^T, \Sigma_{i-1}^T, s_{i-1}^T, V_{i-1}$, we proceed as follows:

1. If $i = 0$, take an initial value $\alpha_0^T = \{\alpha_{0,t}\}_{t=1}^T$. If not,

2. Given α_{i-1}^T, compute $\left\{\alpha_{t|t+1}^{*(i-1)}, P_{t|t+1}^{*(i-1)}\right\}_{t=1}^T$ using the EKS where $\left\{P_{t|t+1}^*\right\}_{t=1}^T$ denotes the covariance matrix of $\left\{\alpha_{t|t+1}^*\right\}_{t=1}^T$.

3. Generate a candidate draw $z^T = \{z_t\}_{t=1}^T$, where for each $p_{*\alpha}(z_t \mid \alpha_{i-1,t}) = N\left(\alpha_{i-1,t}, rP_{t|t+1}^{*(i-1)}\right)$, r is a constant,$t = 1, \ldots, T$. Let $p_{*\alpha}\left(z^T \mid \alpha_{i-1}^T\right) = \prod_{t=1}^T p_{*\alpha}(z_t \mid \alpha_{i-1,t})$.

4. Compute $\theta = \frac{p(z^T) \cdot p_{*\alpha}\left(\alpha_{i-1}^T \mid z^T\right)}{p\left(\alpha_{i-1}^T\right) \cdot p_{*\alpha}\left(z^T \mid \alpha_{i-1}^T\right)}$, where $p(.)$ is the RHS of (24) using the EKS approximation. Draw a $v \sim U(0,1)$. Set $\alpha_i^T = z^T$ if $v < \omega$ and $\alpha_i^T = \alpha_{i-1}^T$ otherwise, where

$$\omega \equiv \begin{cases} \min\{\theta, 1\}, & \text{if } I_\alpha\left(z^T\right) = 1 \\ 0, & \text{if } I_\alpha\left(z^T\right) = 0 \end{cases}$$

and $I_\alpha(.)$ is a truncation indicator.

Finally, steps 2 to 4 in this sub-loop are repeated every time step 3 of the main loop is executed.

2.4.3 Extended Kalman Smoother

To apply the EKS to our system of equations, we first predict the mean and variance at each $t = 1, \ldots, T$:

$$\widehat{a}_{t|t-1} = t_{t-1}\left(\widehat{a}_{t-1|t-1}\right)$$

$$P_{t|t-1} = \widehat{T}_t P_{t-1|t-1} \widehat{T}_t' + \widehat{R}_t Q_t^{\eta} \widehat{R}_t'$$

and compute Kalman gain:

$$\Omega_t = \widehat{Z}_t' P_{t|t-1} \widehat{Z}_t + Q_t^{\varepsilon}$$

$$K_t = P_{t|t-1} \widehat{Z}_t' \Omega_t^{-1}$$

As new information arrives, estimates of α_t and variance are updated according to

$$\widehat{a}_{t|t} = \widehat{a}_{t|t-1} + K_t\left[y_t - z_t\left(\widehat{a}_{t|t-1}\right)\right]$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1} \widehat{Z}_t' \Omega_t^{-1} \widehat{Z}_t P_{t|t-1}'$$

To smooth the estimates, set $\alpha^*_{T|T} = \widehat{a}_{T|T}$, $P^*_{T|T} = P_{T|T}$ and , for $t = T-1, \ldots, 1$, compute

$$\alpha^*_{t|t+1} = \widehat{a}_{t|t} + P_{t|t} \widehat{Z}_t' P_{t+1|t}^{-1}\left(\alpha^*_{t+1|t+2} - \widehat{Z}_t' \widehat{a}_{t|t}\right)$$

$$P^*_{t|t+1} = P_{t|t} - P_{t|t} \widehat{Z}_t'\left[P_{t+1|t} + \widehat{R}_t Q_t^{\eta} \widehat{R}_t'\right]^{-1} \widehat{Z}_t P_{t|t-1}'$$

To start the iterations we use $\widehat{a}_{1|0} = \mathbf{0}_{N\times 1}$ and $P_{0|0} = I_N$. Notice that since the original $t_t(.)$ and $z_t(.)$ are used for computing prediction and updating equations, $\widehat{d}_t$ nor $\widehat{c}_t$ do not directly enter here.

