On the Time Variations of US Monetary Policy: Who is right?

Fabio Canova
Universitat Pompeu Fabra, IGIER and CEPR

Luca Gambetti *
Universitat Pompeu Fabra

First Draft: June 2003
This Draft: April 2004

Abstract

This paper investigates whether monetary policy accounts for the changes in the output and inflation process observed in the US over the last 25 years. It estimates a structural Bayesian TVC-VAR with MCMC methods where sign restrictions are used to identify monetary policy shocks and analyzes the transmission of two types of disturbances: those to the non-systematic and those to the systematic component of monetary policy. Dynamic analysis is implemented using the difference between two conditional expectations of future realizations of the vector of time series differing for a shock in the conditioning sets. We find that there are structural variations in both the coefficients of the model and the variance of the structural shocks but that the transmission of monetary policy disturbances has hardly changed over the last 25 years. Changes in the systematic component of policy have also negligible changes in the dynamic of the system. Changes in inflation persistence have also little connection with the dynamics of monetary policy. Results are robust to a number of alterations to the model.

*We would like to thank the participants of seminars at UPF and the Federal Reserve Bank of Atlanta for comments and suggestions

1 Introduction

There is considerable evidence suggesting that the US economy has changed fundamentally over the last couple of decades. In particular, several authors have noted a marked decline in the variance of real activity and inflation since the early 1980s (see e.g. Blanchard and Simon (2000), McConnelly and Perez Quiroz (2001) and Stock and Watson (2003)). What are the reasons for such a decline? A stream of literature attributes these changes to changes in the mechanisms through which exogenous shocks spread through the economy and propagate over time. Since the transmission mechanism depends on the structure of the economy, the main implication of this point of view is that the underlying characteristics of the economy have evolved over time. Several factors which could be behind these changes, for instance, changes in the behavior of consumers and firms or changes in the policy rules adopted by the government. Particular attention has been paid to monetary policy. Several studies (see e.g. Boivin and Giannoni, (2002a), (2002b), Clarida, Gali and Gertler (2000), Cogley and Sargent (2001) (2003)) have argued that monetary policy was "loose" in the 1970s but had become more aggressive against inflationary pressures since the early 1980s and see in this change the reason for the observed reduction of inflation and output volatility. This view, however, is far from unanimous. For example, Bernanke and Mihov (1998), Orphanides (2001), Leeper and Zha (2003) find little evidence of significant changes in the policy rule used over the last 25 years and in the propagation of monetary shocks to the economy while Hanson (2001) claims that changes in the dynamics of output and inflation are mainly due to changes occurred in the rest of the economy, thus limiting the importance of monetary policy in accounting for these difference. Finally, Sims (2001) and Sims and Zha (2001) find that changes in the variance of exogenous shocks (and, in particular, of monetary policy innovations) account for a substantial portion of the observed changes in output and inflation variance.

This debate is not new. Rational expectations econometricians (e.g. Sargent (1984)) often argued that policy changes involving regime switches dramatically altered the dynamics of private behavior and therefore of the macroeconomic variables and search for historical episodes supporting the arguments (see e.g. Sargent (1999)). VAR econometricians, on the other hand, often denied the empirical relevance of the argument suggesting that the systematic portion of monetary policy has never really switched, at least in the US, and that policy changes were due different draws of the non-systematic part (Sims (1982)). The debate now has been cast into the dual framework of "bad policy" (failure to respond to inflationary pressure) vs. "bad luck" (policy shocks are drawn from a distribution whose second moments vary over time) and found new vigor thanks to the development of tools which can explicitly allow to examine time variations in the structure of the system and in the variance of the exogenous processes and the timing of the changes. Overall, and despite the recent contributions, the question of why the US economy has changed and the role of monetary policy shocks in shaping these changes is still open.

This paper enters the debate and provide new evidence on these issues examining time variations in the US economy from a structural point of view, analyzing the impact of time variations on the propagation mechanism of policy shocks and of shocks to the systematic component of policy (changes in policy preferences) and assessing the robustness of the conclusions to a number of changes in the model.

The reduced form model we employ is a time varying coefficients VAR model (TVC-VAR) similar to the one employed by Cogley and Sargent (2001) and (2003). The model has a useful state space representation model where the vector of endogenous variables, which includes output, inflation, the federal funds rate and nominal balances has a VAR representation and the time varying coefficients, treated as hidden states variables, evolve according to a nonlinear transition equation which puts zero probability to paths whose associated VAR roots are explosive. Cogley and Sargent (2003) add to this structure a stochastic volatility model for the reduced form innovation. We do not follow their approach here: an heteroskedatic representation for the forecast error of the reduced form model is directly produced via time-variations in the coefficients. We choose this approach for two main reasons. First, as shown in Canova (1993) such a framework can account for a variety of non-normalities and second moment non-linearities which the literature has found to be empirically important. Second, the framework retains a conditional linear structure which facilitate the derivation of posterior estimates of the structural coefficients of the model substantially cuts down computational costs of the model.

Our estimation methodology which is based on MCMC methods is similar to the one of Cogley and Sargent but our analysis differ from their in several respects. First, we use a structural version of the model to conduct inference. Second, we explicitly study the variations in the propagation mechanism to monetary policy disturbances as opposed to relating the timing of reduced form changes. Third, we analyze both short run and long run features of the model and quantify the importance of monetary policy to the observed changes.

The structural setup we use is particularly convenient because it allows us to directly address one of the main questions of the debate: are changes due to changes in the systematic component of policy, to changes in the propagation mechanism of policy or to changes in the variance of the shocks hitting the economy. In fact, we are able to distinguish changes the structure in the economy from changes in the variance of the shocks and, more importantly, we can separately investigate the effect of shocks to the systematic part of monetary policy (which affect the economy because coefficients evolve over time) and to the unsystematic part of monetary policy (which may affect the dynamics even when the systematic part of policy is unchanged) and to measure their importance in accounting for the observed changes in the output and inflation process.

Structural time variations in the TVC-VAR model are obtained identifying structural disturbances to the VAR via sign restrictions as in Faust (1998), Uhlig (1999) and Canova and De Nicolo' (2002). While in this paper we concentrate on monetary policy disturbances and therefore refrain to give a name to the other disturbances, the methodology can be used to jointly identify multiple sources of disturbances so long as theory produces robust sign implications for the dynamics of the variables of the system in response to these shocks. It is important to stress that we chose a sign restriction approach to identification, as opposed to a more standard Choleski or short run based approach, for two reasons. First, the zero restrictions used to identify these systems are not often present in the models economists currently use. Second, a Choleski decomposition imposes strongly restriction on the structure of time variations in the structural model. In particular, it forces changes to occur only in the lagged structural coefficient (no changes are allowed in the contemporaneous). In the robustness exercises we conduct, we however show that the our conclusions about time variations in the model do not depend on the identification procedure we use.

Because the presence of time varying coefficients induces important non-linearities in the dynamics of the model, analyses based on standard impulse responses and variance decompositions are inappropriate. In particular, since at each point in time the coefficient vector is hit by a specific shock, assuming that between $T+1$ and $T+K$ no shocks other than the monetary policy disturbance hits the system does not make much sense and can give misleading conclusions. To study the evolution the economy responses to structural shocks we therefore employ a different concept of impulse responses, which shares similarities to the one used in Koop, Pesaran and Potter (1996), Koop (1996) and Gallant, Rossi and Tauchen (1996), and it is more appropriate in a setup like ours where coefficients may vary over time. In particular, impulse response functions are defined as the difference between two conditional expectations which differ in their conditioning sets.

Our results reconcile, in part, existing views and offer important new evidence on the issue. In particular, we find that there are some significant time variations in some of the coefficients of the reduced form model, in particular the inflation and the money equations; that time variations also appear in the structural representation of the system; that the largest time variations occur up to 1986 and that the variance of the structural shocks (and of the structural forecast errors of the model) has changed (declined) dramatically over time.

However, we find little correspondence in the timing of the variations observed in the structural coefficients of the inflation and output equations and those in the monetary policy equation, little correspondence which also carry out to measures of inflation persistence. We also find that the propagation of monetary disturbances has not much changed over the sample under consideration, and that changes in the systematic component of monetary policy hardly account for the observed changes in inflation and output. Taken together, these results indicate that the contribution of monetary policy to the observed structural changes is probably minor, that "bad luck" is probably responsible for a large portion of difficulties experienced by the US economy in the 1970s and that, as in Hanson (2001), changes in the transmission of other structural shocks may explain the pattern we found.

We argue that difference between our results and those existing in the literature primarily to three factors (i) the use of reduced form vs. structural models; (ii) the use of subsamples with fixed coefficients as opposed to explicit TVC models; (iii) the inappropriate use of subsample evidence to infer the nature of the changes. The use of short vs. long run statistics; of univariate vs. multivariate methods and of different model specifications may also partially be responsible for the different outcomes present in the literature.

The rest of the paper is organized as follows. Section 2 presents the reduced form model, describes our identification scheme and the computational approach used to obtain posterior distributions of the structural coefficients of the model Section 3 defines impulse response functions which are valid in our TVC model and describes how to compute dynamics to shocks in the non-systematic and the systematic component of the model. Section 4 presents the results and Section 5 concludes. A series of appendices describes the technical details involved in the computation of posterior distributions and of impulse responses.

