Assessing the Importance of Learning in an Empirical Monetary Model for the U.S.

Eleonora Granziera*

January 20, 2010

Job Market Paper

Abstract

In this paper we use a time varying coefficient vector autoregression to assess the importance of the learning component in the US postwar economy. The random coefficients are assumed to follow a mean reverting process around an unconditional mean that can be interpreted as the estimates of the coefficients from the reduced form of a rational expectation equilibrium model. The deviations from the unconditional mean are attributed to learning of the agents about the value of the coefficients which regulate the economy. We estimate a monetary model for the post WWII U.S. economy including inflation, output growth and the federal funds rate. We document the presence of learning dynamics and find that the importance of the learning mechanism is somewhat limited for real activity but it is substantial in explaining the dynamics of inflation and interest rate.

KEYWORDS: Learning, Time Varying Parameter Vector Autoregression

JEL: C11, C15, E3, E52

---

* Department of Economics, University of Southern California, 3620 South Vermont Avenue, Los Angeles, CA 90089 USA. Email: granzier@usc.edu. I thank my advisor Roger Moon and Frank Schorfheide for their invaluable guidance. I am grateful to James Bullard, Robert Dekle, Graham Elliot, Mike Owyang and seminar participants at USC and at FRB St. Louis for helpful comments and suggestions.

1 Introduction

The time varying nature of the business cycle and the change in dynamics of key macro variables for the U.S. over the last sixty years have been extensively documented. In particular, many studies report a decrease in the mean and variance of inflation (e.g. Stock and Watson, 2007) and output growth (e.g. Blanchard and Simon 2000, Stock and Watson 2003) starting from the early 1980s. A growing literature adopts time varying parameter vector autoregressions as empirical specification to identify the determinants of these changes. Starting with Cogley and Sargent (2001), which suggest it as estimation methodology, vector autoregressions with time varying coefficients have been used in studying the change in the persistence of inflation and unemployment (Cogley and Sargent, 2005) and of inflation gap (Cogley, Primiceri and Sargent, 2008), in spreading light over the bad luck versus bad policy debate (Primiceri, 2005, Canova and Gambetti, 2004) and in analyzing the causes of persistence of inflation (Cogley and Sbordone, 2008). The rationale underlying their specification is provided by *learning* of the policymaker and private agents about the economy. For example, the central bank might adjust its target inflation rate in view of changes in beliefs about the effectiveness of the monetary policy and the agents might slowly learn about the policy change. Learning mechanisms have been introduced also in dynamic stochastic general equilibrium models (DSGE) to generate endogenous persistence in macro series (Milani 2008). To account for learning in these models, the assumption of rational expectations is replaced by adaptive expectations, that is, agents form their expectations on future values of the variables based on past data. To date a clear link between the empirical and the theoretical literature on learning is missing.

The aim of this paper is twofold: first, to illustrate the link between DSGE models with learning and time varying parameter VAR; second to assess the importance of the learning mechanism in explaining the dynamics of the U.S. postwar data.

The existing literature on time varying parameter VAR characterizes the evolution of the time varying coefficients as a driftless random walk. This parsimonious specification is suited to capture sharp changes in the coefficients. However, despite the emphasys on the learning mechanism as engine for time variation in the coefficients, the random walk specification does not allow to distinguish between variations due to structural changes in the underlying economy from variations due to learning dynamics. Being unable to disentangle these two sources of variations in the coefficients, this specification does not provide with a measure of the impact of the learning component on the dynamics of the macro variables under analysis.

We suggest a way to assess the importance of the learning mechanism by assuming that parameters change over time in a continuous fashion but differently from previous studies we assume that the coefficients evolve according to stationary process rather than as random walks. We interpret the unconditional mean as the value of the coefficients that would result from the solution of a rational expectation model, and the deviation of the coefficients from the mean as consequence of learning dynamics.

The proposed specification allows to decompose our vector autoregression model separating the learning component (the deviation of the coefficients from their unconditional mean) from the rational expectation component, i.e. the term related to the unconditional mean of the coefficients. The empirical model is then consistent with the reduced form from a DSGE model with adaptive expectations in which agents know the correct specification but not the 'true' value of the parameters. Agents recognize that the actual law of motion of the model is characterized by time varying parameters and they obtain estimates of these time varying parameters through Kalman Filtering of the data[1]. Instead, in a model solved under the assumption of rational expectations the agents are endowed with knowledge about the law of motion of the economy which would be characterized by constant coefficients. Because we do not map the parameters from the empirical model with

[1]The literature on learning surveyed by Evans and Honkaphoja (2008) choose recursive least square as updating rule for the estimates of the coefficients by the agents. This updating rule implicitly assumes that agents do not take into account that they will learn also in the future or equivalently agents believe that the actual law of motion of the model exhibits time invariant coefficients. Criticisms to this assumption can be found in McGough (2003) and Bullard and Sutra (2009) among others.

structural parameters from a theoretical macro model of learning, we cannot impose restrictions on the values of the parameters in the empirical model in order to guarantee determinacy and learnability. However, by modeling the law of motion of the parameters as a mean reverting process we guarantee that the coefficients do not depart for too long from their unconditional mean, i.e. from their rational expectation equilibrium value. At the same time the drifting coefficients will not converge to the unconditional mean, and so agents are engaged in a perpetual learning[2]. In our model we neglet the possibility of structural changes in the parameters and we focus our attention exclusively on learning dynamics as possible explanation of the changes in dynamics of the macro series.

A primary objective of the paper is then to quantify the learning component in the US post war data and the extent to which learning might explain the changes in dynamics of inflation, output growth and interest rate. In order to do so, we proceed in two ways: first, we suggest to measure the proportion of the variance of each equation that is accounted for by the learning component. Second, we run a counterfactual example in which we simulate the data in the absence of learning. The empirical analysis delivers the following results: through the first exercise, we find that the importance of the learning mechanism is somewhat limited for inflation and output growth but it is substantial in explaining the dynamics of the federal funds rate, as the learning component accounting for about two thirds of the variance of the interest rate. From the counterfactual example it emerges that in the absence of learning, the inflation series would have been consistently higher than the actual one, while the interest rate series would have lied below the actual series. The magnitude of the discrepancies between actual and simulated series varies over time with peaks occurring during the Volker chairmanships: the value of the simulated interest rate for 1982:Q2 is 3.5%, almost 6 percentage points lower than the actual value, and for inflation the simulated series is 2 percentage points higher than the

[2] In the theoretical macro literature on learning, perpetual learning is assumed by imposing a constant gain parameter in the updating rule for the estimated coefficients. This implies that past data is discounted and so agents 'forget' about older data. Perpetual learning has been shown to improve the fit of DSGE models with learning (see Milani 2005).

actual series from 1977:Q3 to 1980:Q1. For output growth, the chosen measure of real acivity, the actual and simulated series overlap for virtually every date. The results from this empirical analysis suggest that learning behaviour of the monetary authority might explain the change in dynamics of the nominal variables, while the learning mechanism is unable to characterize the changes in the real activity.

The outline of the paper is as follows: section 1 presents a New-Keynesian model with learning that serves as example to clarify the interpretation of learning provided in this paper; in section 2 the econometric methodology is illustrated. Section 3 describes our small empirical monetary model. Section 4 applies the methodology discussed in section 2 to assess the importance of the learning mechanism in the US Post WWII economy and discusses the results; section 5 concludes.

2 Motivating Example: a Small Model with Learning

Although this paper does not map the parameters from the empirical model with the structural parameters derived from a theoretical macro model, we present in this section a simple New-Keynesian model to illustrate our definition of learning and to facilitate the interpretation of the learning dynamics in the empirical model.