3 An Application

This section applies our algorithm to study the transmission of monetary policy shocks in an overidentified, non-recursive structural TVC-VAR. We are interested in knowing whether the propagation of policy shocks has changed over time and in identifying the sources of variation. For comparison, we will also examine the conclusions obtained estimating a more standard recursive, just identified TVC system, and an overidentified SVAR with constant coefficients.

$B^{prior} \sim N\left(\overline{\mathbf{B}}, 4 \cdot \overline{\mathbf{VB}}\right) I_B(.)$	$Q^{prior} \sim IW\begin{pmatrix} k_Q^2 \cdot \overline{\mathbf{VB}}, \\ (1+K) \end{pmatrix}$
$\alpha^{prior} \sim N\left(\overline{\boldsymbol{\alpha}}, 4 \cdot diag\left(abs\left(\overline{\boldsymbol{\alpha}}\right)\right)\right) I_\alpha(.)$	$S^{prior} \sim IW\begin{pmatrix} k_S^2 \cdot diag\left(abs\left(\overline{\boldsymbol{\alpha}}\right)\right), \\ 1+\dim\alpha \end{pmatrix}$
$\log\left(\sigma^{prior}\right) \sim N\left(\overline{\boldsymbol{\sigma}}, 10 \cdot I_M\right)$	$W_i^{prior} \sim IW\left(k_W^2, 1+1\right), i = 1, \ldots, M$

Table 2: Prior distributions

3.1 The Data

Data comes from the *International Financial Statistics (IFS)* database at the International Monetary Fund and from the Federal Reserve Board (www.imfstatistics.org/imf/about.asp and www.federalreserve.gov/econresdata/releases/statisticsdata.htm, respectively). The sample is 1959:I - 2005:IV. We stop at this date to avoid the last financial crisis and to compare our results to those of Sims and Zha (2006), who use a Markov switching model over the same sample. The GDP deflator, the unemployment rate, the aggregate Gross Domestic Product index (Volume, base 2005=100), the commodity prices index, and M2 are from IFS, the Federal Funds rate is from the Fed. All the variables are expressed in annual rate changes, i.e. $y_t^* = \log(y_t) - \log(y_{t-4})$, except for the Federal Funds and the unemployment rate, and standardized, that is, $x_t = \left(y_t^* - mean\left(y_t^*\right)\right)/std\left(y_t^*\right)$, to have all the variables on the same scale.

3.2 The structural model, the prior and computation details

The VAR includes six variables and it is estimated with 2 lags. This is what the BIC criteria selects for the constant coefficient version of the model. The structural system is the one in section 2.3, it is non-recursive and overidentified. The priors are in Table 2, are proper, and conjugate for computational convenience.

With these priors, the conditional posteriors will be of Normal, Inverted Wishart type. To calibrate the prior, the first 40 observation are used as a training sample: the reduced-form parameters $\overline{\mathbf{B}}$ and $\overline{\mathbf{VB}}$ are estimated with OLS; the structural parameters $\overline{\boldsymbol{\alpha}}$ and $\overline{\boldsymbol{\sigma}}$ with Maximum Likelihood using 100 different starting points. We set $k_Q^2 = 0.5 \times 10^{-4}$ and $k_S^2 = 1 \times 10^{-3}, k_W^2 = 1 \times 10^{-4}$ and, in line with the literature, we set $k = 7$. A total amount of $150,000$ draws for the Gibbs sampler routine were performed, the first $100,000$ were discarded and one every 100 of the remaining draws was used for inference. Convergence was checked using standard statistics. To insure stationarity of the estimated system, draws for B_t are monitored, the companion form associated with (3) is computed and the draw discarded if it does not satisfy the stability condition. The indicator function $I_\alpha(.)$, which is used to eliminate outlier draws, is uniform over the set $(-20, 20)$. In our application all $150,000$ draws were inside the bounds, so the constraint is not binding. Finally, the acceptance rate for the Metropolis step is 20.3 percent.