2 The Model

Let y_t be a 4×1 vector of time series including (linearly) detrended real output, inflation, the federal funds rate and nominal balances with the representation

$$y_t = A_{0,t} + A_{1,t}y_{t-1} + A_{2,t}y_{t-2} + ... + A_{p,t}y_{t-p} + \varepsilon_t \tag{1}$$

where $A_{0,t}$ is a $n \times 1$ vector, $A_{i,t}$, for $i = 1, ..., p$ are $n \times n$ matrices of coefficients and ε_t is a $n \times 1$ Gaussian white noise process with zero mean and covariance Σ_t. Let $A_t = [A_{0,t}, A_{1,t}...A_{p,t}]$, $x_t' = [1_n, y_{t-1}'...y_{t-p}']$, where 1_n is a row vector of ones of length n, let $vec(\cdot)$ denote the stacking column operator and let $\theta_t = vec(A_t')$. Then y_t can be written as

$$y_t = X_t'\theta_t + \varepsilon_t \tag{2}$$

where $X_t' = (I_n \bigotimes x_t')$ is a $n \times (np+1)n$ matrix, I_n is a $n \times n$ identity matrix, and θ_t is a $(np+1)n \times 1$ vector. If we treat θ_t as a hidden state vector, equation (2) represents the observation equation of a state space model. We assume that θ_t evolves according to the following nonlinear transition equation

$$p(\theta_{t+1}|\theta_t, \Omega_t) \propto I(\theta_{t+1})f(\theta_{t+1}|\theta_t, \Omega_t) \tag{3}$$

where $f(\theta_{t+1}|\theta_t, V_t) = N(\theta_t, \Omega_t)$ and $I(\theta_{t+1})$ is an indicator function discarding explosive paths for y_t. Such an indicator is necessary to make dynamic analysis sensible and, as we will see below, it is easy to implement numerically.

We represent $f(\theta_{t+1}|\theta_t, \Omega_t)$ as

$$\theta_{t+1} = \theta_t + u_{t+1} \tag{4}$$

where u_t is a $(np+1)n \times 1$ Gaussian white noise process with zero mean and covariance Ω_t. We select this specification because more general AR and mean reverting structures produced very similar results.

We make three assumptions. First, we set $\Sigma_t = \Sigma \;\; \forall t$; second we assume $corr(u_t, \varepsilon_t) = 0$, third we specify Ω_t to be diagonal. These assumptions may, at first sight, appear to be restrictive but they are not. For example, the first assumption does not imply that the forecast errors of the model are homoskedastic. In fact, substituting 4 into we have that $y_t = X_t'\theta_{t-1} + e_t$ where $e_t = \varepsilon_t + X_t'u_t$. Hence, an heteroskedastic structure appears without assuming that Σ_t or Ω_t vary over time. Note that the assumed heteroskedastic structure is appealing since variations in the variance of the forecast errors are due to changes in the variance of the regressors. Since we are analyzing a system where the variance of inflation and output are potentially varying and since output and inflation are regressors of the model, this seems a reasonable starting point for the analysis (see Sims and Zha (2002) or Cogley and Sargent (2003) for alternative specifications). The second assumption is somewhat stronger and implies that the dynamics of the model are conditionally linear. [1]. Sargent and Hansen (1998) showed how to relax this assumption by letting the innovations of the measurement equation to be serially correlated. Since in our setup ε is by construction a white noise, the loss of information caused by imposing uncorrelation between the shocks is likely to be small. The third assumption implies that each element of θ_t evolves independently. This assumption dramatically simplifies the computations and it is irrelevant for the outcomes of the estimation process since, as shown below, structural coefficients evolve in a correlated manner.

Let S be the square root of Σ, i.e., $\Sigma = SS'$. Since Σ is time invariant, then S is time invariant. Let H_t be an orthonormal matrix such that $H_t H_t' = I$ and let $K_t^{-1} = H_t' S^{-1}$. K_t is a particular decomposition of Σ which transforms 2 in two ways: it produced uncorrelated innovations; it gives a structural interpretation to the system of equation (2) [2]. Premultiplying y_t by K_t^{-1} we obtain

$$K_t^{-1} y_t = K_t^{-1} A_{0,t} + \sum_j K_t^{-1} A_{j,t} y_{t-j} + e_t \tag{5}$$

where $e_t = K_t^{-1}\varepsilon_t$ satisfies $E(e_t) = 0$, $E(e_t e_t') = I_n$. Equation (5) represents the class of "structural" representations of y_t we are interested in: a standard Choleski representation can be obtained setting S to be lower triangular and $H_t = I_n$. Here H_t is a rotation matrix which implement particular economic restrictions.

Letting $C_t = [K_t^{-1} A_{1t} ... K_t^{-1} A_{pt}]$, and $\gamma_t = vec(C_t')$, (5) becomes

$$K_t^{-1} y_t = X_t' \gamma_t + e_t \tag{6}$$

As in standard fixed coefficient VARs there is a one-to-one relation between γ_t and θ_t since $\gamma_t = (K_t^{-1} \bigotimes I_{np})\theta_t$ where I_{np} is a $(np+1)\times(np+1)$ identity matrix. Whenever $I(\theta_{t+1}) = 1$, we also have

$$\gamma_{t+1} = \gamma_t + \eta_{t+1} \tag{7}$$

where $\eta_t = (K_t^{-1} \bigotimes I_{np})u_t$, the vector of shocks to structural parameters, satisfies $E(\eta_t) = 0$, $E(\eta_t \eta_t') = E((K_t \bigotimes I_{np})u_t u_t'(K_t \bigotimes I_{np})')$.

Hence, the vector of structural shocks $\xi_t' = [e_t', \eta_t']'$ is a white noise process with zero mean and covariance matrix $E\xi_t\xi_t' = \Xi = \begin{bmatrix} I_n & 0 \\ 0 & E((K_t \bigotimes I_{np})u_t u_t'(K_t \bigotimes I_{np})') \end{bmatrix}$. Hence, shocks to structural parameters are no longer independent (each element of γ_t depends on

[1] This means, for instance, that we can not study whether shocks of different sign or of different magnitude have different dyanmic effects on the system.

[2] H_t is time dependent since it can be a stochastic function. In this case we assume that is independent of ε_t

several u_{it} via K_t). In particular, it can be shown that when a shock in some coefficients occurs, all the coefficients of the same variable at the same lag in all the equations move.

It is important to stress that the model (6)-(7) contains two types of structural shocks: disturbances to the observations equations, e_t, and disturbances to structural parameters, η_t. While the former have the same interpretation as those in a fixed coefficients VARs, the second are different and new. To understand their meaning, suppose that the $n-th$ equation of (6) is the monetary policy equation and split it into a systematic component, which is characterized by $\tilde{\gamma}_t = [\gamma_{(n-1)(np+1),t}, ..., \gamma_{n(np+1),t}]'$ and which describes how monetary policy rule responds to the economy, and the non-systematic component, which consist of the policy shock $e_{n,t}$. Then, innovations in the coefficients of the systematic part of the monetary rule can be interpreted as changes in the preferences of the monetary authorities or generic variations in the reaction of monetary authorities to developments in the rest of the economy.

Identifying structural shocks is equivalent to choosing a H_t. Here as in Faust (1998), Canova and De Nicoló (2002) and Uhlig (2001), we select H_t so that the sign of the impulse response functions at $t + j, j = 1, 2, \dots, J$ matches some theoretical restriction. We have experimented with two types of restrictions: a monetary policy shock must generate (i) non-positive effects on output and inflation and a non-negative effect on the interest rate for the first J quarters after the shock; (ii) a liquidity effect (a positive effect on the interest rate and a fall in nominal money) for at least J quarters. Since main trust of the results is independent of which identification we use we only report the last one.

We choose sign restrictions to identify shocks to the observation equation, since more standard identification schemes have an undesirable property. Take, for example, a Choleski decomposition. Since Σ is time invariant, also its Choleski factor S is time invariant. Hence, since $H_t = I$, the contemporaneous effects of a monetary policy shock are time-invariant. That is , the impulse response functions will display constant contemporaneous effects no matter what point in time we choose to compute it. With this setup, time variations in impulse responses can appear only if there are variations in the lagged coefficients of the reduced form model.

Our identification approach does allow for time variations in the contemporaneous effects. In fact, by restricting the sign of the impulse for at least two period we make H_t depend on the conditional distribution of the states one period ahead [3]. For sensitivity analysis we report, in the last section, responses obtained letting policy shocks to be the third element of a Choleski system, i.e. we let monetary policy reacts to output and inflation movements but assume that it has no contemporaneous effects on these variables.

Since we focus on the identification of one shock only, we diagonalize the remaining disturbances without giving them any structural interpretation (see Canova and De Nicoló (2002) for the identification of more than one shock via sign restrictions). Appendix A describes the details of the implementation of sign restrictions in our TVC-VAR model.

[3] If the parameters are time-varying, the conditional means of the coefficients will also be time-varying and therefore H_t must vary over time therefore making the contemporaneous effects K_t vary over time

3 Impulse Responses

Our study of the effects of monetary policy is largely based on impulse response analysis. In a fixed coefficient model, impulse response functions measure the effects of a shock on future values of a variable y_i relative to some benchmark. The most common type of impulse responses involves the difference between two realizations of $y_{i,t+\tau}, \tau = 1, 2, ldots$: these realizations are identical up to time t but one assumes that between $t + 1$ and $t + \tau$ a shock in $e_j$has occurred at time $t + 1$ and the other that no shocks have occurred between $t + 1$ and $t + \tau$.

In TVC-VAR impulse responses computed this way are inappropriate since they disregard the fact that between $t + 1$ and $t + \tau$ the coefficients of the system can also change. Therefore, we need impulse response functions designed measure the effects on $y_{it+\tau}$ of a shock in $e_{j,t+1}$, given an history and the parameters, which allow $\eta_{t+\tau}, k = 1, \ldots$ to be non-zero. Here impulse response functions are the difference of two conditional expectations of $y_{t+\tau}$ one conditional to a particular history, the states, the structural parameters of the transition equation and a shock; the second conditional only a particular history, the states and the structural parameters.

