Università degli Studi di Modena e Reggio Emilia
Dipartimento di Economia Politica

Materiali di discussione

\\ 533 \\

Structural Changes in the US Economy: Bad Luck or Bad Policy?

by

Fabio Canova[1]
Luca Gambetti[2]

March 2006

1 IGIER, UPF e CEPR

2 Università di Modena e Reggio Emilia
Dipartimento di Economia Politica
Via Berengario,51
41100 Modena, Italy
e-mail. gambetti.luca@unimore.it

Structural changes in the US economy: Bad Luck or Bad Policy?

Fabio Canova
IGIER, UPF, and CEPR

Luca Gambetti *
UPF and Università di Modena e Reggio Emilia

First Draft: May 2003
This Draft: November 2004

Abstract

This paper investigates the relationship between changes in output and inflation and monetary policy in the US. There are variations the structural coefficients and in the variance of the structural shocks but only the latters are synchronized across equations. The policy rules in the 1970s and 1990s are similar. The transmission of policy disturbances is unchanged over the last 25 years. Variations in the systematic component of policy have limited effects on the dynamics of the system. Changes in inflation persistence are only partly explained by monetary policy. Results are robust to alterations in auxiliary assumptions.

Key words: Monetary policy, Inflation persistence, Transmission of shocks, Time varying coefficients structural VARs

JEL Classification numbers: E52, E47, C53

*We would like to thank the participants of seminars at IGIER Bocconi, UPF, University of Modena, University of Piacenza, Nuffield College, Stockholm University, the Federal Reserve Bank of Atlanta, the workshops 'Monetary Transmission in the old and new EU members', Central Bank of Hungary, Budapest, 'Empirical methods and applications of DSGE models', Federal Reserve Bank of Cleveland, CGBCR at the University of Manchester for comments and suggestions

1 Introduction

There is considerable evidence suggesting that the US economy has fundamentally changed over the last couple of decades. In particular, several authors have noted a marked decline in the volatility of real activity and inflation since the early 1980s (see e.g. Blanchard and Simon (2000), McConnell and Perez Quiroz (2001) and Stock and Watson (2003)). What are the reasons behind such a decline? A stream of literature attributes these changes to alterations in the mechanisms through which exogenous shocks spread across sectors and propagate over time. Since the transmission mechanism depends on the structure of the economy, the main implication of this viewpoint is that important characteristics of the economy, such as the behavior of consumers and firms or the preferences of policymakers, have changed over time. The recent literature has paid particular attention to monetary policy. Several studies, including Clarida, Gali and Gertler (2000), Cogley and Sargent (2001) (2003), Boivin and Giannoni, (2002), have argued that monetary policy was "loose" in fighting inflation in the 1970s but become more aggressive since the early 1980s and see in this change of attitude 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 while Hanson (2001) claims that the propagation of monetary shocks has been stable over subsamples. Sims (2001) and Sims and Zha (2004), on the other hand, claim that changes in the variance of exogenous shocks are responsible for the observed changes in the process for output and inflation.

This controversy is not new. In the past rational expectations econometricians (e.g. Sargent (1984)) have argued that policy changes involving regime switches dramatically alter private agent decisions and, as a consequence, the dynamics of the macroeconomic variables, and searched for historical episodes supporting this view (see e.g. Sargent (1999)). VAR econometricians, on the other hand, often denied the empirical relevance of this argument suggesting that the systematic portion of monetary policy has rarely been altered and that policy changes are better characterized as random variations for the non-systematic part (Sims (1982)). This long standing debate now has been cast into the dual framework of "bad policy" (failure to respond to inflationary pressure) vs. "bad luck" (shocks are drawn from a distribution whose moments vary over time) and new evidence has been collected thanks to the development of tools which explicitly allow the examination of time variations in the structure of the economy and in the variance of the exogenous processes. Overall, and despite recent contributions, the role that monetary policy had in shaping the observed changes in the US economy is still open.

This paper provides novel evidence on the contribution of monetary policy to the structural changes observed in the US economy over the last 25 years. Our basic framework of analysis is a time varying coefficients VAR model (TVC-VAR), similar to the one employed by Cogley and Sargent (2001), where the coefficients evolve according to a nonlinear transition equation which puts zero probability on paths associated with explosive roots. Cogley and Sargent (2003) add to this basic framework a stochastic volatility model for the reduced form innovations. We also allow the variance of the forecast errors to vary over time but we do this in a simpler and more intuitive manner which retains conditional linearity and links changes in the variance of the coefficients and changes in the variance of the forecast error in an economically meaningful way.

We use Markov Chain Monte Carlo (MCMC) methods to estimate the posterior distributions of the quantities of interest. Contrary to previous studies, we explicitly conduct a structural analysis, analyze both short and long run relationships, and measure the variations in the propagation of monetary policy disturbances and the potential effects that changes in preference of the policymakers would have had in particular historical episodes.

The structural setup we employ is particularly suited to study the main issues in the debate. In fact, we can separately evaluate the magnitude of structural variations produced by i) changes in the systematic component of policy, ii) changes in the propagation of policy shocks, iii) changes in the variance of the monetary policy and other shocks and iv) changes in the rest of the economy. Both reduced form time varying coefficient and structural but constant coefficient approaches are unable to separate the relative importance of i) and ii) in accounting for the observed changes.

We identify structural disturbances by means of sign restrictions. While our focus is on monetary policy disturbances, and therefore arbitrarily orthogonalize the other shocks, the methodology can be employed to jointly identify multiple sources of structural disturbances (see e.g. Canova and De Nicolo' (2002)). We choose to work with sign restrictions for two reasons. First, the contemporaneous zero restrictions conventionally used are often absent in those theoretical models one likes to use to guide the interpretation of the results. Second, standard decompositions impose restrictions on the structure of time variations which have no a-priori justification. In the last section we show that our conclusions are robust to the identification procedure and to a number of other auxiliary assumptions we make.

Because time variations in the coefficients induce important non-linearities, standard statistics summarizing the dynamics in response to structural shocks are inappropriate. For example, since at each t the coefficient vector is perturbed by a shock, assuming that between $t+1$ and $t+\tau$ no shocks other than the monetary policy disturbance hit the system is unappealing and may give misleading results. To trace out the evolution of the economy in response to structural shocks, we therefore employ a different concept, sharing similarities with those used in Koop, Pesaran and Potter (1996), Koop (1996), and Gallant, Rossi and Tauchen (1996). In particular, impulse responses are defined as the difference between two conditional expectations, differing in the arguments of their conditioning sets.

Four main conclusions can be drawn from our investigation. First, we find changes in the estimated structural relationships, but these changes are localized in time and involve only particular coefficients in certain equations. In particular, as in Bernanke and Mihov (1998) and Leeper and Zha (2003), we find that excluding the Volker experiment of the beginning of the 1980s, the monetary policy rule has been quite stable over time. Interestingly, the posterior mean of the (sum of) inflation coefficients fails to satisfy the so-called Taylor principle, not only in the 1970s but also in the 1990s. Second, as in Sims and Zha (2004), we find posterior evidence of a decrease in the uncertainty surrounding the structural disturbances of the system and that the timing of the changes roughly coincides with the timing of the changes observed in the output and inflation equations. Third, consistent with the subsample analysis of Hanson (2001), we show that the transmission of monetary policy shocks has been very stable: both the shape and the persistence of output and inflation responses are very similar over time and the differences in the posterior mean of inflation response statistically negligible. Taken together, this evidence suggests that the "bad luck" hypothesis has considerable more posterior support than the "bad policy" hypothesis in accounting for the observed dynamics of the US economy. Fourth, we find that inflation persistence (measured by the height of the zero frequency of the spectrum of inflation in the structural model) has statistically changed over time. We show that both monetary and non-monetary factors account for its magnitude and that, although the relative contribution of monetary policy fluctuates over time, it is increasing since 1981.

We investigate whether a more aggressive policy response to inflation would have made a difference in the dynamics of output and inflation. We show that such a stance would have reduced inflationary pressures and produced significant output costs in 1979. However, no measurable inflation effects are obtained in the 1980s or 1990s and a perverse outcome would have resulted in the 2000s. Hence, while the Fed could have had some room to improve economic performance at the end of the 1970s, it is unlikely that changes in the preference of the Fed would have produced the changes observed in the US economy.

Our conclusions are robust to a number of changes in the auxiliary assumptions used. In particular, we show that they do not depend on the identification procedure, on the treatment of trends and on the variables included in the VAR.

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. Section 3 defines impulse response functions which are valid in our TVC-VAR model and describes how to compute dynamics to shocks in the non-systematic and the systematic component of policy. Section 4 presents the results and Section 5 concludes. Two appendices describes the technical details involved in the computation of impulse responses and of the posterior distributions.

2 The empirical model

Let y_t be a $n \times 1$ vector of time series 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; $A_{i,t}$, are $n \times n$ matrices, $i = 1, ..., p$, 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 (1) 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

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

where $\mathcal{I}(\theta_{t+1})$ is an indicator function discarding explosive paths of 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 assume that $f(\theta_{t+1}|\theta_t, \Omega_t)$ can be represented 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 simple specification because more general AR and/or mean reverting structures were always discarded in out-of-sample model selection exercises. We assume that $\Sigma_t = \Sigma \ \ \forall t$; that $corr(u_t, \varepsilon_t) = 0$, and that Ω_t is diagonal. At first sight, these assumptions may appear to be restrictive but they are not. For example, the first assumption does not imply that the forecast errors are homoschedastic. In fact, substituting (4) into (2) we have that $y_t = X_t'\theta_{t-1} + v_t$ where $v_t = \varepsilon_t + X_t' u_t$. Hence, one-step ahead forecast errors have a time varying non-normal heteroschedastic structure even assuming time invariant Σ_t and Ω_t. The assumed structure is appealing since it is coefficient variations that impart heteroschedastic movements to the variance of the forecasts errors (see Sims and Zha (2004) or Cogley and Sargent (2003) for alternative specifications). The second assumption is standard but 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 equivalently letting the innovations of the measurement equation to be serially correlated. Since in our setup ε_t is, by construction, a white noise process, 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 but it is irrelevant since structural coefficients are allowed to evolve in a correlated manner.

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

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

where $e_t = J_t^{-1}\varepsilon_t$ 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. For example, a standard Choleski representation can be obtained setting S to be lower triangular and $H_t = I_n$ and more

[1] This means, for instance, that we can not study whether shocks to (2) of different sign or of different magnitude have different dynamic effects on the system.

general patterns result choosing S to be non-triangular and $H_t = I_n$. Here S is arbitrary and H_t implements interesting economic restrictions.

Letting $C_t = [J_t^{-1}A_{1t}...J_t^{-1}A_{pt}]$, and $\gamma_t = vec(C_t')$, (5) can be written as

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

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

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

where $\eta_t = (J_t^{-1} \bigotimes I_{np})u_t$ is the vector of shocks to structural parameters and satisfies $E(\eta_t) = 0$, $E(\eta_t\eta_t') = E((J_t^{-1} \bigotimes I_{np})u_tu_t'(J_t^{-1} \bigotimes I_{np})')$. Hence, the vector of structural shocks $\xi_t' = [e_t', \eta_t']'$ is a white noise process with zero mean and covariance matrix $E\xi_t\xi_t' = \begin{bmatrix} I_n & 0 \\ 0 & E((J_t^{-1} \bigotimes I_{np})u_tu_t'(J_t^{-1} \bigotimes I_{np})') \end{bmatrix}$. Note that, since each element of γ_t depends on several u_{it} via J_t, shocks to structural parameters are no longer independent

The structural model (6)-(7) contains two types of shocks: disturbances to the observations equations, e_t, and disturbances to structural parameters, η_t. While the former have the usual interpretation, the latter are new. To understand their meaning, suppose that the $n-th$ equation of (6) is a monetary policy equation and suppose we split it into a systematic component, summarized by $\tilde{\gamma}_t = [\gamma_{(n-1)(np+1),t}, ..., \gamma_{n(np+1),t}]'$, and describing say, how interest rates respond to the developments in the economy, and the non-systematic component, summarized by the policy shock $e_{n,t}$. Then, $\tilde{\gamma}_t$ capture changes in the preferences of the monetary authorities with respect to developments in the rest of the economy.

In our setup, identifying structural shocks is equivalent to choosing a matrix H_t. Here as in Faust (1998), Uhlig (2001), and Canova and De Nicoló (2002), we select H_t so that the sign of the impulse response functions at $t+k, k = 1, 2, \ldots, K_1$ matches some theoretical restriction. In particular, we assume that a contractionary monetary policy shock must generate a non-positive effects on output, inflation and nominal balances and a non-negative effect on the interest rate for two quarters.

We choose sign restrictions to identify shocks to the observation equation for two reasons. First, the contemporaneous zero restrictions conventionally used to identify VARs are often absent in those theoretical (DSGE) models economists like to use to guide the interpretation of the results. Second, standard decompositions have an undesirable property. Take, for example, a Choleski decomposition. Since Σ is time invariant, its Choleski factor S is time invariant. Hence, since $H_t = I_n$, the contemporaneous effects of a monetary policy shock are time-invariant. That is, if a Choleski decomposition is adopted, responses will be constant over time unless there are variations in the lagged reduced form coefficients. Such a restriction is hard to justify and unduly restricts the pattern of time variations allowed in the structural coefficients. Our identification approach, on the contrary, allows for time variations in both contemporaneous and lagged effects and, as a consequence, in the variance of the shocks. For robustness, in the last section of the paper we report results obtained identifying policy shocks as 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 effects within a quarter on these variables. Since we are interested in recovering the systematic and non-systematic part of monetary policy and in analyzing how the economy responds to their changes over time, we arbitrarily diagonalize the remaining disturbances without giving them a structural interpretation.

3 Impulse Responses

One question we would like to address is whether the transmission of monetary policy shocks has changed over time. In a fixed coefficient model, impulse response functions provide information on how the variables react to policy shocks. Impulse responses are typically computed as the difference between two realizations of $y_{i,t+k}$ which are identical up to time t, but one assumes that between $t+1$ and $t+k$ a shock in e_j occurs only at time $t+1$ and the other that no shocks take place at all dates between $t+1$ and $t+k, k = 1, 2, \ldots,$.

In a TVC model, responses computed this way are inadequate since they disregard the fact that between $t+1$ and $t+k$ the coefficients of the system may also change. Hence, meaningful impulse response functions ought to measure the effects of a shock in e_{it+1} on y_{jt+k}, allowing future shocks to the coefficients to be non-zero. For this reason, our impulse responses are obtained as the difference between two conditional expectations of y_{t+k}. In both cases we condition on the history of the data $(y_1, \ldots, y_t)$, of the states $(\theta_1, \ldots, \theta_t)$, on the structural parameters of the transition equation (which are function of J_t) and all future shocks. However, in the first case we condition on a draw for the current shock, while in the second we condition on current shock being zero.