Since the structural model has $M = 6$, and $\dim(\alpha) = 12$, S_A, s_A are

$$S_A = \begin{bmatrix} \mathbf{0}_{6\times\dim(\alpha)} \\ \begin{bmatrix} \mathbf{0}_{1\times(1-1)} & 1 & \mathbf{0}_{1\times(\dim(\alpha)-1)} \end{bmatrix} \\ \mathbf{0}_{1\times\dim(\alpha)} \\ \begin{bmatrix} \mathbf{0}_{1\times(2-1)} & 1 & \mathbf{0}_{1\times(\dim(\alpha)-2)} \end{bmatrix} \\ \mathbf{0}_{3\times\dim(\alpha)} \\ \begin{bmatrix} \mathbf{0}_{1\times(3-1)} & 1 & \mathbf{0}_{1\times(\dim(\alpha)-3)} \end{bmatrix} \\ \begin{bmatrix} \mathbf{0}_{1\times(4-1)} & 1 & \mathbf{0}_{1\times(\dim(\alpha)-4)} \end{bmatrix} \\ \mathbf{0}_{4\times\dim(\alpha)} \\ \begin{bmatrix} \mathbf{0}_{1\times(5-1)} & 1 & \mathbf{0}_{1\times(\dim(\alpha)-5)} \end{bmatrix} \\ \begin{bmatrix} \mathbf{0}_{1\times(6-1)} & 1 & \mathbf{0}_{1\times(\dim(\alpha)-6)} \end{bmatrix} \\ \mathbf{0}_{2\times\dim(\alpha)} \\ \begin{bmatrix} \mathbf{0}_{1\times(7-1)} & 1 & \mathbf{0}_{1\times(\dim(\alpha)-7)} \end{bmatrix} \\ \begin{bmatrix} \mathbf{0}_{1\times(8-1)} & 1 & \mathbf{0}_{1\times(\dim(\alpha)-8)} \end{bmatrix} \\ \begin{bmatrix} \mathbf{0}_{1\times(9-1)} & 1 & \mathbf{0}_{1\times(\dim(\alpha)-9)} \end{bmatrix} \\ \begin{bmatrix} \mathbf{0}_{1\times(10-1)} & 1 & \mathbf{0}_{1\times(\dim(\alpha)-10)} \end{bmatrix} \\ \mathbf{0}_{3\times\dim(\alpha)} \\ \begin{bmatrix} \mathbf{0}_{1\times(11-1)} & 1 & \mathbf{0}_{1\times(\dim(\alpha)-11)} \end{bmatrix} \\ \begin{bmatrix} \mathbf{0}_{1\times(12-1)} & 1 & \mathbf{0}_{1\times(\dim(\alpha)-12)} \end{bmatrix} \\ \mathbf{0}_{5\times\dim(\alpha)} \end{bmatrix}$$

$$s_A = [e_1, e_2, e_3, e_4, e_5, e_6]'$$

where e_i are vectors in $\mathbb{R}^M$ with

$$e_i = [e_{i,j}]_{j=1}^{M} \text{ such that } e_{i,j} = \begin{cases} 1, & j = i \\ 0, & j \neq i \end{cases}.$$

Finally, the computational time for each version of the Non-recursive-TVC VAR is about 10 hours, roughly the same as the computational time needed to estimate a recursive-TVC-VAR. Computations were performed on an Intel (R) CORE(TM) i5-2400 CPU @ 3.1GHz machine with 16GB of RAM.

3.3 Time variations in structural parameters

We start by describing the time variations that our model delivers. In figure 1 we report the highest 68 percent posterior tunnel for the variability of the monetary

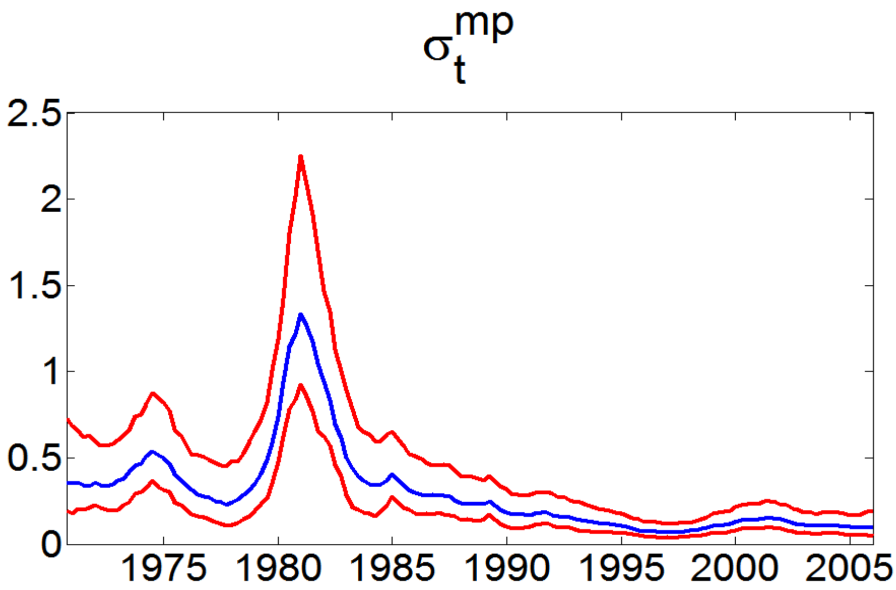

Figure 1: Median and posterior 68 percent tunnel, volatility of monetary policy shock.

policy shock and in figure 2 the highest 68 percent posterior tunnel for the non-zero contemporaneous structural parameters α_t.