Formally speaking, let $y^t = [y_1', ...y_t']'$ be a history for y_{t+1}; $\theta^t = [\theta_1', ...\theta_t']'$ be a trajectory of states up to t. Also, let $y_{t+1}^{t+\tau} = [y_{t+1}', ...y_{t+\tau}']'$ denote a collection of future observations and $\theta_{t+1}^{t+\tau} = [\theta_{t+1}', ...\theta_{t+\tau}']'$ a future trajectory of states from $t + 1$ up to time $t + \tau$. Then, letting $V = (\Sigma, \Omega)$, an impulse response function to a shock $\xi_{i,t+1}, i = 1, \ldots, n$ for each $\tau = 1, \ldots, T$ is defined by:

$$IR_y(\tau, y^t, \theta^t, V, K_t, \xi_{i,t+1}) = E(y_{t+\tau}|y^t, \theta^t, V, K_t, \xi_{i,t+1}) - E(y_{t+\tau}|y^t, \theta^t, V, K_t) \quad (8)$$

Our impulse response function resembles the one used by Gallant et al. (1996), Koop et al. (1996) and Koop (1996). Two important differences need to be noted. First, rather than treating histories as random variables, we condition on a particular realization. Since the scope of the study is to analyze how responses vary over time, we want impulse responses to be history dependent (we do not care if responses at t are representative or not of the typical responses over the sample). Second, our impulse response measures are independent of the sign and the size of the shocks (as it is in the standard case). This is due to our assumption which make shocks to the observation and the transition equations uncorrelated. We could in principle relax such an assumption. However, in our context allowing for asymmetry and size dependence make computations demanding and since we are interested in comparing responses over time, normalization of the size of the shock at each t is not particularly restrictive.

Some features of our impulse responses are worth some discussion. First, in (8) history and the shocks are fixed, but parameters are treated as random variables. That is, IR_y is a random variable since different draws of θ^t, H_t and V produce different realizations of IR_y. Second, IR_y is history dependent and can be make state dependent, for example, conditioning on a particular stretch of a history (a boom or a recession). Third, in (8) we make no explicit assumptions about future. Since conditional expectations average out future shocks, the IR_y we compute is an average of what might happened, given the present and the past. Finally, note that when coefficients are constant, (8) coincides with the standard impulse response definition.

Since there are two types of shocks in our system, we describe how to trace out the effects of shocks of the two types of shock separately.

Let $\xi_{i,t+1} = e_{i,t+1}$. Then, as shown in Appendix B, IR_y are

$$IR_y(\tau, y^t, \theta^t, V, e_{i,t+1}, K_t) = E(\Psi^i_{t+\tau,\tau-1}|y^T, \theta^t, V, K_t)e_{i,t+1} \quad \text{for} \quad \tau = 1, ...T. \tag{9}$$

where $\Psi_{t+\tau,\tau}$ is a nonlinear combination of the structural parameters of the model (precisely defined in Appendix B) and $\Psi^i_{t+\tau,\tau}$ is ith column corresponding to the ith shock. Note that (10) collapses to $IR_y(\tau, y^t, \theta^t, V, e_t) = \Psi^i_\tau$ when coefficients are constant, where Ψ_τ is the coefficient of $e_{t-\tau}$ in the MA representation of the constant coefficient VAR. Clearly IR_y depends on the identifying matrix K_t. Furthermore, since $\Psi_{t+\tau,\tau}$ is the product of matrices whose eigenvalues are non-explosive, IR_y are non-explosive.

Let $\xi_{i,t+1} = \eta_{j,t+1}$ for $j = (n-1)(np+1), ..., n(np+1)$. In Appendix A we show that for $\tau = 1$ IR_y is

$$IR_y(1, y^t, \theta^t, V, \eta_{j,t+1}, K_t) = E(\mathbf{A}_{t+1,1}|y^t, \theta^t, V, \eta_{j,t+1})y_t - E(\mathbf{A}_{t+1,1}|y^t, \theta^t, V)y_t \tag{10}$$

and for $\tau = 2, ..., T$ IR_y is

$$\begin{aligned} IR_y(\tau, y^t, \theta^t, V, \eta_{j,t+1}, K_t) &= E(\tilde{\mathbf{\Phi}}_{t+\tau,\tau}|y^t, \theta^t, V, \eta_{j,t+1})y_t - E(\tilde{\mathbf{\Phi}}_{t+\tau,\tau}|y^t, \theta^t, V)y_t \\ &+ E\left(\sum_{i=1}^{k} \Phi_{t+\tau,j} A_{0,t+\tau-i}|y^t, \theta^t, V, \eta_{j,t+1}\right) \\ &- E\left(\sum_{i=1}^{k} \Phi_{t+\tau,j} A_{0,t+\tau-i}|y^t, \theta^t, V\right) \end{aligned} \tag{11}$$

where $\mathbf{A}$ is the companion matrix of the VAR and Φ is defined in the appendix.

When a shock hit the systematic component of policy, IR_y depend on K_t because $u_t = (K_t^{-1} \bigotimes I_{np})^{-1}\eta_t)$. Also in this case, IR_y are non explosive.

4 Estimation

The model is estimated using Bayesian methods. That is, we specify prior distributions for θ^t, V and H^t and use the data up to t to compute posterior estimates of the structural parameters of the model. Since our sample covers data from 1960 to 2002, we initially estimate the model for the sample 1960:1-1978:3 and then reestimate it moving the terminal date by one quarter from 1978:4 to 2002:4

Posterior distributions for the parameters are not available in a closed form. Therefore, Markov Chain Monte Carlo methods are used to simulate sequences of parameters from the posterior distributions which consistent with the information up to time t. Estimation of reduced form TVC-VAR models with or without time variations in the variance of the shocks to the transition equation is now standard (see e.g. Cogley and Sargent (2001) or Canova (2004)): it simply requires treating the time varying parameters as a block in the Gibbs sampling algorithm. Therefore, in each cycle of the Gibbs sampler, one would run the Kalman filter and the Kalman smoother, conditional on the draw of the other time invariant parameters. In our setup the calculations are complicated by the fact that at each cycle, we need to obtain structural estimates of the time varying features of the model. This means that we need to apply at each step the identification scheme, discarding paths which are explosive and paths which do not satisfy the restrictions we impose. While this complicates the Gibbs sampler loop, and dramatically reduces the number of draws available for inference, it is relatively straightforward to implement and not very time consuming.

Because of the heavy notation and the technicalities involved, we present the details of the estimation in appendix C.

5 The Results

All the data we use is taken from the FREDII data base of the Federal Reserve Bank of San Louis. In our four variable system we use: the log of detrended (linear) real GDP, the log of first difference of GDP deflator, the federal funds rate and M1. We present impulse responses for seven specific dates, 1978:III, 1981:III, 1983:III, 1987:III, 1992:III, 1996:III and 2002:IV, which could represent turning points in the way monetary policy was conducted. The first two dates should account for the Volker monetary targets experiments and the third for the transition period from high to low inflation; 1987 corresponds to the first stock market crash after Greenspan took office; 1992 is the year of Clinton election and 1996 represents the time the substained productivity growth experienced by the US economy became evident.

While the data we used represents revisions as of 2003:2, there is a sense in which the analysis we conduct is in real time. In fact, when we compute the effects of a contractionary policy shock we do so with the data available at that time. In other words, we are not computing responses at intermediate dates in the sample using end-of-the-sample estimates, but at each t, use recursive estimates obtained with information up to that point in time.

We divide the presentation of the results around four general themes: (i) Are there changes in the structural coefficients of the model? When do they occur? Is there a synchronization in the changes of the coefficients of inflation and output with monetary policy? Are there changes in the variances of the structural shocks? When do they occur? (ii) Are there changes in the propagation of disturbances to the monetary policy equations? (iii) Are there changes in the systematic component of policy? Do disturbances to the systematic component of policy account for the observed changes in output and inflation equations? (iv) Are results robust to (a) the horizon we consider, (b) the variables included in the system, (c) the identification restrictions used?

5.1 Are there time variations in the structural model?

The upper panel of figure 1 presents the evolution of the structural coefficients of each equation of the model and in the bottom panel the change in the coefficients at each date in the sample. The first date corresponds to estimates obtained with the information up to time 1978:3, the last one to estimates obtained with the information up to time 2002:4. We report the conditional mean of the posterior distribution of the structural coefficients obtained with the Gibbs sampler and our basic identification scheme (monetary policy shocks are identified as those producing a liquidity effect for at least 2 quarters).

Three features of the picture needs to be emphasized. First, the coefficients of all four equations evolve over time, but the size of the changes differ across equations. For example, the variations observed in the coefficients of the output equation are up to six times larger than the variations observed in the coefficients of the inflation equation and cumulatively variations in the coefficients of output and money equations are of an order of magnitude larger than in the other two equations. Time variations in the point estimates are at times large and both statistically and economically significant, confirming the inherent instability of the structural relationships in the US economy over the 25 years under consideration. Second, although most the variations are concentrated in the first part of the sample (between 1978 and 1986), changes are not synchronized across equations and the largest changes typically occur at different dates, and overall, the pattern of time variations is different. For example, the inflation equation experienced the largest coefficient changes at the beginning of the 1980's; coefficient variations is reduced since 1984 and shows a substantially stable pattern thereafter while the coefficients of the interest rate equation changes through the sample and the maximum changes occur in 1981 and 1996. The output equation, on the other hand, displays two regimes of coefficient variations (high up to 1986 and low thereafter), but within the high volatility regime, the largest coefficient variations occur in 1986. Finally, the money equation (our monetary policy equation) display large changes up to 1986 and a more stable pattern thereafter. Third, and related to the above, the variations in the structural coefficients of the output and inflation equations are asynchronous and do not match those observed in the interest rate equation.

Figure 2 zooms in on few coefficients of the monetary policy equation and highlight the variations in the lagged inflation coefficients. Clearly, from 1979 to 1982 there was a six fold-increase in the first lagged coefficient of inflation, confirming that during this period the systematic component of monetary policy did become more aggressive. However, from figure 1, it is clear that there instability is pervasive not simply related to inflation coefficients. For example, the increase first lagged output coefficient is even larger (red line in figure 1). We will return on this issue later when we examine whether it would have made a difference if the 1982 policy would have been applied to 1978.

Figure 3 presents the evolution of the variance of the forecast errors of the structural model over time and the variations produced by the heteroskedastic component of the error, i.e the product of the estimated innovations in the coefficient times the regressors. Four features stand out: first, forecast error variances are generally humped shaped, with a significant increase from 1978 to 1981 and there is a smooth decline thereafter in three of the four equations (the interest rate equation being the exception). Second, changes in the variances of the output and the inflation equation appear to be sufficiently well synchronized with the changes in the forecast error variance of our estimated monetary policy equation. Third, the variations in the forecast error variance due to changes in the coefficients of the monetary policy equation have a pattern that differ from those observed in the output and inflation equations and seem to match reasonably well the changes observed in the interest rate equation. Finally, the contribution of e_t to the variance of the forecast error is roughly twice as large as the contribution of $\eta_t x_t$ except at the beginning of the sample and primarily for the output equation. Furthermore, as expected the contribution of the latter component is smoothly declining over the sample.

