REGIME CHANGES IN THE EUROPEAN MONETARY POLICY: METHODS AND APPLICATIONS

JUAN F. RUBIO-RAMÍREZ
FEDERAL RESERVE BANK OF ATLANTA

DANIEL WAGGONER
FEDERAL RESERVE BANK OF ATLANTA

TAO ZHA
FEDERAL RESERVE BANK OF ATLANTA

PRELIMINARY AND INCOMPLETE

ABSTRACT. This paper studies whether Euro area monetary policy changed after the introduction of the European Monetary Union (EMU). To do that we consider vector autoregression (VAR) models that allows for regime switching in coefficients and variances and we estimate them using post-1970 Euro data. Our empirical results overwhelmingly support regime changes in shock variances instead of changes in coefficients. These results are robust to different identifications schemes. We also find that monetary shocks generate the liquidity effect and have significant effects on output. These results are robust to different identifications and different regimes.

I. INTRODUCTION

This paper studies whether Euro area monetary policy changed after the introduction of the European Monetary Union (EMU). To do that we consider vector autoregression (VAR) models that allow for regime switching in coefficients and variances and we estimate them using post-1970 Euro area data.

The process towards the EMU was initiated more than 25 years ago. In March 1979, the European Monetary System (EMS) started operating, with the objectives of reducing inflation and preparing for monetary integration. Ten years later the Delors Report set out a plan to introduce EMU over three stages. The first stage (increasing cooperation among Euro area central banks) was launched in 1990. In January 1994, the second stage began with the establishment of the European Monetary Institute (EMI) as the forerunner to the European Central Bank (ECB). Finally, in January 1999 stage three started. The euro became the single currency for the member states of the Euro area and a single monetary policy was introduced under the authority of the (ECB).

Date: March 30, 2005.

We thank Fabio Canova, Jon Faust and Harald Uhlig for helpful discussions. Eric Wang provided excellent research assistance. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

We have observed that in the last decade annual inflation has been under 5 percent (it was well above 10 percent in the late 70's and early 80's), the volatility of output has decreased (while maintaining its average annual growth rate), and short term nominal interest rates have been at a record low.

The coincidence of both events (introduction of the EMU and lower volatility of prices and output) drives us to ask the following questions: Is the decrease in volatility linked to: (1) a change in the Euro area monetary policy? (2) a change in the actions of the private sector?, or (3) a decrease in the volatility of shocks hitting the economy?

A researcher giving an affirmative answer to the first inquiry could argue that monetary policy has been able to better manage the effects of shocks in the Euro economy since the early 90's. Alternatively, a researcher giving a positive answer to the second question could claim that some type of financial innovation has allowed private agents to better handle the shocks. But these are not the only two possible explanations. A researcher giving an affirmative response to the third query could maintain that shocks hitting the Euro area have been less volatile in the last decade.

In order to shed some light on this debate some research has been done. Peersman and Smets (2003) identified a structural VAR for the Euro area from 1980 until 1998. Using a Chow test, they show that the overall macroeconomic effects of a monetary policy are stable over time. Ciccarelli and Rebucci (2003) use an heterogenous time-varying panel VAR model to analyze the monetary transmission mechanism across countries and time. They find that the transmission mechanism changed in the second part of the 90s. Using post-1999 data, Angeloni and Ehrmann (2003) find evidence that the monetary transmission mechanism has become more potent and homogeneous across countries because of the EMU. De Bondt (2002) presents an error-correction model of the interest rate pass-through process for the Euro area. Using pre- and post-1999 data finds a quicker pass-through process since the introduction of the euro.

We see two potential problems with theses approaches. First, most of them (expect Ciccarelli and Rebucci, 2003) do not consider time-varying parameter models. Instead, they use pre- and post-EMU data and search for a structural break. Second, none of the mentioned papers considers time-varying volatility.

Dividing the sample in pre- and post-EMU data amplifies the small sample problems associated with a short sample. In addition, structural break analysis rest on the assumption that the probability of a regime change is either one or zero. An event as institutionally complicated as the effects of the EMU on monetary policy should not be modelled as a discrete one. Taking into account that the probability of such event may be between zero and one is crucial. Finally, Sims and Zha (2004b) remark how important it is to consider heteroscedastic shocks when studying regime switching models. Assuming constant volatility may bias the results towards finding changes in parameters.

In order to solve these shortcomings we build on Sims and Zha (2004b). We identify monetary policy and private sector behaviors in VARs that allow for regime switching in both coefficients and variances using a Bayesian approach. A crucial feature of our approach is that regimes are modelled as outcomes of a hidden Markov chain. Thus, all the regimes can have positive probability at any moment in time . We contemplate four cases of regime switching: (1) no regime switching (i.e., a standard constant-parameter VAR), (2) regime switching in variances only, (3) regime switching in the monetary policy coefficients and variances, and (4) regime switching in the private sector coefficients and variances. Then, we use the posterior odds ratio to choose the regime switching case and the number of regimes that fits the data best.

There are two main novelties in our approach. The first newness is with respect to Sims and Zha (2004b), while the second innovation is technical. We improve with respect to Sims and Zha (2004b) in that we identify monetary policy and private sector behaviors using four different schemes. This will allow us to check the robustness of our results. The first identification uses a standard ordering as in Christiano, Eichenbaum and Evans (1996). Secondly, we isolate the economic shocks following the strategy described in Gordon and Leeper (1994) and Sims and Zha (2004b). Third, we impose the long-run restrictions introduced by Blanchard and Quah (1993) and Galí (1999). Finally, we use sign restrictions as proposed by Faust (1998), Canova and De Nicoló (2002), and Uhlig (forthcoming).

The technical innovation relates to how some of the identification schemes are implemented. Some of these identification schemes are computationally demanding, hence we develop new algorithms to implement them efficiently. Our methods are not only easy to use, but also take little computational time relative to the existing algorithms.

On the other hand, our methodology is not without problems. First, we use aggregated data for the Euro are.[1] This can only be done under the assumption of homogeneity across countries of both monetary policy and monetary policy effects. Several authors have argued that the Euro area monetary transmission process is uneven across countries. Cecchetti (2001) argues that legal differences between countries create asymmetries in the response to policy among Euro area countries. Mihov (2001) provided econometric evidence in support of this conclusion. Kieler and Saarenheimo (1998), Guiso et al. (1999) and Angeloni, Kashyap and Mojon (2003) have shown, on the other hand, that such differences are not robust to changes in empirical methodology and data. Also, in a recent article, Peersman (2004) shows that the effects of monetary policy are relatively uniform across the whole Euro area.

Second, our model is not microfounded. A bigger challenge seems to be the construction of models of the Euro area, with proper microfoundation and a realistic characterisation of the transmission process. Smets and Wouters (2003) have confronted such a challenge DSGE model can fit the data better that constant parameter VARs. Our results indicate that VAR models with time-varying parameters and variances should replace the traditional constant-parameter VARs as a benchmark to gauge how well a DSGE model fits the data. Therefore, the next step should be to build and estimate a DSGE with time-varying parameters and variances.[2]

Our main findings are as follows. First, the posterior odds ratio overwhelmingly favors two regimes in the post-1970 Euro economy. This result holds for the four identifications used in this paper. Second, most of regime change is reflected in the variances of VAR disturbances. Thus, VAR models with time-varying shock variances should replace the traditional constant-parameter VARs as a benchmark to gauge how well a DSGE model fits the data. Contrary to the results produced by the constant-parameter VARs, the

[1]In particular, we use the "synthetic" data for the Euro area constructed by The Econometric Modeling Unit at the ECB. See Fagan, Henry, and Mestre (forthcoming) for details.

[2]See Fernández-Villaverde et. al., 2005, for details

models with time-varying shock variances imply large uncertainty across identifications about the effect of monetary policy shocks on the general price level. But both the liquidity effect and the output effect in response to a monetary policy shock are significant.

The rest of the paper is organized as follows. Section II lays out the general framework. Section III reviews the four popular identification strategies and develops new methods to implement some of these strategies. Detailed proofs of some theorems are provided in appendices. Section IV applies our methods to the Euro economy and discusses the key robust findings. Section V concludes the paper.

II. General Framework

In this section we present a framework to investigate whether Euro area monetary policy changed after the introduction of the European Monetary Union (EMU). First, we present time-varying parameters structural VAR model (SVAR). Second, we portray the reduced-form VAR implied by the structural model. Third, we introduce a class of linear restrictions on contemporaneous parameters of the SVAR that can be used to identify monetary policy and private sector behaviors. Fourth, we write the likelihood function of the identified SVAR. Fifth, we define the priors. Finally, we present the posterior distributions and show how to draw from them.

II.1. **The Structural Model.** Here, we present a model to estimate changes in the Euro area monetary policy. Following Hamilton (1989) and Sims and Zha (**?**), we use time-varying parameters SVAR that allows us to identify monetary authority and private sector behaviors. The structural VAR is as follows:

$$y_t' A_0(s_t) = \sum_{\ell=1}^{p} y_{t-\ell}' A_\ell(s_t) + z_t' C(s_t) + \varepsilon_t', \quad t = 1, \ldots, T. \tag{1}$$

where T is the sample size, h is the number of states, and $s_t \in \{1, \ldots, h\}$ is the state of the economy at time t. s_t follows a Markov chain with the transition probability matrix $\Pi = [\pi_1, \ldots, \pi_h]$ where $\pi_j = \left[\pi_{1,j}, \ldots, \pi_{h,j}\right]'$ and $\pi_{i,j} = \pi(i|j)$, where $\pi(i|j)$ is the probability that s_t equals i given that s_{t-1} was equal to j. $A_0(s_t)$ is an $n \times n$ contemporaneous parameter matrix when the state of the economy is s_t, p is the lag length, $A_\ell(s_t)$ is an $n \times n$ lag ℓ parameter matrix when the state of the economy is s_t, and $C(s_t)$ is a $1 \times n$ vector of constant parameters when the state of the economy is s_t. y_t is an $n \times 1$ vector of endogenous variables at time t, $z_t = 1$ for all t[3], ε_t is an $n \times 1$ vector of structural shocks at time t. ε_t is i.i.d. normally distributed with mean 0 and covariance matrix I_n. The initial conditions, $y_0, \cdots, y_{1-p}$, are taken as given.

Now we provide some definitions that will be useful in the rest of the paper. Let

$$A_0 = [A_0(1)', \ldots, A_0(h)']'$$

and

$$A_+ = [A_+(1)', \ldots, A_+(h)']'$$

[3]It is straightforward to include other exogenous variables in our framework.

where

$$A_+(k) = [A_1(k)', \ldots, A_p(k)', C(k)']'$$

for $k = 1, \ldots, h$.

Also let

$$a_{j,0} = [a_{j,0}(1)', \ldots, a_{j,0}(h)']'$$

be the j^{th} column of A_0 and

$$a_{j,+} = [a_{j,+}(1)', \ldots, a_{j,+}(h)']'$$

be the j^{th} column of A_+, where

$$a_{j,+}(k) = [a_{j,1,+}(k), \ldots, a_{j,p,+}(k), c_j(k)]'$$

for $k = 1, \ldots, h$ and $j = 1, \ldots, n$.

Finally, if we define

$$x_t = [y'_{t-1}, \ldots, y'_{t-p}, z'_t]'$$

then (1) can be compactly written as:

$$y'_t A_0(s_t) = x'_t A_+(s_t) + \varepsilon'_t,\ t = 1, \ldots, T. \tag{2}$$

II.2. **The Reduced-Form Representation and the Impulse Response Functions.** The reduced-form representation implied by the structural model (2) is as follows:

$$y'_t = x'_t B(s_t) + u'_t(s_t),\ t = 1, \ldots, T. \tag{3}$$

where

$$B(s_t) = A_+(s_t) A_0^{-1}(s_t),$$
$$u'_t(s_t) = \varepsilon'_t A_0^{-1}(s_t),$$

and

$$E\left(u_t(s_t) u'_t(s_t)\right) = \left(A_0(s_t) A'_0(s_t)\right)^{-1}$$

for $s_t = 1, \ldots, h$ and $t = 1, \ldots, T$.

Let us now consider the impulse response functions. First, let us define:

$$B(k) = [B_1(k)', \ldots, B_p(k)', B_0(k)']'$$

for $k = 1, \ldots, h$.

Let the state be k, then the impulse response of y_{t+v} to shock u_t is:

$$\Psi_{\nu}(k) = \frac{\partial y_{t+\nu}}{\partial u_t}$$

for $k = 1,\ldots,h$, where $\Psi_{\nu}(k)$, for $\nu \geq 0$, solves the following system of equations:

$$\Psi_{\nu}(k) = B_1'(k)\Psi_{\nu-1}(k) + B_2'(k)\Psi_{\nu-2}(k) + \ldots + B_p'(k)\Psi_{\nu-p}(k) \tag{4}$$

with $\Psi_0(k) = I_n$ and $\Psi_{\nu}(k) = 0$ if $\nu < 0$ for $k = 1,\ldots,h$.

Now let P be an orthogonal matrix and let the reduced-form representation implied by $(A_0(k)P, A_+(k)P)$ be:

$$y_t' = x_t'\widetilde{B}(k) + \widetilde{u}_t'(k),\ t = 1,\ldots,T. \tag{5}$$

and the impulse response of $y_{t+\nu}$ to shock $\widetilde{u}_t$ be:

$$\widetilde{\Psi}_{\nu}(k) = \frac{\partial y_{t+\nu}}{\partial \widetilde{u}_t}$$

for $k = 1,\ldots,h$.

It can be shown that $\widetilde{B}(k) = B(k)$ and $\widetilde{\Psi}_{\nu}(k) = \Psi_{\nu}(k)P^{-1}$ for $k = 1,\ldots,h$ and all $\nu \geq 0$.

Therefore, the reduced-form representation implied by $(A_0(k), A_+(k))$ and $(A_0(k)P, A_+(k)P)$ are identical. Hence, the structural models implied by $(A_0(k), A_+(k))$ and $(A_0(k)P, A_+(k)P)$ would share the likelihood function and be observationally equivalent. On the other hand, since $\widetilde{\Psi}_{\nu}(k) \neq \Psi_{\nu}(k)$, the reduced-form representation implied by $(A_0(k), A_+(k))$ and $(A_0(k)P, A_+(k)P)$ does not share the impulse response functions. Finally, note, that in order to compute the impulse responses implied by (3), we only need $B(k)$ and $A_0(k)$ for $k = 1,\ldots,h$.

II.3. **Identifying Restrictions.** Without restrictions the structural system (2) would not be identified. If P is an orthogonal matrix, the reduced-form representation implied by $(A_0(k), A_+(k))$ and $(A_0(k)P, A_+(k)P)$ would be identical. Therefore, the structural models implied by $(A_0(k), A_+(k))$ and $(A_0(k)P, A_+(k)P)$ would be observationally equivalent. In this section we define a class of linear restrictions imposed on A_0.

For $j = 1,\ldots,n$ and $k = 1,\ldots,h$, let the $q_j \times n$ matrix Q_j, where $q_j \leq n$, define the q_j restrictions over the elements of $a_{j,0}(k)$ such as

$$Q_j\, a_{j,0}(k) = 0 \tag{6}$$

and let U_j be the $n \times q_j$ matrix whose columns form the orthonormal basis for the null space of Q_j. Then, $Q_j\, a_{j,0}(k) = 0$ if and only if $\exists$ a $q_j \times 1$ vector $b_j(k)$ such that

$$a_{j,0}(k) = U_j\, b_j(k). \tag{7}$$

Finally, for $j = 1,\ldots,n$, let

$$b_j = [b_j(1)', \ldots, b_j(h)']',$$

$$b = [b_1, \ldots, b_n,]$$

and

$$U = [U_1', \dots, U_n']'.$$

Note the following three points. First, any set of $a_{j,0}(k)$ and Q_j for $j = 1, \dots, n$ and $k = 1, \dots, h$ implies a set of U_j and $b_j(k)$ for $j = 1, \dots, n$ and $k = 1, \dots, h$ and vice versa. Therefore, it is equivalent to defining the linear restrictions using either $a_{j,0}(k)$ and Q_j or U_j and $b_j(k)$. This implies that we can evaluate the likelihood function either using Q_j and $a_{j,0}(k)$ or U_j and $b_j(k)$. As it will be clear in the next section, we follow the second approach.