Consider a New-Keynesian model summarized by the following log-linearized equations:

$$\pi_t = \kappa x_t + \beta \tilde{E}_t \pi_{t+1} + u_t \tag{1}$$

$$x_t = \tilde{E}_t x_{t+1} - \sigma \left(i_t - \tilde{E}_t \pi_{t+1} \right) + g_t \tag{2}$$

$$i_t = \rho i_{t-1} + (1-\rho) \left[\pi_t^* + \psi_\pi \left(\tilde{E}_t \pi_{t+1} - \pi_t^* \right) + \psi_x \tilde{E}_t x_{t+1} \right] + \eta_t \tag{3}$$

where π_t is inflation, x_t is output gap and i_t is the nominal interest rate, u_t, g_t, η_t are exogenous processes. $\tilde{E}_t$ denotes subjective expectations. Equations (1) through (3) represent a forward looking Phillips curve, a log-linearized Euler equation and a Taylor rule for the monetary authority respectively. Variations of this small monetary model are used for example in Milani (2006, 2008) and Del Negro and Schorfheide (2004). The solution of the model and the notion of equilibrium depend on the formation of the expectations: a rational expectation equilibrium is achieved when the agents take expectations on the distribution of the actual stochastic processes that generate the data. Agents are then endowed with knowledge about the correct structure of the model including the values of the structural parameters. A growing literature, surveyed by Evans and Honkapohja (2008) criticizes the assumption of rational expectations as too restrictive and suggests to substitute it by bounded rationality. This implies that agents know the structure of the rational expectation equilibrium (that is the solution under bounded rationality includes the same variables as the Minimum State Variable solution under rational expectations) but they lack the knowledge of the value of the parameters that govern the economy, and, like econometricians, they *learn* about these parameters by forming estimates based on past data. The implication of this assumption for the New-Keynesian model above is that agents form their expectations using the following 'Perceived Law of Motion':

$$Z_t = \Phi_{0,t} + \Phi_{1,t} Z_{t-1} + \xi_t \tag{4}$$

where $Z_t \equiv \{\pi_t, x_t, i_t\}$ collects the variables included in the Minimum State Variable Solution, $\xi_t \equiv \{u_t, g_t, \eta_t\}$ contains the exogenous processes and the matrices $\Phi_{0,t}$ and $\Phi_{1,t}$ are time varying coefficients, whose estimates, $\hat{\Phi}_{0,t}$ and $\hat{\Phi}_{1,t}$, are updated every period. After having estimated the parameters, the agents use (4) to form their expectations:

$$\tilde{E}_t Z_{t+1} = \hat{\Phi}_{0,t} + \hat{\Phi}_{1,t} Z_t.$$

Substituting back the agents' expectations into the log-linearized model described by equations (1) to (3), yields the Actual Law of Motion of the model, which is therefore characterized by time varying parameters. This result motivates our choice of a time varying coefficient vector autoregression as empirical model. Many papers discuss restrictions on the structural parameters under which the equilibrium under learning converges to the rational expectation equilibrium. In this paper we do not map our reduced form parameters with the parameters from a theoretical model, therefore instead of imposing the learnability or determinacy conditions discussed in Bray and Savin (1986), Evans and Honkapohja (2001) or Bullard and Mitra

(2002)[3], we impose stability by modeling the law of motion of the coefficients as a mean reverting process. Hence, we let the coefficients fluctuate around their unconditional mean, which we interpret as the value taken by the coefficients under rational expectations, but not to depart too much from it. Also, we assume that the variance-covariance matrix of the estimated coefficients is constant over time; this specification is then consistent with perpetual learning, as it implies that the coefficients will never converge to the value they take under rational expectations.

3 Econometric Model

Given the motivation example discussed in section (2), consider the following vector autoregression model with time varying parameters:

$$y_t = \Phi_{0,t} + \Phi_{1,t} y_{t-1} + .. + \Phi_{p,t} y_{t-p} + \varepsilon_t \tag{5}$$

where y_t is an $N \times 1$ vector of endogenous variables, $\Phi_{0,t}$ is a vector of intercepts, $\Phi_{i,t}$, $i = 1, .., p$ are matrices of time varying autoregressive coefficients, ε_t is a vector of normally distributed errors.

Rewrite the model in the following form:

$$y_t = X_t' B_t + \varepsilon_t \tag{6}$$

where X_t' is an $N \times K$ matrix collecting a constant and lagged values of the endogenous variable and B_t is a $K \times 1$ vector of time varying coefficients:

$$X_t' = I_N \otimes \left[\ 1, \ y_{t-1}', \ y_{t-2}', ..., y_{t-p}' \ \right]$$

$$B_t = \left[\boldsymbol{\varphi}_{1,t}^0, \boldsymbol{\varphi}_{1,t}^1, .., \boldsymbol{\varphi}_{1,t}^p, .., \boldsymbol{\varphi}_{n,t}^0, \boldsymbol{\varphi}_{n,t}^1, .., \boldsymbol{\varphi}_{n,t}^p \right]'$$

[3]Note that Bray and Savin (1986), Evans and Honkapohja (2001) and Bullard and Mitra (2002) derive learnability and determinacy conditions under the assumption that the estimates of the parameters are obtained through recursive least square. Our setting is instead consistent with a model in which agents understand that the actual law of motion involves time varying parameters and get estimates of the parameters through the Kalman Filter, as in Bullard (1992) and McGough (2003).

with φ_i^j , $i = 1,..,n,\ j = 1,..p$, being the i-th raw of the matrix of coefficients $\Phi_{j,t}$ and $K = N(Np+1)$. Differently from the previous literature on time varying parameter VAR, which assume that the coefficients of the VAR evolve as driftless random walks, we specify the following law of motion for the vector of coefficients:

$$B_t = (I - \Pi)\bar{B} + \Pi B_{t-1} + v_t \tag{7}$$

with Π a $K \times K$ matrix of coefficients, I an identity matrix of suitable dimention and v_t a vector of errors. B_t is a mean reverting process with unconditional mean $\bar{B}$, provided that the roots of Π lie outside the unit circle. Note that (7) nests the driftess random walk specification if Π equals the identity matrix. The errors from the state and measurement equations are distributed according to: $[\varepsilon_t, v_t]' \sim i.i.d$ $N(0, V)$ and

$$V = \begin{bmatrix} R & 0 \\ 0 & Q \end{bmatrix}$$

so that R is the $N \times N$ covariance matrix for the innovations in equation (6) and Q is the $K \times K$ covariance matrix for the innovations in equation (7). Note that in this specification the errors are assumed to be time invariant and there is no correlation between the errors from the two equations. While the second assumption is standard in the time varying parameter VAR literature, a time variant representation for R is more widely used (Primiceri 2005, Cogley and Sargent 2005), although Cogley and Sargent (2001) and Canova and Gambetti (2004) specify the variance covariance function of the errors to be constant across time.

Define $\tilde{B}_t \equiv B_t - \bar{B}$ as the deviation of B_t from its unconditional mean and rewrite equation (7) as:

$$\tilde{B}_t = \Pi \tilde{B}_{t-1} + v_t \tag{8}$$

The persistence of these deviations is determined by the matrix of autoregressive parameters Π. Then equation (6) can be rewritten as:

$$y_t = X_t' \tilde{B}_t + X_t' \bar{B} + \varepsilon_t. \tag{9}$$

The system of equations (8) and (9) has a Gaussian state space representation where equation (9) is the measurement equation and the equation describing the law of motion for $\tilde{B}_t$ is the state equation. In a time varying framework $\tilde{B}_t$ and B_t are usually called the parameters, while $\bar{B}$, Π, Q and R are called the hyperparameters. We estimate the state space model with Bayesian methods and make use of the Kalman filter to retrieve the value of the time varying coefficients. Our state space representation is non-linear as we impose a stability condition on the roots of the coefficients in the VAR. Following the previous literature, we estimate our model using Bayesian methods.