To summarize, instabilities in the system appear to be pervasive. Both the coefficients of the structural model and the variance of the forecast error in different equations appear to be changing. A typical humped shaped pattern occur where variances and most of the coefficients increase from 1978 to 1981 and decline thereafter. In general, most of the variations appear to take place between 1978 to 1986: changes in the coefficients of the inflation and output show little synchronicity with the changes in the coefficients of the monetary policy equation. In particular, variations in the coefficients of the output and money equations are considerably larger than those in the coefficients of the other two equations and occur at different dates from those in the inflation equation. We find that the coefficient on lagged inflation in the monetary policy equation display a dramatic increase from 1978 to 1981 but this change is part of a more general change occurring in all the coefficients of the equation. There is more synchronization in the changes in the variances of the forecast errors of the model and it appears that the humped shape of the estimated variances is only partially due to coefficient variation. Finally, changes in the variances of shocks to the structural observation equation dominate the pattern of time variations.

Sargent and Cogley (2001) and (2003) look at measures of core inflation to establish their claim that monetary policy is responsible for the observed changes in the structure. Since our analysis so far has been based on short run dynamic analysis it may therefore miss the more longer run type of relationships. In particular, they show that inflation persistence substantially declined over time and that there is some synchronicity in the changes in persistence and in the presumed changes in monetary policy. Pivetta and Reis (2004), using univariate conventional classical methods, claim that differences in persistence over time are statistically insignificant.

The analysis of Cogley and Sargent differs from our in two respects: first, unemployment is used in place of output; second, persistence is measured using reduced form coefficients. What our structural model has to say about the evolution of persistence over time and on its relationship with changes in the monetary policy coefficients?

Since it is not clear what persistence means, the literature has used different way to measure the persistence of a process. Here we report two: the first is the value at frequency zero of the spectral density of inflation, that is, the sum of all autocovariances of estimated inflation process. The second is the half life of inflation responses to monetary shocks, measured here by the time it takes to inflation responses to reach half of the impact effect.

Figure 4 shows the spectrum of inflation computed at each date of the sample, using estimated structural model. To be precise, we report the estimated average spectral density (where average is taken over simulations) computed from the companion form representation of the VAR, conditioning on the information available at various dates in the sample. On the vertical axis we report the size of the spectral density, and on the two horizontal axis the frequency (on the side) and the date (in front). Overall, smooth time variations are pervasive and the peak at the zero frequency has declined over time. In fact, from 1978 to 1996 the mean value at the zero frequency dropped from 0.04 to 0.01; the decline is roughly monotonic and at a constant rate. Also in this case the pattern of changes is not very well synchronized with the changes in the coefficients of the monetary policy equation.

There is also clear evidence that the point estimates of the half life of inflation responses (as measured by the posterior mean) have changes over time. For example, the half life in 1978 is two while in 1992 is about 10. However standard errors are large. In fact, we find the 68% band for the location of the half life in the two cases overlap to a large extent (they are equal to [xxx,xxx], in the first case and [yyy,yyy] in the second). Hence the visual differences in the point estimates turn out to be statistically insignificant.

Hence, no matter what statistic we use we find that the timing of the changes in the monetary policy equation do not definitively match the changes in the coefficients of the inflation process and have hard time to account for the changes in the coefficients of output process.

Therefore, our investigation so far suggest that all the camps have a point. There are structural instabilities in the output and inflation equation, there are instabilities in the monetary policy equation and there are also changes in the variances of the structural shocks hitting the economy. However, based on simple synchronization accounting, it appears that the "luck" component dominates the "bad policy" component as explanation for the observed changes.

Since a synchronicity test however is a somewhat weak tool to examine a causality proposition we next proceed to analyze (i) whether the observed changes in output and inflation can be attributed to changes in the propagation mechanism of monetary policy shocks (ii) whether changes in the systematic component of monetary policy account and to what extent for the observed changes in output and inflation process.

5.2 Are there changes in the propagation of monetary policy shocks?

Figure 5 presents the responses of output and inflation to identified monetary shocks at selected dates. To make pictures legible we present only the mean of the posterior distribution of the responses for horizons from 0 to 20 quarters. Figure 6 presents the posterior 68% band for the difference between the two most extreme responses in each picture (i.e. 1978-1996 for the case of inflation, 1983-1996 in the case of output). We have omitted the responses of interest rates because they are similar across periods and quite standard: the initial increase dissipates quite quickly so that responses become insignificantly different from zero after the 4th quarter for each of the dates we selected.

Several interesting features of the two figures deserve some comment. First, while there quantitative differences in the responses obtained conditioning on the information available at different dates, the shapes of the responses is unchanged. In particular, a contractionary monetary policy shock which produces a liquidity effect significantly reduces output and inflation for about 5-6 quarters, regardless of the time when the responses are computed [4]. Therefore, the changes we have observed in the structural coefficients roughly wash out when the moving average representation is computed. Second, the largest absolute inflation response and the second largest absolute output response occur in 1978 and, as time goes by, the effect of monetary shocks on these two variables declines over the 6 periods where significant responses are recorded. Therefore, it appears that monetary policy shocks were much more effective in displacing output and inflation from their path 20-25 years ago than they are now. Third, differences in impulse responses at all horizons are both statistically and economically very small. In fact, as shown in figure 4, the simulated posterior distributions of the largest discrepancies include zero at all horizons. Furthermore, the largest absolute cumulative difference over 20 horizons in, say, the mean of the conditional distribution of output responses is only 0.96. That is, in the two most extreme cases a one standard error shock in interest rates, would have changed yearly output growth by at most 0.18 percentage points.

The main message one should retain from these figure is that changes in the structural coefficients of the output and inflation equation can not be attributed to changes in the propagation mechanism of monetary policy shocks over time. This agrees with the analyses of Sims and Zha (2001) and Hanson (2001) and stands in sharp contrast with those of Boivin and Giannoni (2002a,b).

5.3 Are there changes in the systematic component of policy?

(Later)

5.4 Are results sensitive to the choices made?

There are a number of choices we have made which may be responsible for the results and for the differences between our conclusions and those present in the literature. In the subsection we therefore analyze the sensitivity of the outcomes to two choices: the identifying restrictions and the variables included in the VAR

We start first from the identification restrictions. So far, monetary shocks have been identified through a liquidity effect. Would result change if we would alternatively use the restriction that contractionary monetary shocks should produce non-positive output and inflation responses or that monetary shocks are identified via a Choleski decomposition? The pattern of impulse responses do change with identification scheme. However, the conclusion that the propagation of monetary shocks has not dramatically changed over time is still valid. To illustrate the point, we plot in figure 7, the responses of output and inflation to contractionary monetary policy shocks identified with a Choleski system. The figure indicates that an inflation puzzle is generated with this identification scheme (inflation

[4] At all the selected dates, standard error bands do not include zero up to 5 or 6 quarters.

increases following a increase in the interest rate) and the output responses are much more muted and die out quicker. However, the conclusion that magnitude differences are relatively small over time still remain. For example, the maximum difference in output responses occurs at lag 3 when we condition on 1983 and 1978 information and it is only 0.1 in absolute value, while the maximum difference in inflation responses occur at lag 2 when we condition on 1978 and 2003 information and the magnitude of the difference is 0.18. For reason of space we only report responses to shocks in the non-systematic component of policy, but it should be clear that the other conclusions we have reached stand unchanged. That is, (i) structural inflation and output coefficients do vary over time and the variance of the forecast errors are time varying; (ii) there is little coincidence in the timing of the changes in the coefficient of the policy and non-policy equations; (iii) there are structural variations in the coefficients of the policy rule but unexpected variations on the preferences of policymakers do not account for the pattern of changes in the process for inflation and output.

Second, we have examined the sensitivity of our conclusions to changes in the variables of the VAR. It is well known that small scale VAR models are appropriate only to the extent that omitted variables exert no influence on the dynamics of the included ones. However, a-priori there is no way to check that our system effectively marginalized the influence of these variables. For this reason we repeated our exercise substituting unemployment rate to detrended output and non-borrowed reserves to money.

(Also later).

6 Comparison with the literature and Conclusions

(Later, later)

References

Bernanke, B. and I. Mihov (1998), "Measuring Monetary Policy", *The Quarterly Journal of Economics*, 113:869-902.

Blanchard, O.J. and J. Simon (2000), "The Long and Large Decline in U.S. Output Volatility", *Brookings Papers on Economic Activity*.

Boivin, J. and M. Giannoni (2002a), "Has Monetary Policy Become Less Powerful?", Columbia Business School, mimeo.

Boivin, J. and M. Giannoni (2002b), "Assessing Changes in the Monetary Transmission Mechanism: A VAR Approach", *Federal Reserve Bank of New York Monetary Policy Review*, 8, n.1:97-111.

Canova, F (1993), "Modeling and Forecasting Exchange Rates Using a Bayesian Time Varying Coefficients Model", *Journal of Economics Dynamics and Control*, 17:233-262.

Canova, F. and M. Ciccarelli (2002), "Panel Index VAR Models: Estimation, Testing and Leading Indicators", Universidad de Alicante, WP-AD 2002-21.

Canova, F. and G, De Nicolo (2002), "Monetary Disturbnces Matter for Business Cycles Fluctuations in the G-7", *Journal of Monetary Economics*

Christiano, L., M. Eichenbaum and C. Evans (1999) "Monetary Shocks: What Have We Learned and to What End?", *Handbook of Macroeconomics*, Ch.2.

Cogley, T. and T.J. Sargent (2001) "Evolving Post-World War II U.S. Inflation Dynamics", *NBER Macroeconomic Annual*.

Cogley, T. and T.J. Sargent (2003) "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S. ", New York University, mimeo.

Doan, T., R. Litterman and C. Sims (1984), "Forecasting and Conditional Projections Using Realistic Priors Distributions", *Econometric Reviews*, 3:1-100.