There are significant changes in the standard deviation of the policy shocks and a large swing in the late 1970s-early 1980s is visible. Given the identification we use, this increase in volatility must be attributed to some unusual and unexpected policy action, which made the typical relationship between interest rates and money growth different. This outcome is consistent with the arguments of Strongin (1995) and Bernanke and Mihov (1998), who claim that monetary policy during the Volker era was run differently that in the 1960s and the 1970s.

Figure 2 indicates the non-policy parameters $[\alpha_{7,t}, \alpha_{8,t}, \alpha_{11,t}]'$ exhibit considerable time variations which are a posteriori significant. Note that it is not only the magnitude that changes; the sign of the posterior tunnel is also affected. Also worth noting is the fact that both the GDP coefficient in the inflation equation and the inflation coefficient in the unemployment equations move from negative to positive, suggesting a generic switch in the slope in the Phillips curve.

The parameter $\alpha_{2,t}$, which controls the reaction of the nominal interest rates to money growth, also displays considerable changes. In particular, while in the 1970s and in the 1980s the relationship between interest rate and money growth was negative and significant, it weakens in the late 1990s, and disappears completely in

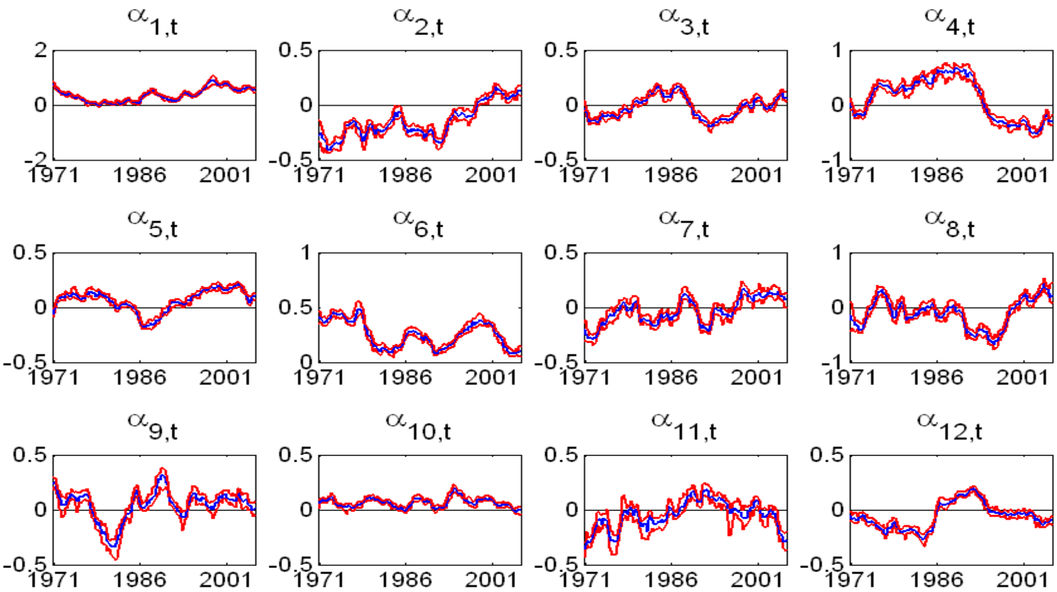

Figure 2: Non-Recursive estimates of α

the early 2000s. Thus, during this last part of the sample, interest rates were no longer used to control the amount of money supply in circulation.

The coefficients of the money demand equation, $[\alpha_{4,t}, \alpha_{6,t}, \alpha_{10,t}]'$ are also unstable. For example, the elasticity of money demand to the nominal interest rate is negative at the beginning of the sample, as theory would suggest; it turns strongly positive up to the middle of the 1980s, and slowly and continuously declines after that. Since the beginning of the 1990s, the coefficient turns negative once again and becomes strongly negative in the 2000s. Also interesting is the fact that the elasticity of money (growth) demand to inflation is very low and fluctuating around zero. Thus, homogeneity of degree one of money in prices does not seem to hold.