Let $B^T = [B_1', .., B_T']'$ denote the history of the coefficients B_t up to time T; we are interested in characterizing the joint posterior distributions of the history of parameters and the posterior distribution for the hyperparameters:

$$p\left(B^T, \bar{B}, \Pi, V \mid Y^T\right).$$

The joint posterior for states and hyperparameters can be simulated through Gibbs-sampling by iterating on the conditional distributions in two steps;

step 1: conditional on data and hyperparameters draw a history of states from:

$$p\left(B^T \mid Y^T, \bar{B}, \Pi, V\right);$$

step2: conditional on data and states draw the hyperparameters from:

$$p\left(\bar{B}, \Pi, V \mid B^T, Y^T\right).$$

3.1 Obtaining a history of states

The evolution of the states given the hyperparameters and the data is characterized as:

$$p\left(B_{t+1} \mid B_t, Y^T, \bar{B}, \Pi, V\right) \propto I\left(B_{t+1}\right) p_U\left(B_{t+1} \mid B_t, Y^T, \bar{B}, \Pi, V\right) \quad (10)$$

where $I\left(B_{t+1}\right)$ is an indicator function that takes the value 1 if the eigenvalues of the companion matrix associated to (6) are within the unit circle and zero otherwise. This is to guarantee that at any date the VAR does not exhibit explosive roots. Allowing for unstable roots in the vector autoregression would imply an infinite variance for the macro series included in the model, and therefore in an unplausible representation of the data. The second term in (10) represents the unrestricted posterior density of B_{t+1}. Given the normality assumption for the errors in the state equation and the law of motion (7), $p_U\left(B_{t+1} \mid B_t, Y^T, \bar{B}, \Pi, V\right) \sim N\left((I-\Pi)\,\bar{B} + \Pi B_{t-1}, Q\right)$. The truncation is implemented in the simulation by disregarding the draws that violate the stability condition: if any draw B_τ, $\tau = 1,..,T$ gives rise to unstable roots the whole history of draws for B^τ is rejected[4].

Inference on the state space model above is implemented by conditioning on the hyperparamets and applying the Kalman filter to the state equation (8) after having initialized the state vector $\tilde{B}_0$.

3.2 Obtaining the Hyperparameters

Conditional on the history of states, one needs to sample from the posterior distribution of the hyperparameters. Note that given our stationary autoregressive specification for the law of motion of the states, we need to estimate two additional hyperparameters with respect to the existing literature on TVP VARs, which assume a random walk law of motion for B_t. In particular, we have to obtain the posterior distributions for the unconditional mean $\bar{B}$ and the matrix of autoregressive coefficients Π. Again, note our specification nests the random walk one and that a posterior distribution for Π centered at the identity matrix would provide evidence toward a random walk characterization for the evolution of B_t. We assume a hierarchical prior and posterior distribution for the hyperparameters:

$$p\left(\bar{B},\Pi, Q, R \mid Y^T, B^T\right) = p\left(\bar{B},\Pi \mid Q, R, Y^T, B^T\right) p(Q, R \mid Y^T, B^T) =$$

$$= p\left(\bar{B},\Pi \mid Q, R, Y^T, B^T\right) p(Q \mid Y^T, B^T) p(R \mid Y^T, B^T)$$

The vectors of time varying coefficients B_t and $\tilde{B}_t$ are of dimension $K \times 1$, where $K = N(Np+1)$. In our small model for the US post WWII economy with three

[4] The validity of this strategy is shown in Cogley and Sargent (2001); refer to Koop and Potter (2008) for a discussion on the merit of this and other algorithms for drawing the history of states.

endogenous variables and two lags $K = 21$, so that both Π and Q are matrices of dimension $K \times K = 21 \times 21$. The high dimensionality of these matrices is then a concern for estimation. To overcome the dimentionality issue we impose diagonality of Π and Q. This implies that the random coefficients evolve independently. We assume a normal-inverted gamma prior for each $i-th$ element of the vector $\bar{B}$ and of the diagonal of Π and Q:

$$\begin{aligned} [\bar{B}_i,\ \Pi_{ii}] \quad | \quad & Q_{ii} \sim N(\lambda, Q_{ii}/\tau) \\ Q_{ii} \quad | \quad & \alpha, \beta, B_{i,t}, .., B_{i,1} \ \sim \Gamma^{-1}(\alpha, \beta_i) \end{aligned}$$

A conjugate prior delivers a conjugate posterior. For each $i-th$ equation in (8) the relationship between the parameters of the the priors (indicated with the superscript 0) and those of the posterior (indicated by the superscript T) are as follows:

$$\lambda'_{i,T} = \left[\frac{\tau_0 \lambda_{\hat{\Pi}_{ii},0} + \tau_T \hat{\Pi}_{ii}}{\tau_0 + \tau_T}, \frac{\tau_0 \lambda_{\bar{B}_i,0} + \tau_T \hat{B}_i}{\tau_0 + \tau_T}\right]'$$

$$\alpha_0 = \tau_0$$

$$\alpha_T = \tau_0 + \frac{\tau_T}{2}$$

$$\beta_{i,T} = \beta_{i,0} + \frac{1}{2}\sum_{t=1}^{T}\left(B_{i,t} - \left(1 - \hat{\Pi}_{ii}\right)\hat{B}_i - \hat{\Pi}_{ii} B_{i,t-1}\right)^2 +$$

$$+\frac{\tau_0 \tau_T}{\tau_0 + \tau_T}\left[\hat{\Pi}_{ii} - \lambda_{\hat{\Pi}_{ii},0},\ \hat{B}_i - \lambda_{\bar{B}_i,0}\right]\left[\begin{array}{c} \hat{\Pi}_{ii} - \lambda_{\hat{\Pi}_{ii},0} \\ \hat{B}_i - \lambda_{\bar{B}_i,0} \end{array}\right]$$

where τ_0 and τ_T are the size of the training sample and of the estimation sample respectively, $\hat{\Pi}_{ii}$ and $\hat{B}_i$ are estimated from $B_{i,t} = (1 - \Pi_{ii})\bar{B}_i + \Pi_{ii} B_{i,t-1} + \nu_{i,t}$, and the parameters $\beta_{i,0}$, $\lambda_{\bar{B}_i,0}$, $\lambda_{\hat{\Pi}_{ii},0}$ are chosen arbitrarily.

The prior and the posterior distributions for the variance covariance matrix of the measurement equation, R, take an Inverted Wishart form:

$$R \sim IW(\tau_R R_T, \tau_R)$$

where the posterior distribution parameters can be derived from the prior parameters as:

$$\tau_R = \tau_0 + \tau_T$$

$$R_T = \frac{\tau_0}{\tau_R} R_0 + \frac{\tau_T}{\tau_R} \hat{R}_T$$

$$\hat{R}_T = \frac{1}{\tau_T} \sum_{t=1}^{\tau_T} \hat{\varepsilon}_t \hat{\varepsilon}_t'$$

where R_0 is an arbitrary $N \times N$ matrix and $\hat{\varepsilon}_t$ are the residuals from the measurement equation.

4 A Small Monetary Model for the US Economy

We consider a small empirical model for the US post WWII data, which focuses on variables relevant for monetary policy analysis: the model includes inflation, output growth and short term interest rate. Inflation is computed as annual percentage change in the consumer price index; output growth is obtained as annual percentage change in real GDP, the Federal Funds Rate is used as measure of short term interest rate. A detailed description of the data is provided in the data appendix. Data are collected at the quarterly frequency from 1955Q3 to 2009Q1. The VAR specification includes 2 lags and an intercept term. The coefficients are estimated through a Gibbs sampling algorithm that involves 45000 iterations with the first 5000 discarded to allow for burn in.

4.1 Priors

We use a training sample of 11 years, (fourty-four observations, from 1954:Q3 to 1965:Q3) in order to obtain priors for $\bar{B}$, Π, Q and R and to initialize the state vector. The prior mean and the variance for $\bar{B}$ are obtained as the MLE estimate and its variance from a vector autoregression with constant coefficients on the pre-sample. We impose a standard prior for the diagonal elements in Q: $Q_{ii} \sim \Gamma^{-1}(\tau, gQ * VB\_OLS_{ii})$ where VB_OLS is the long run variance of the coefficients obtained from an OLS regression of a constant coefficients VAR of order two on the training sample, scaled by a factor $gQ = 0.0001$ and τ is the size of the training sample. The prior mean for R is the variance covariance matrix of the residuals from the constant coefficients VAR on the training sample.

While the priors for the hyperparameters above can be chosen with general consensus, there is no guidance for the choice of the prior Π, which is crucial in determining the persistence of the learning dynamics. The choice of a prior mean for Π is less trivial. We arbitrarily assume that deviations from the unconditional mean are larger for the time varying coefficient describing the effect of lagged inflation on itself, of lagged output on itself and of lagged interest rate on itself so we allow for higher prior mean (0.7) for the correspondent entries in Π. We experimented with different values for Π and we document that choosing values for the prior mean of Π larger than 0.7 results in no draw for the history of B_t that satisfy the stability condition. The prior mean for the autoregressive coefficients which characterize the evolution of the intercepts are set to 0.4. All the other entries in the diagonal of Π equal 0.2. We initialize $\tilde{B}_0$ to zero, i.e. we assume that B_t is equal to its unconditional mean up to the beginning of the estimation sample; we also assume that the initial state $\tilde{B}_0$ and the hyperparameters are independent.