Faust, J (1998), "The Robustness of Identified VAR Conclusion about Money", *Carnegie Rochester Conference Series in Public Policy*, 49:207-244.

Clarida, R., J. Gali and M. Gertler (2000) "Monetary Policy Rule and Macroeconomic Stability: Evidence and Some Theory", *Quarterly Journal of Economics*

Gallant, R., P. Rossi and G. Tauchen (1993), "Non Linear Dynamic Structures", *Econometrica*, 61:871-907.

Gelman, A.,J.B. Carlin, H.S. Stern and D.B. Rubin (1995) *Bayesian Data Analysis*, London: Chapman and Hall.

Hanson, M. (2001), "Varying Monetary Policy Regimes: a Vector Autoregressive Investigation", Wesleyan University, mimeo.

Kim, C. and C.R. Nelson (1999), *State Space Models with Regime Switching*, MIT Press: Cambridge, Massachussets.

Koop, G (1996), "Parameter Uncertainty and Impulse Response Analysis", *Journal of Econometrics*, 72:135-149.

Koop, G., H. Pesaran and S. Potter (1996), "Impulse Response Analysis in Non-Linear Multivariate Models", *Journal of Econometrics* 74:119-147

Marsiglia, G. (1972), "Choosing a Point from the Surface of a Sphere", *Ann. Math. Stat.* 43:645-646.

Orphanides, A. (2001), "Monetary Policy Rules, Macroeconomic Stability, and Inflation: A View from the Trenches", *Boards of Governors of the Federal Reserve System* Discussion Paper.

Sims, C. A. (1992), "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy", *European Economic Review*, 36:975-1011.

Sims, C. A. (1993), "A 9 Probabilistic Macroeconomic Forecasting Model", in J.H. Stock and M.W. Watson (Eds.), *Business Cycles, Indicators and Forecasting*, NBER Studies in Business Cycles Vol. 28:179-214.

Sims, C.A. and T. Zha (2002), "Macroeconomic Switching", Princeton University, mimeo.

Stock, J.H. and M.W. Watson (1996), "Evidence on Structural Instability in Macroeconomic Time Series Relations", *Journal of American Statistical Association*, 93:349-358.

Uhlig, H. (2001), "What Are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure", Tilburg University, mimeo.

Appendix

A. Imposing sing restrictions

We implement our sign restriction approach as follow. Let h_t be the first column of H_t. Assume that h_t is stochastic distributed independently of ε_t on a unit sphere $\mathcal{S}_n$ Let $\mathcal{M}_{t,j}$ be the set of impulse response functions satisfying the restriction for period j. Since this set is dense, there exists several h_t over the unit sphere generating impulse response functions satisfying the restrictions. Call this set of points $\mathcal{H}_t$ All the h_t belonging to $\mathcal{H}_t$ generate representations which are consistent with the definition of monetary policy shock - there may be more than one since nothing insures that $\mathcal{H}_t$ is a singleton.

(TO BE FIXED)

B. Impulse Response Functions

Substituting recursively into the system for ks we obtain

$$\mathbf{y}_{t+k} = \mathbf{A}_{0,t+k}+\sum_{j=1}^{k-1}\left(\prod_{i=0}^{j-1}\mathbf{A}_{t+k-j}\right)\mathbf{A}_{0,t+k-j}+\prod_{j=0}^{k-1}\mathbf{A}_{t+k-j}\mathbf{y}_t+\sum_{j=1}^{k-1}\left(\prod_{i=0}^{j-1}\mathbf{A}_{t+k-i}\right){}^2{}_{t+k-j}+{}^2{}_{t+k}$$

Rewrite the above representation as

$$\mathbf{y}_{t+k} = \mathbf{A}_{0,t+k} + \sum_{j=1}^{k-1}\mathbf{\Phi}_{t+k,j}\mathbf{A}_{0,t+k-j} + \mathbf{\Phi}_{t+k,k-1}\mathbf{y}_t + \sum_{j=0}^{k-1}\mathbf{\Phi}_{t+k,j}{}^2{}_{t+k-j}$$

where $\mathbf{\Phi}_{t+k,j} = \prod_{i=0}^{j-1}\mathbf{A}_{t+k-j}$ for $j = 1,2,...,$ $\mathbf{\Phi}_{t+k,0} = I$. Let $ext_{(h,k)}(X)$ be an extractor, a function which extract the h-rows and the k-columns of the matrix X. Since $y_t = ext_{(n,1)}(\mathbf{y}_\text{t})$ we have

$$y_{t+k} = \tilde{A}_{0,t+k} + \tilde{\mathbf{\Phi}}_{t+k,k-1}\mathbf{y}_t + \sum_{j=0}^{k-1}\Phi_{t+k,j}SH_{t+k-j}e_{t+k-j}$$

where $\tilde{A}_{0,t+k} = A_{0,t+k} + \sum_{j=1}^{k-1}\Phi_{t+k,j}A_{0,t+k-j}$ for $k > 1$ and $\tilde{A}_{0,t+1} = A_{0,t+1}$, $\tilde{\mathbf{\Phi}}_{t+k,k-1} = ext_{(n,n^2p)}(\mathbf{\Phi}_{t+k,k-1})$ and $\Phi_{t+k,j} = ext_{(n,n)}(\mathbf{\Phi}_{t+k,j})$. Thus

$$y_{t+k} = \tilde{A}_{0,t+k} + \tilde{\mathbf{\Phi}}_{t+k,k-1}\mathbf{y}_t + \sum_{j=0}^{k-1}\Psi_{t+k,j}e_{t+k-j} \tag{12}$$

where $\Psi_{t+k,j} = \Phi_{t+k,j}SH_{t+k-j}$ $\Psi_{t+k,0} = \Phi_{t+k,0}SH_t = SH_t$ and . Let us now consider a partitioned version of equation (). Let $e_t = (e_{i,t}|e_{-i,t})$ where $e_{i,t}$ is an element of e_t and $e_{-i,t}$ is the vector containing the other $n-1$ elements of e_t. Let $H_t = (h_t|h_t^{-i})$, where h_t is a column of H_t corresponding to $e_{i,t}$ and h_t^{-i} is the matrix formed by remaining $n-1$ columns of H_t. Then we can rewrite equation (13) as

$$y_{t+k} = \tilde{A}_{0,t+k} + \tilde{\mathbf{\Phi}}_{t+k,k-1}\mathbf{y}_t + \sum_{j=0}^{k-1}\Phi_{t+k,j}Sh_te_{i,t+k-j} + \sum_{j=0}^{k-1}\Phi_{t+k,j}Sh_t^{-1}e_{-i,t+k-j}$$

or

$$y_{t+k} = \tilde{A}_{0,t+k} + \tilde{\mathbf{\Phi}}_{t+k,k-1}\mathbf{y}_t + \sum_{j=0}^{k-1} \Psi^i_{t+k,j} e_{i,t+k-j} + \sum_{j=0}^{k-1} \Psi^{-i}_{t+k,j} e_{-i,t+k-j} \tag{13}$$

where $\Psi^i_{t+k,j} = \Phi_{t+k,j} Sh_t$ and $\Psi^{-i}_{t+k,j} = \Phi_{t+k,j} Sh^{-i}_t$.

Time-Varying Coefficients VAR: Non-Systematic Component

Let us consider equation (34). Fix $e_{i,T+1}$ to be a monetary policy shock to the non-systematic component occurring at time $T+1$. Taking conditional expectations

$$\begin{aligned} E\left(y_{T+k}|y^T, \theta^T, V, e_{i,T+1}\right) &= E(\tilde{A}_{0,T+k}|y^T, \theta^T, V, e_{i,T+1}) + E\left(\tilde{\mathbf{\Phi}}_{T+k,k}\mathbf{y}_T|y^T, \theta^T, V, e_{i,T+1}\right) + \\ &\quad + E\left(\sum_{j=0}^{k-1} \Psi^i_{T+k,j} e_{i,T+k-j}|y^T, \theta^T, V, e_{i,T+1}\right) \\ &\quad + E\left(\sum_{j=0}^{k-1} \Psi^{-i}_{T+k,j} e_{-i,T+k-j}|y^T, \theta^T, V, e_{i,T+1}\right) \end{aligned} \tag{14}$$

First notice that

$$\begin{aligned} E\left(\sum_{j=0}^{k-1} \Psi^{-i}_{T+k,j} e_{-i,T+k-j}|y^T, \theta^T, V, e_{i,T+1}\right) &= \sum_{j=0}^{k-1} E\left(\Psi^{-i}_{T+k,j}|y^T, \theta^T, V, e_{i,T+1}\right) \times \\ &\quad \times E\left(e_{-i,T+k-j}|y^T, \theta^T, V, e_{i,T+1}\right) \\ &= \sum_{j=0}^{k-1} E\left(\Psi^{-i}_{T+k,j}|y^T, \theta^T, V\right) \times \\ &\quad \times E\left(e_{-i,T+k-j}|y^T, \theta^T, V\right) \\ &= 0 \end{aligned} \tag{15}$$

where the first line derive from the assumption $E(u_t \varepsilon'_t) = 0$, the second line from $E(e_{i,t} e_{-i,t}) = 0$ and the third line from $E(e_t) = 0$. Thus we have

$$\begin{aligned} E\left(y_{T+k}|y^T, \theta^T, V, e_{i,T+1}\right) &= E(\tilde{A}_{0,T+k}|y^T, \theta^T, V, e_{i,T+1}) + E\left(\tilde{\mathbf{\Phi}}_{T+k,k}\mathbf{y}_T|y^T, \theta^T, V, e_{i,T+1}\right) + \\ &\quad + E\left(\sum_{j=0}^{k-1} \Psi^i_{T+k,j} e_{i,T+k-j}|y^T, \theta^T, V, e_{i,T+1}\right) \\ &= E(\tilde{A}_{0,T+k}|y^T, \theta^T, V, e_{i,T+1}) + E\left(\tilde{\mathbf{\Phi}}_{T+k,k}\mathbf{y}_T|y^T, \theta^T, V, e_{i,T+1}\right) + \\ &\quad + E\left(\sum_{j=0}^{k-2} \Psi^i_{T+k,j} e_{i,T+k-j}|y^T, \theta^T, V, e_{i,T+1}\right) + \\ &\quad + E\left(\Psi^i_{T+k,k-1} e_{i,T+1}|y^T, \theta^T, V, e_{i,T+1}\right) \end{aligned} \tag{16}$$