Second, any identification scheme defined as exclusion restrictions on the contemporaneous parameters belongs to this class of linear restrictions. For example, if we let $n = 3$, the exclusion restrictions identification scheme described by a lower triangular matrix is characterized by the following set of Q_j for $j = 1, \dots, 3$:

$$Q_1 = [(0,1,0)', (0,0,1)']'$$
$$Q_2 = [(0,0,1)']'$$

Finally, if P is an orthogonal matrix, in general, it would not be the case that $Q_j a_{j,0}(k)P = 0$ holds. In fact, if the model is exactly identified, for any matrix $A_0(k)$, there is a unique P such that $Q_j a_{j,0}(k)P = 0$ holds. We show these results in the following subsection.

II.4. **Normalization.** Because linear restrictions do not uniquely determine the sign of any equation, a SVAR with linear restrictions cannot be globally identified. A normalization rule is need in addition to the linear restrictions. There are many different normalization rules and we follow the one defined in Waggoner and Zha (2003b). As was pointed out in that paper, the choice of normalization rule is important particularly with respect to inferences concerning impulse responses. However, the theory developed in this paper will work for any choice of normalization rule as long as for any set of parameters the rule uniquely determines a choice of sign for each equation in the system.

At the same time, for Markov switching models, any permutation of the states will result in an observationally equivalent set of parameters. So as with SVARs, a normalization rule for determining the naming of the states is required. We follow the Wald normalization as described in Hamilton, Waggoner, and Zha (2003). As for the normalization rule for the SVAR, the theory developed in this paper will work for any choice of normalization as long as for any set of parameters the rule uniquely determines a choice for the naming of the states. We shall implicitly assume that all of our Markov switching SVAR models are normalized concerning both the signs of the impulse response functions and the naming of the states.

II.5. **Is the Model Exactly Identified?** An important question when dealing with SVAR is to know if the set of restrictions characterized by (6) exactly identifies the SVAR. Let us first define what we mean by a SVAR with linear restrictions given by (6) to be exactly identified.

Definition 1. A SVAR with linear restrictions given by (6) is exactly identified if and only if for every reduced form parameter $(B(k), \Sigma(k))$, except perhaps on a set of measure zero, and for every $k = 1, \dots, h$

there exists a unique set $(A_0(k), A_+(k))$, where $B(k) = A_+(k)A_0(k)^{-1}$ and $\Sigma(k) = (A_0(k)A_0(k)')^{-1}$, such that satisfies the restrictions.

Rothenberg (1971) gives a necessary condition for exact identification, which requires $n(n-1)/2$ restrictions. However, Rothenberg's (1971) condition is not sufficient.[4] The following theorem 2 gives us a necessary and sufficient condition for the SVAR system to be exactly identified.

Theorem 2. A SVAR with linear restrictions given by (6) is exactly identified if and only if there exists a permutation σ of $\{1, \cdots, n\}$ such that $rank\left(Q_{\sigma(i)}\right) = n - i$.

Proof. The proof is provided in Appendix D. □

Theorem 2 allows us to check if the SVAR is exactly identified. The next theorem tell us how to find such an identification.

Theorem 3. A SVAR with linear restrictions given by (6) is exactly identified if and only if for almost all values of the structural parameters $(A_0(k), A_+(k))$, such that $B(k) = A_+(k)A_0(k)^{-1}$ and $\Sigma(k) = (A_0(k)A_0(k)')^{-1}$, there exists the unique orthogonal matrix $P(k)$ such that $(A_0(k)P(k), A_+(k)P(k))$ satisfies the identifying restrictions in the form of (6).

Proof. The proof is provided in Appendix D. □

Theorem 3 implies that if restrictions described by (6) exactly identify the SVAR, then we can find a unique $P(k)$, such that $(A_0(k)P(k), A_+(k)P(k))$ satisfies the identifying restrictions for any set $(A_0(k), A_+(k))$ such that $B(k) = A_+(k)A_0(k)^{-1}$ and $\Sigma(k) = (A_0(k)A_0(k)')^{-1}$. Thus, we can always start with a recursive framework and then find $P(k)$ such that $(A_0(k)P(k), A_+(k)P(k))$ satisfies (6).

The final question is how to find such a $P(k)$ for $k = 1, \ldots, h$. In subsection III.3 we show how to find $P(k)$ for $k = 1, \ldots, h$ for SVARs exactly identified using short-run and long-run restrictions.

II.6. **The Likelihood Function.** In this section we describe how to evaluate the likelihood for the time-varying parameters SVAR defined by (2) and identified using the class of linear restrictions just described in section II.3. We first define:

$$d_j(k) = a_{j,+}(k) - \overline{S}a_{j,0}(k), \tag{8}$$

for $k = 1, \ldots, h$ and $j = 1, \ldots, n$, where

$$\overline{S} = [I'_{n\times n}, 0'_{(m-n)\times n}]'$$

Now let

$$d_j = [d_j(1)', \ldots, d_j(h)']'$$

for $j = 1, \ldots, n$.

[4]See Sims and Zha (1999) for a counter example.

Finally, let

$$d = [d_1, \dots, d_n].$$

Note that any A_0 and d imply a matrix A_+. Therefore, for any given U, the matrices b and d imply the matrices A_0 and A_+. Thus, we can write the likelihood function either using A_0 and A_+ or b and d. We choose the first option.

Now, if we define

$$Y^t = [y_1 \dots y_t]', \text{ and}$$

for all t we can write the following theorem.

Theorem 4. Given the restriction matrix U, the conditional likelihood function, $\pi\left(y_t|Y^{t-1}, s_t, b, d\right)$, is:

$$\begin{aligned}\pi\left(y_t|Y^{t-1}, s_t, b, d\right) \propto\ & \det\left|\left[U_1 b_1(s_t) \dots U_n b_n(s_t)\right]\right| \exp\left[-\frac{1}{2}\sum_{j=1}^{n} b_j'(s_t)\, U_j' S_t U_j b_j(s_t)\right] \\ & \exp\left[-\frac{1}{2}\sum_{j=1}^{n}\left(d_j(s_t) + \left(\overline{S} - P_t\right) U_j b_j(s_t)\right)' H_t \left(d_j(s_t) + \left(\overline{S} - P_t\right) U_j b_j(s_t)\right)\right],\end{aligned}$$

where

$$H_t = x_t' x_t,$$

$$P_t = H_t^{-1} x_t' y_t,$$

and

$$S_t = y_t' y_t - P_t' H_t P_t.$$

Proof. The proof is given in Appendix C. □

Following Kim and Nelson (), we can write the likelihood function $\pi(Y_T|b, d)$.

Corollary 5. Given the restriction matrix U, the likelihood function, $\pi\left(Y^T|b, d, \Pi\right)$, is:

$$\pi(Y^T|b, d, \Pi) \propto \prod_{t=1}^{T}\left\{\sum_{s_t=1}^{h}\left[\pi(y_t|Y^{t-1}, s_t, b, d)\Pr(s_t|Y^{t-1}, b, d, \Pi)\right]\right\}$$

where

$$\Pr(s_t|Y^{t-1}, b, d, \Pi) = \sum_{s_t=1}^{h}\pi(s_t|s_{t-1})\Pr(s_{t-1}|Y^{t-1}, b, d, \Pi)$$

and $\Pr(s_{t-1}|Y^{t-1}, b, d, \Pi)$ is updated using the Bayes rule.[5]

[5]We initialize the system setting $\Pr(s_0|Y^0, b, d, \Pi) = \Pr(s_0|b, d, \Pi) = 1/h$

II.7. **Priors: Modelling Regimes.** If we let all the parameters vary across regimes, b and d can be estimated independently across regimes. Therefore, we could use the methods by Chib (1996) and Sims (1999) to perform the model estimation. The problem is that a VAR with four to seven endogenous variables and one-year lag length would suffer the over-parameterization problems associated with few degrees of freedom. Hence, we define three set of priors that restrict the variation of parameters across regimes. First, we consider priors that impose constant parameters model, i.e., no cross-regime variation. Second, we contemplate priors which only allow for variances to change across regimes. Finally, we also use priors that imply that both parameters and variances can change across regimes. The actual priors for each of the cases are defined in Appendix B. In this section we just highlight the main differences among the three set of priors and their implications for cross-regime variation. In order to do that we first rewrite the parameters defining model (1) in the following way:

$$a_{i,j,0}(k) = \overline{a}_{i,j,0}\xi_j(k)\phi_{i,j}(k),$$

$$d_{i,j,\ell}(k) = \overline{d}_{i,j,\ell}\xi_j(k)\lambda_{i,j}(k),$$

and

$$c_j(k) = \overline{c}_j\xi_j(k)\mu_j(k)$$

for $i, j = 1, \ldots, n$ and $k = 1, \ldots, h$. Notice that writing the parameters this way already imposes a restriction on cross-regime variation. We restrict the cross-regime variation of d since we do not allow for variation between lags (i.e., $(d_{i,j,\ell}(k) = d_{i,j,\ell'}(k)$ for $\ell', \ell = 1, \ldots, p)$). This restriction is common to the three cases considered here.

- *Case I: Constant Parameters Priors.* These priors impose $\xi_j(k) = 1$, $\phi_{i,j}(k) = 1$, $\lambda_{i,j}(k) = 1$, and $\mu_j(k) = 1$ for $i, j = 1, \ldots, n$ and $k = 1, \ldots, h$. Therefore $a_{i,j,0}(k) = \overline{a}_{i,j,0}(k)$, $d_{i,j,\ell}(k) = \overline{d}_{i,j,\ell}(k)$, and $c_j(k) = \overline{c}_j$ for $i, j = 1, \ldots, n$ and $k = 1, \ldots, h$. This case corresponds to the constant-parameters VARs widely used in the literature.
- *Case II: Regime-Varying Variances Priors.* These priors impose $\phi_{i,j}(k) = 1$, $\lambda_{i,j}(k) = 1$m and $\mu_j(k) = 1$ for $i, j = 1, \ldots, n$ and $k = 1, \ldots, h$. Therefore, we can write $a_{i,j,0}(k) = \overline{a}_{i,j,0}\xi_j(k)$, $d_{i,j,\ell}(k) = \overline{d}_{i,j,\ell}\xi_j(k)$, and $c_j(k) = \overline{c}_j\xi_j(k)$ for $i, j = 1, \ldots, n$ and $k = 1, \ldots, h$. These priors imply that structural equations in model 1 are proportional across regimes. These priors also imply that the reduced-form parameters are constant across regimes, i.e., $B(k) = B(k')$ for all k, k', while the variances of u_t' vary across regimes.
- *Case III: Regime-Varying Variances and Parameters Priors.* These priors impose $\xi_j(k) = 1$ and $\overline{a}_{j,0} = 1$ for $i, j = 1, \ldots, n$ and $k = 1, \ldots, h$. Therefore $a_{i,j,0}(k) = \phi_{i,j}(k)$, $d_{i,j,\ell}(k) = \overline{d}_{i,j,\ell}\lambda_{i,j}(k)$, and $c_j(k) = \overline{c}_j(k)\mu_j(k)$ for $i, j = 1, \ldots, n$ and $k = 1, \ldots, h$. These set of priors imply that structural equations in model 1 move freely across regimes, with the only restriction that they do not change across lags. These priors also imply that the reduced-form parameters and variances change across regimes.

Given that our priors are defined over columns of $A_0(k)$ and $A_+(k)$, we can always mix cases II and III. For example, we can specify priors such that a set of columns of $A_0(k)$ and $A_+(k)$ follows case II, while the rest of the columns of $A_0(k)$ and $A_+(k)$ follow case III. Given that the columns of our SVAR (see equation (1)) are structural equations, mixing cases II and III imply that we have structural equations with only regime-varying variances (case II) and structural equations with both regime-varying variances and regime-varying parameters (case III).

II.8. **Posterior.** Sections II.6 and II.7 describe the likelihood function of model (1) and the three sets of prior distributions that we use in this paper. Since our priors restrict the variability of parameters across regimes, we cannot use the methods developed by Hamilton (1989) and Chib (1996). Instead we use the method described in Sims and Zha (2004a). Because of space considerations we refer the reader to Sims and Zha (2004a) for a detailed discussion of the posterior distributions implied by prior cases I-III and how to draw from them. Suffices to say that we are interested in the following posterior distributions:

$$\begin{gathered}
\pi(S^T|Y^T,b,\overline{d},\phi,\lambda,\mu,\Pi) \\
\pi(\Pi|Y^T,b,\overline{d},\phi,\lambda,\mu,S^T) \\
\pi(\phi,\lambda,\mu|Y^T,b,\overline{d},S^T,\Pi) \\
\pi(b|Y^T,\overline{d},\phi,\lambda,\mu,S^T,\Pi),
\end{gathered}$$

where $S^T=(s_1,\ldots,s_T)$ and $\overline{d}$, ϕ, λ, and μ are defined in appendix B, and we use standard McMc to draw from these posterior distributions and the modified harmonic mean (MHM), described in Gelfand and Dey (1994), to compute the marginal likelihood.

III. Identification Schemes

The general framework described in section II allows us to consider various identification schemes. We use four popular identification strategies. First we use a recursive scheme as in Christiano, Eichenbaum, and Evans (1996).[6] Second, we use a non-recursive method as in Gordon and Leeper (1994) and Sims and Zha (2004b). Third, we identify shocks using long-run restrictions as in Blanchard and Quah (1993) and Galí (1992). Finally, we use sign restrictions as in Faust (1998), Canova and De Nicoló (2002), and Uhlig (forthcoming). The first two identification methods can be summarized as linear restrictions on $A_0(k)$ for $k=1,...,h$, while the two last cannot. Therefore, the first two identification schemes can be directly mapped into the framework described in section II. In this section, we show how the last two can be derived from a recursive identification scheme, hence we can also use the methods described in section II to analyze them. The fact that we can consider four different identification schemes will be important to check the robustness of our results.

[6]SVARs identified using this scheme we will be sometimes called triangular systems.

	Fed	Inf	MD	PS	PS
Y	$\times$	$\times$	0	$\times$	$\times$
P	$\times$	$\times$	0	$\times$	0
R	$\times$	$\times$	$\times$	0	0
M	$\times$	$\times$	$\times$	0	0
Ex	0	$\times$	0	0	0

TABLE 1. Identification Scheme for "GLSZ"

It is also important to note that both sign and long run restrictions are computationally demanding. We develop new methods to implement them efficiently. Our methods are not only easy to use, but also take little computational time relative to the existing methods.

III.1. **CEE Identification.** In an influential paper, Christiano, Eichenbaum, and Evans (1996) propose a recursive identification strategy to identify monetary policy. We call this identification "CEE". They assume $A_0(k)$ to be a triangular system for $k = 1, \ldots, h$. Our, VAR identified using "CEE", includes output (Y), output deflator (P), nominal short-term interest rate (R), M3 (M), and Euro/dollar exchange rate (Ex).[7] Since in this identification scheme the order of variables matters, we follow Christiano, Eichenbaum, and Evans and order the variables in the following way: Y, P, R, M, and Ex. Hence, a structural shock to output will only affect output, a structural shock to inflation will affect output and inflation, etc. Using theorem 2 it can be checked the this SVAR is exactly identified.