5 Importance of Learning

We investigate the importance of learning dynamics in three ways: first we look at the impact of a shock to monetary policy and we ask whether this effect is the same for each period. Any heterogeneity in the impulse response functions is imputed to deviations of the coefficients from their unconditional mean and therefore, given our interpretation, it is evidence in favor of learning dynamics. Second, we quantify the importance of learning by computing the contribution of the learning component to the variance of the variables in the VAR. Last, we run a counterfactual example by simulating the data in absence of learning and we compare the actual and the simulated data.

5.1 Impulse Responses

We study the effect of a tightening of the monetary policy on inflation, output growth and short term interest rate. Identification of the monetary shock is achieved through sign restrictions in the spirit of Uhlig (2005). We assume that after a contractionary monetary policy the federal funds rate increases and both inflation and output decrease. We repeat the analysis imposing the restrictions for up to H= 2 and H=4 horizons after the shock. In a time varying framework, in order to assess the importance of learning dynamics, rather than looking at a representative impulse response function over the sample, we derive impulse response function for each date in the estimation sample. Canova and Gambetti (2004) provide with a formal definition of impulse response functions in the case of time varying coefficients VAR; following their definition, the impulse response functions are constructed taking into account future projections of the time varying coefficients. Figure (1) through (3) show the impulse responses of inflation, output growth and federal funds rate to a shock of one-standard deviation in size for each date for horizon 1 through 20 (the bounds of the credible set are not plotted to make the figures legible) when the sign restriction is imposed only for the first 2 periods after the shock; the responses obtained by imposing the restrictions for up to horizon 4 are analogous to the ones shown. In order to highlight differences in the uncertainty around the median impulse response over time, figure (4) through (6) display the median and the lower and upper bound of the 90th-percent highest posterior density interval of the impulse response functions for each date at selected horizons: on impact, after four quarters, after 2 years and after five years.

From theoretical models we expect a tightening of the monetary policy to increase the short term interest rate, decrease prices and reduce real output. Because of the sign restrictions imposed, our empirical results confirm the predictions for the variables up to the second horizon, but they also show an anomaly for the response of the output growth at longer horizons. A contractionary monetary shock has a significant negative effect on output growth on impact. However, starting from the third quarter after the shock output growth increases sharply for few quarters and it peaks at about 1 percent after one year from the shock; it then declines rapidly. Finally the effect of the shock fades away after 10 quarters. Contrary to theoretical predictament then, output growth is positive for one year. An increase in real activity after a tightening of the monetary policy has been documented also in Uhlig (2005) for a constant coefficient VAR. The contractionary monetary shock decreases the inflation rate on impact, but this effect is short lived, and from the second quarter after the shock inflation sluggishly goes back to its initial value. After 2 years the effect of the monetary shock on inflation disappears. The federal funds rate increases on impact by about 0.6 percentage points and rapidly goes back to its initial value. We just described some features of the impulse response functions common throughout the sample. The object of the analysis however is to highlight the differences in the impulse responses across time. From the tridimentional figures (1) to (3) as well as from figures (4) through (6) it emerges that for all the variables considered the median of the impulse response function exhibits little heterogeneity across the sample. This finding is consistent with Primiceri (2005), which derives trivariates impulse response functions of inflation and unemployment to a contractionary monetary policy shock. Using a time varying parameter VAR in which the coefficients follow a random walk specification and the variance covariance matrix of the errors is assumed to be time varying he constructs impulse response functions for inflation and unemployment rate at three dates in the estimation sample (75:Q1, 81:Q3 and 96:Q1) and finds that the responses are very similar, particularly for unemployment series. In figure (7) through (9) are depicted the median as well as the lower and upper bound of the 90th-percent highest posterior density interval of the sampled impulse response functions of the variables for a horizon of up to 5 years after the shock for 1981:Q3 and 2007:Q3. The first date coincides with the peak in the interest rate series, and it is also a NBER business cycle peak date, while 2007:Q3 is the last observation before the beginning of the current recession. The differences across periods are limited for all the series. The medians are relatively stable, while the confidence bounds show more heterogeneity across periods: In particular, for all variables they are wider.for 81:Q3, suggesting that more uncertainty about the behavior of the variables is associated with that period.

To summarize, the impulse response functions show some heterogeneity across the sample. Any difference in the impulse response across time is due to the deviations of the estimated coefficients from their unconditional mean and are therefore attributable to learning dynamics. Therefore our findings seems to suggest that learning dynamics do play a role in the behavior of the U.S. postwar series under analysis.

5.2 Marginal Effect

In the previous section, we documented thorugh a graphical analysis of impulse responses that learning dynamics are limited in the data. We now consider a different strategy to investigate the extent of learning in the data and we quantify the importance of learning by computing how much of the overall sample variance of the variables under analysis is explained by the learning component. In order to do so, we need to disentangle the contribution of the learning component from the contribution of the rational expectation component to the variance of y_t. Recall from (9) that the equation that describes the evolution of y_t can be decomposed into two separate components: one that is interpreted as the rational expectation component, $X_t'\bar{B}$, and the learning component, $X_t'\tilde{B}_t$. Despite this break-down of the measurement equation, the correlation between $X_t'\tilde{B}_t$ and $X_t'\bar{B}$ does not allow to break down the variance of y_t as sharply. We propose to isolate the marginal effect of the learning component on the variance of y_t by proceding in two steps. Denote $\tilde{Z}_t = X_t'\tilde{B}_t$, $\bar{Z}_t = X_t'\bar{B}$ and denote Y, $\tilde{Z}$, and $\bar{Z}$ as the matrices that stack y_t', $\tilde{Z}_t'$, and $\bar{Z}_t'$, respectively.

1. The first step is to regress Y on $\bar{Z}$ and regress $\tilde{Z}$ on $\bar{Z}$, and obtain the residuals: $M_{\bar{Z}}Y$ and $M_{\bar{Z}}\tilde{Z}$, where $M_{\bar{Z}} = I_T - \bar{Z}(\bar{Z}'\bar{Z})^{-1}\bar{Z}'$.

2. In the second step, we compute the sample correlation between $M_{\bar{Z}}Y$ and $M_{\bar{Z}}\tilde{Z}$.

This is the R^2 of the regression of $M_{\bar{Z}}Y$ on $M_{\bar{Z}}\tilde{Z}$, and we may interpret it as a marginal R^2 of $\tilde{Z}_t$ on y_t, that is the proportion of the variation of y_t explained marginally by $\tilde{Z}_t$. Note that the measure we propose purges out the effect of the interaction between $X_t'\tilde{B}_t$ and $X_t'\bar{B}$ and hence the contribution of the learning component to the variance of y_t could actually be underestimated.

Table (1) shows the marginal effect of the learning component on each of the three equations included in the vector autoregression. The learning component contributes for slightly less than 17% of the variation of inflation and for about 22% of the variation in output but it plays a much bigger role in the equation for the federal funs rate, accounting for more than two-thirds of the variation in our measure for the short term interest rate. This analysis suggests that learning dynamics are important in explaining the evolution of the variables and that the equation capturing the behaviour of the monetary authority is the one more subject to learning from agents.