The parameters of the information equation are also time varying. However, changes look more like serially correlated variations around a constant mean. In fact, the mean level at the beginning and at the end of the sample is similar.

One additional features of figure 2 needs to be mentioned. Time variations in elements of α_t are correlated (see, in particular, α_{2t} and α_{4t} or α_{2t} and α_{7t}). Thus our setup, in which the matrix S is not necessarily diagonal, captures the idea that changes in policy and private sector parameters are related.

In sum, in agreement with the DSGE evidence of Justiniano and Primiceri (2008) and Canova and Ferroni (2012), time variations appear in the variance of the mon-

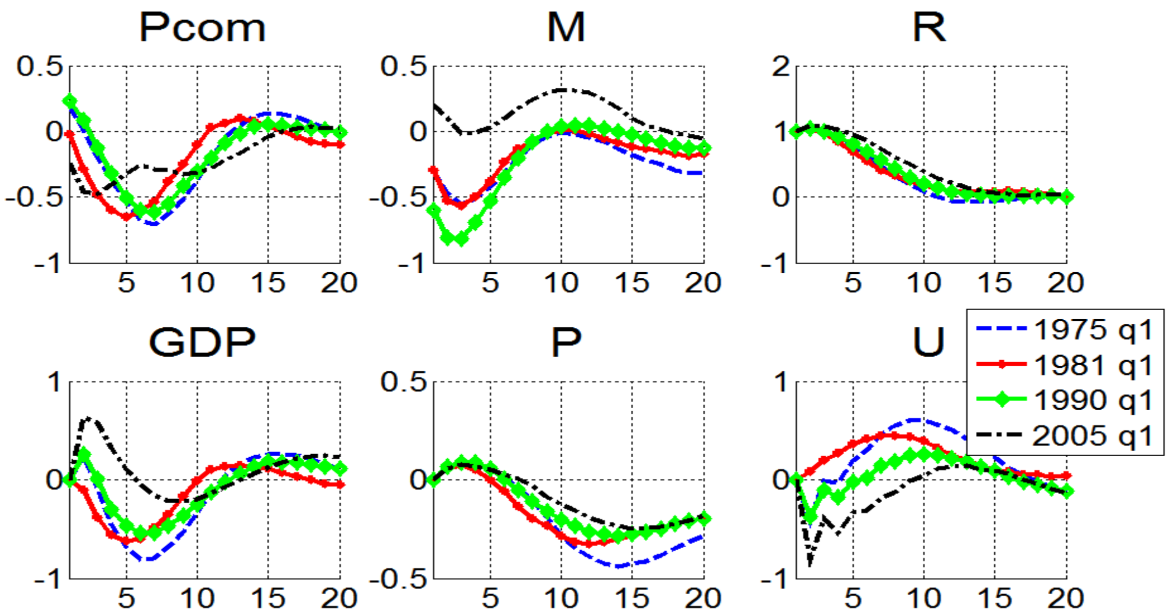

Figure 3: Dynamics following a monetary policy shock, different dates.

etary policy shock and in the contemporaneous policy and non-policy coefficients.

3.4 The transmission of monetary policy shocks

Next, we would like to see how the time variations we have described affect the transmission of monetary policy shocks. To control from the fact that σ_t^{mp} is time-varying, we normalize the impulse to be one at all t. Thus, the variations we describe below are due to changes in the propagation but not in the size of the shocks. We compute responses as the difference between two conditional projections, one with the structural shock normalized to one and one with the structural shock normalized to zero. In both cases, the structural parameters are allowed to be random.

In theory, a surprise increase in interest rates, should make money growth, output growth and inflation fall, while unemployment should go up. Such a pattern is present in the data in the early part of the sample, but disappears as time goes by. As figure 3 indicates, monetary policy shocks have the largest effects in 1990; the pattern is similar but weaker in 1975 and 1981. In 2005, the liquidity effect has disappeared (interest rate increases imply positive although insignificant responses of the growth rate of money) while output and unemployment effects are perverse (output growth significantly increase and unemployment significantly fall). Note that differences in the responses of the variables between, say, 1990 and 2005 are a-posteriori significant. Thus, it appears that one of the main mechanisms though which of monetary policy affects the real economy (see e.g. Gordon and Leeper, 1994) has considerably weakened over time. Perhaps, the informal inflation targeting practices that the Fed has follows in the 2000s are responsible for these changes.