III.2. **GLSZ Identification.** Gordon and Leeper (1994) and Sims and Zha (2004b) propose another identification strategy. We call this identification "GLSZ". Their identification focuses on the interpretation of the structural equations themselves. In particular, they separate the monetary policy equation from the money demand equations and other non-policy equations. The restrictions used to achieve such identification typically require simultaneous (non-recursive) relationships between the financial variables like the interest rate and money. We use the same variables used in "CEE". The identification is described in Table 1. An X in Table 1 indicates unrestricted parameters in $A_0(k)$ for all $k = 1, \ldots h$ and the blank spaces indicate the parameters that are restricted to be zero. The "Fed" column in Table XX represents the Federal Reserve contemporaneous behavior, the "Inf" column describes the financial sector, the "MD" represents the money demand equation, and the block consisting of the last two columns represents the production sector, whose variables are arbitrarily ordered in an upper triangular form. Using theorem 2 it can be proved that the this SVAR is over-identified.

III.3. **BGQ Identification.** In a seminal work Blanchard and Quah (1993) propose to use restrictions on the long-run impulse responses to achieve exact identification of a VAR model (henceforth we call it "BGQ" identification). For a given regime or state $k = 1, \ldots, h$, the long run responses can be expressed

[7] A full description of the data is given in Appendix A

as the inverse of the matrix $L(k)$, where:

$$L(k) = A_0(k) - \sum_{\ell=1}^{p} A_\ell(k).$$

We often do not have sufficient long-run restrictions that are economically justifiable to achieve exact identification. Galí (1992) suggests a combination of short-run and long-run restrictions to get the VAR model identified. The short-run restrictions used by Galí (1992) are linear restrictions imposed on $A_0(k)$ or $A_0^{-1}(k)$ or both (see also Blanchard and Watson 1986 and Bernanke and Mihov 1998). This set of restrictions forms a system of nonlinear equations to be solved to get the maximum likelihood estimates or the posterior estimates if a prior is used.[8] Solving a system of nonlinear equations (or minimizing a nonlinear function) for each posterior draw is time-consuming and not feasible for the regime switching model studied here. In this section we describe some new methods to deal efficiently with a combination of long-run and short-run restrictions in exactly-identified SVARs.

For $k = 1, \ldots, h$, define

$$X_{(A_0(k),\, A_+(k))} = \begin{bmatrix} \left(A_0^{-1}(k)\right)' \\ \left(L_0^{-1}(k)\right)' \end{bmatrix} \tag{9}$$

The restrictions considered by Blanchard and Quah (1993), Galí (1992), and others are special cases of linear restrictions on each column of $X_{(A_0(k),\, A_+(k))}$. This means that they can be defined as a set of matrices Q_j for $j = 1, \ldots, n$ such that:

$$Q_j X_{(A_0(k),\, A_+(k))} e_j = 0, \tag{10}$$

where e_j is the j^{th} column of the $n \times n$ identity matrix. The rank of Q_j is the number of linear restrictions on the j^{th} column of $X_{(A_0(k),\, A_+(k))}$.[9]

Theorem 2 allows us to check if the SVAR identified using (10) is exactly identified. If that is the case, theorem 3 tells us that for any $(A_0(k),\, A_+(k))$ a unique orthogonal matrix $P(k)$, for $k = 1, \ldots, h$, exists, such that:

$$Q_j X_{(A_0(k)P(k),\, A_+(k)P(k))} e_j = 0,$$

for $j = 1, \ldots, n$ and $k = 1, \ldots, h$.

The question is how to find the orthogonal matrix $P(k)$. In the following algorithm we describe an efficient way to find $P(k)$:

Algorithm 1.
Assume the SVAR model is exactly identified. Let σ be the permutation with the property that $rank\left(Q_{\sigma(i)}\right) = n - i$ for $i = 1, \ldots, n$. Let $(A_0(k), A_+(k))$, for $k = 1, \ldots, h$, be the set of structural parameters coming from the recursive identification.

[8] The 2SLS estimate, as used by Galí (1992), is an approximation to the maximum likelihood estimate. How well the approximation is depends on how good the instruments are in the first stage of the estimation.

[9] In addition to the linear restrictions, we also have to impose a normalization rule to uniquely determine the sign of each equation in the system. See Waggoner and Zha (2003b) for details.

(1) Let $k = 1$.

(2) Let $i = 1$. Let $p_i(k)$ be any unit-length n-dimensional vector such that $X_{(A_0(k),A_+(k))}p_i(k)$ satisfies the restrictions on column $\sigma(i)$, i.e. $Q_{\sigma(i)}X_{(A_0(k),A_+(k))}p_i(k) = 0$. Such vector exists because $rank\left(Q_{\sigma(i)}X_{(A_0(k),A_+(k))}\right) \leq rank\left(Q_{\sigma(i)}\right) = n-1 < n$. The vector can be found using the LU decomposition of $Q_{\sigma(i)}X_{(A_0(k),A_+(k))}$. Set $i = i+1$.

(3) Form the matrix:

$$\tilde{Q}_{\sigma(i)} = \begin{bmatrix} Q_{\sigma(i)}X_{(A_0(k),A_+(k))} \\ p_1(k)' \\ \vdots \\ p_{i-1}(k)' \end{bmatrix}$$

(4) Let $p_i(k)$ be a unit length vector such that $\tilde{Q}_{\sigma(i)}p_i(k) = 0$. Such vector exists because $rank\left(Q_{\sigma(i)}\right) = n-i$ and hence $rank\left(\tilde{Q}_{\sigma(i)}\right) < n$. The vector can be found using the LU decomposition of $\tilde{Q}_{\sigma(i)}$.

(5) If $i < n$ go to (3).

(6) If $k < h$ go to (1), otherwise stop.

The above algorithm produces the matrices

$$P(k) = \begin{bmatrix} p_{\sigma^{-1}(1)}(k) & \cdots & p_{\sigma^{-1}(n)}(k) \end{bmatrix},$$

for $k = 1,\ldots,h$, which are the required matrices.[10]

III.4. **Exclusion Restrictions.** Most long-run and short-run restrictions used in the literature are of exclusion nature. If these restrictions meet certain conditions, we have an even more efficient algorithm for determining the matrix P. Such conditions are described by the following definition.

Definition 6. Identifying restrictions of the form of (9) are *triangularizable* if the following condition holds: $Q_jX_{(A_0(k),\,A_+(k))}e_j = 0$ if and only if there is a permutation $P_1(k)$ of the rows of $X_{(A_0(k),A_+(k))}$ and a permutation $P_2(k)$ of the columns of $X_{(A_0(k),A_+(k))}$, such that the permuted matrix $P_1(k)X_{(A_0(k),A_+(k))}P_2(k)$ is lower triangular.

If exclusion restrictions are triangularizable, algorithm 1 can be further improved, so that the orthogonal matrix given by theorem 3 can be found using a single QR decomposition as described in the following theorem.

Theorem 7. Suppose the identifying restrictions are triangularizable. For $k = 1,\ldots,h$, let $P_1(k)$ and $P_2(k)$ be the permutation matrices. Let $(A_0(k),A_+(k))$, for $k = 1,\ldots,h$, be the set of structural parameters coming from the recursive identification. Using the QR decomposition on $\left(P_1(k)X_{(A_0(k),A_+(k))}\right)'$, write $P_1(k)X_{(A_0(k),A_+(k))} = T_L(k)P_3(k)$ where $P_3(k)$ is an orthogonal matrix and $T_L(k)$ is lower triangular. The structural parameters $(A_0(k)P(k),A_+(k)P(k))$ for $P(k) = P_3(k)'P_2(k)'$ satisfy the restrictions.

[10] Note that by construction P is an orthonormal matrix

Three long-run restrictions
Aggregate demand shocks have no long-run effect on output
Monetary policy shocks have no long-run effect on output
Exchange rate shocks have no long-run effect on output

Three short-run restrictions
Monetary policy shocks have no contemporaneous effect on output
Exchange rate shocks have no contemporaneous effect on output
Exchange rate shocks have no contemporaneous effect on the interest rate

TABLE 2. Identifying restrictions under BGQ

III.5. **Data Description and Identification Assumptions.** Finally, let us describe the data and the combination of long-run and short-run restrictions that we use to identify our SVAR. We follow Peersman and Smets (2003) and consider a four-variable VAR system combining both long-run and short-run restrictions to achieve a particular BGQ identification. The four endogenous variables are quarterly output growth (Δy), quarterly inflation (ΔP), the nominal short-term interest rate (R), and quarterly change of the exchange rate (ΔEx). There are four structural shocks in this system: an aggregate supply shock (ε^s), an aggregate demand shock (ε^d), a monetary policy shock (ε^p), and an exchange rate shock (ε^e). In the notation of (1) we have

$$y_t = [\Delta y_t \; \Delta P_t \; R \; \Delta E x_t]',$$
$$\varepsilon_t = [\varepsilon_t^s \; \varepsilon_t^d \; \varepsilon_t^p \; \varepsilon_t^e]'.$$

Peersman and Smets' (2003) long-run and short-run restrictions are summarized in table 2. We use the same set of restrictions. Table 2 identification restrictions imply three exclusion restrictions in $L^{-1}(k)$ and three exclusion restrictions in $A_0^{-1}(k)$. Then, we have:

$$A_0^{-1}(k) = \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & 0 & \times \end{bmatrix}, \quad L^{-1}(k) = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{bmatrix}, \tag{11}$$

where the symbol $\times$ means no restriction imposed and 0 means the exclusion restriction. Checking if the model is identified or not should always be the first step. If we use theorem 2, we can see that (11) exactly identifies the model.[11]

[11]From the discussion prior to theorem 2, it should be clear that we cannot count the number of exclusion restriction to verify if the model is or not exactly identified. Here it is an example: If, instead of the assumption that an exchange rate shock has no contemporaneous effect on output, we assume that a demand shock has no contemporaneous effect on output, we have

III.6. **CDFU Identification.** The above described identification schemes are based on zero restrictions either on the contemporaneous coefficients of the SVAR or the long run responses of certain variables to shocks. This class of approaches is easy to implement, but sometimes they do not generate the impulse responses that fit economists' prior beliefs. Faust (1998), Canova and De Nicoló(2002), and Uhlig (forthcoming) propose an alternative approach. Their basic idea is to use sign restrictions directly on impulse responses themselves to identify SVARs. For example, in response to a contractionary monetary shock the interest rate should rise, while money, prices, and output should all fall. We call their approach "CDFU" identification. Although Faust (1998), Canova and De Nicoló(2002), and Uhlig (forthcoming) start from the same idea, they implement it in different ways. In this section, we first briefly describe the approaches of Faust, Canova and De Nicoló, and Uhlig, highlighting the problems of applying them to our switching model. Then, we describe our algorithm. Finally, we describe the variables considered in the VAR and state the sign restrictions we use to identify structural shocks.

III.6.1. *Faust Method.* Faust (1998) presents a way to check the robustness of any claim from a SVAR. All possible identifications are checked searching for the one that is worst for the claim, subject to the restriction that the identified VAR produces the correct impulse response functions.

Faust (1998) shows that this problem is equivalent to solving a eigenvalue problem $\sum_{i=0}^{M} \frac{R!}{i!(R-i)!}$ times, where R is the number of sign restrictions and $M = \max(n-1, R)$. As Faust (1998) recognizes, this method may not be feasible for large problems, like the one analyzed here. Finally, we see Faust's approach as a way to check claims on contributions of identified shocks to the forecast error variance, not as a way to identify SVARs.

III.6.2. *Canova and De Nicoló Method.* Canova and De Nicoló (2002) also identify SVARs using impulse response sign restrictions. Their method is based on the following theorem:

Theorem 8. Let P $(n \times n)$ be an orthogonal matrix. Then a unique series $\{\{\theta_{i,j}\}_{j=i+1}^{n}\}_{i=1}^{n-1}$ exists, where $0 \le \theta_{i,j} < 2\pi$ if $j = i+1$ and $-\pi/2 \le \theta_{i,j} \le \pi/2$ if $j > i+1$, such that:[12]

$$P = \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} Q_{i,j}\left(\theta_{i,j}\right)$$

or

$$P = S \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} Q_{i,j}\left(\theta_{i,j}\right)$$

---

the following pattern of restrictions on $A_0^{-1}(k)$ and $L^{-1}(k)$:

$$A_0^{-1}(k) = \begin{bmatrix} 0 & \times & \times & \times \\ \times & \times & \times & \times \\ 0 & \times & \times & \times \\ \times & \times & 0 & \times \end{bmatrix}, \quad L^{-1}(k) = \begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{bmatrix}, \tag{12}$$

If we compare (11) and (12), we observe that both have the same number of exclusion restriction. On other hand, if we use theorem 2 we find that the restriction implied by (12) do not exactly identify the system.

[12]In Canova and De Nicoló (2002), the notation $Q_{i,j}(\theta)$ is used where θ is implicitly assumed to vary with different i and j.

where

$$S = \begin{bmatrix} 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ 0 & \cdots & 0 & -1 \end{bmatrix}$$

and

$$Q_{i,j}\left(\theta_{i,j}\right) = \begin{array}{c} \begin{array}{ccccccc} & & \text{col } i & & \text{col } j & & \\ & & \downarrow & & \downarrow & & \end{array} \\ \begin{array}{r} \\ \\ \text{row } i \rightarrow \\ \\ \text{row } j \rightarrow \\ \\ \\ \end{array} \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & \cos\left(\theta_{i,j}\right) & \cdots & -\sin\left(\theta_{i,j}\right) & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & \sin\left(\theta_{i,j}\right) & \cdots & \cos\left(\theta_{i,j}\right) & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \end{array}.$$

Proof. The proof follows from Algorithm 5.2.2 of Golub and Van Loan (**?**). □

Using theorem 8, Canova and De Nicoló (2002) identify SVARs with the following algorithm:

Algorithm 2.

(1) Begin with a triangular SVAR system.

(2) Draw the system parameters $A_0(k)$ and $B(k)$ from the posterior distribution.

(3) Determine a grid on the set of all orthogonal matrices.

(4) Perform a grid search to find an orthogonal matrix $P(k)$, such that the impulse responses generated from $A_0(k)P(k)$ and $B(k)$ satisfy all the sign restrictions.

Theorem 8 allows for different ways to design a grid, but because the space of all orthogonal $n \times n$ matrices is a $n(n-1)/2$ dimensional space, any grid that divides the interval $[-\pi/2, \pi/2]$ in M points[13] implies a search over $M^{n(n-1)/2}$ points in the space of all orthogonal $n \times n$ matrices. Thus, it is not feasible to perform this grid search for large values of n.

III.6.3. *Uhlig's Methods.* Uhlig's (forthcoming) proposes another method to identify SVARs based on impulse response sign restrictions. His method draws from the set of posterior orthonormal matrices, such that the impulse response sign restrictions hold, using the following algorithm:

Algorithm 3.

(1) Begin with a triangular SVAR system.

(2) Draw the system parameters $A_0(k)$ and $B(k)$ from the posterior distribution.

[13] Or the interval $[-\pi, \pi]$ in $2M$ points.

(3) Draw n independent standard normal vectors of length n and recursively orthonormalize them. Call $P(k)$ the resulting orthonormal matrix.

(4) Generate the impulse responses from $A_0(k)P(k)$ and $B(k)$.

(5) If these impulse responses do not satisfy the sign restrictions, keep the draw. Otherwise discard it.

This method is feasible for large models like the one we are dealing with in this paper. In fact, the method we propose in the following subsection is just a more efficient version of Uhlig's approach.