Notice that $E(\tilde{A}_{0,T+k}|y^T, \theta^T, V, e_{i,T+1}) = E(\tilde{A}_{0,T+k}|y^T, \theta^T, V)$, $E\left(\Phi_{T+k,k}\mathbf{y}_T|y^T, \theta^T, V, e_{i,T+1}\right) = E\left(\Phi_{T+k,k}\mathbf{y}_T|y^T, \theta^T, V\right)$ since $\Phi_{T+k,k}$ and $\tilde{A}_{0,T+k}$ are functions of the sequence u_T^{T+k} and

$Ee_T u'_{T+k} = 0$ for $k = 0, 1, ...$
For $j \neq k$ we have $E\left(\Psi_{T+k,j} e_{T+k-j} | y^T, \theta^{T+1}, V, e_{i,T+1}\right) = E\left(\Psi_{T+k,j} e_{T+k-j} | y^T, \theta^T, V\right)$ because $Ee_{i,T} e'_{i,T+k} = 0$ for $k \neq 0$.
Moreover $E\left(\Psi_{T+k,j} e_{T+k-j} | y^T, \theta^T, V\right) = E\left(\Psi_{T+k,j} | y^T, \theta^T, V\right) E\left(e_{T+k-j} | y^T, \theta^T, V\right) = 0$. Therefore

$$
\begin{aligned}
E\left(y_{T+k} | y^T, \theta^T, V, e_{i,T+1}\right) &= E(\tilde{A}_{0,T+k} | y^T, \theta^T, V) + E\left(\tilde{\mathbf{\Phi}}_{T+k,k} \mathbf{y}_T | y^T, \theta^T, V\right) + \\
&\quad + E\left(\Psi^i_{T+k,k-1} e_{i,T+1} | y^T, \theta^T, V, e_{i,T+1}\right)
\end{aligned} \tag{17}
$$

Now let us consider the other realization of y_{T+k}.

$$
\begin{aligned}
E\left(y_{T+k} | y^T, \theta^T, V\right) &= E(\tilde{A}_{0,T+k} | y^T, \theta^T, V) + E\left(\tilde{\mathbf{\Phi}}_{T+k,k} \mathbf{y}_T | y^T, \theta^T, V\right) + \\
&\quad + E\left(\sum_{j=0}^{k-1} \Psi^i_{T+k,j} e_{i,T+k-j} | y^T, \theta^T, V\right) \\
&\quad + E\left(\sum_{j=0}^{k-1} \Psi^{-i}_{T+k,j} e_{-i,T+k-j} | y^T, \theta^T, V\right)
\end{aligned} \tag{18}
$$

The third and the fourth term of (39) are equal to zero because of the independence of VAR and coefficients innovations. Thus

$$
E\left(y_{T+k} | y^T, \theta^T, V\right) = E(\tilde{A}_{0,T+k} | y^T, \theta^T, V) + E\left(\tilde{\mathbf{\Phi}}_{T+k,k} \mathbf{y}_T | y^T, \theta^T, V\right) \tag{19}
$$

Now taking the difference between (37) and (40) we obtain

$$
E\left(y_{T+k} | y^T, \theta^T, V, e_{i,T+1}\right) - E\left(y_{T+k} | y^T, \theta^T, V\right) = E\left(\Psi^i_{T+k,k-1} | y^T, \theta^T, V\right) e_{i,T+1} \tag{20}
$$

Time-Varying Coefficients VAR: Systematic Component

Let $\eta_{i,T+1}$ for $i = 15, ..., 21$ be shock in the systematic component of the monetary policy. Taking conditional expectation we have

$$
\begin{aligned}
E\left(y_{T+k} | y^T, \theta^T, V, \eta_{i,T+1}\right) &= E(\tilde{A}_{0,T+k} | y^T, \theta^T, V, \eta_{i,T+1}) + E\left(\tilde{\mathbf{\Phi}}_{T+k,k} \mathbf{y}_T | y^T, \theta^T, V, \eta_{i,T+1}\right) + \\
&\quad + E\left(\sum_{j=0}^{k-1} \Psi_{T+k,j} e_{T+k-j} | y^T, \theta^T, V, \eta_{i,T+1}\right)
\end{aligned} \tag{21}
$$

The third term is equal to zero because the independence assumption between e_t and η_t. Consider the case $k > 1$. Substituting the definition of $\tilde{A}_{0,T+k}$ we obtain

$$
\begin{aligned}
E\left(y_{T+k} | y^T, \theta^T, V, \eta_{i,T+1}\right) &= E(A_{0,t+k} | y^T, \theta^T, V, \eta_{i,T+1}) + \\
&\quad + E\left(\sum_{j=1}^{k-1} \Phi_{t+k,j} A_{0,t+k-j} | y^T, \theta^T, V, \eta_{i,T+1}\right) \\
&\quad + E\left(\tilde{\mathbf{\Phi}}_{T+k,k} \mathbf{y}_T | y^T, \theta^T, V, \eta_{i,T+1}\right)
\end{aligned} \tag{22}
$$

Notice that since the coefficients in $A_{0,t+k}$ are independent from the coefficients of lagged variable in the monetary policy equation, $E(A_{0,t+k}|y^T, \theta^T, V, \eta_{i,T+1}) = E(A_{0,t+k}|y^T, \theta^T, V)$. Thus

$$\begin{aligned} E\left(y_{T+k}|y^T, \theta^T, V, \eta_{i,T+1}\right) &= E(A_{0,t+k}|y^T, \theta^T, V) + E\left(\sum_{j=1}^{k-1} \Phi_{t+k,j} A_{0,t+k-j}|y^T, \theta^T, V, \eta_{i,T+1}\right) \\ &\quad + E\left(\tilde{\mathbf{\Phi}}_{T+k,k}|y^T, \theta^T, V, \eta_{i,T+1}\right)\mathbf{y}_T \end{aligned} \tag{23}$$

Now consider $k = 1$. We have

$$E\left(y_{T+1}|y^T, \theta^T, V, \eta_{i,T+1}\right) = E(A_{0,T+1}|y^T, \theta^T, V) + E\left(\mathbf{A}_{T+1}|y^T, \theta^T, V, \eta_{i,T+1}\right)\mathbf{y}_T$$

Now consider the other realization of y_{T+k}

$$\begin{aligned} E\left(y_{T+k}|y^T, \theta^T, V\right) &= E(\tilde{A}_{0,T+k}|y^T, \theta^T, V) + E\left(\tilde{\mathbf{\Phi}}_{T+k,k}\mathbf{y}_T|y^T, \theta^T, V,\right) + \\ &\quad + E\left(\sum_{j=0}^{k-1} \Psi_{T+k,j} e_{T+k-j}|y^T, \theta^T, V\right) \end{aligned} \tag{24}$$

Again the last term is equal to zero. Considering the case $k > 1$ and substituting the definition of $\tilde{A}_{0,T+k}$ we obtain

$$\begin{aligned} E\left(y_{T+k}|y^T, \theta^T, V\right) &= E(A_{0,t+k}|y^T, \theta^T, V) + E\left(\sum_{j=1}^{k-1} \Phi_{t+k,j} A_{0,t+k-j}|y^T, \theta^T, V\right) \\ &\quad + E\left(\tilde{\mathbf{\Phi}}_{T+k,k}|y^T, \theta^T, V\right)\mathbf{y}_T \end{aligned} \tag{25}$$

For $k = 1$ we have

$$E\left(y_{T+1}|y^T, \theta^T, V\right) = E(A_{0,T+1}|y^T, \theta^T, V) + E\left(\mathbf{A}_{T+1}|y^T, \theta^T, V\right)\mathbf{y}_T \tag{26}$$

Taking the difference for the two realization for $k > 1$ we obtain

$$\begin{aligned} E\left(y_{T+k}|y^T, \theta^T, V, \eta_{i,T+1}\right) - E\left(y_{T+k}|y^T, \theta^T, V\right) &= E\left(\sum_{j=1}^{k-1} \Phi_{t+k,j} A_{0,t+k-j}|y^T, \theta^T, V, \eta_{i,T+1}\right) \\ &\quad + E\left(\tilde{\mathbf{\Phi}}_{T+k,k}|y^T, \theta^T, V, \eta_{i,T+1}\right)\mathbf{y}_T - \\ &\quad - E\left(\sum_{j=1}^{k-1} \Phi_{t+k,j} A_{0,t+k-j}|y^T, \theta^T, V\right) - \\ &\quad - E\left(\tilde{\mathbf{\Phi}}_{T+k,k}|y^T, \theta^T, V\right)\mathbf{y}_T \end{aligned} \tag{27}$$

while for $k = 1$ we obtain

$$\begin{aligned} E\left(y_{T+1}|y^T, \theta^T, V, \eta_{i,T+1}\right) - E\left(y_{T+1}|y^T, \theta^T, V\right) &= E\left(\mathbf{A}_{T+1}|y^T, \theta^T, V, \eta_{i,T+1}\right)\mathbf{y}_T - \\ &\quad E\left(\mathbf{A}_{T+1}|y^T, \theta^T, V\right)\mathbf{y}_T \end{aligned} \tag{28}$$

Fixed Coefficients VAR

Consider equation (?) and assume that the coefficients are constant. Thus the time varying coefficients become $\Phi_{t+k,k} = \Phi_k = \mathbf{A}^k$ and $\Psi_k = \mathbf{A}^k SH$ for all k which correspond to traditional impulse response function for structural shocks in fixed coefficients VARs. Thus IR for shock i will be Ψ_k^i.