Despite these noticeable variations, the proportion of the forecast error variance of output, prices and unemployment due to policy shocks is consistently small. Monetary policy shocks explain little of the forecast error variance of inflation at all times and about 10-15 percent of the variability of output growth and the unemployment rate, with a maximum of 20 percent in the early 1980s. Thus, as in Uhlig (2005) or Sims and Zha (2006), monetary policy has modest real effects.

Our results are very much in line with those of Canova et al. (2008), even though they use sign restrictions to extract structural shocks, and of Boivin and Giannoni (2006), who use sub-sample analysis to make their points, and of Boivin et al. (2010). They differ somewhat from those reported in Sims and Zha (2006) primarily because they do not allow time variations in the instantaneous coefficients. They also differ from Fernández-Villaverde et al. (2010), who allow for stochastic volatility and time variations only in the coefficients of the policy rule.

3.5 Comparing the results to traditional models

This subsection compares our results with those obtained in a constant coefficient overidentified structural model (henceforth, overidentified SVAR) and with a TVC model where the monetary shock is identified with standard recursive restrictions. Given that an overidentified, non-recursive structural model is more complicated than the alternatives, one would like to know whether the economic interpretation would change if one would follow standard approaches.

To start with we compare the fit of various specifications using marginal likelihood (ML) computed using an harmonic mean estimator. The recursive TVC-SVAR has, perhaps unsurprisingly, the highest ML, (-299), followed by the overidentified system (-316). Thus the restrictions implied by the model are rejected. Finally, the model with fixed coefficients is clearly inferior to both, ML$= -574$.

Next, we examine the time variations present in the structural parameters α in the recursive and non-recursive TVC model and compare them to the estimates obtained in the overidentified SVAR. Here, we report only the parameters of the money demand equation which are common to the two time varying specifications.

The elasticity of money growth to interest rates has very different features. In the model with fixed coefficients is strongly positive, while in the two TVC specifications, it fluctuates around zero. Furthermore, the posterior 68 percent tunnel in the two time varying specifications has quite different behavior and, in the last part of

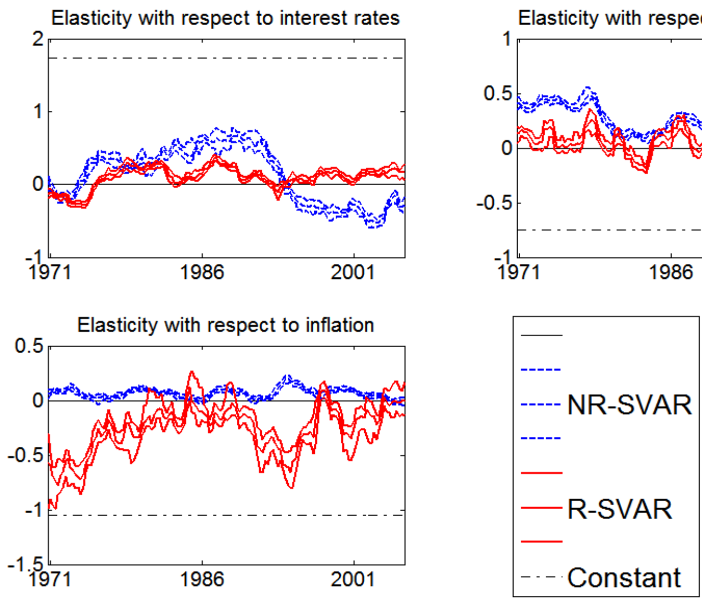

Figure 4: Coefficients in the money demand equation, different systems

the sample, it is strongly negative with the non-recursive specification and positive with the recursive one. There are smaller differences in the properties of the elasticity of money growth demand to output. Nevertheless, fluctuations over time in the posterior estimates obtained with the recursive system are smaller. In addition, while in the recursive system the estimate fluctuates around zero, the mean value is positive and large in the non-recursive one. Finally, there are also considerable differences both in the level and in the time variations of the elasticity of money growth demand to inflation: the overidentifed SVAR has a very negative and significant elasticity; in the recursive system the tunnel starts quite negative but then fluctuates around zero with some prevalence of negative values; in the non-recursive system, instead, the tunnel is generally positive and much less fluctuating.