III.6.4. *Our Algorithm.* In this subsection we propose a modified version of Uhlig's method to draw from the posterior distribution of orthonormal matrices such that a given set of impulse response sign restrictions hold. The main difference between Uhlig's and our approach is that while Uhlig recursively orthonormalize $P(k)$, we use the following theorem to directly draw an orthonormal matrix $P(k)$.

Theorem 9. Let X be an $n \times n$ random matrix with each element having an independent standard normal distribution. Let $X = QR$ be the QR decomposition of X with the diagonal of R normalized to be positive. Then Q has the uniform (or Haar) distribution.

Proof. The proof follows directly from Stewart (**?**). □

Theorem 9 gives us an easy and fast way to implement random selections of orthonormal matrices in order to get the impulse responses that satisfy a set of sign restrictions as described below.[14]

Algorithm 4.

(1) Begin with a triangular SVAR system.

(2) Draw the system parameters $A_0(k)$ and $B(k)$ from the posterior distribution.

(3) Draw an independent standard normal $n \times n$ matrix X and let $X = QR$ be the QR decomposition of X with the diagonal of R normalized to be positive.

(4) Let $P(k) = Q$ and generate the impulse responses from $A_0(k)P(k)$ and $B(k)$.

(5) If these impulse responses do not satisfy the sign restrictions, then return to step (3).

In theory, this algorithm is not guaranteed to terminate. In practice, we set a maximum number of iterations in which steps (3) through (5) should be repeated. If the maximum is reached, the algorithm should move to step (1) to draw another set of parameter values.[15]

The main differences with Uhlig's method are: (1) We do not discard any posterior draw and (2) we directly draw from uniform (or Haar) distribution while Uhlig does it recursively. These two differences make our algorithm more efficient and faster–two important features given the large system we consider.

---

[14] Stewart (**?**) has even more efficient algorithms for generating uniform random orthogonal matrices, but they are less straightforward and more difficult to implement.

[15] In the applications discussed in Section IV, we set the maximum number to be 1000 and this maximum was never reached in our millions of posterior draws.

III.7. **Data Description and Identification Assumptions.** Finally, let us describe the data and the sign restrictions we use to identify our SVAR. Our SVAR, identified using sign restrictions, includes the same data that "CEE" and "GLSZ" include, i.e., output (Y), output deflator (P), nominal short-term interest rate (R), M3 (M), and Euro/dollar exchange rate (Ex). We use the following sign restrictions:

- An expansionary monetary policy shock implies an interest rate decrease and an increase of M3 for two periods.
- A positive shock to money demand implies an interest rate and M3 increase for two periods.
- A positive demand shock implies an increase in output and prices for two periods.
- A positive supply shock implies an increase in output and a decrease in prices for two periods.
- A positive external shock implies an exchange rate devaluation and an increase in output for two periods.

IV. Empirical Results

In this section we identify a set of five-lag SVAR using the identification schemes discussed in section III. As mentioned in section II.7, our priors specification allows us to mix cases II and III. We consider five different specifications:

- *All-constant specification.* No regime change is allowed.
- *Variance-only specification.* All the structural equations are Case II.
- *Monetary-policy specification.* All the structural equations except monetary policy equation are Case II, while the monetary policy equation is Case III.
- *Private-sector specification.* All the structural equations except monetary policy equation are Case III, while the monetary policy equation is Case II.
- *All-change specification.* All the structural equations are Case III.

The all-constant specification does not consider regime change. We will take this as the benchmark specification. In the variance-only specification all the structural equations have case II priors. Therefore, only the variance of the structural equations changes across regimes. With this specification we consider the case where no change in behavior has occurred and all the change in observed volatility is explained by changes in the volatility of the structural shocks. In the monetary-policy specification we allow for the parameters of the monetary policy equation to change across regimes, while only the variance of the rest of the structural equations is allowed to change across regimes. With this specification we consider the case that the monetary authority has changed its behavior while the private sector has not. In the private-sector specification the parameters of all the structural equations, but those of the monetary policy equation, are allowed to change across regimes, while only the variance of the monetary policy equation can change across regimes. With this specification we consider the case that the private sector has changed its behavior while the monetary authority has not. Finally, in the all-change specification, all the parameters of all the structural equations can change.

Constant	2271.0			
	Variance Only	Monetary Policy	Private Sector	All Change
2 states	2291.8	2284.5	2274.4	2278.8
3 states	2287.3	DEG	DEG	DEG
4 states	2281.9	DEG	DEG	DEG
5 states	2279.1	DEG	DEG	DEG
6 states	2272.9	DEG	DEG	DEG

TABLE 3. Marginal log likelihoods for the five specifications and different number of states under the CEE identification scheme.

IV.1. **Results for the CEE identification.** The CEE identification scheme considers output (Y), output deflator (P), nominal short-term interest rate (R), M3 (M), and Euro/dollar nominal exchange rate (Ex) and uses a recursive scheme to identify the structural shocks. The order of the variables is the following: Y, P, R, M, and Ex. Hence, R contemporaneously responds to changes in Y and P, but Y does not contemporaneously respond to changes to any other variables.

Table 3 reports the marginal log likelihoods for the five specifications and different number of states under the CEE identification scheme.[16] [17] A variance-only specification, with two or three states, is overwhelmingly favored by the data in comparison to the constant VAR model and any other time-varying VAR models. From these results we interpret that, when identified using the CEE scheme, SVARs with more than 3-state tend to overfit the data, and thus are penalized by the marginal likelihood. The log marginal likelihood difference between the 2-state and 3-state variance-only specifications is less than four. This evidence is strong, but not conclusive, in favor of the 2-state variance-only specification. Therefore, because of space considerations, we only analyze the 2-state variance-only specification. Table 3 also shows that the data strongly favor the 2-state monetary policy specification over the constant, 2-state private sector, and 2-state all-change specifications. We interpret this result as evidence in favor of a change in the Euro-area monetary policy authority behavior during the studied period.

Figure 1 displays the posterior probability of each state for the 2-state variance-only specification model under the CEE identification. We call the state with very high and persistent probability after 1992 the EMS regime, while the other is called Non-EMS regime. Although the EMS regime remains with high probability for the some years before 1993, it periodically switches to the other state, probably reflecting

[16]All the marginal likelihoods reported in this paper are computed with 6 million MCMC draws. Using repeated runs, the computed maximum of numerical standard errors for all marginal likelihoods is less than 0.7 in log value. Using the Newey-West (1987) approximation procedure, the numerical standard errors give even smaller values. The marginal likelihood for the constant VAR model is computed using the algorithm described by Chib (1996) and Waggoner and Zha (2003a). The Matlab code can be downloaded from home.earthlink.net/ tzha02/programCode.html. Because the Markov chain Monte Carlo algorithm for the time-varying VAR models is not a Gibbs sampler, the marginal likelihoods for these models are computed using the modified harmonic means procedure discussed by Geweke (1999).

[17]DEG stands for “degenerate”, meaning that for these models there is no posterior probability for at least one state.

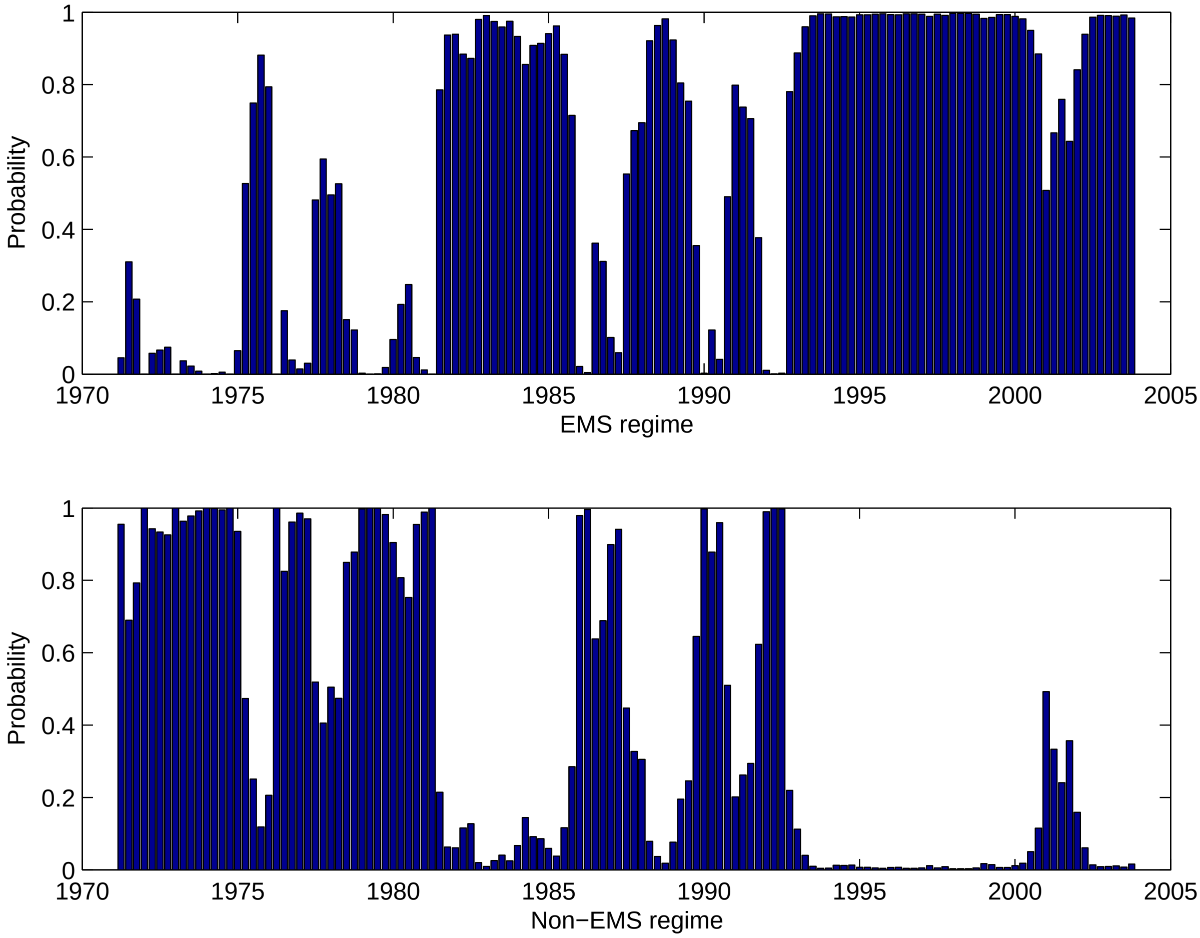

FIGURE 1. Posterior probabilities of states for the 2-state variance-only specification model under the CEE identification.

the uncertain periods of implementing the new monetary system (Ungerer et all **?**). We notice three things in figure 1: (1) the coincidence between the 1993 regime change and the institutional evolution of the EMS occurred in August of 1993, after 1992-93 crisis (2) the non-EMS regime has very high and persistent probability before 1980, and (3) between 1980 and 1993 the probability of any of the regimes is quite volatile. This high volatility reflects the uncertainty associated with the intuitional change taking place in the Euro area. It is also important to notice that the fact that the probability of the EMS regime increases after 1980 should not be surprising, since most of the inflation decline in the Euro area occurred during the 80's.

Table 4 reports the variance of the structural shocks for each variable under the two regimes, along with the relative variance across regimes, for the 2-state variance-only specification model under the CEE identification. The EMS regime is associated with much smaller volatility of structural shocks of all the variables. In particular, most of the fall in volatility is due to the fall in the variance of shocks to R and P. This phenomenon is also reflected in the impulse responses to a monetary shock displayed in Figure 2. The first column graphs the impulse response associated with a monetary policy shock for the constant

Variables	EMS	Non-EMS	Relative volatility
Y	0.014E-03	0.042E-03	3.039
P	0.003E-03	0.016E-03	4.235
R	0.015E-03	0.064E-03	4.063
M	0.015E-03	0.028E-03	1.854
Ex	0.553E-03	1.087E-03	1.963

TABLE 4. Residual variance of the shocks for the 2-state variance-only model under CEE

specification model under the CEE identification. The last two columns graph the impulse response associated with a monetary policy shock for the 2-state variance-only specification. The non-EMS regime impulse responses are larger than those for the EMS regime, while the responses for the constant model are in between those of the two regimes. In response to a contractionary shock to monetary policy, the interest rate rises and money falls (the liquidity effect), output falls, but the price level rises somewhat.[18] The increase in the price level, although statistically significant, is not economically important as compared to the other models.

Thus, if we identify the SVAR using the CEE scheme, the data favor a 2-state variance-only specification. One of the regimes mainly occurs after 1993 and is associated with lower volatility of structural shocks to R and P. We call this regime EMS regime. After a contractionary shock to monetary policy, we estimate a liquidity effect, a drop in output, and an increase in the price level. The estimated price puzzle is significant, but weak.

IV.2. **Results for the GLSZ identification.** As in the CEE identification scheme, the GLSZ scheme considers output (Y), output deflator (P), nominal short-term interest rate (R), M3 (M), and Euro/dollar nominal exchange rate (Ex) and uses a recursive scheme to identify the structural shocks. The GLSZ identification differs from the CEE identification in two ways. First, it does not treat M and R recursively, but models them simultaneously as money demand and money supply, where the money demand equation includes the variables M, R, Y, and P and the monetary policy equation includes only the two variables M and R.[19] Second the resulting structural model is over-identified.

Table 5 reports the marginal log likelihoods for the five specifications and different number of states under the GLSZ identification scheme. We conclude two things: (1) the 2-state variance-only specification dominates all the other specifications for the GLSZ identification scheme and (2) the marginal log likelihood slightly favors the GLSZ identification scheme over the CEE. Hence, as it was the case with the CEE scheme, the data favors the 2-state variance-only specification, although there is an important

[18] The price puzzle exists even when we include commodity prices in the VAR and when we reorder the variables (for example, letting R responds to commodity prices or Ex or both).

[19] We choose this identification on the basis of a information-delay assumption that the central bank cannot observe real GDP and the GDP deflator within the quarter.