Formally speaking, let $y^t = [y_1', ...y_t']'$ be a history for y_t; $\theta^t = [\theta_1', ...\theta_t']'$ be a trajectory for the states up to t, $y_{t+1}^{t+k} = [y_{t+1}', ...y_{t+k}']'$ be a collection of future observations and $\theta_{t+1}^{t+k} = [\theta_{t+1}', ...\theta_{t+k}']'$ be a future trajectory of states. Let $V_t = (\Sigma, \Omega_t)$; recall that $\xi_t' = [e_t', \eta_t']'$ and let $\zeta_t' = [u_t', \epsilon_t']$. Let $\xi_{i,t+1}^{\delta}$ be a particular realization of size δ in $\xi_{i,t+1}$ and let $I_t^1 = \{y^t, \theta^t, V_t, J_{t+1}, \xi_{i,t+1}^{\delta}, \xi_{-i,t+1}, \zeta_{t+2}^{t+\tau}\}$ and $I_t^2 = \{y^t, \theta^t, V_t, J_{t+1}, \xi_{t+1}, \zeta_{t+2}^{t+\tau}\}$ be conditioning sets where $\xi_{-i,t+1}$ indicates all shocks excluding the one in the i-th component. Then an impulse response function to a shock $\xi_{i,t+1}^{\delta}$, $i = 1, \ldots, n$ is defined as: [2]

$$IR_y(t,k) = E(y_{t+k}|I_t^1) - E(y_{t+k}|I_t^2) \qquad k = 1, 2, \ldots \tag{8}$$

(8) resembles the impulse response function suggested by Gallant et al. (1996), Koop et al. (1996) and Koop (1996). Three important differences need to be noted. First, rather than treating histories as random variables, we condition on a particular realization of y^t. Since we want to analyze how responses vary over time, history dependence is a must. Second, responses to shocks to the measurement equation are independent of the

[2] An alternative definition of impulse responses is obtained averaging out future shocks. Our definition is preferrable for two reasons: it is easier to compute and produces numerically more stable distributions; it produces impulses responses which are similar to those generated by constant coefficient impulse responses when shocks to the measurement equations are considered. However, since future shocks are not averaged out, our impulse responses will tend to display larger variability.

sign and the size of the shocks (as it is in a fixed coefficient case). This is not the case for shocks to the transition equation. Third, we do not condition on a particular realization of θ_{t+1} (for example, its conditional mean) but instead treat θ_{t+1} as a random variable and integrate it out when calculating impulse responses. This implies that $IR_y(t,k)$ are random variables. Integrating θ_t out of impulse responses allows us to concentrate attention on time differences which depend on the history of y_t but not on the size of the sample. Finally, note that $IR_y(t,k)$ can be made state dependent, if we condition on a particular stretch of a history (a boom or a recession), and that (8) coincides with standard impulse responses when coefficients are constant.

Since there are two types of shocks, we describe how to trace out the dynamics effects of each of them separately. Let $\xi_{i,t+1} = e_{i,t+1}$. Then

$$\begin{aligned} IR_y(t,1) &= J_t^{-1,i} e_{i,t+1} \\ IR_y(t,k) &= \Psi^i_{t+k,k-1} e_{i,t+1} \qquad k = 2,3,\ldots \end{aligned} \tag{9}$$

where $\Psi_{t+k,k-1} = \mathcal{S}_{n,n}[(\prod_{h=0}^{k-1} \mathbf{A}_{t+k-h}) \times J_{t+k-(k-1)}]$, $\mathbf{A}$ is the companion matrix of the VAR; $\mathcal{S}_{n,n}$ is a selection matrix which extracts the first $n \times n$ block of $[(\prod_{h=0}^{k-1} \mathbf{A}_{t+k-h}) \times J_{t+k-(k-1)}]$ and $\Psi^i_{t+k,k-1}$ is the column of $\Psi_{t+k,k-1}$ corresponding to the $i-$th shock.

When the coefficients are constant, $\prod_h \mathbf{A}_{t+k-h} = \mathbf{A}^k$ and $\Psi_{t+k,k-1} = \mathcal{S}_{n,n}(\mathbf{A}^k \times J)$ for all k. Hence (9) collapses to traditional impulse response function to unitary structural shocks. Clearly, IR_y depends on the identifying matrix J_t and is non-explosive, since $\Psi_{t+k,k-1}$ is the product of matrices whose eigenvalues are non-explosive.

When $\xi_{i,t+1} = \eta_{j,t+1}$ for $j = (n-1)(np+1), ..., n(np+1)$, appendix A shows that

$$\begin{aligned} IR_y(t,1) &= \left[E(\mathbf{A}_{t+1,1}|I_t^1) - E(\mathbf{A}_{t+1,1}|I_t^2)\right] y_t \\ IR_y(t,k) &= E(\tilde{A}_{0,t+k}|I_t^1) - E(\tilde{A}_{0,t+k}|I_t^2) + E\left(\tilde{\mathbf{\Phi}}_{t+k,k}|I_t^1\right)\mathbf{y}_t - E\left(\tilde{\mathbf{\Phi}}_{t+k,k}|I_t^2\right)\mathbf{y}_t \\ &+ E\left(\sum_{j=0}^{k-1} \Psi_{t+k,j}|I_t^1\right) e_{t+k-j} - \left(\sum_{j=0}^{k-1} \Psi_{t+k,j}|I_t^2\right) e_{t+k-j} \quad k = 2,3,\ldots \end{aligned} \tag{10}$$

where $\tilde{\mathbf{\Phi}}_{t+k,k}$ and $\tilde{A}_{0,t+k}$ are defined in appendix A. There are three terms in (10): the first two show how the shocks spread through the system through the intercept; the next two how they spread through the lags of y_t and the last two how they spread through future shocks to the structural equations. Note that when a shock hits the systematic component of policy, $IR_y(t,k)$ depends on J_{t+1} only because $\eta_{t+1} = (J_{t+1}^{-1} \otimes I_{np})u_{t+1}$. Also in this case, IR_y are non-explosive.

4 Estimation

The model (6)-(7) is estimated using Bayesian methods. That is, we specify prior distributions for $\theta_0, \Sigma, \Omega_t$, and H_0 and use data up to t to compute posterior estimates of the structural parameters and of continuous functions of them. Since our sample goes from

1960:1 to 2003:2, we initially estimate the model for the sample 1960:1-1977:2 and then reestimate it 103 times moving the terminal date by one quarter up to 2003:2.

Posterior distributions for the structural parameters are not available in a closed form. MCMC methods are used to simulate posterior sequences with the information available 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)): it requires treating the parameters which are time varying as a block in a Gibbs sampler algorithm. Therefore, at each t and in each Gibbs sampler cycle, one runs 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, in each cycle, we need to apply the identification scheme, discarding paths which are explosive and paths which do not satisfy the restrictions we impose. The computational costs are compounded because we need to run the Gibbs sampler more than a 100 times. Convergence was checked using a standard CUMSUM statistic. The results we present are based on 15,000 draws for each t [3].

Because of the heavy notation involved in the construction of posterior distributions and the technicalities needed to produce draws from these posteriors, we defer the presentation of the details of the estimation to appendix B.

5 The Results

The data we use is taken from the FREDII data base of the Federal Reserve Bank of San Louis. In our basic exercise we use the log of (linearly) detrended real GDP, the log of first difference of GDP deflator, the log of (linearly) detrendend M1 and the federal funds rate in that order. Systems containing other variables are analyzed in the next section.

We organize the presentation around four general themes: (i) Do reduced form coefficients display significant changes? (ii) Are there synchronized changes in the structural coefficients of different equations and/or in the structural variances of the model? (iii) Are there changes in the propagation of monetary policy disturbances in the short and the long run? (iv) Would it have made a difference for macroeconomic performance if monetary policy were more aggressive in fighting inflation, in particular, at the end of the 1970's?

5.1 The evolution of reduced form coefficients

Figure 1 plots the evolution of the mean of the posterior distribution of the reduced form coefficients in each of the four equations (top panel) and their change (bottom panel). The first date corresponds to estimates obtained with the information available up to time 1977:3, the last one to estimates obtained with the information up to time 2003:2.

Several interesting aspects of the figure deserve some comments. First, consistent with the evidence of Sargent and Cogley (2001) and (2003) all equations display some coefficient

[3] Total computational time for each specification on a Pentium IV machine was about 100 hours.

variation. In terms of size, the money (third) and interest rate (fourth) equations are those with the largest changes, while variations in the coefficients of the inflation (second) equation are the smallest of all. Second, while changes appear to be stationary in nature, there are few coefficients which display a clear trend over time. For example, in output (first) equation the coefficient on the first lag of money is drifting downward from 0.6 in 1977 to essentially zero at the end of the sample; while in the money equation, the first lagged money coefficient is drifting upward from roughly zero in 1977 to about 0.9 in 2002. In general, and excluding for the 1979-1986 period, coefficients drift is smooth and relatively small. Perhaps more importantly, there is little evidence of a once-and-for-all structural break in the coefficients of the output and inflation equation (i.e. coefficients do not jump at some date and stays there afterward). Third, the majority of the changes appear to be concentrated at the beginning of the sample. The period 1979-1982 is the one which displays the most radical variations; there is some coefficient drift up to 1986, and after that date variations appear to be random and small. These changes seem to involve primarily the coefficients on the first lag of money (this is the case in three equations) or of interest rates (one equation). Finally, centered 68% posterior bands for the coefficients at the beginning (1977:3) and at the end of the sample (2003:2) overlap in many cases. Therefore, barring few relevant exceptions, instabilities appear to be associated with the Volker (1979-1982) experiment and the adjustments following it. Furthermore, they are temporary and mean reverting in nature.

To go beyond the documentation of patterns of time variations in reduced form coefficients and study whether monetary policy is responsible for the changes, we next examine the dynamics of structural coefficients.

5.2 Structural time variations

As figure 1, the upper panel of figure 2 presents the evolution of the posterior mean of lagged structural coefficients of each equation and the bottom panel their changes at each date in the sample. The first date corresponds again to estimates obtained with the information up to time 1977:3, the last one to estimates obtained with the information up to time 2003:2.

It is immediate to notice that changes in the structural coefficients are typically larger and more generalized than those in the reduced form coefficients. The output and the monetary policy equations are those displaying the largest absolute coefficient changes - these are up to 4 times as large as the largest absolute changes present in the other two equations. The coefficients of the structural inflation equation are still the most stable ones. Furthermore, except for the money (demand) equation, most the variations are concentrated in the first part of the sample, are large in size, statistically and often economically significant. Consistent with the conventional wisdom the money (demand) equation displays trending coefficients (its own first lagged one goes from -0.6 in 1977 to 0.7 in 2003) and large swings in the output and interest rate coefficients from 1991 on. More interestingly from our point of view, there is a pattern in the structure of time variations. The output equation displays two regimes of coefficient variations (one with high variations up to 1986 and one with low variations thereafter) and, within the high volatility regime, the largest coefficient variations occur in 1986. The inflation equation shows the largest coefficient changes up to 1982 and, barring few exceptions, a more stable pattern resulted since then. Finally, our identified monetary policy equation displays large and erratic coefficient changes up to 1986 and coefficients variation is considerably reduced after that. Since the timing of the variations in the structural coefficients of the output and inflation equations are somewhat asynchronous with those of the monetary policy equation, figure 2 casts some doubts on a causal interpretation of the observed changes running from changes in the policy equation to changes in the dynamics of output and inflation.

Figure 3 zooms in on the evolutions of the coefficients of the monetary policy equation (which is normalized to be the last one of the system) and examines in more details the relationship with the structural changes in the rest of the system. Three facts stand out. First, posterior mean estimate of all contemporaneous coefficients are humped shaped: they significantly increase from 1979 to 1982 and smoothly decline afterwards. Second, although all contemporaneous coefficients are higher at the end than at the beginning of the sample, they are typically lower than the conventional wisdom would suggest. In particular, the contemporaneous inflation coefficient peaks at about 1.2 in 1982 and then declines to a low 0.3, on average, in the 1990s. This pattern is also shared by the two lagged inflation coefficients: they both peak in 1982 and smoothly decline afterward. Interestingly, excluding the beginning of 1980s, the sum of coefficients on current and lagged inflation fails to satisfy the so-called Taylor principle. In this sense, Alan Greenspan's regime was only marginally more effective than Arthur Burns's in insuring inflation stability: interest rate responses to inflation movements were barely more aggressive in the 1990s than they were in the 1970s. Note also that, again excluding the beginning of the 1980's, the estimated monetary policy rule displayed considerable stability, in line with the subsample evidence presented, e.g. by Bernanke and Mihov (1998). Hence, the fact that the macroeconomic performance was considerably different in the two time periods seems to suggest that the size and characteristics of the shocks hitting the US economy in the two periods were different. We will elaborate on this issue later on.

Our estimated policy rule displays a six fold-increase in all contemporaneous and first lagged coefficients from 1979 to 1982. Interestingly, this increase is not limited to the inflation coefficients, but also involve output and the money coefficients. The high responsiveness of interest rates to economic conditions is consistent with the idea that by targeting monetary aggregates the Fed forced interest rates to jump to equilibrate a "fixed" money supply with a largely varying money demand - the period was characterized by a number of important financial innovations. The pervasive instability characterizing this period and the subsequent three years adjustments contrasts with the substantial stability of the coefficients of the monetary policy rule in the rest of the sample. Hence, excluding the "Volker experiment", the systematic component of monetary policy has hardly changed over time and if, any change must be noted, it is more toward a decline in the responsiveness of interest rates to economic conditions. Therefore this outcome is consistent with the "business as usual" characterization over the last 30 years put forward by Leeper and Zha (2003) and with the time profile of the policy rule obtained recursively estimating a small scale DSGE model with Bayesian methods (see e.g. Canova (2004)).

The evidence we have so far collected seems to give little credence to the "bad policy" hypothesis. If policy mistakes were made in the late 1970's, they seemed to have been repeated in the 1990's. Still, the dynamics and the volatility of output and inflation were quite different. The "back luck" hypothesis suggests that policy has little to do with the observed changes and that instead it is alterations in the distribution of the shocks hitting the economy that is responsible for the improved macroeconomic outcome.