5.3 Counterfactual Experiment

An alternative way to evaluate the importance of learning is to run a counterfactual experiment in which the data are simulated as if learning dynamics did not take place. The experiment is implemented as follows: first, for each date in the estimation sample we compute the residuals from the model $y_t = X_t'\tilde{B}_t + X_t'\bar{B} + \varepsilon_t$; the simulated data, $\bar{y}_t$ are obtained by iterating on $\bar{y}_t = \bar{X}_t'\bar{B} + \hat{\varepsilon}_t$ with $\hat{\varepsilon}_t$ being the residuals computed in the first step and $\bar{X}_t'$ including lagged values of $\bar{y}_t$: $\bar{X}_t' = I_N \otimes \left[\begin{array}{c} 1, \ \bar{y}_{t-1}', \ \bar{y}_{t-2}', ..., \bar{y}_{t-p}' \end{array} \right]$. Figures (10) through (12) plot the actual and the simulated series of inflation, output growth and federal funds rate for the sample 1966:Q1 - 2009:Q1 in the upper panel and the difference between the two series in the lower panel. For most of the sample, the simulated inflation series lies above the actual series, implying that without learning, the inflation rate would have been higher. The discrepancy is remarkably large in the first part of the sample, where the simulated series is more than 2 percentage point higher than the actual series. Except for the subsample 68Q1-69Q1, the actual series of interest rate is higher than the simulated one. Again, the difference between the series is accentuated in the first part of the sample, peaking at 3.6 percentage points in 1983Q1. This behavior is consistent with our findings for the inflation series: learning dynamics induce the monetary authority to set the interest rate higher than it would in absence of learning, and this in turn keeps inflation lower than in absence of learning. The actual and simulated series for gdp growth are almost overlapping, suggesting that learning dynamics did not affect the evolution of our measure for real activity. However, the loose monetary policy of the late '70s would imply a drop in output growth of up to 0.6 percentage points in the sample 1976Q1 to 1980Q1. Table (2) reports some descriptive statistics of the simulated and of the actual data for the whole sample and for two subsamples of equal size: 66:Q1-87:Q4 and 88Q1-09:Q1. The table confirms that the differences between the actual and simulated series are more marked for inflation and federal funds rate. For the overall sample, the mean of the actual inflation rate is about 80% of the mean of the simulated data, while for the interest rate the mean of the simulated data is 15% smaller than the mean for the actual series. The actual inflation is less volatile than the simulated series, while the converse holds for the federal funds rate. Mean and standard deviations are about the same for the output growth actual and simulated data. A comparison of the statistics over sub-samples confirms that in the second part of the sample the mean and the standard deviations of the simulated series are closer to the mean and standard deviations of the actual series for all the three variables considered.

6 Conclusions

In this paper we use a time varying coefficient vector autoregression to assess the importance of the learning component in the US postwar economy and its ability to explain the changes in dynamics of key macroeconomic series. The random coefficients are assumed to follow a mean reverting process around an unconditional mean that can be interpreted as the estimates of the coefficients from the reduced form of a rational expectation equilibrium model. The deviations from the unconditional mean are attributed to the learning behavior of the agents about the value of the coefficients which regulate the economy. The proposed specification allows to decompose our vector autoregression model separating the learning component (the deviation of the coefficients from their unconditional mean) from the rational expectation component, i.e. the term related to unconditional mean of the coefficients.

We estimate a monetary model for the post WWII U.S. economy including inflation, output growth and the federal funds rate. We document the presence of learning dynamics and we assess their importance by measuring the proportion of the variance of each equation that is accounted for by the learning component and by running a counterfactual example in which the data are generated in the absence of learning. We find that the importance of the learning mechanism is somewhat limited for inflation and output growth but it is substantial in explaining the dynamics of the federal funds rate. Our results suggest that learning behaviour of the monetary authority might explain the change in dynamics of the nominal variables, while the learning mechanism is unable to characterize the changes in the real activity.

References

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

[2] Bray, M. and N. Savin, 'Rational Expectations, Learning and Model Specification', Econometrica (1986), 1129-1160

[3] Bullard, J. 'Time-Varying non Convergence to Rational Expectations Equilibrium under Least Squares Learning', Economics Letter 40 (1992), 159-166

[4] Bullard, J. and J. Suda, 'The Stability of Macroeconomics Systems with Bayesian Learners', Federal Reserve Bank, of St. Louis, wp. 2008-043B

[5] Bullard, J. and K. Mitra, 'Learning about Monetary Policy Rules', Journal of Monetary Economics 49 (2002), 1105-1129

[6] Canova, F. and L. Gambetti, 'On Time Variations of U.S. Monetary Policy: Who is Right?', Money Macro and Finance (MMF) Research Group Conference, 96 (2004)

[7] Cogley, T., Primiceri G. E. and T. J. Sargent, 'Inflation-Gap Persistence in the U.S.', NBER w.p. 13749 (2008)

[8] Cogley, T. and T. J. Sargent, 'Evolving Post WWII U.S. Inflation Dynamics', NBER Macroeconomics Annual (2001)

[9] Cogley, T. and T. J. Sargent, 'Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.', Review of Economic Dynamics 8 (2005), 262-302

[10] Cogley, T. and A. Sbordone, 'Trend Inflation and Inflation Persistence in the New Keynesian Phillips Curve', American Economic Review 98 (2008), 2101-26

[11] Del Negro, M. and F. Schorfheide, 'Priors from General Equilibrium Models for VARs', International Economic Review, 45(2) (2004), 190-217

[12] Evans, G. and S. Honkapohja, 'Expectations, Learning and Monetary Policy: an Overview of Recent Research', CDMA Working Paper Series (2008)

[13] Koop, G.M. and S. Potter, 'Time Varying VARs with Inequality Restrictions', Working Papers University of Strathclyde, (2008)

[14] McGough, B., 'Statistical Learning with Time-Varying Parameters', Macroeconomic Dynamics, 7 (2003), 119-139

[15] Milani, F. 'The Evolution of the Fed's Inflation Target in an Estimated Model under RE and Learning', (2006), mimeo

[16] Milani, F., 'Learning, monetary policy rules, and macroeconomic stability', Journal of Economics, Dynamics and Control, 32 (2008), 3148-3165

[17] Primiceri, G. E., 'Time Varying Structural Vector Autoregressions and Monetary Policy', The Review of Economic Studies, 72 (2005), 821-852

[18] Sargent, T. J. and N. Williams, 'Impacts of Priors on Convergence and Escapes from Nash Inflation', Review of Economic Dynamics, 8(2) (2005), 360-391

[19] Sims, C.A., 'Are Forecasting Models Usable for Policy Analysis', Minneapolis Federal Reserve Bank Quarterly Review, (1986), 2-16

[20] Sims, C.A., 'Interpreting the Macroeconomic Time Series Facts: the Effect of Monetary Policy', European Economic Review, 36 (1992), 975-1011

[21] Slobodyan S., and Wouters, 'Estimating a Medium-Scale DSGE Model with Expectations Based on Small Forecasting Models', mimeo (2008)

[22] Stock,J. H. and Watson, M. W.,'Has the Business Cycle Changed and Why?', NBER Macroeconomics Annual, 17 (2003)

[23] Stock,J. H. and Watson, M. W., 'Has Inflation Become Harder to Forecast?', Journal of Money, Credit and Banking, 39 (2007), 3-34

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

7 Data Appendix

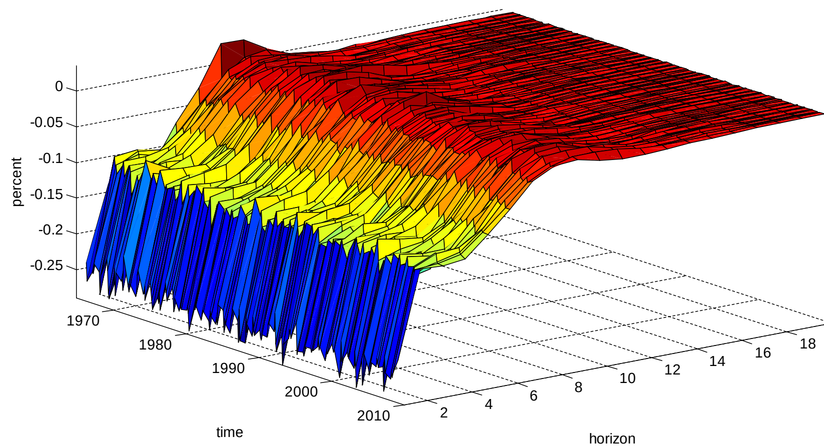
We consider a trivariate vector autoregression for inflation, output growth and short term interets rate. Data for all series are taken from the FRED database from the Federal Reserve Bank of Saint Louis.