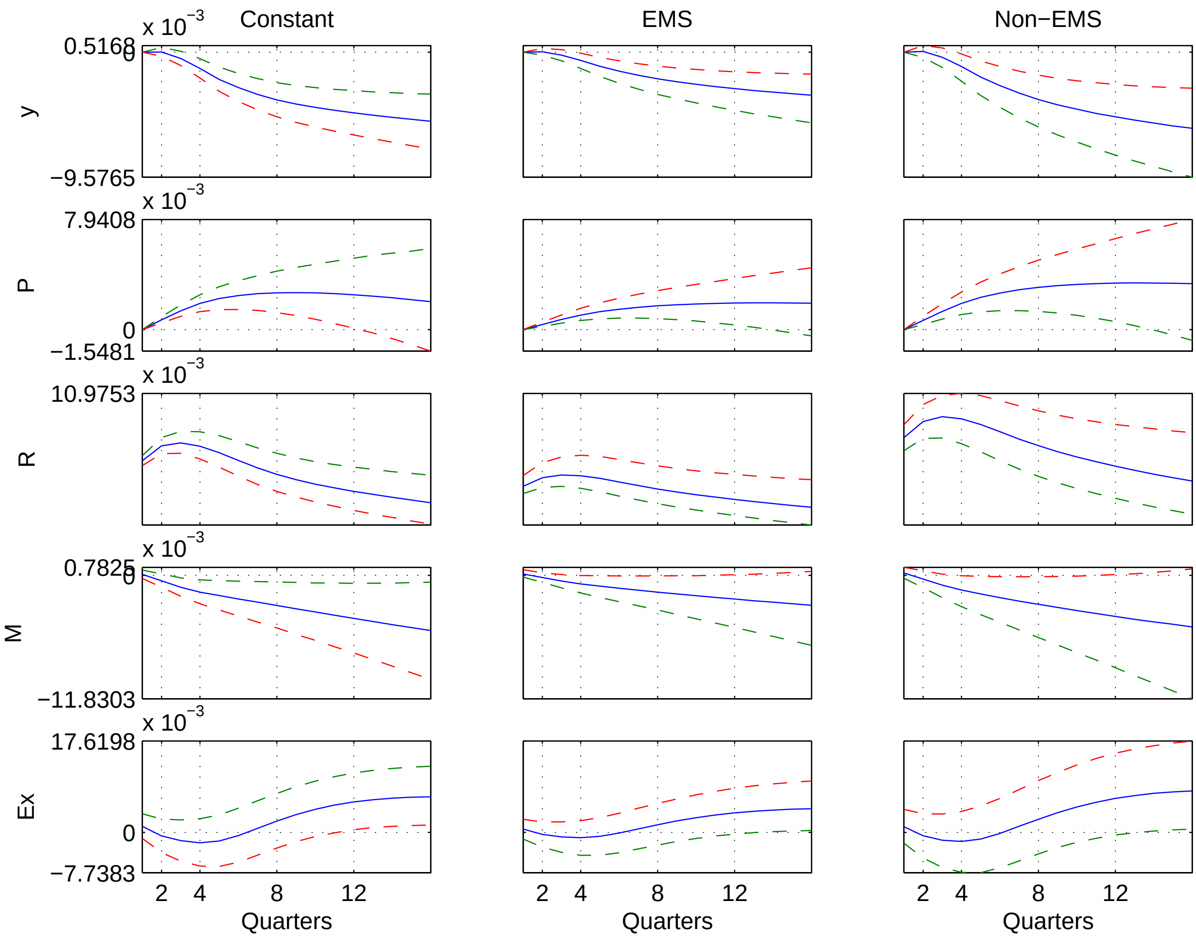

FIGURE 2. Impulse responses to a one-standard-deviation monetary policy shock under the CEE identification. The solid line represents the posterior median estimate and the two dashed lines contain the 68 percent probability based on 500,000 MCMC draws.

Constant	2273.6			
	Variance Only	Monetary Policy	Private Sector	All Change
2 states	2297.8	2290.8	2277.6	2279.7
3 states	2290.6	DEG	DEG	DEG
4 states	2283.3	DEG	DEG	DEG
5 states	2281.0	DEG	DEG	DEG
6 states	2277.8	DEG	DEG	DEG

TABLE 5. Marginal log likelihoods for the five specifications and different number of states under the GLSZ identification scheme.

difference. When identifying the SVAR using GLSZ, the evidence is conclusive in favor of the 2-state model over the 3-state one.

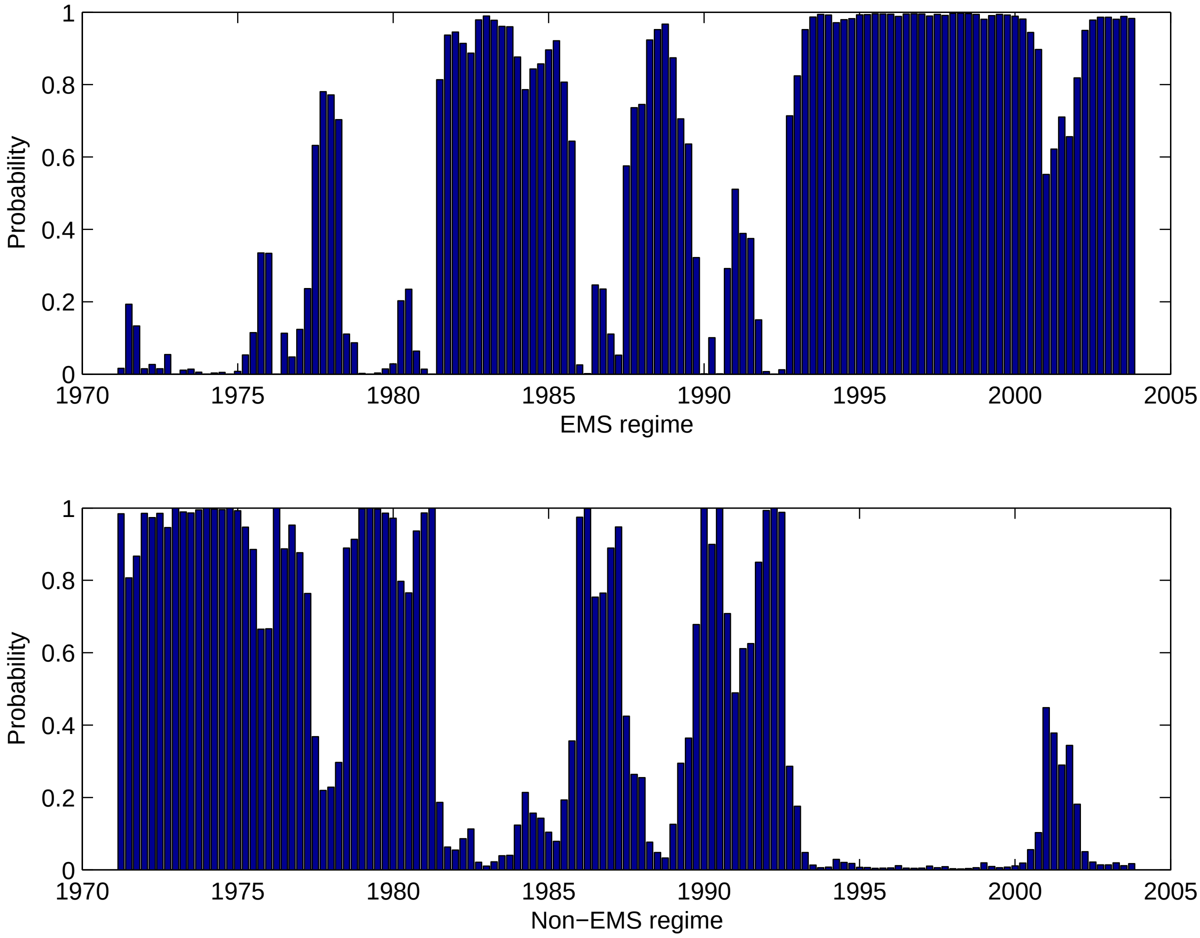

FIGURE 3. Posterior probabilities of states for the 2-state variance-only specification model under the GLSZ identification.

Figure 3 displays the posterior probability of each state for the 2-state variance-only specification model under the GLSZ identification. Figure 1 and figure 3 are remarkably similar, with two very distinguishable regimes. The EMS regime has very high and persistent probability after 1993, while the non-EMS regime has very high and persistence probability before 1980. Between 1980 and 1993 the probability of any of the regimes is quite volatile.

Table 6 reports the variance of the structural shocks for each variable under the two regimes, along with the relative variance across regimes, for the 2-state variance-only specification model under the GLSZ identification. The EMS regime is associated with smaller volatility of structural shocks of all the variables. As in CEE scheme, most of the fall in volatility is concentrated in the shocks to R and P.

Figure 4 reports the impulse responses to a monetary policy shock for the GLSZ identification scheme. The first column graphs the impulse response associated with a monetary policy shock for the constant specification, while the last two columns graph the impulse response associated with a monetary policy shock for the 2-state variance-only specification. An important thing to notice is the uncertainty about the dynamic responses. All the 68 percent confidence intervals are wider than in the CEE scheme. As in the

Variables	EMS	Non-EMS	Relative volatility
Y	0.013E-03	0.039E-03	2.797
P	0.004E-03	0.014E-03	3.459
R	0.018E-03	0.052E-03	2.755
M	0.012E-03	0.028E-03	2.340
Ex	0.551E-03	1.200E-03	2.175

TABLE 6. Residual variance of the shocks for the 2-state variance-only model under GLSZ

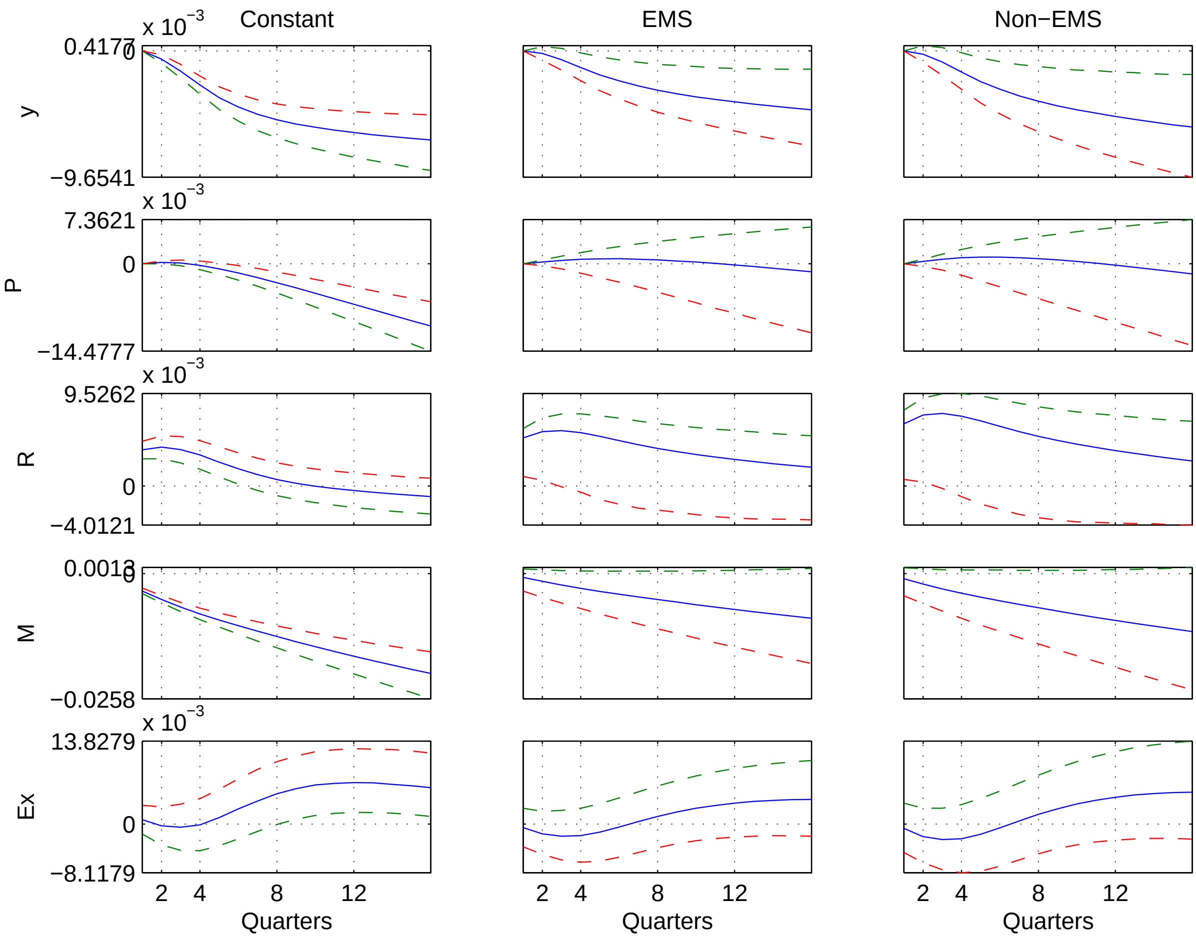

FIGURE 4. Impulse responses to a one-standard-deviation monetary policy shock under the GLSZ identification. The solid line represents the posterior median estimate and the two dashed lines contain the 68 percent probability based on 500,000 MCMC draws.

CEE scheme, we find that, in response to a contractionary shock to monetary policy, the interest rate rises and money falls (the liquidity effect), output falls, but the price level rises somewhat. The main difference with the CEE scheme is that, since the confidence intervals are so wide, the prize puzzle is not statistically significant.

Hence, we conclude that, for the 2-state variance-only specification, the GLSZ and the CEE schemes produce very similar results, although the CEE scheme produces a more accurate estimation and the marginal likelihood strongly supports the GLSZ scheme.

IV.3. **Results for the BGQ Identification.** As discussed in Section III, another widely-used identification scheme is to use long-run restrictions (Blanchard and Quah 1993 and Galí 1992). A typical finding is that the estimated impulse responses of output to a monetary shock are small, even in the short run (see Galí, 1992). Using the identification scheme of Galí (1992) and Peersman and Smets (2003), in this section we show that our estimated output responses to a monetary policy shock are also small, consistent with the previous estimates in the literature. However, the error bands are so wide and skewed that they change the implications of the point estimates. The wide error bands also raise a question of how informative this particular identification is.[20] [21] We use the following variables: output growth (ΔY), quarterly inflation (ΔP), nominal short-term interest rate (R), and quarterly change of the exchange rate (ΔEx). Using the restrictions reported in table 2, we identify four structural shocks: an aggregate supply shock (ε^s), an aggregate demand shock (ε^d), a monetary policy shock (ε^p), and an exchange rate shock (ε^e).

As mentioned in section III.3, in order to implement the BGQ scheme, we start from a recursive system (a CEE type of scheme). Then, for each posterior draw we find the rotation of $A_0(k)$ and $A_+(k)$ such that the restrictions reported in table 2 hold. This implies that there are subtleties in interpreting the five types of time-varying specifications. An orthonormal rotation of $A_0(k)$ will, in general, violate the time-varying restrictions on the original form of $A_0(k)$. Hence, all the four columns in Table **??** relate to all-change specifications in the sense that the parameters in each column of $A_0(k)$, after the rotation, vary across states beyond a scaling factor. The differences among these columns reflect different parsimonious ways of parameterizing the time-varying coefficients, and accordingly we label these columns as "TV I", "TV II", etc., where TV stands for time-variation.

For this four-variable VAR system, the 3-state TV-I specification gives the highest marginal likelihood.[22] Figure 5 displays the posterior probabilities of each regime for the 3-state TV-I specification for the BGQ scheme. We call the first state the early-EMS regime, which is concentrated between late 70's and early 90's. We call the second state the EMS regime, which concentrates after 1993. If we compare

---

[20]Faust and Leeper (**?**) make this same point.

[21]Unrealistic wide error bands can be caused by not properly normalizing the signs of responses for each posterior draw. Such problem can be solved by appropriate normalization (see Waggoner and Zha, 2003a, and Hamilton, Waggoner and Zha, 2003, for detailed discussions). Indeed, when the impulse responses are properly normalized, one may be able to obtain sensible error bands of impulse responses for certain SVAR models with long-run restrictions imposed (e.g., Sims and Zha 1999 and Evans and Marshall **?**). The impulse responses of output to a monetary policy shock under the long-run restrictions of Evans and Marshall are long-lived (more than 4 years) and strong.

[22]Because three out of the four variables are in first differences, we make the overall tightness of the hyperparameters very loose by setting $\lambda_0 = 2, \mu_5 = \mu_6 = 0.1$. The log value of the highest marginal likelihood is 1736.8 and the log value of the marginal likelihood for the constant model is 1697.1. We also experimented with different prior hyperparameter values and our results are not sensitive to these changes despite the fact that three series are differenced and presumably stationary.