Figure 4 presents some evidence on this issue. In the top panel we report the evolution of the posterior mean estimate of the variance of the structural forecast errors and, in the bottom panel, the variations produced by its heteroschedastic component, i.e the variations induced by product of the estimated innovations in the coefficient and the regressors of the model. Three features of the figure 4 are of interest. First, the forecast error variance in three of the four equations is humped shaped: it shows a significant increase from 1979 to 1982 followed by a smooth decline. As it happened with structural coefficients, the posterior mean estimate of the variance of the shocks in 2003 is roughly similar in magnitude to the posterior mean estimate obtained in 1978. Second, the time profile of the changes in the forecast error variances of the output and the inflation equations are sufficiently well synchronized with the variations in the forecast error variance of our estimated monetary policy equation. In particular, since 1981, the point estimate of the variance of the forecast error of these three equations displays a common and significantly declining trend. Perhaps more importantly, the declining trend coincide with the decline in the variability observed in the output and inflation coefficients. Third, the contribution of changes in the coefficients to the forecast error variance is much larger in the output and inflation equations than in the other two equations up to 1982. After that date, the contribution is similar across equations. Shocks to the model contribute most to the variability of the forecast error between 1979 and 1982: for example, they account for about 50% of the variance in the output and inflation equation. In general, the proportion of the variance of the forecast error due to the variance of the structural shocks has generally declined after 1982 and the decline is stronger in the inflation equation.

In sum, the variance of the forecast errors in three equations has declined and this decline is equally due to smaller shocks to the model and to a smaller contribution of the heteroschedastic component of the forecast error. Since the decline is, to a large extent, simultaneous, it is likely that the probability distribution from which shocks to the model and to the coefficients were drawn has become less dispersed over time.

5.3 Changes in the propagation of monetary policy disturbances?

Figure 5 reports the posterior mean responses of output and inflation to identified monetary policy shocks for each date at the sample for horizons from 1 to 12 quarters. We do not report interest rate responses because they are similar over time and quite standard in shape and magnitude: after the initial impulse, the increase dissipates rather quickly and becomes insignificantly different from zero after the 3th quarter for each date in the sample. Figure 6 presents 68% posterior confidence bands for the difference in output and inflation responses at selected dates.

The shape of both output and inflation responses is quite similar over time. Output responses are U-shaped; a though response occurs after about 3 quarters and there is a smooth convergence to zero after that. Inflation responses are also slightly U-shaped; the effect at the one quarter horizon is typically the largest, and responses smoothly converge toward zero afterwards. Note also that, consistent with a-priori expectations, output responses are persistent and significant for about 5 quarters at every date of the sample.

Some quantitative difference over time in the mean posterior responses of output and inflation is present but it is small. For output, the posterior mean of the instantaneous response is always centered around -0.15 and the size of the through responses at lag 3 varies in the range (-0.20,-0.05). For inflation, minor magnitude differences occur at lag one (posterior mean varies between -0.07 to -0.16) while in 1978 responses are slightly more persistent than all the others at horizons ranging from 3 to 8.

Differences in inflation responses are both statistically and economically small. In fact, the posterior 68% confidence band for the largest discrepancy (the one at lag 1) includes zero at almost all horizons and if we exclude the initial three years, the time path of inflation responses appears to be unchanged over time. The posterior 68% confidence band for the largest discrepancy in output responses (the one at lag 3) does at times exclude zero. In particular, the trough response in 1982 appear to be significantly deeper than the though response in 1978 and 1979 and at most dates after 1992. Note also that differences are economically insignificant: the maximum difference in the cumulative output multiplier twelve quarters ahead is only 0.5%. In other words, a one percent increase in interest rates produced output responses which differ over time on average by 0.04% points at each horizon.

Overall, the dynamics induced by monetary policy shocks have not significantly changed over time. In particular, and in agreement with the results of section 5.2 responses in the end of the 1990's look very similar in shape and size to those of the end of the 1970's.

5.4 Inflation Dynamics and Monetary Policy

Cogley and Sargent (2001) and (2003) have examined measures of core inflation to establish their claim that monetary policy is responsible for the observed changes in inflation dynamics. They define core inflation as the persistent component of inflation, statistically measured by the zero frequency of the spectrum (that is, by the sum of all autocovariances of the estimated inflation process), and show i) persistence has substantially declined over time and ii) there is synchronicity between the changes in persistence and a narrative account of monetary policy changes. Pivetta and Reis (2004), using univariate conventional classical methods, dispute the first claim showing that differences over time in two measures of inflation persistence are statistically insignificant. Since our study has so far concentrated on short/medium run frequencies, we next investigate whether the longer run relationship between inflation and monetary policy. In particular, we are curious as to whether different frequencies of the spectrum carry different information and whether our basic conclusions on the role of monetary policy for the observed changes in inflation dynamics is robust.

Our analysis differs from existing ones in two important respects: we use output in place of unemployment in the estimated system; we measure persistence using the estimated structural model. While the first difference is minor, the second is not. In fact, thanks the orthogonality of the structural shocks and of the ordinates of the spectrum, our setup allows us not only to describe the evolution of the spectrum of inflation over time, but also to measure of proportion of the spectral power at frequency zero due to monetary policy shocks and describe its evolution over time. From the structural MA representation of the system we have that $\pi_t = \sum_{i=1}^{n} \phi_{it}(\ell) e_{it}$ where e_{it} is orthogonal to $e_{i't}$. Hence the spectrum of inflation at Fourier frequencies ω is $S_{\pi}(\omega) = \frac{1}{2\pi} \sum_{i=1}^{n} |\phi_{it}(\omega)|^2 \sigma_i^2$ and the component at frequency zero due to monetary policy shocks is $S_{\pi}^{*}(\omega = 0) = \frac{1}{2\pi} |\phi_{nt}(\omega = 0)|^2 \sigma_n^2$.

The top panel of figure 7 shows the time evolution of the posterior mean of the spectrum of inflation. On the vertical axis we report the size of the spectral density, and on the two horizontal axis the frequency and the date. The estimate of the zero frequency displays an initial increase in 1978-1980 followed by a sharp decline the year after; since 1981 the estimated posterior mean of the zero frequency of the spectrum has been very stable (with the exclusion of 1991). The initial four fold jump and the following ten fold decrease are visually large and statistically significant. In fact, the bottom panel of figure 7 indicates that the 68% posterior band for the differences between the log spectrum in 1979 and 1996 (the date with the lowest estimates) does not include zero at the zero frequency, contrary to Pivetta and Reis's conclusion. At all other frequencies differences over time both in terms of size and shape are negligible. Hence, except for the zero frequency, the posterior distribution of the spectrum of inflation has also been relatively stable over time.

Figure 8 provides visual evidence on the role that monetary policy had in shaping the inflation persistence. Three important conclusions can be derived. First, the two graphs in the top panel track each other reasonably well suggesting that, in at least in terms of timing, monetary policy shocks are important in determining inflation dynamics. Second, the contribution of monetary policy to inflation persistence varies over time: fluctuations are large and the percentage explained ranges from about 20 to about 75 percent over the 1978-2003 sample. Interestingly, there is a significant trend increase since 1981 and the percentage found in 2002 is roughly the same as it was in 1978. Third, there is a substantial portion of inflation persistence (roughly, 50 percent on average) which has nothing to do with monetary policy shocks. Determining what are the forces which can explain this large percentage is beyond the scope of this paper. Nevertheless, one can conjecture that real and financial factors could account for these variations. AS mentioned, the years between 1978 and 1982 were characterized by financial innovations and high nominal interest rate variability. The pattern present at frequency zero in 1978-1981 is consistent with these two features while the subsequent decline is consistent with the reduction of interest rate volatility produced by a substantially more stable macroeconomic environment and the return to the pre-1979 orthodoxy.

In conclusions, as in Sargent and Cogley, we find visual and statistical evidence of instabilities in the posterior mean of our measure of inflation persistence. Changes in the point estimates of the posterior of inflation persistence go hand in hand with changes in the contribution of monetary policy shocks. Perhaps more importantly, we find that the contribution of monetary policy shocks to variations in the posterior means of inflation persistence is smaller than expected, that factors other than monetary policy are crucial to understand the evolution of inflation persistence, and that the relative contribution of monetary policy has increased since 1981.

5.5 What if monetary policy would have been more aggressive?

It is common in the literature to argue by mean of counterfactuals that monetary policy failed to perform an inflation stabilization role in the 1970s (see e.g Clarida, Gali and Gertler (2000) or Boivin and Giannoni (2002)) and that, had it followed a more aggressive stance against inflation, dramatic changes in the economic performance would have resulted. While exercises of this type are meaningful only in dynamic models with clearly stated microfundations, our structural setup allows us to approximate the ideal type of exercise without falling into standard Lucas-critique type of traps. In fact, to the extent that the monetary policy equation we have identified is structural we can examine what would have happened, e.g., if a shock to the preferences of the Fed would have made the policy response to inflation significantly stronger. Given the estimated distribution of coefficients, we interpret "significantly stronger" as a (permanent) two standard deviations increased in the inflation coefficients above the estimate posterior mean. Figure 9 plots the percentage output and inflation changes from the baseline value of the year which would have been produced in this case at selected dates in the sample. To appropriately interpret the numbers note, e.g., that the maximum inflation response in 1979 (-5 percent) correspond to a 1.0 point absolute decline annual inflation recorded at that date (around 19 percent) and that a 15 percent decline in 2003, at the annual rate of 2.5 percent, corresponds to an absolute fall of less than a 0.4 points.

Overall, a permanent more aggressive stance on inflation would have had important inflation effects in 1979, primarily in the medium run. However, at all dates in the 1980s and 1990s, the effect would have been statistically negligible. Interestingly, if such a policy would have been applied in 2003, it would have produced a small but significant medium run increase in inflation. Such a policy stance however, would not have been painless: important output effects would have been generated. In fact, in 1979 output would have significantly fallen for about four years while a 7 percent fall would have been recorded in 2003 for quite a long time. The Phillips curve trade-off, measured here by the conditional correlation between output and inflation in response to the change, displays an interesting pattern: it is positive and significant in 1979, it is zero in 1983, and it is negative and significant in 1992 and 2003, with the last being statistically significant. While there are many reasons which can account for the change in sign of the tradeoff, a better control of inflation expectations and an improved credibility in the policy environment are clearly consistent with this pattern.

In general, while there was some room for stabilizing inflation in the end of the 1970 and it is not clear that a tougher inflation stance would have been costly in terms of output. There is a sense in which the conventional view is right: being tough on inflation in the end of the 1970s would have produced a different macroeconomic outcome than in the end of the 1990s. However, the reasoning seems to be wrong: being tough on inflation appears to be counterproductive when the mean inflation is low since at low inflation the slope of the Phillips curve trade-off may be different from the standard one. All in all, also in this case, our analysis fails to support the conclusions that monetary policy was to a large extent responsible for the improved macroeconomic performance of the US economy observed since the middle of the 1980s.

6 Robustness analysis

There is a number of specification choices we have made which may affect the results. In the section we analyze the sensitivity of our conclusions to variations in the identification method, in the treatment of trends, and in the variables included in the VAR.

All the results we have presented so far have been produced identifying monetary policy shocks using sign restrictions on the dynamics of money, inflation and output. This allows both contemporaneous and lagged relationship to change with time. Would the results be different if another identification scheme is used? In particular, would the pattern of time variations, the estimated policy rule and the time profile of impulse responses be altered if only lagged coefficients are allowed to vary over time? Figure 10 shows the evolution of the variance of the forecast error and output and inflation responses. obtained identifying policy shocks using a Choleski decomposition. Since here contemporaneous coefficients are time invariant, the evolution of structural coefficients reproduces the pattern of time variations present in the reduced form coefficients (they are simply multiplied by a constant). Therefore, the arguments of subsection 5.1 apply here without a change. Overall, the main conclusions we have derived hold also restricting contemporaneous coefficients to be time invariant. In particular, there are time variations in the coefficients but they are not synchronized across equations; the sum of the inflation coefficients in the interest rate equation is less than one at every date in the sample (maximum value is 0.8); the evolution of the estimated forecast error variance reproduces the one present in figure 4. Finally, impulse responses are broadly similar across time. Clearly, there are changes in pattern of responses relative to our baseline case - inflation increases for at least a year after an interest rate shock. Nevertheless, it is still true, for example, that the maximum discrepancy in the output responses occurs at the third lag and now the posterior 68% band for the difference between 1982 and any other date includes zero. Similarly, the 68% posterior band for the differences between inflation responses between 1978 and any other date includes zero.

Some feel uncomfortable with dynamic exercises conducted in a system where linearly detrended output and linearly detrended money are used. One argument against this choice is that after these transformations these two variables are still close to be integrated and are not necessarily cointegrated. Hence, the dynamics we trace out may be spurious. A second argument, put forward in Orphanides (2001) has to do with the fact that measures of the output gap obtained linearly filtering the data are plagued by measurement error. This measurement error is presumably reduced when output growth is considered. To verify if arguments of this type alter our conclusions we have repeated estimation using the growth rate of output and of M1 in place of the detrended values of output and M1. A sample of the results appears in figure 11 where we plot the evolution of contemporaneous policy coefficients, the variances of the forecasts errors and the time profile of output and inflation responses to a policy shock, identified using sign restrictions. Once again, our basic conclusions remain. In particular, the variance of all four forecast errors decline significantly over time: the variability of GDP and inflation forecast errors in the 1990's is about half what it was in the 1980's and 1970's. Furthermore, the sum of coefficients on inflation in the policy equation is less than one, and now for the entire sample. Finally, the transmission of monetary policy shocks displays remarkable stability and numerical difference emerge only in the response of inflation in the medium run, which was stronger at the beginning of the sample than at the end. Interestingly, figure 11 suggests that policy coefficients are even more stable than in our baseline exercise.

We have examined the sensitivity of our conclusions also 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. A-priori it is hard to know what variables are more important and to check if our system effectively marginalized the influence of all relevant variables. We have therefore repeated our exercise substituting the unemployment rate to detrended output. Figure 12 reports the evolution of the posterior mean of the contemporaneous policy coefficients, of the variances of the forecast errors and the responses of unemployment and inflation to policy shock identified via sign restrictions. Also in this case, our conclusions appear to be robust: excluding the coefficient on money, variations in the policy coefficients are small; there is a significant decline in the uncertainty surrounding the economy since the beginning of the 1980s; the transmission of monetary policy shocks is similar over the entire time period.

Finally, it is now common to examine monetary policy in empirical and theoretical models which exclude money. While we believe that such an exercise is dangerous in a system like ours since omission of money may cause identification problems (demand and supply of currency can not be disentangled) and since money was a crucial ingredient in the considerations that shaped monetary policy decisions, at least up to the end of the 1980's. Commentators have argued that the inclusions of money in the policy rule may lead to an improper characterization of the policy decision of the Fed, especially during the Greenspan's tenure. Figure 13 presents a sample of the results obtained with a trivariate system which excludes money. Also in this case, our basic conclusions are robust. Policy coefficients are stable over time; forecast error variances decline; with the exclusion of inflation in 1978, output and inflation responses have similar size and shape over time. Interestingly, the contemporaneous output coefficient obtained in policy rule is counterintuitively negative and significant, suggesting that an important misspecification is likely to be present in such a system.