C. Estimation

Priors

Let T be the end of the estimation sample. We choose prior specifications for the unknowns which gives us analytic expressions for the conditional posteriors of subvectors of the unknowns. Let $F(\mathcal{M}_t)$ be an indicator function which assumes value one if the identifying restrictions are satisfied, that is it is one if $(\Psi_{t+1,1}^i, ..., \Psi_{t+K,K}^i) \in \mathcal{M}_t$ and zero otherwise. Let $F(\mathcal{M}^T) = \prod_{t=1}^T F(\mathcal{M}_t)$. Let the joint prior for θ^T, V and h^T be proportional to the joint prior of θ^T and V whenever the identifying restrictions are satisfied, that is

$$p(\theta^T, V, h^T) = p(\theta^T, V)F(\mathcal{M}^T)p(h^T) \tag{29}$$

where the first term, the joint prior for states and variances, can be factored as $p(\theta^T, V) = p(\theta^T|V)p(V)$. Here $p(\theta^T|V)$ is the conditional density of θ^T which is given by $p(\theta^T|V) \propto I(\theta^T)f(\theta^T|V)$ where $f(\theta^T|V) = f(\theta_0|V)\prod_{t=0}^{T-1} f(\theta_{t+1}|\theta_t, V)$ and $I(\theta^T) = \prod_{t=0}^T I(\theta_t)$. Hence, the conditional density of θ^T is normal times the indicator function.

We assume that the covariance matrices Σ and Ω are inverse-Wishart distributed with scale matrix Σ_0^{-1}, Ω_0^{-1} and degree of freedom T_{01} and T_{02}. We also assume that the prior for the initial state is a Gaussian truncated random variable independent of Σ and Ω, i.e. $p(\theta_0) \propto I(\theta_0)N(\bar{\theta}, \bar{P})$. Finally, we assume that the h_t are independent across time so that $p(h^T) = \prod_{t=0}^T p(h_t)$ and each h_t is uniformly distributed over the unit sphere $\mathcal{S}_n = \{h_t \in \Re^n : \|h_t\| = 1\}$. Such a prior specification is similar to the one used by Uhlig (2001) in a fixed coefficients VAR. The uniform prior is justified by the fact that all the trajectories satisfying the restrictions are a-priori equally likely.

Collecting the pieces the joint prior is:

$$p(\theta^T, V, h^T) \propto I(\theta^T)F(\mathcal{M}^T)f(\theta_0)\prod_{t=0}^{T-1} f(\theta_{t+1}|\theta_t, V)p(\Sigma)p(\Omega) \tag{30}$$

Note that for a Choleski scheme $H_t = I_n$, and the priors for θ^T, V remain unchanged. Thus, the prior reduces to

$$p(\theta^T, V, h^T) = I(\theta^T)f(\theta_0)\prod_{t=0}^{T-1} f(\theta_{t+1}|\theta_t, V)p(\Sigma)p(\Omega) \tag{31}$$

We calibrate the prior by estimating a fixed coefficients VAR using data from 1960:I up to 1969:I. We set $\bar{\theta}$ equal to the point estimates of the coefficients and $\bar{P}$ to the estimated covariance matrix. We set Σ_0 equal to the VAR innovations covariance matrix and $\Omega_0 = \varrho \bar{P}$. The parameter ϱ measure how much the time variation is allowed in coefficients. Although as t grows likelihood tends to dominate the prior, results are somewhat sensitive to the choice of ϱ. In particular, specifying a small ϱ we restrict the time variation in the coefficients while specifying a large ϱ reduces the probability of finding draws which generate non-explosive paths. We end up choosing ϱ on the basis of the sample size i.e. for the sample 1969:I-1983:III $\varrho = 0.0025$, 1969:I-1981:III $\varrho = 0.003$, 1969:I-1983:III $\varrho = 0.0035$, for 1969:I-1987:III $\varrho = 0.004$, 1969:I-1992:III $\varrho = 0.007$, 1969:I-1996:I $\varrho = 0.008$, 1969:I-2002:IV $\varrho = 0.01$. The range of values of ϱ implies a quiet conservative prior coefficient variations: in fact, time variation accounts from a 0.35 and a 1 percent of the total coefficients standard deviation.

Our primary goal is to compute impulse response functions, which depend on $\Phi_{t+k,k}$'s, the square factor S and the matrix H_t. Therefore, we characterize first the posterior distributions of these parameters and then describe a sampling approach from these posteriors to construct a draw for the impulse responses. Note that in the sign restriction case H_t is a random variable (while under the Choleski identification scheme H_t is a matrix of constants) Thus in the second case, need only to characterize the posterior distribution of θ^{T+K} and V (no need to worry about h^{T+k}.

Posteriors

To draw the relevant quantities we need to obtain $p(h_{T+1}^{T+K}, h^T, \theta_{T+1}^{T+K}, \theta^T, V|y^T)$, which is analytically intractable. However, it is can be decomposed into simpler components. First, note that such a distribution is proportional to the unrestricted posterior predictive distribution times the two indicator variables. In fact

$$\begin{aligned} p(h_{T+1}^{T+K}, h^T, \theta_{T+1}^{T+K}, \theta^T, V|y^T) &= p(h^{T+K}, \theta^{T+K}, V|y^T) \\ &\propto p(y^T|h^{T+K}, \theta^{T+K}, V)p(h^{T+K}, \theta^{T+K}, V) \end{aligned} \tag{32}$$

the first term of the right hand side in the second line is the likelihood and is invariant for any orthogonal rotation thus $p(y^T|h^{T+K}, \theta^{T+K}, V) = p(y^T|\theta^{T+K}, V)$. The second term is the joint prior $p(h^{T+K}, \theta^{T+K}, V) = p(\theta^{T+K}, V)F(\mathcal{M}^{T+K})p(h^{T+K})$. Thus we have

$$p(h^{T+K}, \theta^{T+K}, V|y^T) \propto p(\theta^{T+K}, V|y^T)F(\mathcal{M}^{T+K})p(h^{T+K}) \tag{33}$$

where $p(\theta^{T+K}, V|y^T) \propto p(y^T|\theta^{T+K}, V)p(\theta^{T+K}, V)$ is the posterior distribution for θ^{T+K} and V (i.e. the posterior for reduced form parameters). Such a posterior can be factored as

$$\begin{aligned} p(\theta^{T+K}, V|y^T) &= p(\theta_{T+1}^{T+K}, \theta^T, V|y^T) \\ &= p(\theta_{T+1}^{T+K}|y^T, \theta^T, V)p(\theta^T, V|y^T) \end{aligned} \tag{34}$$

where the first term of the right hand side represents beliefs about the future and the second term represents the posterior density for states and hyperparameters. First notice that $p(\theta_{T+1}^{T+K}|y^T,\theta^T,V) = p(\theta_{T+1}^{T+K}|\theta^T,V)$ and because of the Markov assumptions future states can be factored as

$$p(\theta_{T+1}^{T+K}|\theta^T,V) = \prod_{k=1}^{K} p(\theta_{T+k}|\theta_{T+k-1},V) \tag{35}$$

and θ_{T+k} is conditionally normal with mean θ_{T+k-1} and variance Ω times the indicator variable. Therefore we can write

$$\begin{aligned} p(\theta_{T+1}^{T+K}|\theta^T,V) &= I(\theta_{T+1}^{T+K}) \prod_{k=1}^{K} f(\theta_{T+k}|\theta_{T+k-1},V) \\ &= I(\theta_{T+1}^{T+K}) f(\theta_{T+1}^{T+K}|\theta_T,V) \end{aligned} \tag{36}$$

The posterior density for the hyperparameters and the states can be factored as

$$p(\theta^T,V|y^T) \propto p(y^T|\theta^T,V)p(\theta^T,V) \tag{37}$$

The first term is the likelihood function which, given the states, has Gaussian innovations and then $p(y^T|\theta^T,V) = f(y^T|\theta^T,V)$. The second term is the joint prior for states and hyperparameters. This second term can be factored into a conditional for the states and a marginal for the hyperparameters

$$p(\theta^T,V|y^T) \propto f(y^T|\theta^T,V)p(\theta^T|V)p(V) \tag{38}$$

The conditional density for the states can be written as $p(\theta^T|V) \propto I(\theta^T)f(\theta^T|V)$ where $f(\theta^T|V) = f(\theta_0|V)\prod_{t=1}^{T} f(\theta_t|\theta_{t-1},V)$ and $I(\theta^T) = \prod_{t=0}^{T} I(\theta_t)$, thus we obtain

$$p(\theta^T,V|y^T) \propto I(\theta^T)f(y^T|\theta^T,V)f(\theta^T|V)p(V) \tag{39}$$

But $f(y^T|\theta^T,V)f(\theta^T|V)p(V)$ is the posterior density resulting if no restrictions were imposed, $p_u(\theta_T,V|y^T)$. Thus we have

$$p(\theta^T,V|y^T) \propto I(\theta^T)p_u(\theta^T,V|y^T) \tag{40}$$

and

$$I(\theta^T) = \prod_{t=0}^{T} I(\theta_t) \tag{41}$$

Collecting the pieces the posterior predictive distribution is

$$\begin{aligned} p(h_{T+1}^{T+K},h^T,\theta_{T+1}^{T+K},\theta^T,V|y^T) \propto\ & F(\mathcal{M}^{T+K})\left[I(\theta^{T+K})f(\theta_{T+1}^{T+K}|\theta^T,V)p_u(\theta^T,V|y^T)\right] \\ & p(h^{T+K}) \end{aligned} \tag{42}$$

Note that for a Choleski identification

$$p(\theta_{T+1}^{T+K},\theta^T,V|y^T) = I(\theta^{T+K})f(\theta_{T+1}^{T+K}|\theta^T,V)p_u(\theta^T,V|y^T) \tag{43}$$

Drawing from the Posterior of structural parameters

To draw θ^{T+K} and V from the posterior density we proceed as follows.

1. Draw from the unrestricted posterior, $p_u(\cdot)$, computed with the Gibbs sampler (see below), a vector of θ^T.

2. Apply the filter $I(\theta^T)$: discard explosive paths.

3. Draw a sequence of future states θ_{T+1}^{T+K}, i.e. obtain K draws of u_t from $N(0,\Omega)$ and iterate in $\theta_{T+i} = \theta_{T+i-1} + u_{T+i}$ K times (θ_{T+i} is conditionally normal with mean θ_{T+i-1} and variance Q (???)). Discard explosive paths.