Variables	Early EMS	EMS	Non-EMS
ΔY	0.056E-03	0.013E-03 (0.242)	0.028E-03 (0.509)
ΔP	0.009E-03	0.003E-03 (0.355)	0.064E-03 (7.011)
R	0.053E-03	0.013E-03 (0.248)	0.081E-03 (1.505)
ΔEx	1.216E-03	0.663E-03 (0.545)	1.439E-03 (1.182)

TABLE 7. Residual variance of the shocks for the 2-state TV-I model under BGQ

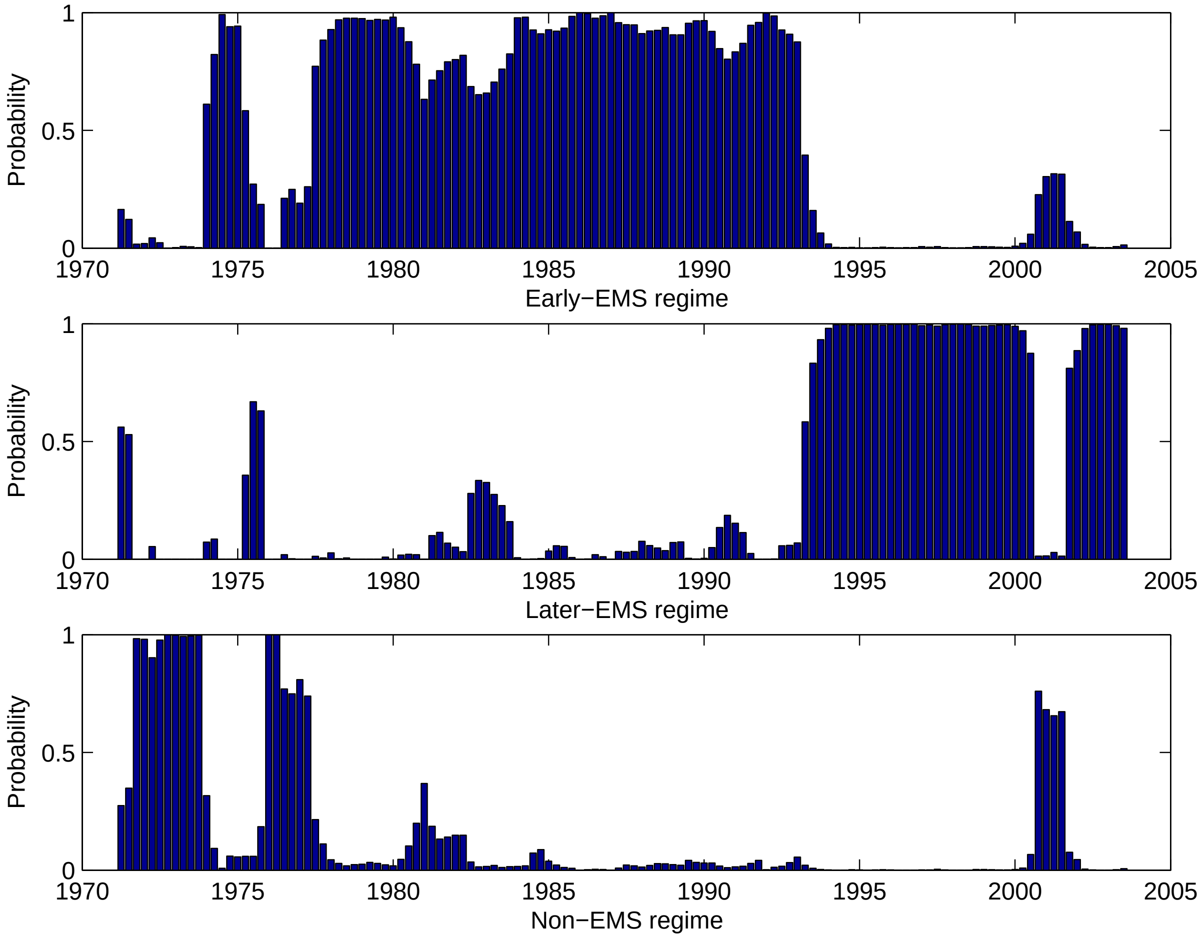

FIGURE 5. Posterior probabilities of states for the 3-state BGQ model.

the regimes we obtained using the CEE and the GLSZ schemes, with the ones estimated here, we find that the EMS regime is consistent throughout the three identification schemes and the new (early-EMS) regime concentrates during the 80's disinflation period.

Table 7 reports the residual variances for the four variables of the 3-state TV-I specification. The variances of all structural shock variances are smaller for the EMS. This result is consistent with the finding reported for the CEE and GLSZ schemes. Figure 6 displays the impulse responses to a contractionary monetary policy shock for the constant and 3-state TV-I specification. The point estimates, represented by the solid lines, say that, in response to a contractionary monetary policy shock, the interest rate rises and both output and prices decline. The point estimate indicates that the output effect is quite small, but the error bands show that there is a fat tail of the probability distribution skewed toward a large output loss for the three regimes. In other words, there is a substantial probability for a large output drop after a contractionary monetary policy shock. The point estimate also indicates that there is no price puzzle, but the error bands are so wide that there is a non-trivial probability of a price increase after a contractionary monetary policy shock. Overall, the error bands reported in Figure 6 seem unusually wide and ill-determined.[23] They are nonetheless important because they imply that the point estimates could be misleading. The methods developed in Section III.3 not only generalize the types of restrictions and applications one can use, but also provide a convenient and efficient algorithm for obtaining accurate posterior distributions to assess both quantitative and qualitative implications derived from the point estimates.

We conclude that the BGQ scheme finds that a 3-state TV-I specification fits the data better. As it was the case for both the CEE and the GLSZ schemes, there is an EMS regime after 1993 characterized by lower volatility of the structural shocks. On one hand, the point estimates for the impulse response functions show that, in response to a contractionary monetary policy shock, the interest rate rises and both output and prices decline (no price puzzle). On the other hand, the estimated error bands are so wide that none of the point estimates implications are conclusive.

IV.4. **Results for the CDFU identification.** The last identification scheme we consider is the CDFU scheme. As in the CEE and GLSZ identification schemes, we consider a SVAR with output (Y), output deflator (P), nominal short-term interest rate (R), M3 (M), and Euro/dollar nominal exchange rate (Ex). We restrict the signs of the impulse responses such that:

- An expansionary monetary policy shock implies an interest rate decrease and an increase of M3 for two periods.
- A positive shock to money demand implies an increase in interest rate and M3 for two periods.
- A positive demand shock implies an increase in output and prices for two periods.
- A positive supply shock implies an increase in output and a decrease in prices for two periods.
- A positive external shock implies an exchange rate devaluation and an increase in output for two periods.

As it was the case for the BGQ scheme, in order to implement the CDFU scheme, we start from a recursive system.[24] Then, we rotate each posterior draw of $A_0(k)$ and $A_+(k)$, such that the signs restrictions hold. As before, this rotation implies that all considered specifications are all-change specifications in the sense that the parameters in each column of $A_0(k)$, after the rotation, vary across states beyond a scaling

[23] The error bands reported by Peersman and Smets are much better behaved. Note that they have a different sample period and their bands are generated by only 100 draws. We find that this particular identification is quite fragile. For example, when the data in 2003 were taken out of our sample, the characteristics of the estimated impulse responses were completely changed.

[24] Since the CEE scheme and the CDFU scheme share variables, we start form the CEE scheme reported in section IV.1

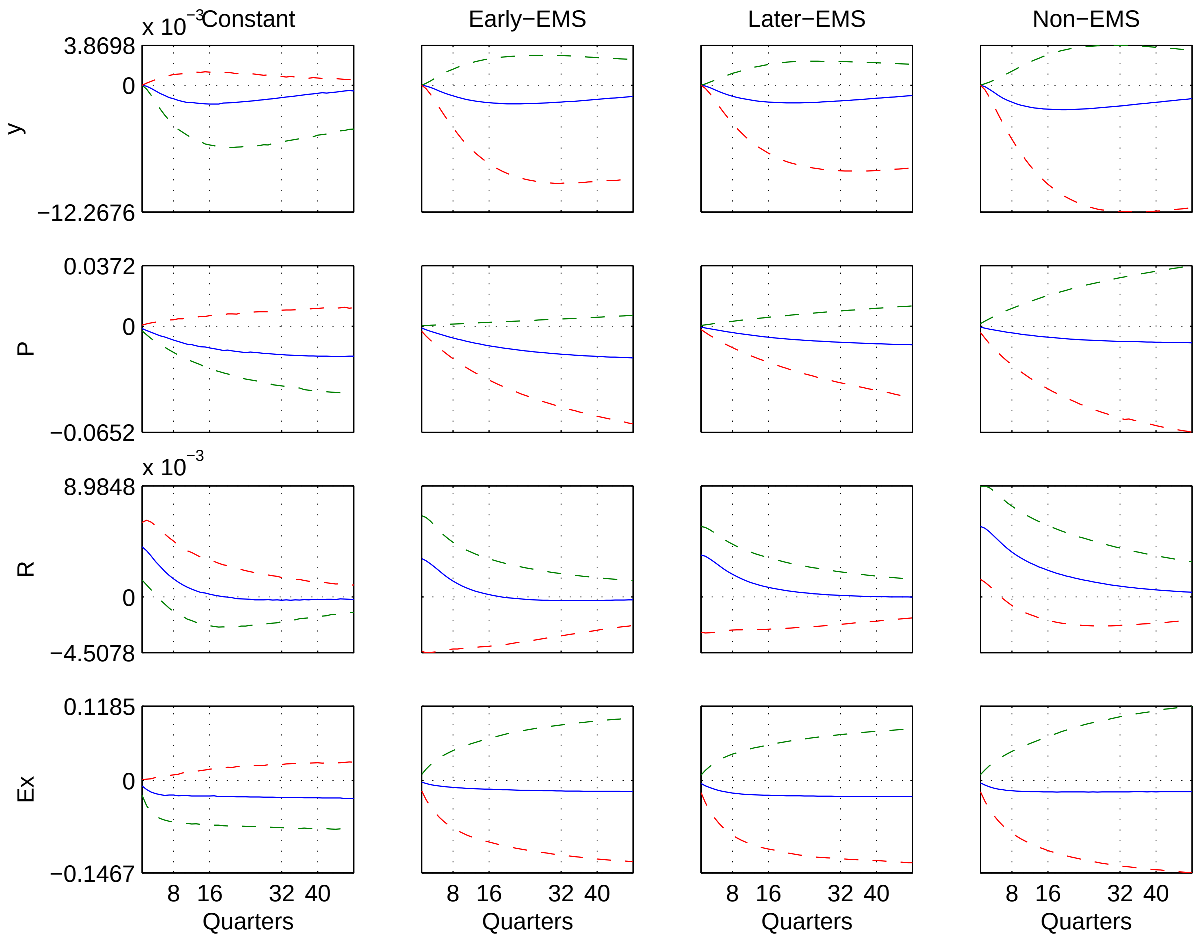

FIGURE 6. Impulse responses to a one-standard-deviation monetary policy shock under the BGQ identification. The solid line represents the posterior median estimate and the two dashed lines contain the 68 percent probability based on 500,000 MCMC draws.

factor. The differences between specifications reflect different parsimonious ways of parameterizing the time-varying coefficients, and we label these columns accordingly as “TV I”, “TV II”, etc., where TV stands for time-variation.[25]

Table 8 reports the marginal log likelihood for different specifications and number of the states for a SVAR identified using the CDFU scheme.[26] Data favors a 2-state TV-I specification.

It can be shown that the rotation does not change the variances of these residuals in any state. Therefore, the shock variances for the 2-state TV-I specification identified using the CDFU scheme are the same as reported in Table 4. Similarly, the posterior probabilities of states are the same as displayed in figure 1.

[25] We also experimented with other time-varying combinations of Case II and Case III for the different equations before applying the sign restrictions on $A_0^{-1}(k)$, and the resultant marginal likelihoods have substantially lower values than the 2-state TV-I model.

[26] Since we start form the CEE scheme reported in section IV.1, table 3 and tables 8 are identical.

Constant	2271.0			
	TV-I	TV-II	TV-III	TV-IV
2 states	2291.8	2284.5	2274.4	2278.8
3 states	2287.3	DEG	DEG	DEG
4 states	2281.9	DEG	DEG	DEG
5 states	2279.1	DEG	DEG	DEG
6 states	2272.9	DEG	DEG	DEG

TABLE 8. Marginal log likelihoods for the five specifications and different number of states under the CDFU identification scheme.

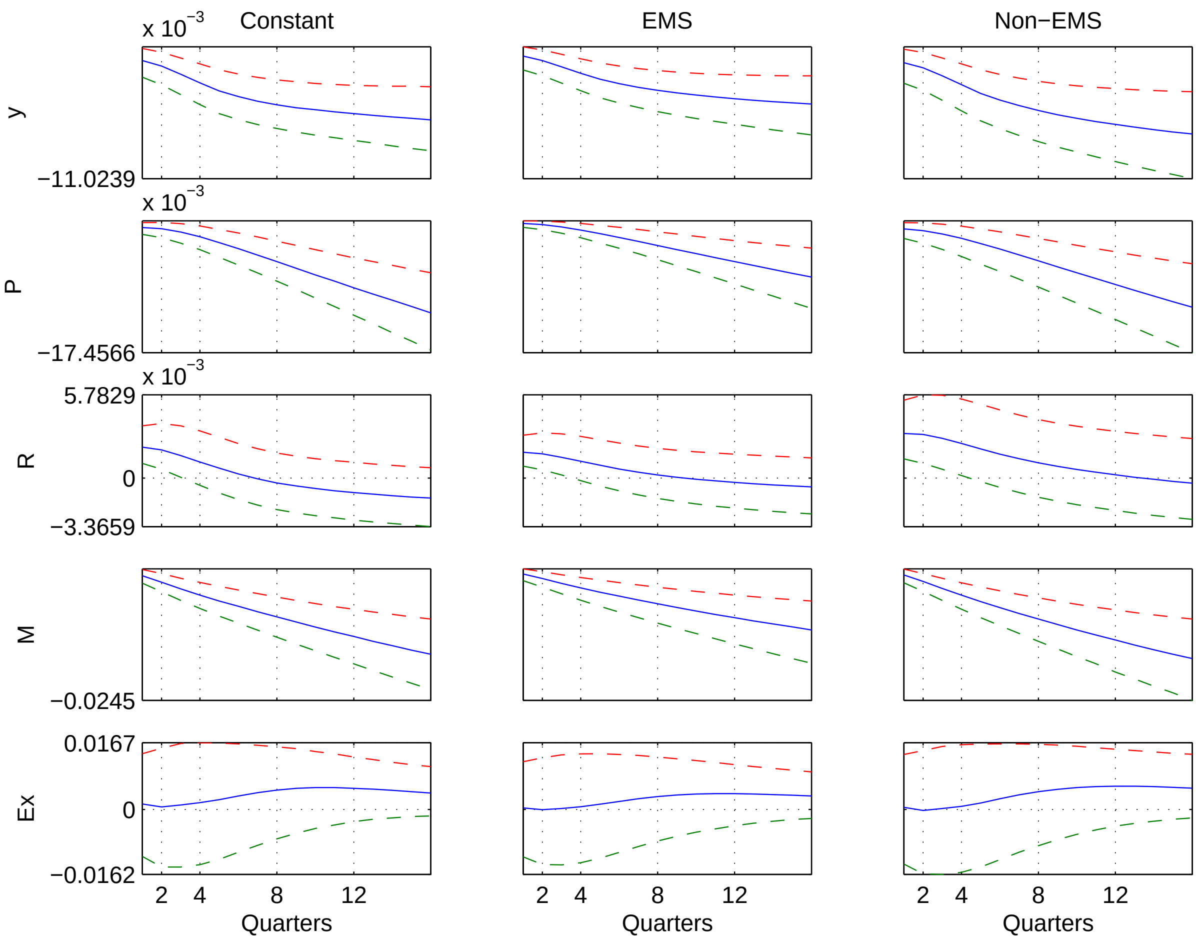

FIGURE 7. Impulse responses to a one-standard-deviation monetary policy shock under the CDFU identification. The solid line represents the posterior median estimate and the two dashed lines contain the 68 percent probability based on 500,000 MCMC draws.