What do these differences imply for the transmission of monetary policy shocks? For illustration, we report in figure 5, the responses of money growth to a unexpected interest rate impulse at four dates (1975, 1981, 1990, 2005) in the three systems we analyze. Three features are evident. First, in the overidentified SVAR the liquidity effect is strong and money growth falls quite a lot when interest rates are surprisingly increased. Second, when TVC are allowed for, the liquidity effect is generally weaker. Third, the size of the money growth responses is different in the recursive and in the non-recursive system. In the former time variations are small and the effect is

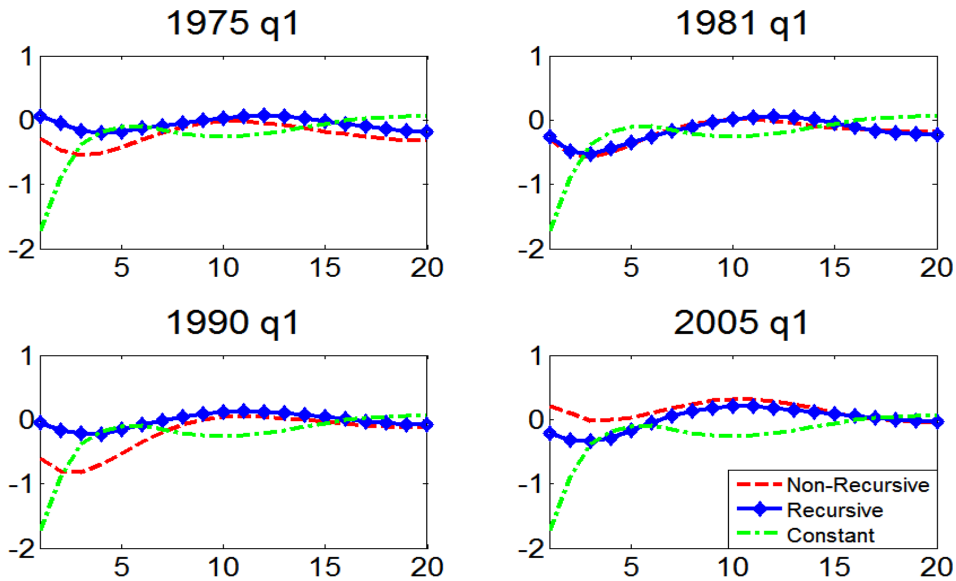

Figure 5: Responses of money growth, different time periods.

significant only in 1981 and 2005; in the latter time variations are larger and while the response was strongly, persistently and significantly negative in 1990, it turns positive in 2005. Thus, inference concerning the relevance of the liquidity effect of an interest rate shock depends on the whether fixed or time varying coefficients are considered and on whether the system is just or overidentified.

4 Conclusions

This paper proposes a unified framework to estimate structural VARs. The methodology can handle time varying coefficient or time invariant models, identified with recursive or non-recursive restrictions, and that could be just identified or overidentified. Our algorithm adds a Metropolis step to a standard Gibbs sampling routine but nest the model into a general non-linear state space. We do so, since this setup allows us to impose general identification restrictions. Thus, it greatly expands the set of structural models we can deal with, within the same estimation framework.

We apply the methodology to the estimation of a monetary policy shock in a non-recursive overidentified model used with fixed coefficients in the literature by Robertson and Tallman (2001), Waggoner and Zha (2003). We show that there are important time variations in the variance of the monetary policy shock and in the estimated non-zero contemporaneous relationships. These time variations translate in important changes in the transmission of monetary policy shocks to the variables in the economy. We also show, that one would have got a different characterization of the liquidity effect of an interest rate shock and of the properties of the money demand function had one used an overidentified but fixed coefficient VAR or a time varying coefficient VAR identified with recursive restrictions.

The range of potential applications of the methodology is large. For example, one could use the same setup to identify fiscal shocks or externally generated shocks in models which theory tightly parametrizes. One could also use the same methodology to identify shocks imposing magnitude restrictions on impulse responses or long run restrictions, as in Rubio-Ramírez et al. (2010). The computational complexity is important but it is not overwhelming and all the calculations can be easily performed on a standard PC with sufficient RAM memory.

References

[1] Amisano, G. and C. Giannini (1997). Topics in Structural VAR Econometrics. Springer and Verlag.

[2] Bernanke, B. S. and A. S. Blinder (1992). The federal funds rate and the channels of monetary transmission. American Economic Review 82(4), 901–921.

[3] Bernanke, B. S. and I. Mihov (1998). Measuring monetary policy. Quarterly Journal of Economics 113(3), 869–902.