Inflation is constructed as annual percentage change of the price index, observed at a quarterly frequency. The CPI all items seasonally adjusted series, CUSR0000SA0, is used. Monthly data are converted into quarterly by point sampling. Data for GDP are taken from the Bureau of Economic Analysis (Gross Domestic Product, Seasonally Adjusted at Annual Rate, billions of 2000 chained Dollars). Gdp growth rates are computed as $\Delta gdp_t = 400 * (\log gdp_t - \log gdp_{t-1})$ where gdp_t is the level of output at time t. For the short term interest rate series we use Federal Funds Rate; monthly data are converted into quarterly by point-sampling the first month of each quarter.

The sample runs from 1954.Q3 to 2009.Q1, for a total of 219 observations.

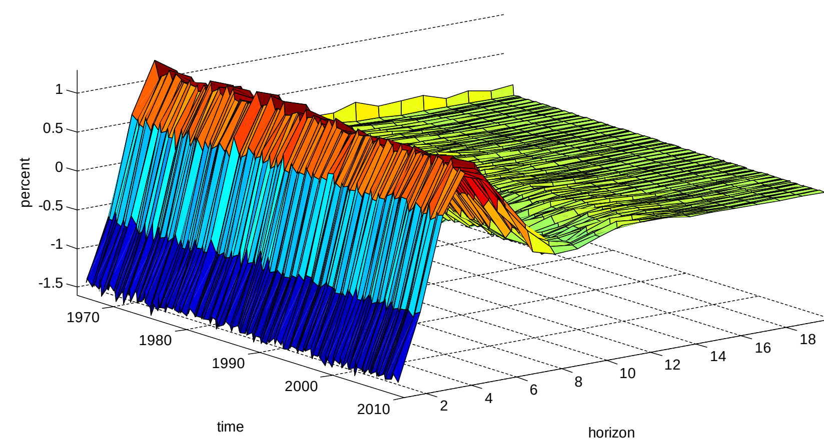
8 Appendix for Tables and Figures

Fig. 1. IRF of Inflation to Monetary Shock



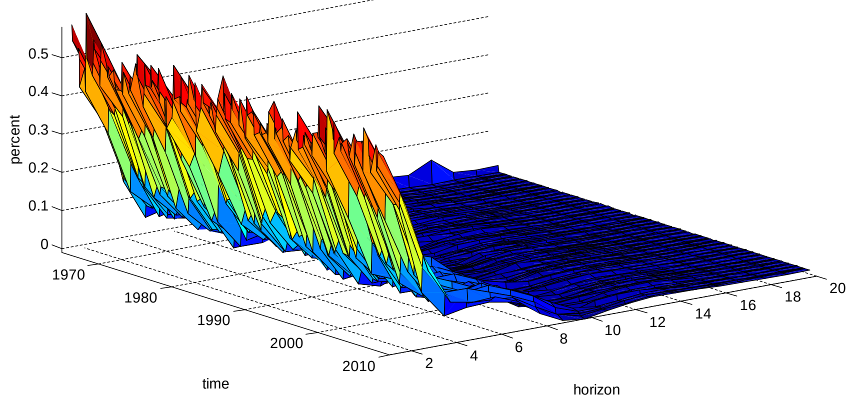
IRF of inflation to a contractionary monetary shock from 1966Q1 to 2009Q1, for horizon 1-20.

Fig. 2. IRF of Output Growth to Monetary Shock.



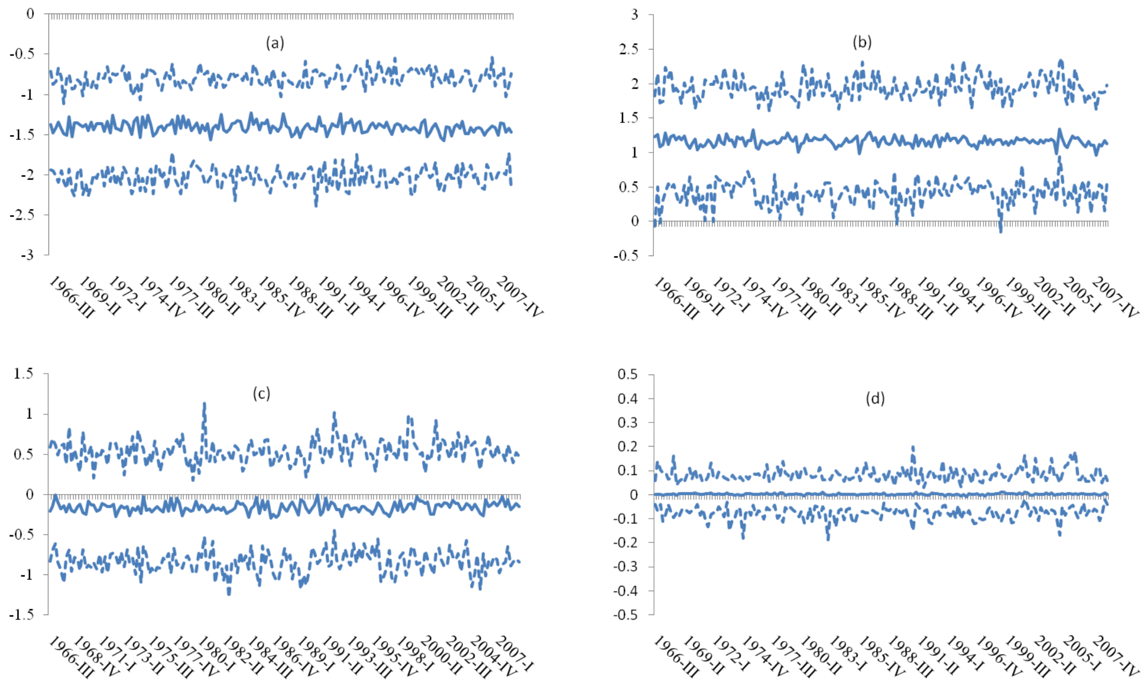
IRF of gdp growth to a contractionary monetary shock from 1966Q1 to 2009Q1, for horizon 1-20.

Fig. 3. IRF of Interest Rate to Monetary Shock.



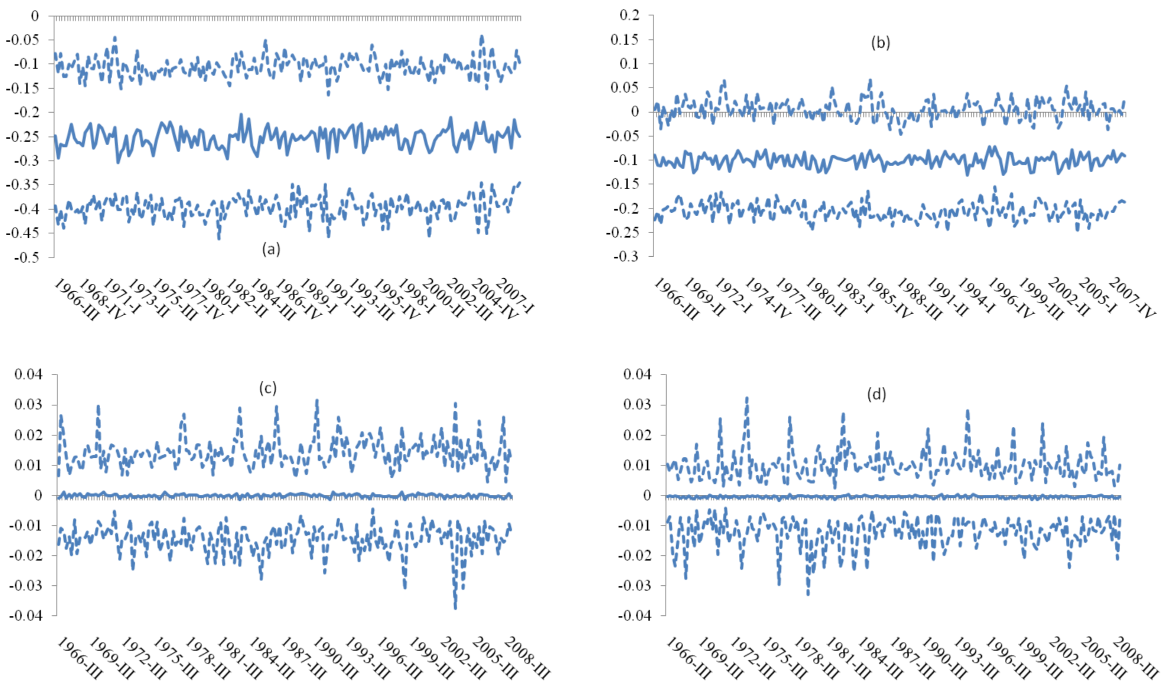
IRF of ffr to a contractionary monetary shock from 1966Q1 to 2009Q1, for horizon 1-20.