7 Conclusions

This paper provides novel evidence on the contribution of monetary policy to the structural changes observed in the US economy over the last 25 years. We use a time varying structural VAR model to evaluate the magnitude of the changes in the coefficients of different equations and in the variance of the forecast errors, and the synchronicity in the timing of the changes. Our framework also allows us to assess how much time variation there is in the propagation of policy shocks both in the short and in the long run and to run some counterfactuals, to understand whether changes in the systematic component of policy would have significantly altered macroeconomic performance.

Structural disturbances are identified using sign restrictions, but the main trust of the results is independent of the identification scheme and the specification of the VAR. Because time variations in the coefficients induce important non-linearities, we provide and implement a new definition of impulse responses, based on the difference of two conditional expectations with different arguments in their conditioning set.

Our results indicate that while there are changes in the estimated structural relationships, they tend to be localized in time and involve particular coefficients in certain equations. We show that, if we exclude the beginning of the 1980s, the monetary policy rule has been quite stable over time and that the posterior mean of the (sum of) inflation coefficients fails to satisfy the so-called Taylor principle, not only in the 1970s but also in the 1990s. We also find evidence of a generalized decrease in the uncertainty surrounding the US economy over time whose timing of the changes roughly coincides with those in the output and inflation equations. Taken together, these facts suggest that the "bad luck" hypothesis has considerable more posterior support than the "bad policy" hypothesis in accounting for the observed changes in the US economy.

We show that the transmission of monetary policy shocks to output and inflation has also been very stable over time. Furthermore while there is posterior evidence that inflation persistence has changed over time, and that changes are related to the magnitude of monetary policy shocks, we also find that non-monetary factors account for a considerable percentage of inflation persistence and that the relative contribution of monetary policy is increasing since 1981.

Finally, we investigate the claim that a more aggressive stance on inflation would have made a difference on output and inflation dynamics. We find that such a policy would have reduced inflationary pressures in the medium run in 1979 but not afterwards and that the output costs of such a policy would not have been negligible. Hence, while there was room for improvement in the conduct of monetary policy at the end of the 1970s, it is unlikely that the observed changes in the Us economy comes from changes in the preferences of the Fed. not that clear that the economy would have benefitted from a tougher inflation stance except perhaps in the 1990s.

Since our conclusions go against several preconceived notions existing in the literature, it is important to highlight what are the features of our analysis which may be responsible for the differences. As repeatedly emphasized, our analysis uses a structural model. Previous studies which used the same level of econometric sophistication (such as Cogley and Sargent (2001) (2003)) have concentrated on reduced form estimates and were forced to use the timing of the observed changes to infer the contribution of monetary policy to changes in output and inflation. Our approach allows not only informal tests based but also to quantify a-posteriori the relationship between monetary policy and output and inflation dynamics. Relative to earlier studies such as Bernanke and Mihov (1998), Hanson (2001) or Leeper and Zha (2003), which use subsample analyses to characterize the changes over time in structural VAR coefficients, we are able to precisely track the evolution of the coefficients over time and produce a more complete and reliable picture of the relatively minor changes in the monetary policy stance in the US.

Our results agree with those obtained recursively estimating a small scale DSGE model with Bayesian methods (see Canova (2004)) and contrast with those of Boivin and Giannoni (2002) who use an indirect inference principle to estimate the parameters of a DSGE model over two subsamples. We conjecture that differences in the estimation method could be the responsible for the difference since the latter method may have problems exploring flat likelihood functions. Finally, our results are in line with those of Sims and Zha (2004), despite the fact that in that paper time variations in both the coefficients and the variance are accounted for with a Markov switching approach. Relative to their work, our analysis emphasizes that factors other than monetary policy could be more important in explaining the structural changes witnessed in the US economy over the last 25 years and provides recursive impulse response measures.

While the decline in the variance of the shocks hitting both the economy and the coefficients of its structural representation seems to suggest to exogenous reasons for the changes in the US economy, it is important to emphasize that our conclusions are consistent with the analysis of McConnell and Perez Quiroz (2001) and with the idea that a more transparent policy process has reduced the volatility of agent's expectations over time. It is therefore important to extend the current study, enlarging the number of variables included in the structural model, identifying other sources of shocks and disentangling the factors which may be behind the decline in the volatility of structural shocks. Also, we have repeatedly mentioned that the monetary policy rule failed to satisfy the Taylor principle both in the 1970s in the 1990s. While Davig, et. al (2003) have argued that the optimality of the Taylor principle is model dependent and, e.g. is turned around in a model where both fiscal and monetary policy can switch over time, the similarities in the policy rules in the 1970s and 1990s suggest important questions. Why is it that inflation did not follow the same pattern as in the 1970? What is the contribution of technological changes to this improved macroeconomic framework? We plan to study these and related issues in future work.

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*, 1, 135-147.

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

Canova, F. (2004) "Monetary Policy and the evolution of the Us economy: 1948-2002", available at www.igier.uni-bocconi.it/ Canova.

Canova, F. and G, De Nicolo (2002), "Monetary Disturbances Matter for Business Cycles Fluctuations in the G-7", *Journal of Monetary Economics*, 29, 1131-1159.

Cogley, T. and T.J. Sargent (2001) "Evolving Post-World War II U.S. Inflation Dynamics", *NBER Macroeconomic Annual*, 16,

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

Davig, T., Leeper, E. and Chung, H. (2003), "Monetary and Fiscal Policy Switching" available at http:/mypage.iu.edu/%7Eeleeper/Papers

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*, CXV, 147-180.

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

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

Hansen, L. and Sargent, T. (1998) *Linear Methods for Dynamic Economies*, Princeton, NJ: Princeton University Press.

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

Leeper, E. and Zha, T. (2003) "Small Policy Interventions", *Journal of Monetary Economics*, 50, 1673-1700.

McConnell, M. and Perez Quiroz, G. (2000) "Output fluctuations in the US: what has changed since the early 1980s? ", *American Economic Review*, 90, 1464-1476.

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 62, forthcoming, *Journal of Money, Banking and Credit*

Pivetta, F. and Reis, R.(2004) "The Persistence of Inflation in the US", Harvard University, mimeo

Sargent, T. (1984) "Vector Autoregression, Policy and Advice",*American Economic Review*, 74(2), 408-415.

Sargent, T. (1999) *The Conquest of American Inflation*, Princeton University Press, Princeton, N.J.

Sims, C. (1982) "Policy Analysis with econometric models ", *Brookings Papers on Economic Activity*, 1, 107-164.

Sims, C. (2001) Comment to "Evolving post WWII US inflation dynamics", NBER Macroeconomic Annual, 16,

Sims, C.A. and T. Zha (2004), "Were there Regime switches in US Monetary Policy", Federal Reserve Bank of Atlanta, working paper 2004-14.

Stock, J. and Watson, M (2003) "Understanding Changes in Business Cycle Dynamics", Harvard University, mimeo.

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

Appendix

A. Impulse Responses

The structural model is

$$y_t = A_{0,t} + A_{1,t}y_{t-1} + A_{2,t}y_{t-2} + ... + A_{p,t}y_{t-p} + S \times H_t \times e_t \tag{11}$$

and its companion form is

$$\mathbf{y}_t = \mathbf{A}_{0,t} + \mathbf{A}_t\mathbf{y}_{t-1} + \epsilon_t \tag{12}$$

where $\mathbf{A}_{0,t} = [A_{0,t}, 0, ..., 0]$, $\mathbf{y}_t = [y_t', y_{t-1}', ..., y_{t-p}']'$, $\epsilon_t = [(S \times H_t \times e_t)', 0, ..., 0]'$ are $n(np+1) \times 1$ vectors and

$$\mathbf{A}_t = \begin{pmatrix} A_{1t} & A_{2t} & \cdots & A_{p-1t} & A_{pt} \\ I_k & 0 & \cdots & 0 & 0 \\ 0 & I_k & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & I_k & 0 \end{pmatrix}$$

is an $n(np+1) \times n(np+1)$ matrix. Substituting into the companion form we obtain

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

where $\mathbf{\Phi}_{t+k,j} = \prod_{i=0}^{j-1} \mathbf{A}_{t+k-i}$ for $j = 1, 2, ...$, $\mathbf{\Phi}_{t+k,0} = I_{np}$. Let $\mathcal{S}_{(h,h')}(X)$ be a selection matrix extracting h-rows and h'-columns of the matrix X. Since $y_t = \mathcal{S}_{(n,1)}(\mathbf{y_t})$, setting $\tilde{\mathbf{\Phi}}_{t+k,k} = \mathcal{S}_{(n,n^2p)}(\mathbf{\Phi}_{t+k,k})$ and $\Phi_{t+k,j} = \mathcal{S}_{(n,n)}(\mathbf{\Phi}_{t+k,j})$ we have

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

where $\tilde{A}_{0,t+1} = A_{0,t+1}$; $\tilde{A}_{0,t+k} = A_{0,t+k} + \sum_{j=1}^{k-1} \Phi_{t+k,j}A_{0,t+k-j}$ for $k > 1$; $\Psi_{t+k,j} = \Phi_{t+k,j} \times S \times H_{t+k-j}$ and $\Psi_{t+k,0} = \Phi_{t+k,0} \times S \times H_{t+k} \equiv S \times H_{t+k} = J_{t+k}$.

Partition $e_t = (e_{i,t}, e_{-i,t})$, 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, and $H_t = (h_{it}, h_{-it})$, where h_{it} is a column of H_t corresponding to $e_{i,t}$ and h_{-it} is the matrix formed by remaining $n-1$ columns. Then equation (13) is

$$y_{t+k} = \tilde{A}_{0,t+k} + \tilde{\mathbf{\Phi}}_{t+k,k}\mathbf{y}_t + \sum_{j=0}^{k-1} \Psi^i_{t+k,j}e_{i,t+k-j} + \sum_{j=0}^{k-1} \Psi^{-i}_{t+k,j}e_{-i,t+k-j} \tag{14}$$

where $\Psi^i_{t+k,j} = \Phi_{t+k,j} \times S \times h_{it+k-j}$ and $\Psi^{-i}_{t+k,j} = \Phi_{t+k,j} \times S \times h_{-it+k-j}$.

Shocks to the Non-Systematic Component

Suppose $e_{i,t+1}$ is a shock to the i-th (structural) measurement equation occurring at t+1. Define the conditioning sets $I^1_t = \{y^t, \theta^t, V_t, J_{t+1}, \xi^\delta_{i,t+1}, \xi_{-i,t+1}, \zeta^{t+k}_{t+2}\}$ and $I^2_t = \{y^t, \theta^t, V_t, J_{t+1}, \xi_{-i,t+1}, \zeta^{t+k}_{t+2}\}$.

Taking conditional expectations we obtain

$$\begin{aligned} E\left(y_{t+k}|I_t^l\right) &= E(\tilde{A}_{0,t+k}|I_t^1) + E\left(\tilde{\mathbf{\Phi}}_{t+k,k}\mathbf{y}_t|I_t^l\right) + \\ &+ E\left(\sum_{j=0}^{k-1}\Psi^i_{t+k,j}e_{i,t+k-j}|I_t^l\right) + E\left(\sum_{j=0}^{k-1}\Psi^{-i}_{t+k,j}e_{-i,t+k-j}|I_t^l\right) \end{aligned} \quad (15)$$

$l = 1, 2$. Since e_t and η_t are orthogonal, taking the difference between the two conditional expectations we have

$$\begin{aligned} E\left(y_{t+k}|I_t^1\right) - E\left(y_{t+k}|I_t^2\right) &= E(\Psi^i_{t+k,k-1}e_{i,t+k-k+1}|I_t^1) - E\left(\Psi^i_{t+k,k-1}e_{i,t+k-k+1}|I_t^2\right) \\ &= \Psi^i_{t+k,k-1}e_{i,t+k-k+1} - \Psi^i_{t+k,k-1}E(e_{i,t+k-k+1}|I_t^2) \\ &= \Psi^i_{t+k,k-1}e_{i,t+1} \end{aligned} \quad (16)$$

Shocks to the Systematic Component

Let $\eta_{i,t+1}$ be a shock to the systematic component of the i-th equation and set $\eta_{-i,t+1} = 0$. Taking expectations with respect to I_t^l, $l = 1, 2$ we have

$$E\left(y_{t+k}|I_t^l\right) = E(\tilde{A}_{0,t+k}|I_t^l) + E\left(\tilde{\mathbf{\Phi}}_{t+k,k}|I_t^l\right)\mathbf{y}_t + E\left(\sum_{j=0}^{k-1}\Psi_{t+k,j}|I_t^l\right)e_{t+k-j} \quad (17)$$

Taking the difference between conditional expectations we obtain equation (10).

B. Estimation

Priors

We choose prior densities for the unknowns which gives us analytic expressions for the conditional posteriors of subvectors of the unknowns. Let T be the end of the estimation sample and let K_1 be the number of periods for which the identifying restrictions must be satisfied. Let $H_T = \rho(\omega_T)$ be a rotation matrix whose columns represents orthogonal points in the hypershere and let ω_T be a vector in R^6 whose elements are $U[0, 1]$ random variables. Let $\mathcal{M}_T$ be the set of impulse response functions at time T satisfying the restrictions and let $F(\mathcal{M}_T)$ be an indicator function which is one if the identifying restrictions are satisfied, that is, if $(\Psi^i_{T+1,1}, ..., \Psi^i_{T+K_1,K_1}) \in \mathcal{M}_T$, and zero otherwise. Let the joint prior for θ^{T+K_1}, Σ, Ω_T and ω_T be

$$p(\theta^{T+K_1}, \Sigma, \Omega_T, \omega_T) = p(\theta^{T+K}|\Sigma, \Omega)p(\Sigma, \Omega_T)F(\mathcal{M}_T)p(\omega_T) \quad (18)$$

Here $p(\theta^{T+K}|\Sigma, \Omega_T) \propto I(\theta^{T+K})f(\theta^{T+K}|\Sigma, \Omega_T)$ where $f(\theta^{T+K}|\Sigma, \Omega_T) = f(\theta_0)\prod_{t=0}^{T+K-1} f(\theta_{t+1}|\theta_t, \Sigma, \Omega_t)$ and $I(\theta^{T+K}) = \prod_{t=0}^{T+K} I(\theta_t)$. Hence, $p(\theta^{T+K}|\Sigma, \Omega_T)$ is truncated normal.