[4] Blanchard, O. J. and D. Quah (1989). The dynamic effects of aggregate demand and supply disturbances. American Economic Review 79(4), 655–673.

[5] Boivin, J. and M. P. Giannoni (2006). Has monetary policy become more effective? The Review of Economics and Statistics 88(3), 445–462.

[6] Boivin, J., M. T. Kiley, and F. S. Mishkin (2010). Chapter 8 - how has the monetary transmission mechanism evolved over time? Handbook of Monetary Economics 3, 369–422.

[7] Canova, F., M. Ciccarelli, and E. Ortega (2012). Do institutional changes affect business cycles? evidence from Europe. Forthcoming, Journal of Economic Dynamics and Control.

[8] Canova, F. and G. De Nicoló (2002). Monetary disturbances matter for business fluctuations in the G-7. Journal of Monetary Economics 49, 1131–1159.

[9] Canova, F. and F. Ferroni (2012). The dynamics of US inflation: Can monetary policy explain the changes? Journal of Econometrics 167, 47–60.

[10] Canova, F. and L. Gambetti (2009). Structural changes in the US economy: Is there a role for monetary policy? Journal of Economic Dynamics and Control 33, 477–490.

[11] Canova, F., L. Gambetti, and E. Pappa (2008). The structural dynamics of U.S. output and inflation: What explains the changes? Journal of Money, Credit and Banking, 40(2-3), 369–388.

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

[13] Carter, C. K. and R. Kohn (1994). On Gibbs sampling for state space models. Biometrika 81(3), 541–553.

[14] Cogley, T. and T. J. Sargent (2005). Drifts and volatilities: Monetary policies and outcomes in the post WWII U.S.. Review of Economic Dynamics 8(2), 262–302.

[15] Fernández-Villaverde, J., P. A. Guerron-Quintana, and J. F. Rubio-Ramírez (2010). Fortune or virtue: Time-variant volatilities versus parameter drifting in U.S. data. CEPR DP7813.

[16] Geweke, J. and H. Tanizaki (2001). Bayesian estimation of state-space models using the Metropolis-Hastings algorithm within Gibbs sampling. Computational Statistics and Data Analysis 37, 151–170.

[17] Gordon, D. B. and E. M. Leeper (1994). The dynamic impacts of monetary policy: An exercise in tentative identification. Journal of Political Economy 102(6), 1228–1247.

[18] Hamilton, J. D. (1994). Time Series Analysis. Princeton University Press.

[19] Justiniano, A. and G. E. Primiceri (2008). The time-varying volatility of macroeconomic fluctuations. American Economic Review 98(3), 604–641.

[20] Kim, C.-J. and C. R. Nelson (1999). State-Space Models with Regime-Switching: Classical and Gibbs-Sampling Approaches with Applications. MIT Press.

[21] Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. The Review of Economic Studies 65(3), 361–393.

[22] Koop, G. (2003). Bayesian Econometrics. John Wiley and Sons Ltd.

[23] Lakdawala, A. (2011). Changes in Federal Reserve preferences. Manuscript, UCSD.

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

[25] Leeper, E., C. Sims, and T. Zha (1996). What does monetary policy do? Brookings Papers on Economic Activity 2.

[26] Primiceri, G. (2005). Time varying structural vector autoregressions and monetary policy. Review of Economic Studies 72, 821–852.

[27] Robertson, J. C. and E. W. Tallman (2001). Improving federal-funds rate forecasts in VAR models used for policy analysis. Journal of Business and Economic Statistics 19(3), 324–330.

[28] Rubio-Ramírez, J. F., D. F. Waggoner, and T. Zha (2010). Structural vector autoregressions: Theory of identification and algorithms for inference. Review of Economic Studies 77, 665–696.

[29] Sims, C. A. and T. Zha (1998). Bayesian methods for dynamic multivariate models. International Economic Review 39(4), 949–968.

[30] Sims, C. A. and T. Zha (2006). Were there regime switches in u.s. monetary policy? American Economic Review 96(1), 54–81.

[31] Strongin, S. (1995). The identification of monetary policy disturbances: Explaining the liquidity puzzle. Journal of Monetary Economics 35, 463–497.

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

[33] Waggoner, D. and T. Zha (2003). A Gibbs sampler for structural vector autoregressions. Journal of Economic Dynamics and Control 28, 349–366.

[34] Waggoner, D. F. and T. Zha (1999). Conditional forecasts in dynamic multivariate models. The Review of Economics and Statistics 81(4), 639–651.