Fig. 4. IRF of Inflation to Monetary Shock at selected horizons



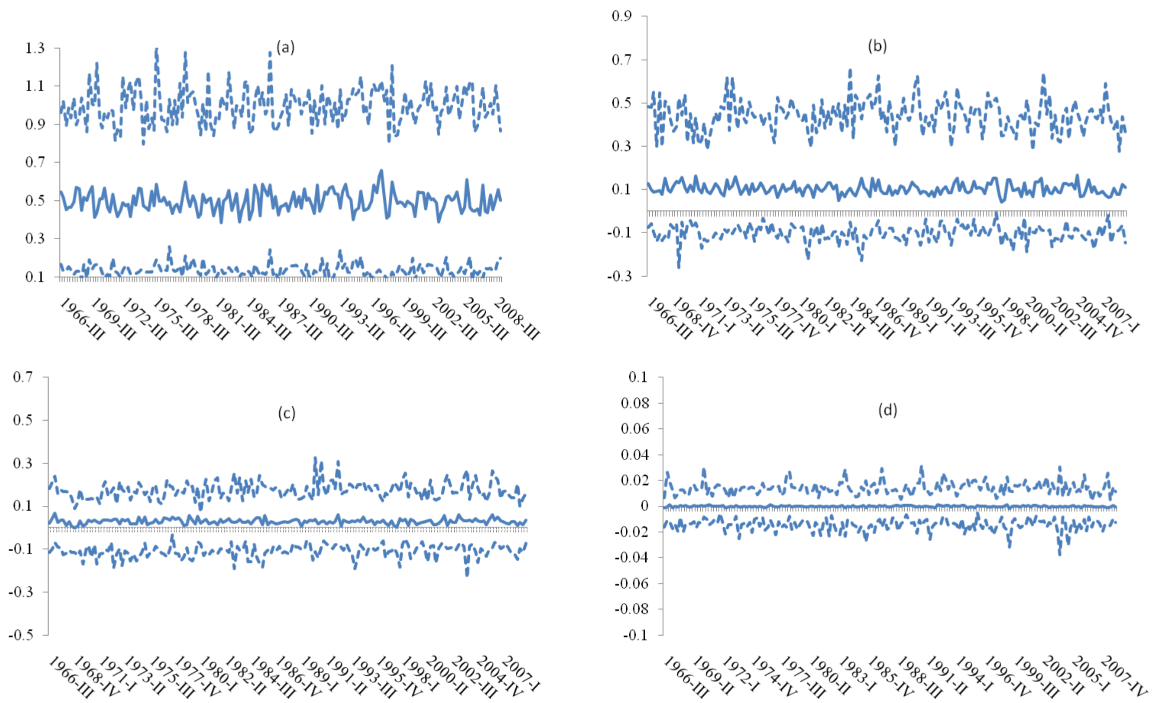
IRF of inflation to a contractionary monetary shock: (a) on impact, (b) after 4 quarters, (c) after 2 years, (d) after 5 years

Fig. 5. IRF of Output Growth to Monetary Shock at selected horizons



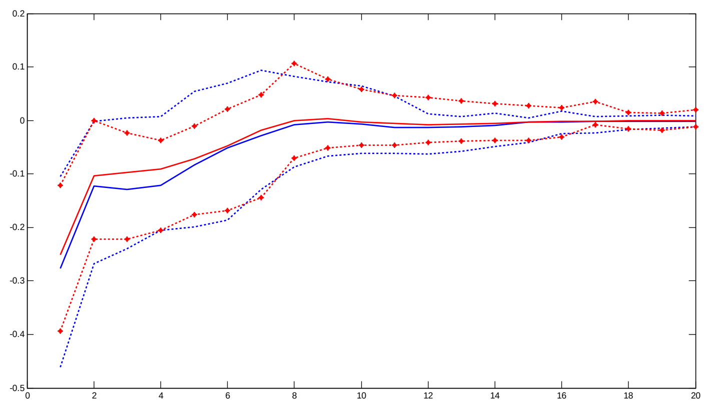
IRF of output to a contractionary monetary shock: (a) on impact, (b) after 4 quarters, (c) after 2 years, (d) after 5 years.

Fig. 6. IRF of Interest Rate to Monetary Shock at selected horizons



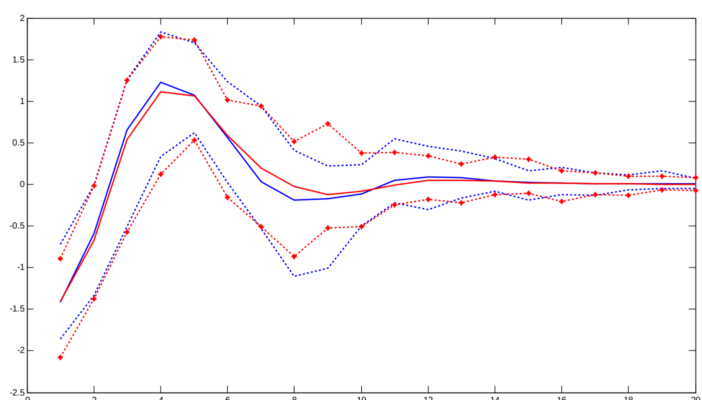
IRF of federal funds rate to a contractionary monetary shock: (a) on impact, (b) after 4 quarters, (c) after 2 years, (d) after 5 years.

Fig. 7. IRF of Inflation to Monetary Shock, 1981:Q3 and 2007:Q3



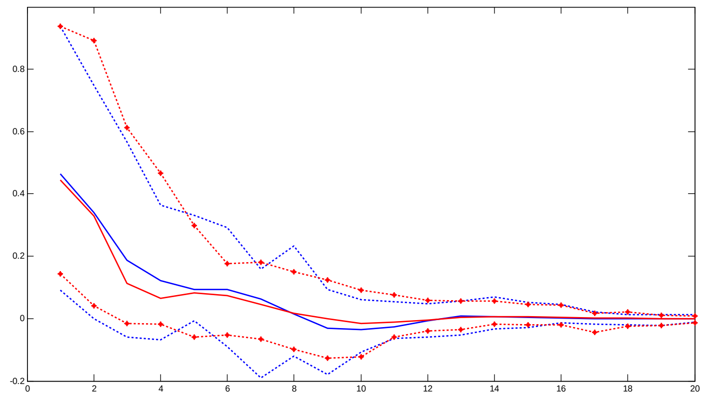
IRF of Inflation to Monetary Shock in 1981:Q3 (blue), and 2007:Q3 (red with marker); The solid line is the median IRF while dotted lines are the lower and upper bound of the 90% HPDI.