Figure 7 reports the impulse responses for both the constant model and the 2-state TV-I model. Again, the responses under the EMS regime are smaller than those under the non-EMS regime and those with the constant VAR model, but the general dynamic shapes are the same. We find a significant liquidity effect and a significant output and price drop after a contractionary monetary policy shock.[27]

V. CONCLUSION

This paper has examined changes of the Euro monetary regime using the VAR methodology. We find that SVARs that consider regime regime switching fit the data much better than constant parameter SVARs. Thus, making the regime switching feature explicit is essential to obtaining a VAR benchmark against which the fit of DSGE models is compared. We robustly find that the regime associated with most years of the EMS has considerably smaller shock variances for all the variables studied in this paper. We find the liquidity and output effects of a monetary policy shock and we show that these results are robust to different identifications and different regimes. For the CEE and GLSZ schemes we find a price puzzle, while that is not the case for the BGQ and CDFU schemes.

On the technical side, we have developed a set of new efficient methods that allow us to consider different identification strategies. We have also shown how these methods can be used to accurately obtain the error bands of impulse responses associated with different identification schemes. These error bands make us question some of the point estimate dynamic responses.

APPENDIX A. THE DATA

We use quarterly data form 1970:1 to 2003:4 from the Area-wide Model (AWM) database released by the European Central Bank. See Fagan, Henry, and Mestre (forthcoming) for details. The variables are listed below, along with the variable symbols used by the AWM database.

- Y: Real GDP in millions of euros with base year 1995. (YER)
- P: Real GDP deflator with base year 1995=100. (YED)
- M: M3 measure of money stock in millions of euros.[28]
- R: The nominal short-term interest rate. (STN)
- Ex: The nominal exchange rate (Euro/$). (EEN)

APPENDIX B. THE PRIORS

In this Appendix we specify the details of the priors used in the paper. First, we describe the priors on Π, common to the three cases. Then, we described the priors on the parameters that differ across the three cases.

The prior of the transition matrix, Π, takes a Dirichlet form as suggested by Chib (). For the kth column of Π, π_k, the prior density is $\pi(\pi_k) = \pi(\pi_{1k}, \dots \pi_{nk}) \propto \pi_{1k}^{\alpha_{1k}-1} \dots \pi_{nk}^{\alpha_{nk}-1}$. We choose α_{ij} for $i, j = 1, \dots, n$ as described in Sims and Zha (2004a).

[27] Hence, we do not find any evidence of price puzzle.

[28] This variables is not included in the Area-wide Model (AWM) database. We obtain this variable from the reference series on monetary aggregates reported by the ECB

Now let us describe the priors on the parameters that differ across the three cases. Before proceeding, we introduce a few new notations. Let ς_n be a column vector of n ones. Let

$$\overline{A}_0 = [\overline{a}_{1,0}, \ldots, \overline{a}_{n,0}],$$

where $\overline{a}_{j,0}$ is a $n \times 1$ vector of the form:

$$\overline{a}_{j,0} = [\overline{a}_{1,j,0}, \ldots, \overline{a}_{n,j,0}]' \text{ for all } j.$$

Now let

$$\xi = [\xi_1, \ldots, \xi_n],$$

where ξ_j is a $h \times 1$ vector of the form:

$$\xi_j = \left[\xi_j(1), \ldots, \xi_j(h)\right]' \text{ for all } j.$$

Let

$$\phi = [\phi_1, \ldots, \phi_n],$$

where ϕ_j is a $nh \times 1$ vector of the form:

$$\phi_j = \left[\phi_j'(1), \ldots, \phi_j'(h)\right]', \text{ for all } j,$$

where $\phi_j(k)$ is a $n \times 1$ vector of the form:

$$\phi_j(k) = \left[\phi_{1,j}(k), \ldots, \phi_{n,j}(k)\right]', \text{ for all } k \text{ and all } j.$$

Define also

$$\overline{d} = [\overline{d}_1', \ldots, \overline{d}_n']',$$

where $\overline{d}_j$ is a $m \times 1$ vector of the form:

$$\overline{d}_j = [\overline{d}_{j,1}', \ldots, \overline{d}_{j,p}', \overline{c}_j]' \text{ for all } j,$$

where $\overline{d}_{j,\ell}$ is a $n \times 1$ vector of the form:

$$\overline{d}_{j,\ell} = \left[\overline{d}_{1,j,\ell}, \ldots, \overline{d}_{n,j,\ell}\right]' \text{ for all } \ell \text{ and all } j.$$

Let

$$\lambda = [\lambda_1, \ldots, \lambda_n],$$

where λ_j is a $nh \times 1$ vector of the form:

$$\lambda_j = \left[\lambda_j'(1), \ldots, \lambda_j'(h)\right]' \text{ for all } j,$$

where $\lambda_j(k)$ is a $n \times 1$ vector of the form:

$$\lambda_j(k) = \left[\lambda_{1,j}(k), \ldots, \lambda_{n,j}(k)\right]' \text{ for all } j \text{ and all } k.$$

Let

$$\mu = [\mu_1, \ldots, \mu_n],$$

where μ_j is a $h \times 1$ vector of the form:

$$\mu_j = \left[\mu_j(1), \ldots, \mu_j(h)\right]' \text{ for all } j.$$

Then we can write

$$a_{j,0} = \Phi_j \left(\xi_j \otimes \overline{a}_{j,0}\right),$$

where

$$\Phi_j = diag\left(\left\{\Phi_j(k)\right\}_{k=1}^{h}\right)$$

and

$$\Phi_j(k) = diag\left(\left\{\phi_{i,j}(k)\right\}_{i=1}^{n}\right).$$

Finally, we can also write

$$d_j = \Lambda_j \left(\xi_j \otimes \overline{d}_j\right),$$

where

$$\Lambda_j = diag\left(\left\{\Lambda_j(k)\right\}_{k=1}^{h}\right),$$

$$\Lambda_j(k) = \begin{bmatrix} I_p \otimes \Delta_j(k) & 0_{np\times 1} \\ 0_{1\times np} & \mu_j(k) \end{bmatrix},$$

and

$$\Delta_j(k) = diag\left(\left\{\lambda_{i,j}(k)\right\}_{i=1}^{n}\right).$$

We are now ready to specify the priors corresponding to Cases I–III. We begin with Case III and work backward to Case I.

B.1. **Case III.** Let $\xi_j = \varsigma_h$ and $\overline{a}_{j,0} = \varsigma_n$ for all j, then

$$a_{j,0} = \Phi_j \varsigma_{hn} = \phi_j \text{ for all } j$$

and

$$d_j = \Lambda_j \left(\varsigma_h \otimes \overline{d}_j\right) \text{ for all } j.$$

Let now the priors on the contemporaneous parameters of the model, $a_{j,0}$, be:

$$\pi\left(a_{j,0}\right) = \pi\left(\phi_j\right) = \aleph\left(0, I_h \otimes H_{j,0}\right) \text{ for all } j.$$

Since

$$\phi_j = \left(I_h \otimes U_j\right) b_j, \text{ for all } j,$$

that imply priors on b_j of the form:

$$\pi\left(b_j\right) = \aleph\left(0, \overline{H}_{j,0}\right),$$

where

$$\overline{H}_{j,0} = \left(U_j' \left(I_h \otimes H_{j,0}^{-1}\right) U_j\right)^{-1}.$$

Let now the priors on the lagged and constant parameters of the model, $\overline{d}_j$, be:

$$\pi\left(\overline{d}_j\right) = \aleph\left(0, H_{j,+}\right) \text{ for all } j,$$

$$\pi\left(\lambda_j\right) = \aleph\left(0, \left(I_h \otimes I_n\right)\sigma_\lambda^2\right) \text{ for all } j,$$

and

$$\pi\left(\mu_j\right) = \aleph\left(0, I_h \otimes \sigma_\mu^2\right) \text{ for all } j.$$

B.2. **Case II.** Let $\phi_j = \varsigma_n$, $\lambda_j = \varsigma_{nh}$, and $\mu_j = 1$ for all j, then

$$a_{j,0} = \xi_j \otimes \overline{a}_{j,0} \text{ for all } j$$

and

$$d_j = \xi_j \otimes \overline{d}_j \text{ for all } j.$$

Let now the priors on the contemporaneous parameters of the model, $\overline{a}_{j,0}$, be:

$$\pi\left(\overline{a}_{j,0}\right) = \aleph\left(0, H_{j,0}\right) \text{ for all } j.$$

Since

$$\xi_j \otimes \overline{a}_{j,0} = \left(I_h \otimes U_j\right) b_j, \text{ for all } j,$$

that imply priors on b_j of the form:

$$\pi\left(b_j|\xi_j\right) = \aleph\left(0, \widetilde{H}_{j,0}\right),$$

where

$$\widetilde{H}_{j,0} = \widetilde{\Upsilon}_{j,h} \otimes \left(U_j' H_{j,0}^{-1} U_j\right)^{-1},$$

and

$$\widetilde{\Upsilon}_{j,h} = \begin{bmatrix} \xi_j(1)^2 & \xi_j(1)\xi_j(2) & \dots & \xi_j(1)\xi_j(h) \\ \xi_j(2)\xi_j(1) & \xi_j(2)^2 & \dots & \xi_j(2)\xi_j(h) \\ \vdots & \vdots & \ddots & \vdots \\ \xi_j(h)\xi_j(1) & \xi_j(h)\xi_j(2) & \dots & \xi_j(h)^2 \end{bmatrix}.$$

Let now the priors on the lagged and constant parameters of the model, $\overline{d}_j$, be:

$$\pi\left(\overline{d}_j\right) = \aleph\left(0, H_{j,+}\right) \text{ for all } j.$$

Finally, let priors on $\xi_j(k)$ be defined over $\zeta_j(k) = \xi_j^2(k)$ as:

$$\pi\left(\zeta_j(k)\right) = \Gamma\left(\alpha_\zeta, \beta_\zeta\right) \text{ for all } k \text{ and } j.$$

B.3. **Case I.** Let $\xi_j = \varsigma_h$, $\phi_j = \varsigma_n$, $\lambda_j = \varsigma_{nh}$, and $\mu_j = 1$ for all j, then

$$a_{j,0} = \varsigma_h \otimes \overline{a}_{j,0} \text{ for all } j$$

and

$$d_j = \varsigma_h \otimes \overline{d}_j \text{ for all } j.$$

Let now the priors on the contemporaneous parameters of the model, $\overline{a}_{j,0}$, be:

$$\pi\left(\overline{a}_{j,0}\right) = \aleph\left(0, H_{j,0}\right) \text{ for all } j.$$

Since

$$\varsigma_h \otimes \overline{a}_{j,0} = \left(I_h \otimes U_j\right) b_j, \text{ for all } j,$$

that imply priors on b_j of the form:

$$\pi\left(b_j\right) = \aleph\left(0, \widehat{H}_{j,0}\right),$$

where

$$\widehat{H}_{j,0} = \widehat{\Upsilon}_{j,h} \otimes \left(U_j' H_{j,0}^{-1} U_j\right)^{-1},$$

and

$$\widehat{\Upsilon}_{j,h} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix}.$$

Let now the priors on the lagged and constant parameters of the model, $\overline{d}_j$, be:

$$\pi\left(\overline{d}_j\right) = \aleph\left(0, H_{j,+}\right) \text{ for all } j.$$

Appendix C. Proof of Theorem 4

Proof. Let Y_t and s_t. Using compact notation (2), model (1) implies the following conditional likelihood function:

$$\pi\left(y_t | Y^{t-1}, s_t, b, d\right) \propto \det\left|A_0\left(s_t\right)\right| \exp\left[-\frac{1}{2}\left(y_t' A_0\left(s_t\right) - x_t' A_+\left(s_t\right)\right)\left(y_t' A_0\left(s_t\right) - x_t' A_+\left(s_t\right)\right)'\right]. \tag{A1}$$

Now, let

$$H_t = x_t' x_t,$$

$$P_t = H_t^{-1} x_t' y_t,$$

and

$$S_t = y_t' y_t - P_t' H_t P_t.$$

Note now that we rewrite the right hand side of (A1) as:

$$\det|A_0(s_t)|\exp\left[-\frac{1}{2}y_t'A_0(s_t)A_0'(s_t)y_t-2y_t'A_0(s_t)A_+'(s_t)x_t+x_t'A_+(s_t)A_+'(s_t)x_t\right]=$$

$$\det|A_0|\exp\left[-\frac{1}{2}\sum_{j=1}^{n}\left(y_t'a_{j,0}(s_t)\right)\left(a_{j,0}'(s_t)y_t\right)-2\left(y_t'a_{j,0}(s_t)\right)\left(a_{j,+}'(s_t)x_t\right)+\left(x_t'a_{j,+}(s_t)\right)\left(a_{j,+}'(s_t)x_t\right)\right]=$$

$$\det|A_0(s_t)|\exp\left[-\frac{1}{2}\sum_{j=1}^{n}\left(a_{j,0}'(s_t)y_t\right)\left(y_t'a_{j,0}(s_t)\right)-2\left(a_{j,+}(s_t)'x_t\right)\left(y_t'a_{j,0}(s_t)\right)+\left(a_{j,+}'(s_t)x_t\right)\left(x_t'a_{j,+}(s_t)\right)\right]=$$

$$\det|A_0(s_t)|\exp\left[-\frac{1}{2}\sum_{j=1}^{n}a_{j,0}'(s_t)y_t'y_ta_{j,0}(s_t)-2a_{j,+}(s_t)'x_t'y_ta_{j,0}(s_t)+a_{j,0}'(s_t)x_t'x_ta_{j,0}(s_t)\right].$$

Then

$$\begin{aligned}\pi\left(y_t|Y^{t-1},s_t,\theta_i\right) &\propto \det|A_0(s_t)|\exp\left[-\frac{1}{2}\sum_{j=1}^{n}a_{j,0}'(s_t)S_ta_{j,0}(s_t)\right]\\ &\exp\left[-\frac{1}{2}\sum_{j=1}^{n}\left(a_{j,+}(s_t)-P_ta_{j,0}(s_t)\right)'H_t\left(a_{j,+}(s_t)-P_ta_{j,0}(s_t)\right)\right]\end{aligned}$$

But equation (8) implies that $a_{j,+}(s_t)=d_j(s_t)+\overline{S}a_{j,0}(s_t)$, then

$$\begin{aligned}\pi\left(y_t|Y^{t-1},s_t,\theta_i\right) &\propto \det|A_0(s_t)|\exp\left[-\frac{1}{2}\sum_{j=1}^{n}b_j'(s_t)U_j'S_tU_jb_j(s_t)\right]\\ &\exp\left[-\frac{1}{2}\sum_{j=1}^{n}\left(d_j(s_t)+\left(\overline{S}-P_t\right)U_jb_j(s_t)\right)'H_t\left(d_j(s_t)+\left(\overline{S}-P_t\right)U_jb_j(s_t)\right)\right]\end{aligned}$$

Finally, note that relationship (7) implies:

$$\det|A_0(s_t)|=\det\left|[a_{1,0}(s_t)\dots a_{n,0}(s_t)]\right|=\det|\ [U_1b_1(s_t)\dots U_nb_n(s_t)]|.$$

□

APPENDIX D. PROOF OF THEOREMS 2, 3, AND 7

D.1. **Definition of Exact Identification in SVARs.** We follow Rothenberg (1971) in defining global and local identification. In the general discussion that follows, we require that the reduced form system be globally identified. For the systems under consideration here, some non-colinearity condition must be imposed on the variable z_t. For instance, one could assume that for $t\geq m$ the matrix $[y_1\cdots y_t]$ is of full rank with positive probability. We are not concerned with the particulars of distributions of the z_t, so we simply assume that the reduced form system is globally identified.