We assume that Σ and Ω_T have independent inverse Wishart distributions with scale matrices Σ_0^{-1}, Ω_0^{-1} and degrees of freedom ν_{01} and ν_{02}. We also assume that the prior for θ_0 is truncated Gaussian independent of Σ and Ω_T, i.e. $f(\theta_0) \propto I(\theta_0)N(\bar{\theta}, \bar{P})$. Note that the uniform prior $p(\omega_T)$ is justified by the fact that all the trajectories satisfying the restrictions are a-priori equally likely. Collecting the pieces the joint prior is:

$$p(\theta^{T+K_1}, \Sigma, \Omega_T, \omega_T) \propto I(\theta^{T+K}) F(\mathcal{M}_T) f(\theta_0) \prod_{t=0}^{T+K-1} f(\theta_{t+1}|\theta_t, \Sigma, \Omega_T) p(\Sigma) p(\Omega_T) \qquad (19)$$

Note that when $H_t = I_n$,the prior reduces to

$$p(\theta^{T+K_1}, \Sigma, \Omega_T, H_T) = I(\theta^{T+K}) f(\theta_0) \prod_{t=0}^{T+K-1} f(\theta_{t+1}|\theta_t, \Sigma, \Omega_T) p(\Sigma) p(\Omega_T) \qquad (20)$$

We "calibrate" the prior by estimating a fixed coefficients VAR using data from 1960:1 up to 1969:1. We set $\bar{\theta}$ equal to the point estimates of the coefficients and $\bar{P}$ to the estimated covariance matrix. Σ_0 is equal to the estimated covariance matrix of VAR innovations and $\Omega_0 = \varrho \bar{P}$. The parameter ϱ measures how much the time variation is allowed in coefficients. Although as t grows the likelihood dominates, the choice of ϱ matters in finite samples. We choose ϱ on the basis of the sample size i.e. for the sample 1969:1-1981:2 $\varrho = 0.0025$, 1969:1-1983:2 $\varrho = 0.003$, 1969:1-1987:2 $\varrho = 0.0035$, for 1969:1-1989:2 $\varrho = 0.004$, 1969:1-1995:4 $\varrho = 0.007$, 1969:1-1999:1 $\varrho = 0.008$, 1969:1-2003:2 $\varrho = 0.01$. This range of values implies a quiet conservative prior coefficient variations: in fact, time variation accounts between 0.35 and a 1 percent of the total coefficients standard deviation.

The primary goal of this paper is to compute impulse response functions, which depend on $\Phi_{T+k,k}$, S and H_T. Therefore, we characterize first the posterior of $\theta^{T+K}, \Sigma, \Omega_T$ which are used to construct $\Phi_{T+k,k}$ and S and then describe an approach to sample from them. Note that when identification is achieved via sign restrictions H_T is a matrix of random variable while under standard schemes H_T is a matrix of constants.

Posteriors

To draw posterior sequences we need $p(\omega_T, \theta_{T+1}^{T+K}, \theta^T, \Sigma, \Omega_T | y^T)$, which is analytically intractable. Luckily we can decompose it into simpler components. First, note that

$$\begin{aligned} p(\omega_T, \theta_{T+1}^{T+K}, \theta^T, \Sigma, \Omega_T | y^T) &\equiv p(\omega_T, \theta^{T+K}, \Sigma, \Omega_T | y^T) \\ &\propto p(y^T | \omega_T, \theta^{T+K}, \Sigma, \Omega_T) p(\omega_T, \theta^{T+K}, \Sigma, \Omega_T) \end{aligned} \qquad (21)$$

Second, since the likelihood is invariant to any orthogonal rotation $p(y^T|\omega^T, \theta^{T+K}, \Sigma, \Omega_T) = p(y^T|\theta^{T+K}, \Sigma, \Omega_T)$. Third, $p(\omega_T, \theta^{T+K}, \Sigma, \Omega_T) = p(\theta^{T+K}, \Sigma, \Omega_T) F(\mathcal{M}_T) p(\omega_T)$. Thus

$$p(\omega_T, \theta^{T+K}, \Sigma, \Omega_T | y^T) \propto p(\theta^{T+K}, \Sigma, \Omega_T | y^T) F(\mathcal{M}_T) p(\omega_T) \qquad (22)$$

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

$$p(\theta^{T+K}, \Sigma, \Omega_T | y^T) = p(\theta_{T+1}^{T+K} | y^T, \theta^T, \Sigma, \Omega_T) p(\theta^T, \Sigma, \Omega_T | y^T) \qquad (23)$$

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. Note that

$p(\theta_{T+1}^{T+K}|y^T,\theta^T,\Sigma,\Omega_T) = p(\theta_{T+1}^{T+K}|\theta^T,\Sigma,\Omega_T) = \prod_{k=1}^{K} p(\theta_{T+k}|\theta_{T+k-1},\Sigma,\Omega_T)$ because the states are Markov. Finally, since θ_{T+k} is conditionally truncated normal with mean θ_{T+k-1} and variance Ω_T, we can write

$$\begin{aligned} p(\theta_{T+1}^{T+K}|\theta^T,\Sigma,\Omega_T) &= I(\theta_{T+1}^{T+K})\prod_{k=1}^{K} f(\theta_{T+k}|\theta_{T+k-1},\Sigma,\Omega_T) \\ &= I(\theta_{T+1}^{T+K}) f(\theta_{T+1}^{T+K}|\theta_T,\Sigma,\Omega_T) \end{aligned} \tag{24}$$

The second term in (23) can be factored as

$$p(\theta^T,\Sigma,\Omega_T|y^T) \propto p(y^T|\theta^T,\Sigma,\Omega_T)p(\theta^T,\Sigma,\Omega_T) \tag{25}$$

The first term is the likelihood function which, given the states, has a Gaussian shape so that $p(y^T|\theta^T,\Sigma,\Omega_T) = f(y^T|\theta^T,\Sigma,\Omega_T)$. The second term is the joint posterior for states and hyperparameters and can be written as:

$$p(\theta^T,\Sigma,\Omega_T|y^T) \propto f(y^T|\theta^T,\Sigma,\Omega_T)p(\theta^T|\Sigma,\Omega_T)p(\Sigma,\Omega_T) \tag{26}$$

The conditional density for the states can be written as $p(\theta^T|\Sigma,\Omega_T) \propto I(\theta^T)f(\theta^T|\Sigma,\Omega_T)$ where $f(\theta^T|\Sigma,\Omega_T) = f(\theta_0|\Sigma,\Omega_0)\prod_{t=1}^{T} f(\theta_t|\theta_{t-1},\Sigma,\Omega_t)$ and $I(\theta^T) = \prod_{t=0}^{T} I(\theta_t)$, thus

$$p(\theta^T,\Sigma,\Omega_T|y^T) \propto I(\theta^T)f(y^T|\theta^T,\Sigma,\Omega_T)f(\theta^T|\Sigma,\Omega_T)p(\Sigma,\Omega_T) = I(\theta^T)p_u(\theta^T,\Sigma,\Omega_T|y^T) \tag{27}$$

where $p_u(\theta_T,\Sigma,\Omega_T|y^T) = f(y^T|\theta^T,\Sigma,\Omega_T)f(\theta^T|\Sigma,\Omega_T)p(\Sigma,\Omega_T)$ is the posterior density obtained if no restrictions are imposed. Collecting pieces we finally have

$$\begin{aligned} p(\omega_T,\theta_{T+1}^{T+K},\theta^T,\Sigma,\Omega_T|y^T) \propto\ & F(\mathcal{M}_T)p(\omega_T) \\ & \Big[I(\theta^{T+K})f(\theta_{T+1}^{T+K}|\theta^T,\Sigma,\Omega_T)I(\theta^T)p_u(\theta^T,\Sigma,\Omega_T|y^T)\Big] \end{aligned} \tag{28}$$

Note that for $H_t = I$, $p(\theta_{T+1}^{T+K},\theta^T,\Sigma,\Omega_T|y^T) = I(\theta^{T+K})f(\theta_{T+1}^{T+K}|\theta^T,\Sigma,\Omega_T)p_u(\theta^T,\Sigma,\Omega_T|y^T)$.

Drawing from the posterior of structural parameters

Given (28) draws for the structural parameters can be obtained as follows

1. Draw $(\theta^T,\Sigma,\Omega_T)$ from the unrestricted posterior $p_u(\theta^T,\sigma,\Omega_T|y^T)$ via the Gibbs sampler (see below). Apply the filter $I(\theta^T)$.

2. Given $(\theta^T,\Sigma,\Omega_T)$, draw a sequence of future states θ_{T+1}^{T+K}, i.e. obtain draws of u_{T+k} from $N(0,\Omega_{T+k})$ and iterate in $\theta_{T+k} = \theta_{T+k-1} + u_{T+k}$, K times. Apply the filter $I(\theta^{T+K})$.

3. Draw $\omega_{i,T}$ for $i = 1,...,6$ from a $U[0,1]$. Construct $H_T = \rho(\omega_T)$.

4. Given Σ, find the matrix S, such that $\Sigma = SS'$. Construct J_t^{-1}.

5. Compute $(\Psi_{T+1,1}^{i,\ell},...,\Psi_{T+K,K}^{i,\ell})$ and $F(\mathcal{M}_T)^\ell$ for each replication ℓ. Apply the filter $F(\mathcal{M}_T)$ and keep draw if the identification restrictions are satisfied.

Drawing from the posteriors of reduced form parameters

The Gibbs Sampler we use to compute the posterior for the reduced form parameters iterate on two steps. The implementation is identical to Cogley and Sargent (2001).

• Step 1: States given hyperparameters
Conditional on y^T, Σ, Ω_T, the unrestricted posterior of the states is normal and $p_u(\theta^T|y^T, \Sigma, \Omega_T) = f(\theta_T|y^T, \Sigma, \Omega_T) \prod_{t=1}^{T-1} f(\theta_t|\theta_{t+1}, y^t, \Sigma, \Omega_t)$. All densities on the right end side are Gaussian they their conditional means and variances can be computed using the Kalman smoother. Let $\theta_{t|t} \equiv E(\theta_t|y^t, \Sigma, \Omega_t); P_{t|t-1} \equiv Var(\theta_t|y^{t-1}, \Sigma, \Omega_t); P_{t|t} \equiv Var(\theta_t|y^t, \Sigma, \Omega_t)$. Given $P_{0|0}$, $\theta_{0|0}$, Ω_0 and Σ, we compute Kalman filter recursions

$$\begin{aligned} P_{t|t-1} &= P_{t-1|t-1} + \Sigma \\ K &= (P_{t|t-1}X_t)(X_t' P_{t|t-1} X_t + \Omega_t)^{-1} \\ \theta_{t|t} &= \theta_{t-1|t-1} + K_t(y_t - X_t'\theta_{t-1|t-1}) \\ P_{t|t} &= P_{t|t-1} - K_t(X_t' P_{t|t-1}) \end{aligned} \tag{29}$$

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

$$\begin{aligned} \theta_{t|t+1} &= \theta_{t|t} + P_{t|t}P_{t|t+1}^{-1}(\theta_{t+1} - \theta_{t|t-1}) \\ P_{t|t+1} &= P_{t|t} - P_{t|t}P_{t+1|t}^{-1}P_{t|t} \end{aligned} \tag{30}$$

where $\theta_{t|t+1} \equiv E(\theta_t|\theta_{t+1}, y^t, \Sigma, \Omega_t)$ and $P_{t|t+1} \equiv Var(\theta_t|\theta_{t+1}, y^t, \Sigma, \Omega_t)$ are the conditional means and variances of the remaining terms in $p_u(\theta^T|y^T, \Sigma, \Omega_t)$. Thus $f(\theta_t|\theta_{t+1}, y^t, \Sigma, \Omega_t) = N(\theta_{t|t+1}, P_{t|t+1})$. 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
Since (Σ, Ω_t) are independent, we can sample them separately. 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 in an inverse-Wishart posterior, so that $p(\Sigma|\theta^T, y^T) = IW(\Sigma_{1T}^{-1}, \nu_{11}); p(\Omega_T|\theta^T, y^T) = IW(\Omega_{1T}^{-1}, \nu_{12})$ where $\Sigma_{1T} = \Sigma_0 + \Sigma_T$, $\Omega_{1T} = \Omega_0 + \Omega_T$, $\nu_{11} = \nu_{01} + T$, $\nu_{12} = \nu_{02} + T$ and Σ_T and Ω_T are proportional to the covariance estimator $\frac{1}{T}\Sigma_T = \frac{1}{T}\sum_{t=1}^T \varepsilon_t\varepsilon_t'; \frac{1}{T}\Omega_T = \frac{1}{T}\sum_{t=1}^T u_t u_t'$.

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

In our exercises T varies from 1979:1 to 2003:2 (these are the dates at which we compute impulse responses). For each of these T, 10000 iterations 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 well behaved and none is multimodal.

We keeping one every four of the remaining 8000 draws and discard all the draws generating explosive paths. The autocorrelation function of the 2000 draws which are left is somewhat persistent. We could reduce it by taking draws more largely spaced (but this comes at the price of reducing the number of draws which satisfy the identification restrictions and therefore substantially reduce the precision of the estimates. In fact, only about 10% of the draws satisfy the identification restrictions. In the end, we have about 250-300 draws for each date to conduct structural inference.

Computation of impulse responses

Given a draw from the posterior of the structural parameters, the calculation of impulse responses to shocks to the non-systematic component is straightforward. In fact, given the an draw for $(\theta^{T+K}, \Sigma, \Omega, H_T)$ we calculate $\Psi_{T+k,k}$ at each draw, compute the posterior median and the 68% central credible set at each horizon k. To compute responses to shocks to the systematic component we proceed as follows.