Fig. 8. IRF of GDP Growth to Monetary Shock, 1981:Q3 and 2007:Q3



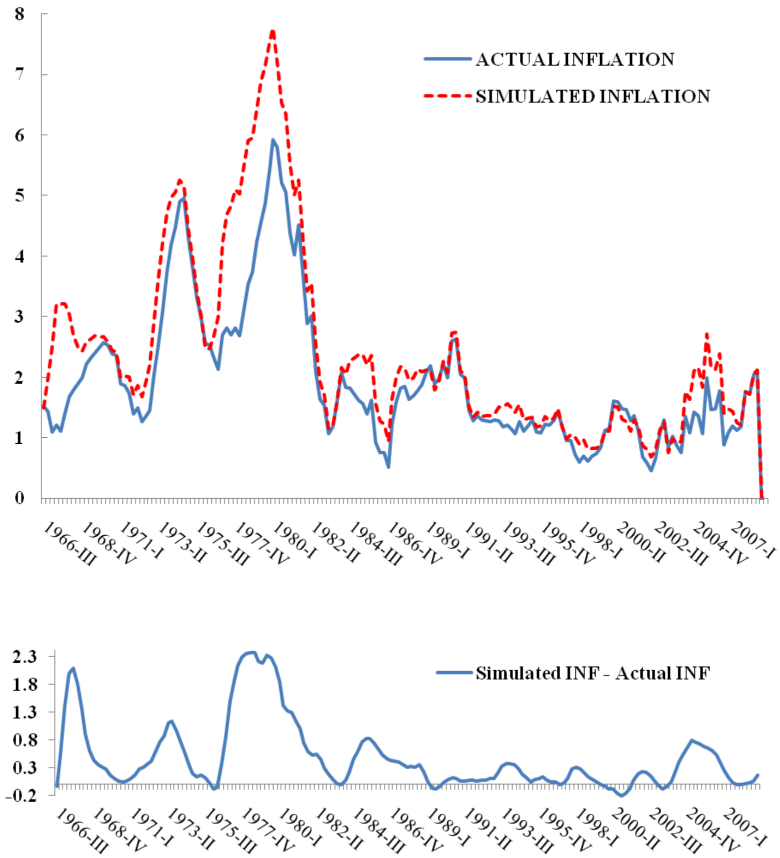
IRF of GDP growth to Monetary Shock in 1981:Q3 (blue), and 2007:Q3 (red with marker); The solid line is the median IRF while dotted lines are the lower and upper bound of the 90% HPDI.

Fig. 9. IRF of Interest Rate to Monetary Shock, 1981:Q3 and 2007:Q3



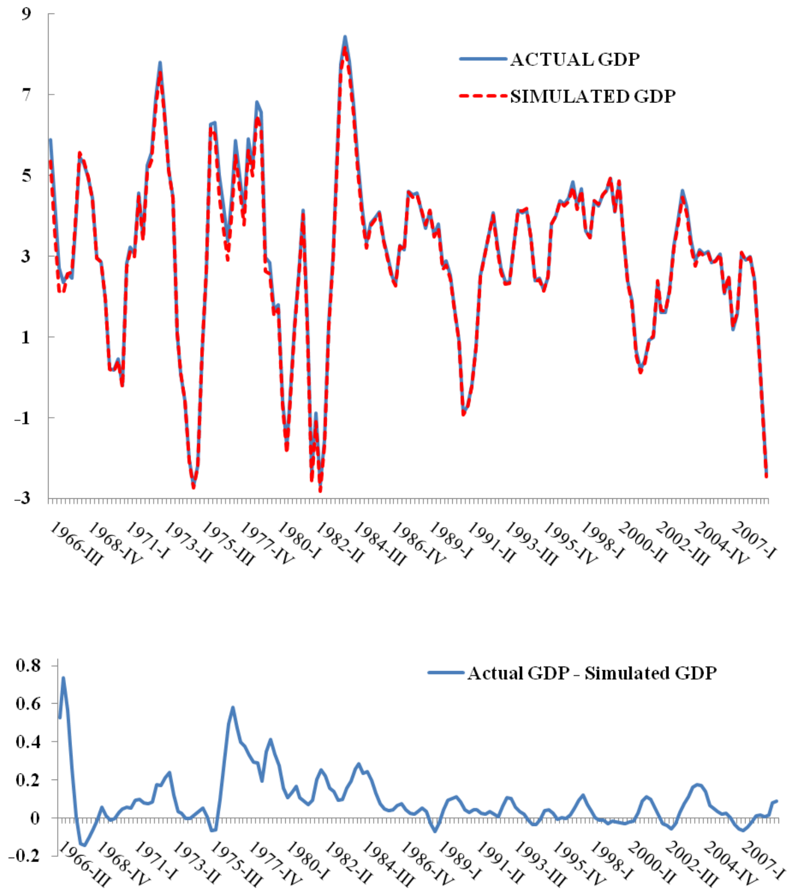
IRF of FFR to Monetary Shock in 1981:Q3 (blue), and 2007:Q3 (red with marker); The solid line is the median IRF while dotted lines are the lower and upper bound of the 90% HPDI.

Fig. 10. Actual and Simulated Inflation Series.



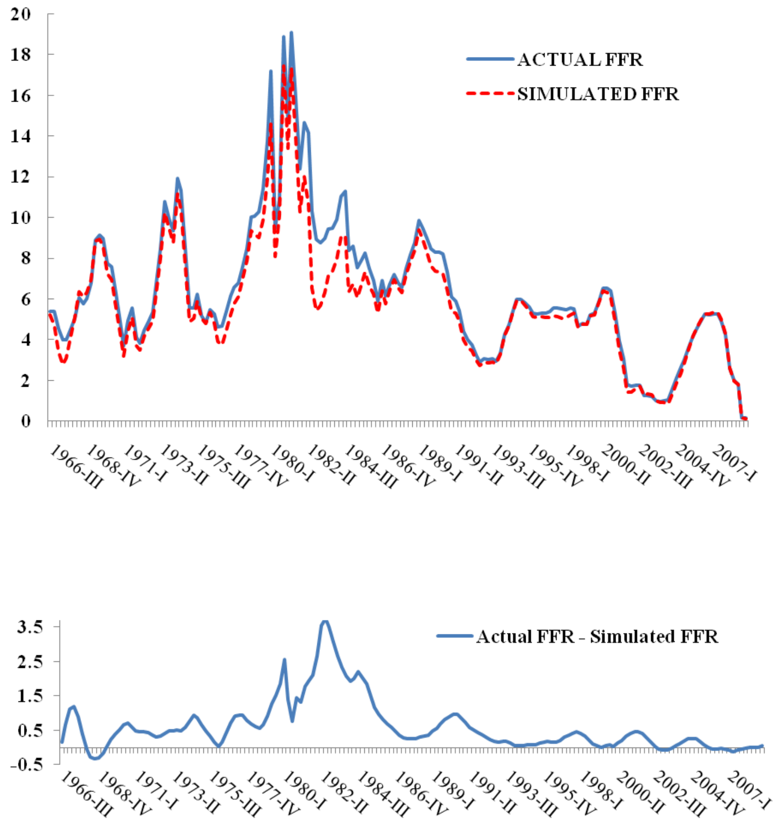
Upper panel: Actual (solid line) and simulated (dotted line) inflation series for the sample (1966:Q3-2009:Q1); simulated data are generated subtracting the learning component to the actual series. Lower panel: difference between simulated and actual series.

Fig. 11. Actual and Simulated Output Growth Series.



Upper panel: Actual (solid line) and simulated (dotted line) gdp growth series for the sample (1966:Q3-2009:Q1); simulated data are generated subtracting the learning component to the actual series. Lower panel: difference between actual and simulated series.

Fig. 12. Actual and Simulated Interest Rate Series.



Upper panel: Actual (solid line) and simulated (dotted line) federal funds rate series for the sample (1966:Q3-2009:Q1); simulated data are generated subtracting the learning component to the actual series. Lower panel: difference between actual and simulated series.

Table 1. Marginal Effect of Learning

Sample 66:Q1-09:Q1	
Equation	M.E.
Inflation	0.1687
GDP growth	0.2202
FFR	0.6850

Marginal Effect of Learning for each equation in the time varying VAR.

Table 2. Descriptive Statistics for Actual and Simulated Data

Simulated			66:Q1-09:Q1	Actual		
INF	GDP	FFR		INF	GDP	FFR
2.366	2.891	5.366	Mean	1.934	2.997	6.370
1.541	2.131	2.737	Std. Dev	1.178	2.163	3.387

Simulated			66:Q1-87:Q4	Actual		
INF	GDP	FFR		INF	GDP	FFR
3.290	3.015	6.615	Mean	2.545	3.210	8.109
1.667	2.602	2.792	Std. Dev.	1.307	2.650	3.399

Simulated			88:Q1-09:Q1	Actual		
INF	GDP	FFR		INF	GDP	FFR
1.431	2.765	4.102	Mean	1.292	2.791	4.561
0.483	1.521	2.015	Std. Dev	0.509	1.515	2.261

Summary statistics for simulated and actual data for the whole sample 66Q1-09Q1 and for the subsamples 66:Q1-87Q4 and 88:Q1-09:Q1.