There are two conflicting conditions in the notion of exact identification. First, there should be enough restrictions so that the system is globally identified. On the other hand, there should be few enough restrictions so that the reduced form specification remains unconstrained. This is formalized in the following definition.

Definition 10. A SVAR with linear restrictions given by **??** is just identified if and only if for every reduced form parameter (B,Σ), except perhaps on a set of measure zero, there exists exactly 2^n structural parameters (A_0,A_+) with $B = A_+A_0^{-1}$ and $\Sigma = \left(A_0A_0'\right)^{-1}$ that satisfies the restrictions.

Because the reduced form system is assumed to be globally identified, if two structural parameters (A_0,A_+) and $\left(\tilde{A}_0,\tilde{A}_+\right)$ are observationally equivalent then they must map to the same reduced form parameter (B,Σ). Thus the definition implies that the restricted structural system is globally identified, except on a set of measure zero[29]. On the other hand, since almost all reduced form parameters correspond to some restricted structural parameters, the reduced form parameters are not constrained by the restrictions on the structural parameters. For SVARs, allowing for exceptional behavior on a set of measure zero is important. If this exception were not allowed, then the only linear restrictions that would result in exact identification would be those that are recursive.

Finally, the issue of normalization must be addressed. For linear restrictions of the kind described in this paper, changing the sign of a column of A_0 and A_+ has no effect on whether or not the restrictions are satisfied.

Lemma 11. For $1 \le i \le k \le n$, let V_i be a subspace of $\mathbb{R}^n$. If for every invertible $n\times n$ matrix A, there exists an orthonormal set $\{v_1,\cdots,v_k\}$ in $\mathbb{R}^n$ such that $v_i \in AV_i$, then there exists a j with $1 \le j \le k$ and $\dim(V_i) \ge k$.

Proof. Given a subspace W of $\mathbb{R}^n$ and $\varepsilon \in \mathbb{R}$, let $A_{W,\varepsilon}$ be the linear transformation that fixes W and maps each u in the perpendicular component of W to εu. If $\dim(V_i) < k$ for $1 \le i \le k$, then using the following three statements a W and $\varepsilon > 0$ can be constructed such that $A_{W,\varepsilon}$ violates the conditions of the lemma. So it suffices to prove the following.

(1) If $\dim(V_i) < k$, then there exists a subspace subspace U of $\mathbb{R}^n$ of dimension $n-k+1$ such that $U \cap V_i = \{0\}$.

(2) Let W be a $k-1$ dimensional subspace of $\mathbb{R}^n$. There exists a $\delta > 0$ such that there cannot be k orthonormal vectors in the set

$$S_{W,\delta} = \{w+u \in \mathbb{R}^n \,|\, w \in W \text{ and } \|u\| < \delta\}.$$

(3) Let U and V be subspaces of $\mathbb{R}^n$ such that $U \cap V = \{0\}$ and let W be the perpendicular complement of U. For every $\delta > 0$ there exists a $\gamma > 0$ such that for all $\varepsilon < \gamma$ if $v \in A_{W,\varepsilon}V$ and $\|v\| = 1$, then $v \in S_{W,\delta}$.

(1) If $\dim(V_i) < k \le n$, then each V_i is of measure zero in $\mathbb{R}^n$, as will be the union of the V_i. So there exists a u_1 that is not contained in any V_i. If $k = n$, then the one dimensional subspace generated

[29]If S_1 is the set of all (B,Σ) such that exists exactly one structural parameter (A_0,A_+) with $B = A_+A_0^{-1}$ and $\Sigma = (A_0A_0')^{-1}$ that satisfies the restrictions, then the complement of S_1 is assumed to be of measure zero. Let S_2 be the set of all structural parameters (A_0,A_+) such that the reduced form parameters $\left(A_+A_0^{-1},(A_0A_0')^{-1}\right)$ is in the complement of S_1. It can be shown that the measure of S_2 is zero.

by u_1 is the required subspace. If $k < n$, then let $\tilde{V}_i$ be the subspace generated by V_i and u_1. Since $\dim\left(\tilde{V}_i\right) < k+1 \leq n$, by the same measure argument as before, there will exist a u_2 that is not contained in the union of the $\tilde{V}_i$. If $k = n-2$, then the two dimensional subspace generated by u_1 and u_2 is the required subspace. This argument can be continued until a basis $u_1, \cdots, u_{n-k+1}$ has been constructed for the required subspace.

(2) Suppose there were $v_1, \cdots, v_k$ in $S_{W,\delta}$ that were orthonormal. Since the v are in $S_{W,\delta}$, write $v_i = w_i + u_i$ where $w_i \in W$ and $\|u_i\| < \delta$. Let X be the $n \times k$ matrix $\begin{bmatrix} w_1 & \cdots & w_k \end{bmatrix}$ and let Y be the $n \times k$ matrix $\begin{bmatrix} v_1 & \cdots & v_k \end{bmatrix}$. Because the w are in a $k-1$ dimensional space, the matrix $X'X$ is singular and because the v are orthonormal, $Y'Y$ is the $k \times k$ identity matrix. Because δ can be chosen arbitrarily small, $X'X$ can be made to be arbitrarily close to the identity matrix, which is a contradiction.

(3) If this were not true, then there would exist a $\delta > 0$ and sequence of v_i and ε_i such that the ε_i tend to zero and $v_i \in A_{W,\varepsilon_i}V$, $\|v_i\| = 1$, and $v_i \notin S_{W,\delta}$. Because U and W are perpendicular components, v_i can be uniquely written as $v_i = u_i + w_i$ where $u_i \in U$ and $w_i \in W$. Since $\|v_i\| = 1$ and u_i and w_i are orthogonal, $\|w_i\| \leq 1$. Since $v_i \notin S_{W,\delta}$, $\|u_i\| > \delta$. Since $v_i \in A_{W,\varepsilon_i}V$, $\frac{1}{\varepsilon_i}u_i + w_i \in V$. Dividing by the norm, we see that

$$\frac{u_i + \varepsilon_i w_i}{\sqrt{\|u_i\|^2 + \varepsilon_i^2 \|w_i\|^2}} \in V$$

Since this is a bounded sequence, some subsequence must converge. Since $\|u_i\|$ is bounded away from zero, $\|w_i\|$ is bounded above, and V is closed, the convergent subsequence must converge to a non-zero element of $U \cap V$, which is a contradiction. $\square$

Theorem 12. For $1 \leq i \leq k \leq n$, let V_i be a subspace of $\mathbb{R}^n$. The following statements are equivalent.
(1) For every invertible $n \times n$ matrix A there exists an orthonormal set $\{v_1, \cdots, v_k\}$ such that $v_i \in AV_i$.
(2) There exists a permutation σ of $\{1, \cdots, k\}$ such that $\dim\left(V_{\sigma(i)}\right) \geq i$.

Proof. $(1) \Longrightarrow (2)$. Proceed by finite induction on k. When $k = 1$, the result is trivially true. Now suppose that $(1) \Longrightarrow (2)$ for some $k < n$. Let $(V_1, \cdots, V_{k+1})$ be subspaces such that (1) holds. By Lemma (11), we know that there exists a j with $1 \leq j \leq k+1$ and $\dim\left(V_j\right) \geq k+1$. Without loss of generality, assume that $j = k+1$. Since (1) holds for $(V_1, \cdots, V_{k+1})$, (1) will also hold for $(V_1, \cdots, V_k)$. This implies that there exists a permutation σ of $\{1, \cdots, k\}$ such that $\dim\left(V_{\sigma(i)}\right) \geq i$. This combined with the fact that $\dim(V_{k+1}) \geq k+1$ shows that (2) holds.

$(2) \Longrightarrow (1)$. Assume that (2) holds and let A be any invertible $n \times n$ matrix. Since $\dim\left(AV_{\sigma(1)}\right) \geq 1$, there exists a vector $v_{\sigma(1)} \in AV_{\sigma(1)}$ of unit length. Now suppose that an orthonormal set $\left\{v_{\sigma(1)}, \cdots, v_{\sigma(j)}\right\}$ has been chosen so that $v_{\sigma(i)} \in AV_{\sigma(i)}$. Let U be the $n-j$ dimensional subspace of $\mathbb{R}^n$ consisting of vectors orthogonal to $\left\{v_{\sigma(1)}, \cdots, v_{\sigma(j)}\right\}$. Since $\dim\left(AV_{\sigma(j+1)}\right) \geq j+1$, the intersection of U and $AV_{\sigma(j+1)}$ contains a non-zero vector. Let $v_{\sigma(j+1)}$ be any element of $U \cap AV_{\sigma(j+1)}$ of unit length. Then $\left\{v_{\sigma(1)}, \cdots, v_{\sigma(j+1)}\right\}$ is a set of orthonormal vectors with $v_{\sigma(i)} \in AV_{\sigma(i)}$. So (1) holds. $\square$

Corollary 13. For $1 \leq i \leq n \leq m$, let e_i be the i^{th} column of the $n \times n$ identity matrix and let Q_i be a matrix with m columns. Let X be a full rank $m \times n$ matrix. The following are equivalent.

(1′) For every investable $n \times n$ matrix A there exists a $n \times n$ orthogonal matrix P such that $Q_i XAPe_i = 0$.
(2′) There exist a permutation σ of $\{1, \cdots, n\}$ such that $rank\left(Q_{\sigma(i)}\right) \leq n - i$.

Proof. The corollary a simple restatement of Theorem 12 when $k = n$. If

$$V_i = \{v \in \mathbb{R}^n \,|\, Q_i X v = 0\},$$

then, $\dim(V_i) = n - rank(Q_i X) \geq n - rank(Q_i) \geq i$. So (2′) is equivalent to (2) of Theorem 12. Since $Q_i XAPe_i = 0$ if and only if $Pe_i \in A^{-1}V_i$, (1′) is equivalent to (1). □

Theorem **??** follows easily from the corollary by noting that since $X(A(L))$ is a surjection, if the system is not over identified then (1) will hold for all full rank $m \times n$ matrices.

Proof. It can be shown that $X_{(A_0P, A_+P)} = X_{(A_0, A_+)}P$. So

$$\begin{aligned} P_1 X_{(A_0P, A_+P)} P_2 &= P_1 X_{(A_0, A_+)} P P_2 \\ &= T_L P_3 P_3' P_2' P_2 = T_L \end{aligned}$$

which implies that the rotated parameters (A_0P, A_+P) satisfy the restrictions. □

REFERENCES

ANGELONI, I., AND M. EHRMANN (2003): "Monetary policy transmission in the euro area: any changes after EMU?," *Economic Policy*, 37, 470–492.

BERNANKE, B., AND I. MIHOV (1998): "Measuring Monetary Policy," *Quarterly Journal of Economics*, CXIII.

BLANCHARD, O. J., AND D. QUAH (1993): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *aer*, 83(3), 655–673.

BLANCHARD, O. J., AND M. W. WATSON (1986): *Are Business Cycles All Alike?* NBER and Chicago Press, Chicago.

CANOVA, F., AND G. D. NICOLO (2002): "Monetary disturbances matter for business fluctuations in the G-7," *jme*, 49(6), 1131–1159.

CECCHETTI, S. (2001): *Legal Structure, Financial Structure and the Monetary Policy Transmission Mechanism*. Palgrave.

CHIB, S. (1996): "Calculating Posterior Distributions and Model Estimates in Markov Mixture Models," *Journal of Econometrics*, 75, 79–97.

CICCARELLI, M., AND A. REBUCCI (2003): "Measuring contagion with a Bayesian, time-varying coefficient model," ECB Working Paper No. 263.

DE BONDT, G. (2002): "Retail bank interest rate pass-through: new evidence at the euro area level," ECB Working Paper No. 136.

FAUST, J. (1998): "The Robustness of Identified VAR Conclusions about Money," *Carnegie-Rochester Conference Series on Public Policy*, 48, 207–244.

GABRIEL FAGAN, J. H., AND R. MESTRE (forthcoming): "An area-wide model (AWM) for the euro area," *Economic Modelling*.

GALÍ, J. (1992): "How Well Does the IS-LM Model Fit Postwar U.S. Data?," *qje*, 107(2), 709–738.

GALÍ, J. (1999): "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *The American Economic Review*, 89(1), 249–271.

GELFAND, A. E., AND D. K. DEY (1994): "Bayesian Model Choice: Asymptotics and Exact Calculations," *Journal of the Royal Statistical Society (Series B)*, 56, 501–514.

GEWEKE, J. (1999): "Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication," *Econometric Reviews*, 18(1), 1–73.

GORDON, D. B., AND E. M. LEEPER (1994): "The Dynamic Impacts of Monetary Policy: An Exercise in Tentative Identification," *jpe*, 102(6), 1228–1247.

HAMILTON, J. D. (1989): "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57(2), 357–384.

HAMILTON, J. D., D. F. WAGGONER, AND T. ZHA (2003): "Normalization in Econometrics," Manuscript, University of California (San Diego) and Federal Reserve Bank of Atlanta.

I. ANGELONI, A. K., AND B. MOJON (2003): *Monetary Policy Transmission in the Euro Area*. Cambridge University Press.

JESÚS FERNÁNDEZ-VILLAVERDE, J. F. R.-R., AND T. J. SARGENT (2005): "A,B,C's (and D)'s for Understanding VARs," mimeo.

KIELER, M., AND T. SAARENHEIMO (1998): "Differences in monetary policy transmission? A case not closed," European Commission Economic Papers No. 132.

L. GUISO, A. KASHYAP, F. P., AND D. TERLIZZESE (1999): "Will a Common European Monetary Policy have Asymmetric Effects?," *Federal Reserve Bank of Chicago Economic Perspectives*, 23(4), 56–75.

LAWRENCE J. CHRISTIANO, M. E., AND C. EVANS (1996): "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds," *The Review of Economics and Statistics*, 78(1), 16–34.

MIHOV, I. (2001): "Monetary policy implementation and the transmission in the European Monetary Union," *Economic Policy*, 33, 369–402.

NEWEY, W. K., AND K. K. WEST (1987): "A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

PEERSMAN, G. (2004): "The transmission of monetary policy in the Euro area: Are the effects different across countries?," *Oxford Bulletin of Economics and Statistics*, 66(3), 285–308.

PEERSMAN, G., AND F. SMETS (2003): *The monetary transmission mechanism in the euro area: more evidence from VAR analysis*. Cambridge University Press, Cambridge.

ROTHENBERG, T. J. (1971): "Identification in Parametric Models," *Econometrica*, 39(3), 577–591.

SIMS, C. A. (1999): "Drift and Breaks in Monetary Policy," Manuscript, Princeton University.

SIMS, C. A., AND T. ZHA (1999): "Error Bands for Impulse Responses," *Econometrica*, 67(5), 1113–1155.

——— (2004a): "MCMC Method for Markov Mixture Simultaneous-Equation Models: A Note," Federal Reserve Bank of Atlanta Working Paper 2004-15.

——— (2004b): "Were There Regime Switches in U.S. Monetary Policy?," Federal Reserve Bank of Atlanta Working Paper 2004-14.

SMETS, F., AND R. WOUTERS (2003): "An estimated stochastic dynamic general equilibrium model of the euro area," *Journal of European Economic Association*, 1(5), 1123–1175.

UHLIG, H. (forthcoming): "What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure," *jme*.

WAGGONER, D. F., AND T. ZHA (2003a): "A Gibbs Sampler for Structural Vector Autoregressions," *Journal of Economic Dynamics and Control*, 28(2), 349–366.

——— (2003b): "Likelihood Preserving Normalization in Multiple Equation Models," *Journal of Econometrics*, 114(2), 329–347.