1. Draw θ_T, Σ, Ω_T and H_T from their posterior distribution.
2. Compute S and draw sequences for ϵ_{T+1}^{T+K} and u_{T+2}^{T+K}.
3. Fix $\eta_{i,T+1}^{\delta}$, draw $\eta_{-i,T+1}$ from the conditional distribution of $(\eta_{-i,T+1}|\eta_{i,T+1} = \delta)$ and form the vector η_{T+1}^{δ}. Compute $u_{T+1}^{\delta} = (J_{T+1}^{-1} \bigotimes I_{np+1})^{-1}\eta_{T+1}^{\delta}$ and define $u_{T+1}^{\delta,T+k} = \{u_{T+1}^{\delta}, u_{T+2}^{T+K}\}$
4. Using $u_{T+1}^{\delta,T+k}$, ϵ_{T+1}^{T+K}, θ_T compute θ_{T+1}^{T+K} compute $\tilde{\Phi}_{T+k,k}$, $\tilde{A}_{0,T+k}$, $\sum_{j=0}^{k-1} \Phi_{T+k,j}\epsilon_{T+k-j}$ for $k = 1, ..., K$. This is a draw for $E(y_{T+k}|I_t^1)$..
5. Fix $\eta_{i,T+1}^{0}$, draw $\eta_{-i,T+1}$ from the conditional distribution of $(\eta_{-i,T+1}|\eta_{i,T+1} = 0)$ and form the vector η_{T+1}^{0}. Compute $u_{T+1}^{0} = (J_{T+1}^{-1} \bigotimes I_{np+1})^{-1}\eta_{T+1}^{0}$ and define $u_{T+1}^{0,T+k} = \{u_{T+1}^{0}, u_{T+2}^{T+K}\}$
6. Using $u_{T+1}^{0,T+k}$, ϵ_{T+1}^{T+K}, θ_T compute θ_{T+1}^{T+K} compute $\tilde{\Phi}_{T+k,k}$, $\tilde{A}_{0,T+k}$ $\sum_{j=0}^{k-1} \Phi_{T+k,j}\epsilon_{T+k-j}$ for $k = 1, ..., K$ and y_{T+k}^{0}. This is a draw for $E(y_{T+k}|I_t^2)$..
7. Take the difference of the realizations in 4. and 6.

Figures

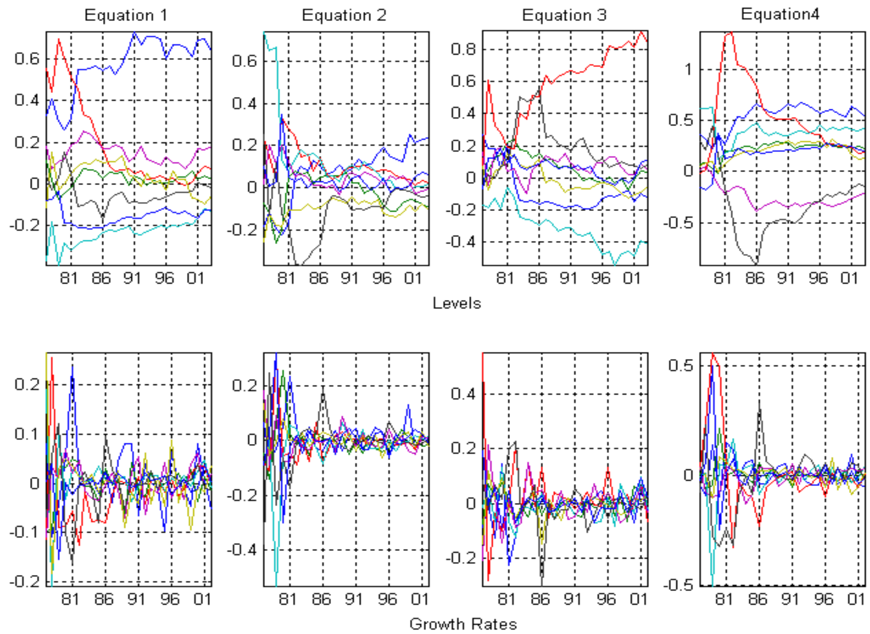


Figure 1: Reduced form coefficients

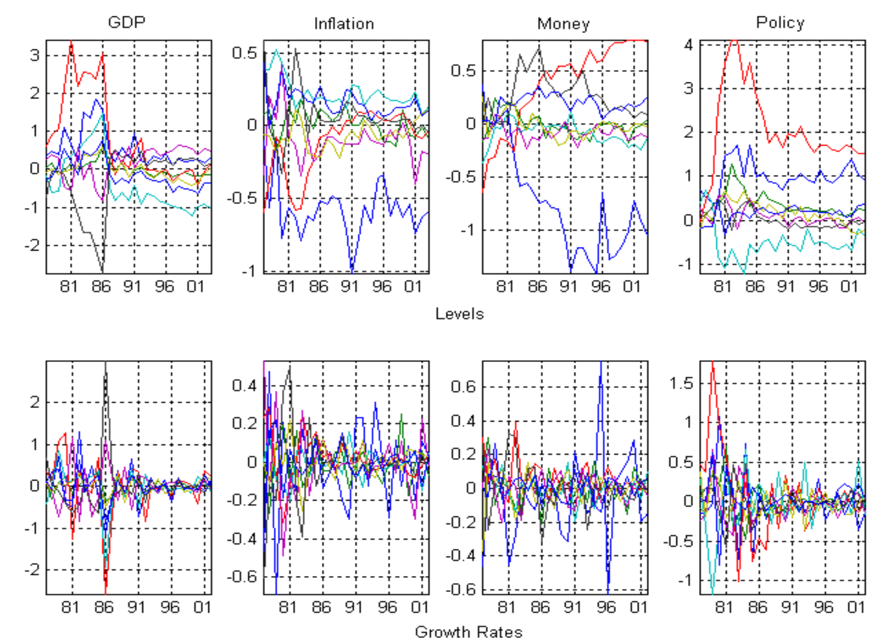


Figure 2: Structural form coefficients

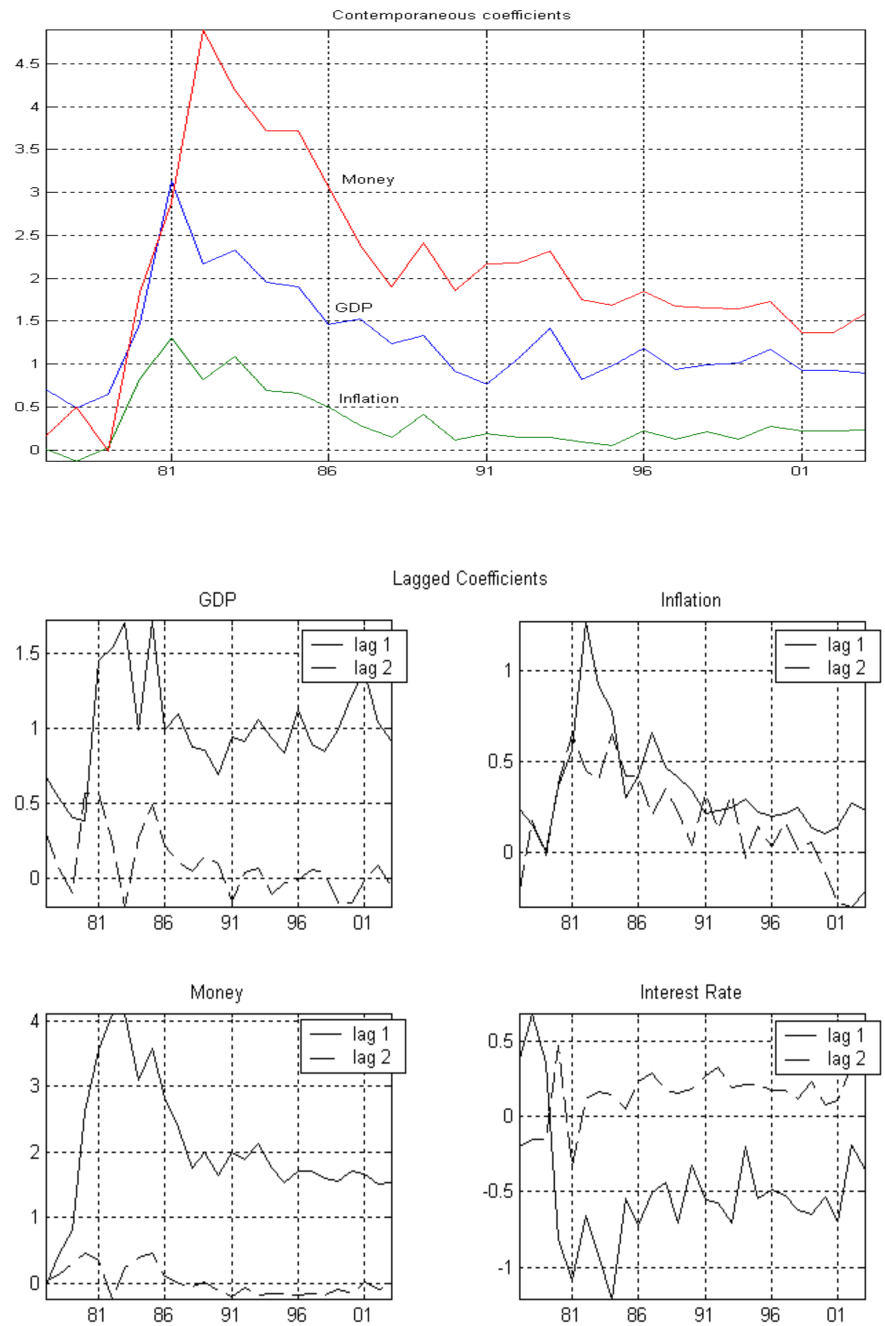


Figure 3: Monetary policy equation

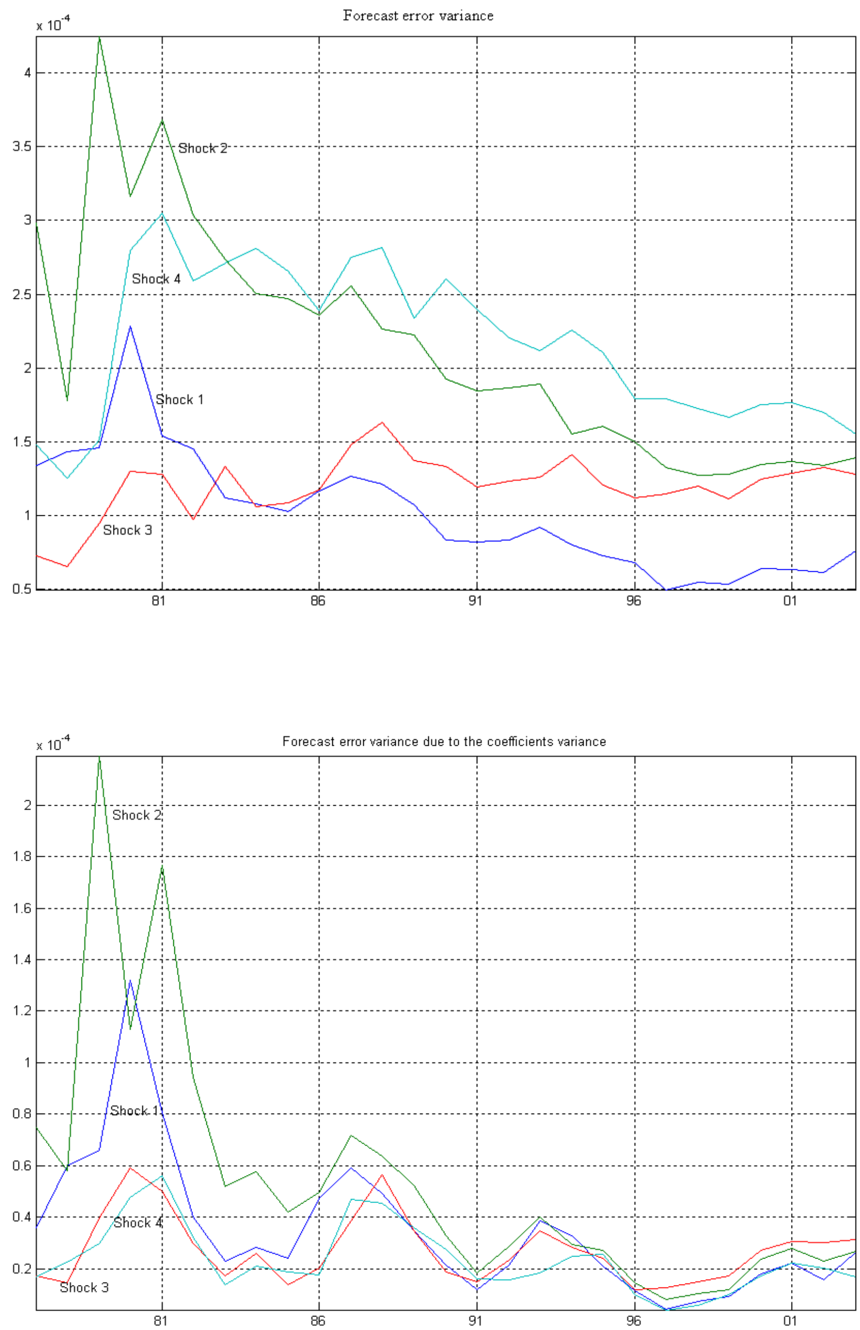
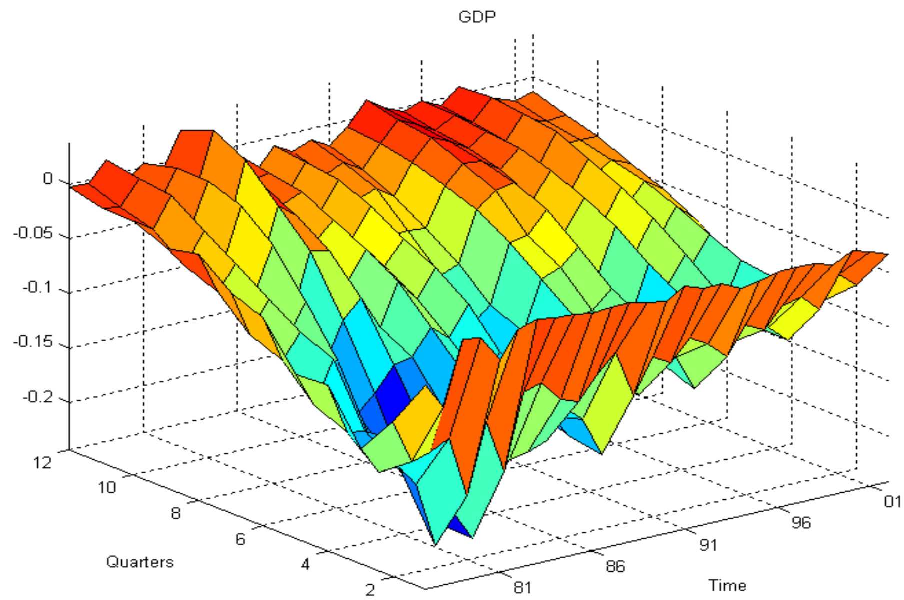