4. Draw h^{T+K} drawing independently $T+K$ times h_t from a uniform distribution over the unit sphere. To draw h_t we draw 4 independent random variables $\bar{h}_{i,t}$, for $i = 1, ..., 4$, from $N(0,1)$, then compute the vector $h_t = \frac{1}{||h_t||}[\bar{h}_{1,t}\ \bar{h}_{2,t}\ \bar{h}_{3,t}\ \bar{h}_{4,t}]'$. (see Marsiglia)

5. Apply the filter $F(\mathcal{M}^T)$: discard draws if identifying restrictions are violated.

Computing Posteriors of reduced form parameters: the Gibbs Sampler

The Gibbs Sampler we use iterate on two steps. The implementation we use is identical to Cogley and Sargent (2001)

• Step 1: States given hyperparameters

Conditional on hyperparameters and the data, the unrestricted posterior of the states is normal and $p_u(\theta^T|y^T,V) = f(\theta_T|y^T,V)\prod_{t=1}^{T-1} f(\theta_t|\theta_{t+1},y^t,V)$ All the density in the right end side are Gaussian they their conditional means and variances can be computed using the backward recursion of the Kalman filter. Define

$$\begin{aligned}\theta_{t|t} &\equiv E(\theta_t|y^t,V)\\ P_{t|t-1} &\equiv Var(\theta_t|y^{t-1},V)\\ P_{t|t} &\equiv Var(\theta_t|y^t,V)\end{aligned}$$

Given some initial $P_{0|0}$, $\theta_{0|0}$, Ω and Σ, we compute forward Kalman filter recursions

$$\begin{aligned}P_{t|t-1} &= P_{t-1|t-1} + \Sigma\\ K &= (P_{t|t-1}X_t)(X_t'P_{t|t-1}X_t + \Omega)^{-1}\\ \theta_{t|t} &= \theta_{t-1|t-1} + K_t(y_t - X_t'\theta_{t-1|t-1})\\ P_{t|t} &= P_{t|t-1} - K_t(X_t'P_{t|t-1})\end{aligned}$$

The last time iteration gives $\theta_{T|T}$ and $P_{T|T}$ which are the conditional means and variance of $f(\theta_t|y^T,V)$. Hence $f(\theta_T|y^T,V) = N(\theta_{T|T},P_{T|T})$. The other $T-1$ densities can be computed using the backward Kalman filter recursions, i.e

$$\theta_{t|t+1} = \theta_{t|t} + P_{t|t}P^{-1}_{t|t+1}(\theta_{t+1} - \theta_{t|t-1}) \quad (44)$$
$$P_{t|t+1} = P_{t|t} - P_{t|t}P^{-1}_{t+1|t}P_{t|t} \quad (45)$$

where $\theta_{t|t+1} \equiv E(\theta_t|\theta_{t+1}, y^t, V)$ and $P_{t|t+1} \equiv Var(\theta_t|\theta_{t+1}, y^t, V)$ are the conditional means and variances of the remaining terms in $p_u(\theta^T|y^T, V)$. Thus $f(\theta_t|\theta_{t+1}, y^t, V) = N(\theta_{t|t+1}, P_{t|t+1})$ Therefore, to sample θ^T from the conditional posterior we proceed backward, sampling θ^T from $N(\theta_{T|T}, P_{T|T})$ and θ^t from $N(\theta_{t|t+1}, P_{t|t+1})$ for all $t < T$.

• Step 2: Hyperparameters given states

To sample V, notice that we can sample separately Σ and Ω because of the independence assumption. Conditional on the states and the data ε_t and u_t are observable and Gaussian. Combining a Gaussian likelihood with an inverse-Wishart prior results is an inverse-Wishart posterior, so that

$$\begin{aligned} p(\Sigma|\theta^T, y^T) &= IW(\Sigma_1^{-1}, T_{11}) \\ p(\Omega|\theta^T, y^T) &= IW(\Omega_1^{-1}, T_{12}) \end{aligned} \quad (46)$$

where $\Sigma_1 = \Sigma_0 + \Sigma_T$, $\Omega_1 = \Omega_0 + \Omega_T$ and $T_{11} = T_{01} + T$ $T_{12} = T_{02} + T$ and Σ_T and Ω_T are proportional to the covariance estimator

$$\begin{aligned} \frac{1}{T}\Sigma_T &= \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t\varepsilon_t' \\ \frac{1}{T}\Omega_T &= \frac{1}{T}\sum_{t=1}^{T}u_t u_t' \end{aligned}$$

To draw sample from inverse-Wishart distributions, we draw T n-dimensional vectors $\tilde{\varepsilon}_t$ from $N(0, \Sigma_1)$ and construct, e.g., $\Sigma^{-1} = \sum_{t=1}^{T}\tilde{\varepsilon}_t\tilde{\varepsilon}_t'$. Σ^{-1} is Wishart distributed and therefore Σ is inverse-Wishart.

Under some regularity conditions (see Tierney (1994)) and after a burn-in period, iterations on these two steps produce draw from $p_u(\theta^T, V|y^T)$.

For each date we are interested in computing impulse responses, 10000 iteration of the Gibbs sampler are made. We have constructed CUMSUM graphs to check for convergence and found that the chain had converged roughly after 2000 draws for each date in the sample. The densities for the parameters obtained with the remaining draws are always well behaved and none is multimodal. We keeping one every two of the remaining draws and discard all the draws generating explosive paths. The autocorrelation function of draws which are left is somewhat persistent and autocorrelations twenty draws apart are still significantly different from zero. This is somewhat of a problem. We could reduce the autocorrelation taking draws more largely spaced (say, one every 5) but this come at the price of reducing the number of draws which satisfy the identification restrictions and therefore substantially reduce the precision of the exercise. With our approach we ended up having approximately 300 draws for each date to conduct structural inference. All these properties are very similar at each date in the sample. we consider.

Computation of IR

• Shocks to the Non-Systematic Component

In order to compute (11) we proceed as follows

1. Draw $\theta_{T+1}^{T+K,\ell}, V$ following steps 1-4 described in section 6.2.
2. Compute the Choleski factor S^ℓ of Σ^ℓ. For Choleski identification scheme compute $\Psi_{T+k,k}^\ell$ and stop, go to step 5. For the second identifying approach go to step 3.
3. Exploiting independence across time of h^{ℓ_T}, draw h^{ℓ_T} from a uniform distribution over the unit sphere. To draw h_t^ℓ we draw 4 independent random variables $\bar{h}_t$, for $i = 1, ..., 4$, from $N(0,1)$, then compute the vector $h_t^\ell = \frac{1}{||h_t||}[\bar{h}_{1,t} \;\; \bar{h}_{1,t} \;\; \bar{h}_{3,t}]'$ (see Marsiglia, 19).
4. Compute $(\Psi_{T+1,1}^{i,\ell}, ..., \Psi_{T+K,K}^{i,\ell})$ and $F(\mathcal{M}_T)^\ell$. If the draw satisfies identification restrictions keep the draw otherwise discard it and repeat 1-4.
5. Repeat 1-4 L times.
6. Compute posterior mode and the 68% confidence band.

• Shocks to the Systematic Component

IR in this case can be computed only for the first identification scheme. In order to compute (13)-(14) we have to compute the two realizations of y_{T+k}. We proceed as follows

1. Draw $\theta_{T+1}^{T+K,\ell}$ and V using passages 1-4 described in section 6.2 and compute a draw for $\Phi_{T+k,i}^\ell$, and $\sum_{i=1}^k \Phi_{T+k,j}^\ell A_{0,T+k-i}^\ell$ for $k = 1, ..., K$ and $i = 1, ...K-1$.
2. Collect the draw of $u_{T+1}^{T+K,\ell}$.
3. Compute $\eta_t^\ell = (S^{\ell-1} \bigotimes I_{np+1})u_t^\ell$. Fix a shock $\eta_{j,T+1} = \delta$. From section 2 we know that such a shock will have non zero correlation with two other shocks. Sample these two other shocks from the conditional normal distribution given $\eta_{j,T+1} = \delta$. Substitute $\eta_{j,T+1} = \delta$ and the other two new shocks in the vector η_{T+1}^ℓ and call this vector $\tilde{\eta}_{T+1}^\ell$. Compute $\tilde{u}_{T+1}^\ell = (S^{\ell-1} \bigotimes I_{np+1})^{-1}\tilde{\eta}_{T+1}^\ell$. Substitute u_{T+1}^ℓ with $\tilde{u}_{T+1}^\ell$ and call the new vector $\tilde{u}_{T+1}^{T+K,\ell}$; compute $\tilde{\theta}_{T+1}^{T+K,\ell}$ using $\tilde{u}_{T+1}^{T+K,\ell}$. Using $\tilde{\theta}_{T+1}^{T+K,\ell}$ compute $\tilde{\Phi}_{T+k,i}^\ell$ and $\sum_{i=1}^k \tilde{\Phi}_{T+k,j}^\ell \tilde{A}_{0,T+k-i}^\ell$ for $k = 1, ..., K$ and $i = 1, ...K-1$.
4. Repeat 1-5 of the non-systematic component IR procedure previously described using the new vector $\tilde{u}_{T+1}^{T+K,\ell}$.
5. The draw in step 1 is a draw for

$$E(\Phi_{T+k,k}|y^T, \theta^T, V), \quad E\left(\sum_{i=1}^k \Phi_{T+k,j} A_{0,T+k-i}|y^T, \theta^T, V\right)$$

while the draw in step 2 is a draw for

$$E(\Phi_{T+k,k}|y^T, \theta^T, V, \eta_{j,T+1}), \quad E\left(\sum_{i=1}^{k} \Phi_{T+k,j} A_{0,T+k-i}|y^T, \theta^T, V, \eta_{j,T+1}\right)$$

Take the difference of the two realizations.

6. Repeat 1-5 L times ($\ell = 1, ..., L$).
7. Compute the posterior mode and the 68% confidence band.

the Choleski factor of to 1. otherwise keep the draw in 1.