Figure 4: Forecast error variance



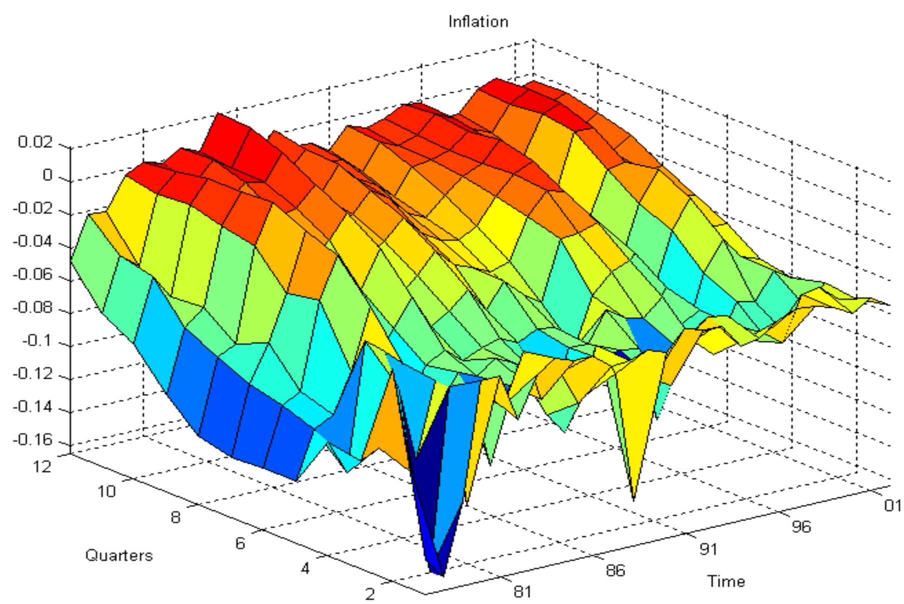


Figure 5: Structural impulse response to monetary policy shocks

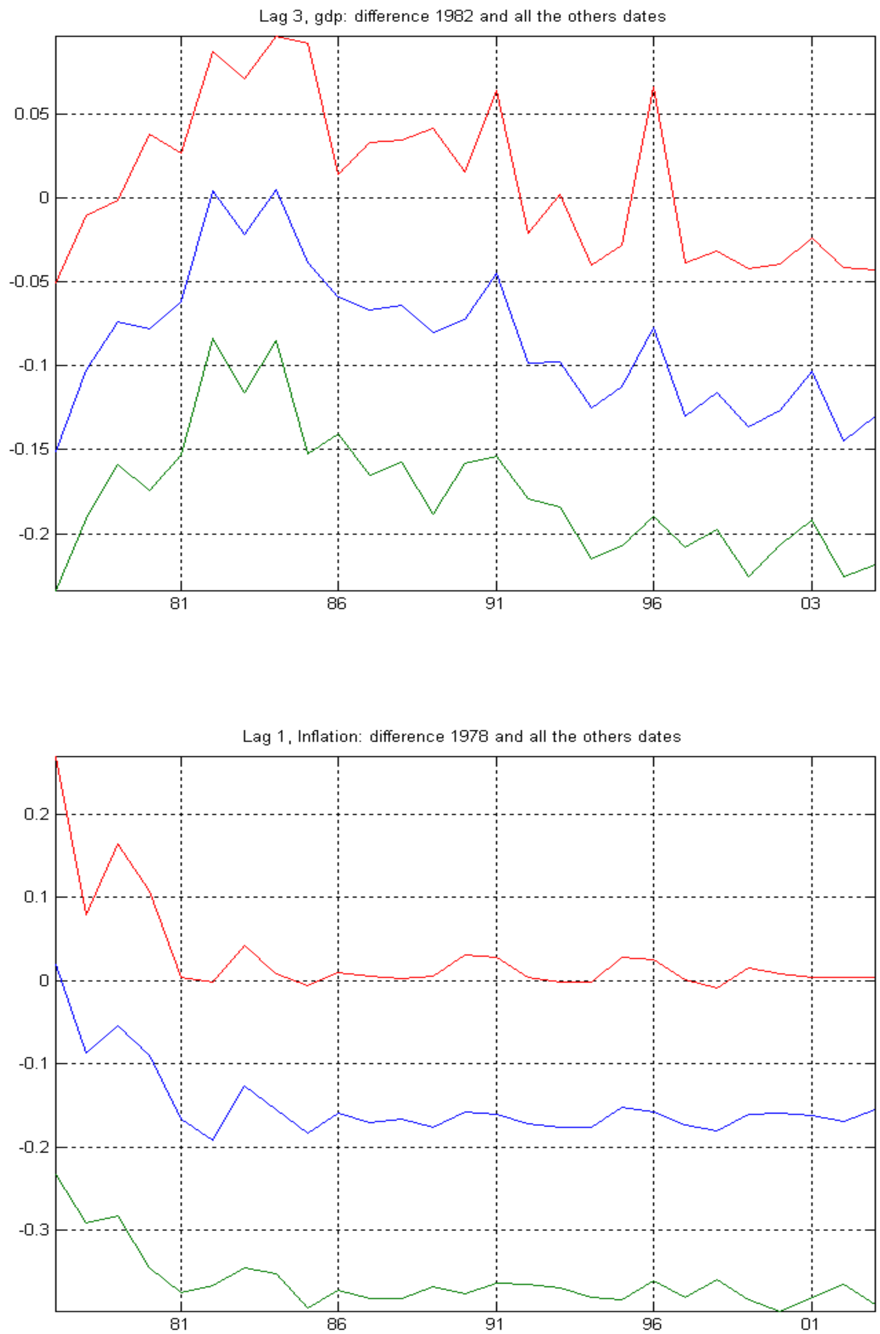
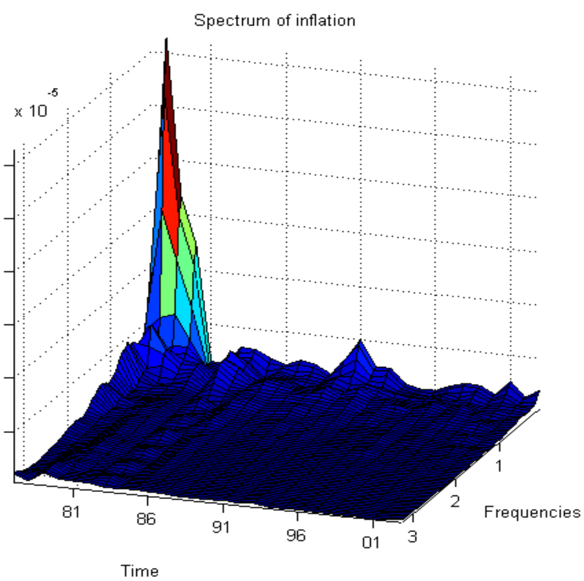


Figure 6: Posterior 68% bands, difference in impulse response functions.



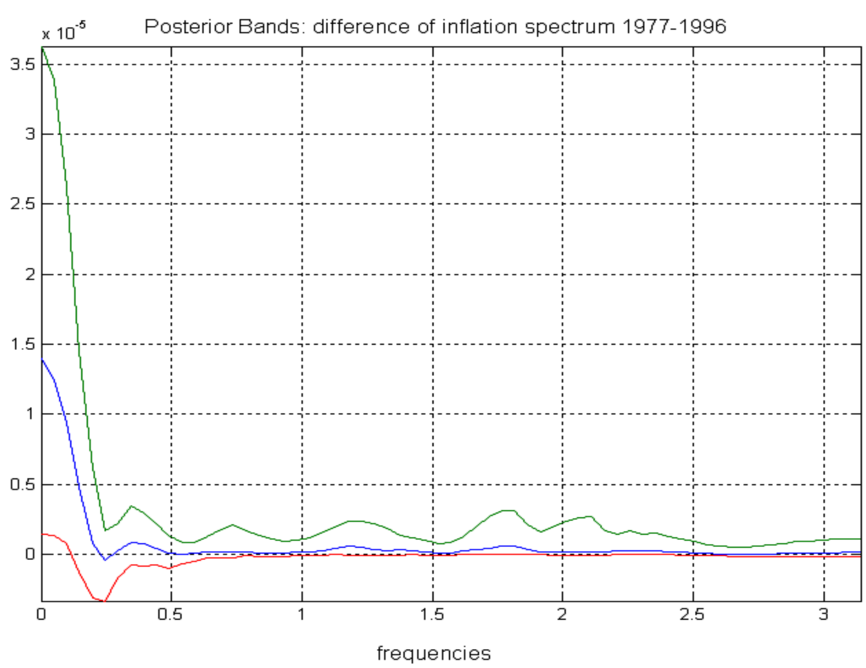


Figure 7: Spectrum of inflation

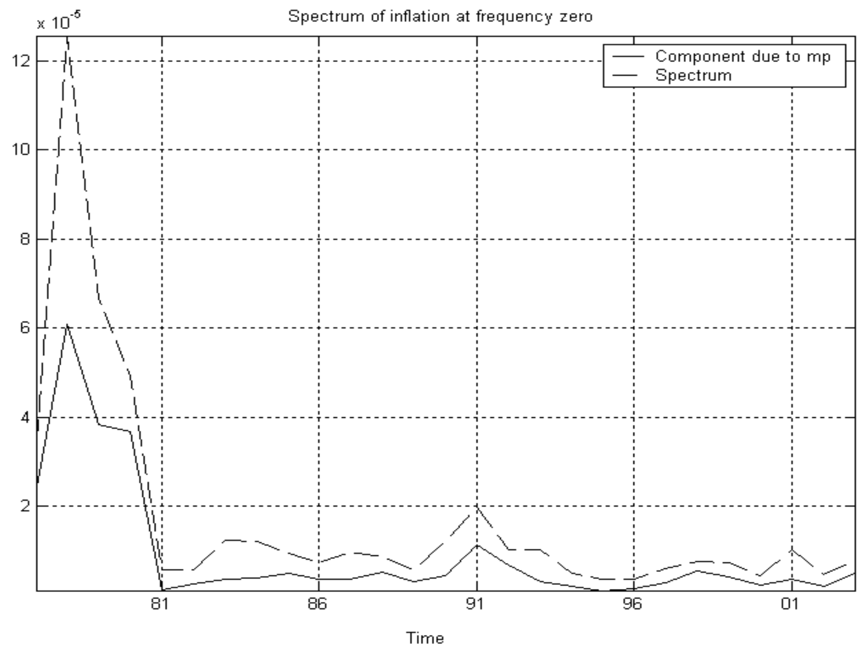


Figure 8: Contribution of monetary policy shocks to inflation persistence

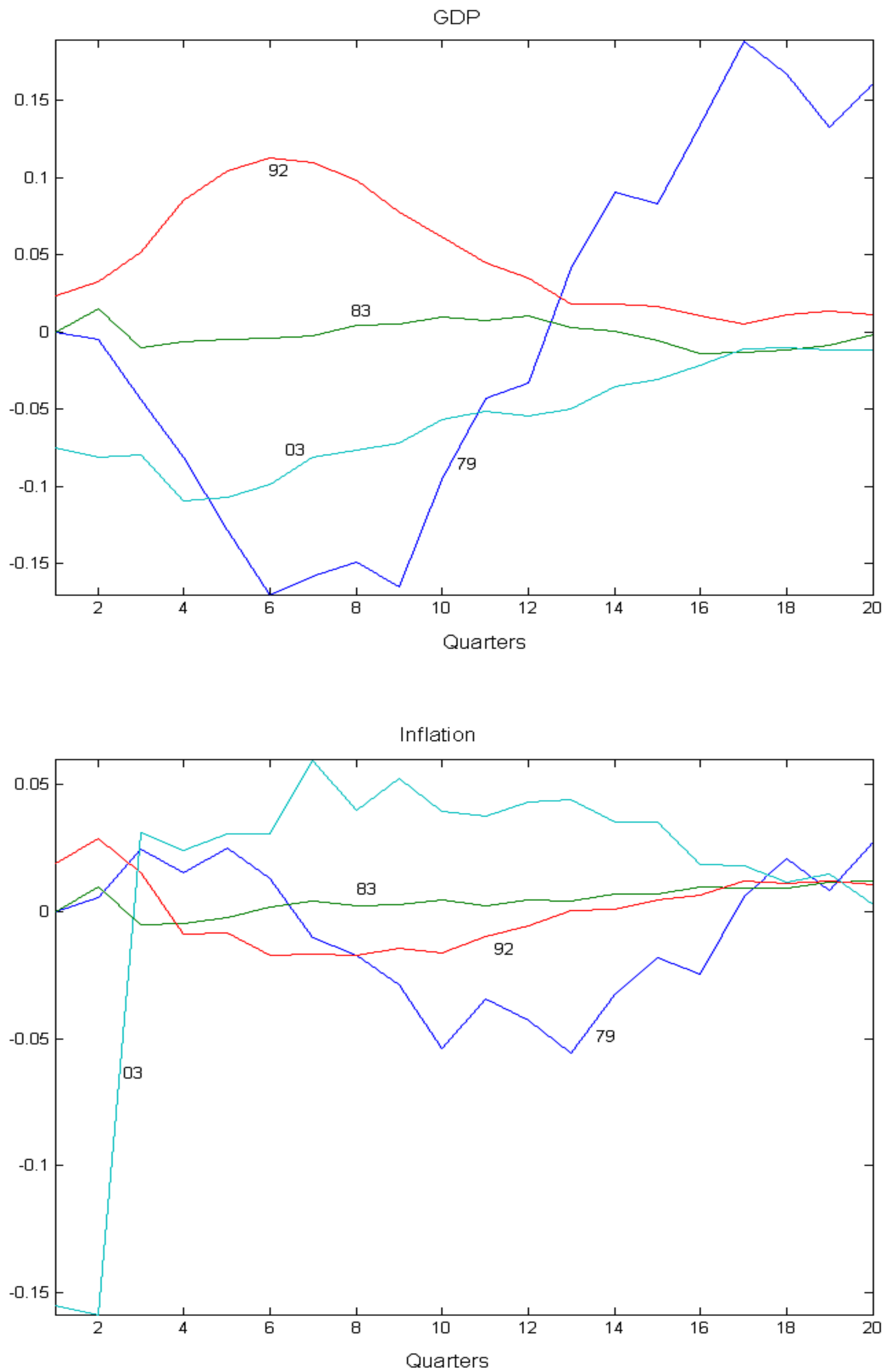


Figure 9: Impulse response functions: more aggressive stance on inflation

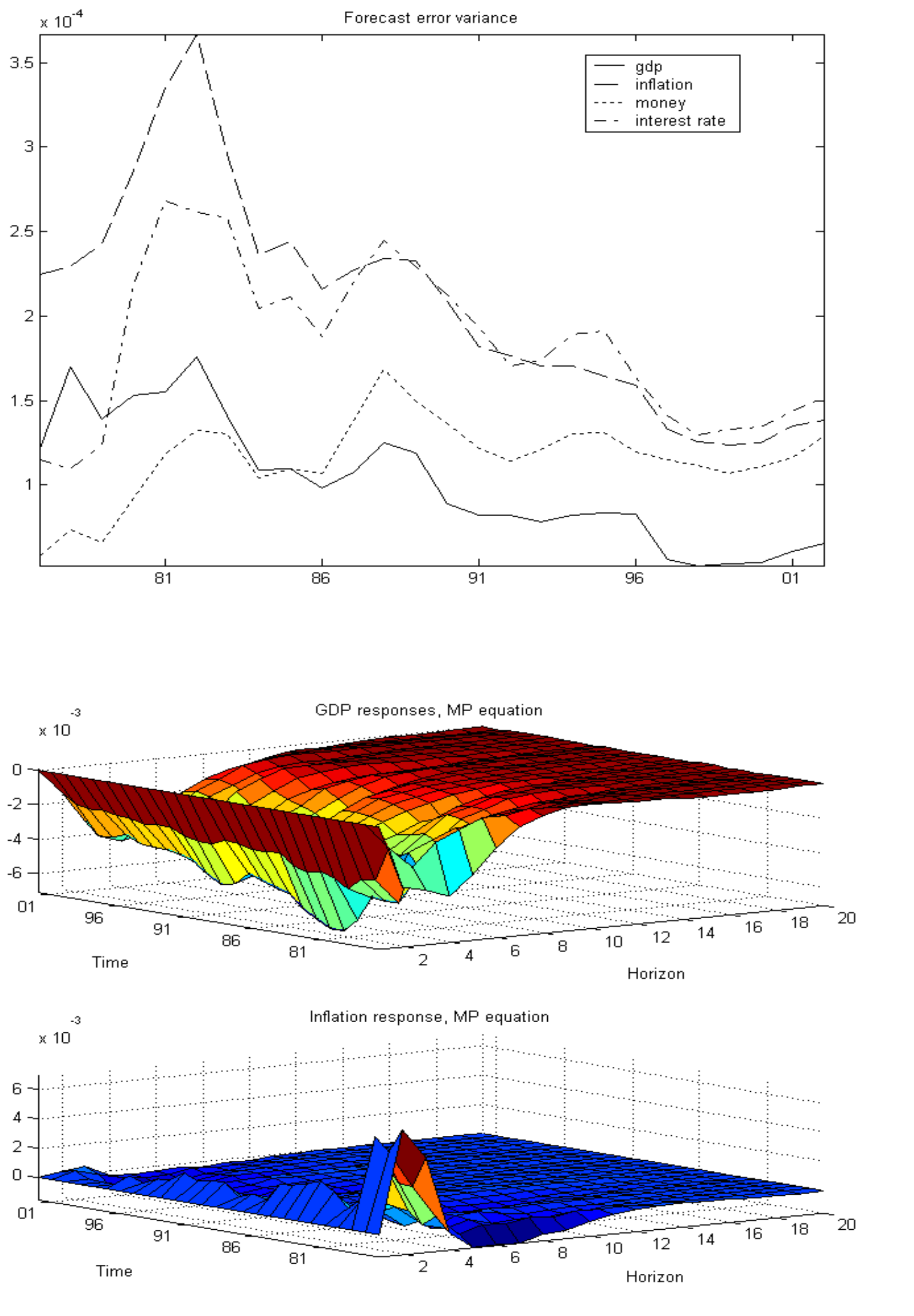


Figure 10: Forecast error variance and impulse response, Cholesky identification.

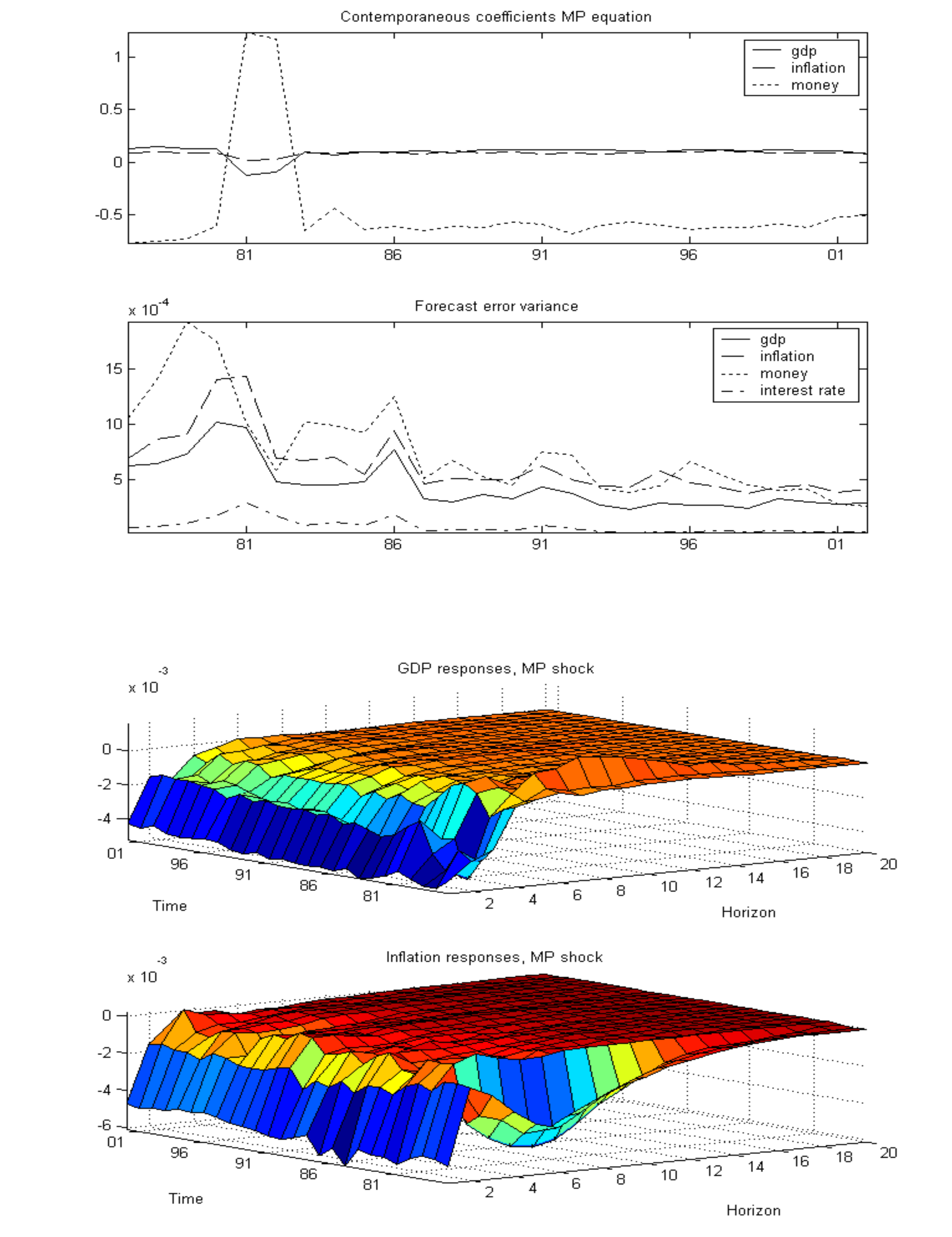


Figure 11: Contemporaneous coefficients, forecast error variance and impulse response functions, output and money in growth rates.

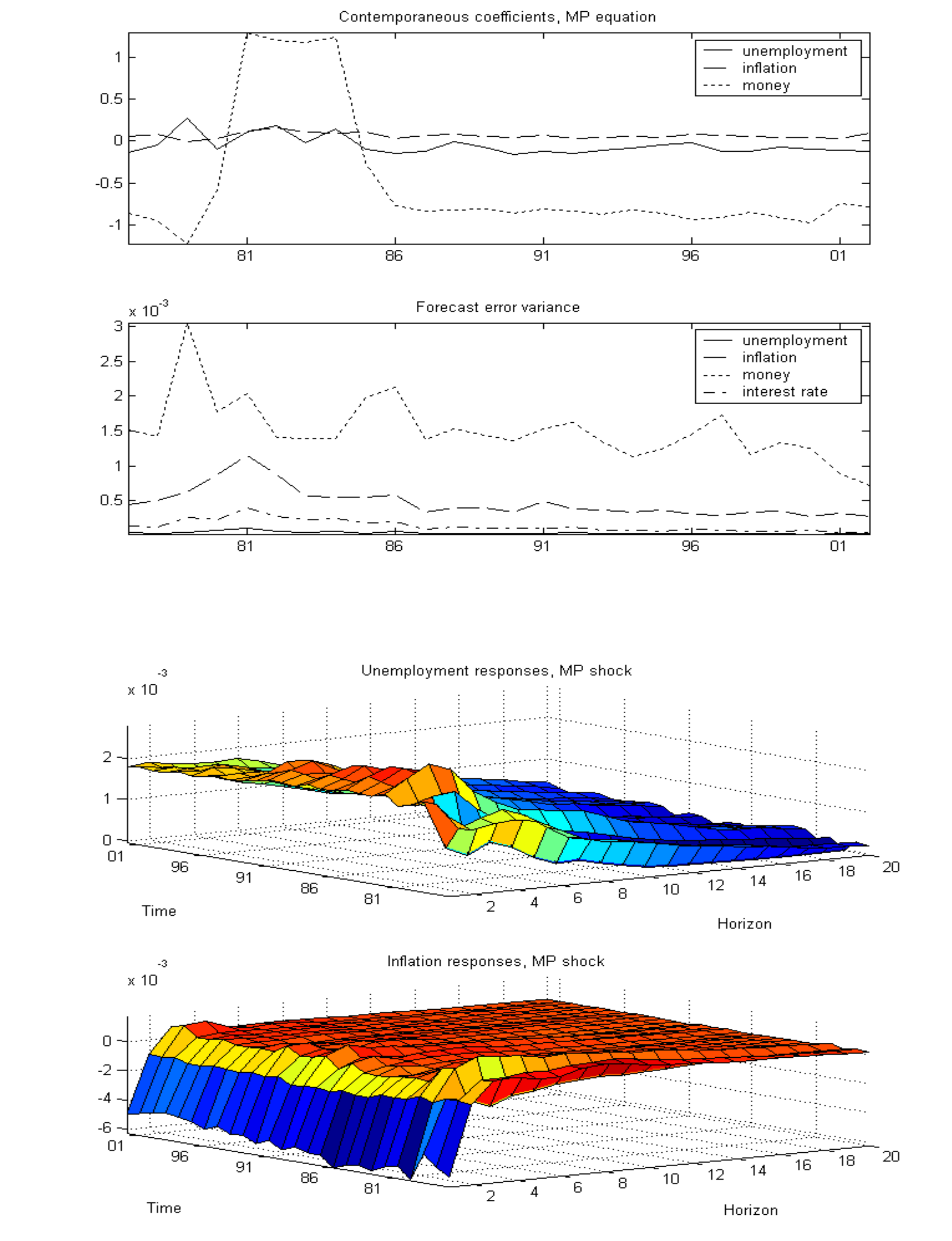


Figure 12: Contemporaneous coefficients, forecast error variance and impulse response functions, unemployment instead of output.

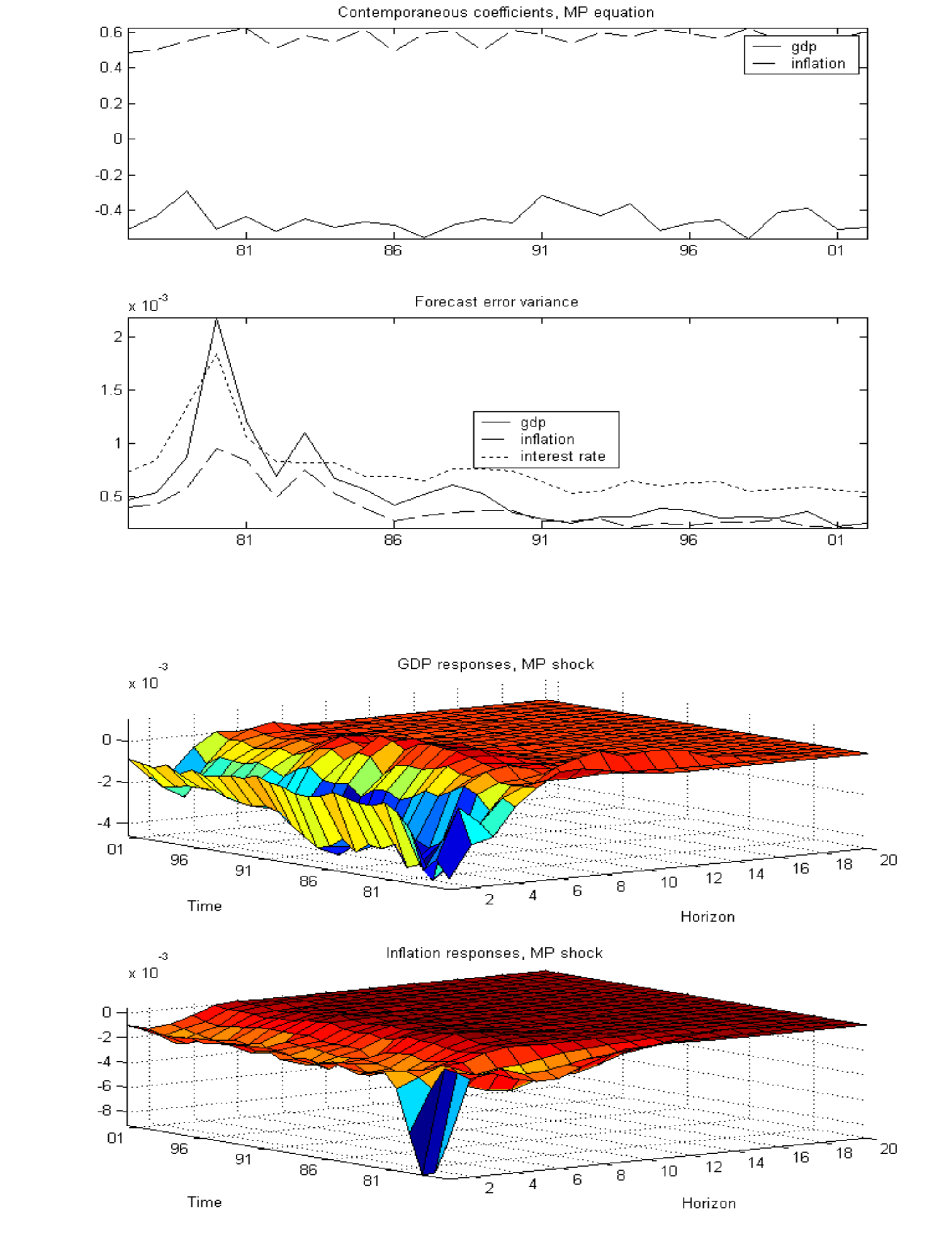


Figure 13: Contemporaneous coefficients, forecast error variance and impulse response functions, system without money.