LEARNING, MONETARY POLICY RULES, AND MACROECONOMIC STABILITY

FABIO MILANI

Princeton University

Abstract. This paper estimates a DSGE model with learning to re-examine the evidence on time variation in post-war U.S. monetary policy. Several papers document a regime switch, by showing that policy changed from 'passive' and destabilizing in the pre-1979 period to 'active' and stabilizing in the following decades. These papers typically work with DSGE models with rational expectations.

This paper relaxes the assumption of rational expectations and it allows for learning instead. Economic agents form expectations from simple models and update the parameters through constant-gain learning.

I estimate the model by Bayesian methods. The constant gain coefficient is jointly estimated with the structural and policy parameters of the system.

I find that the feedback coefficient to inflation was well above 1 also in the 1960s and 1970s and therefore policy was not leading to macroeconomic instability. The results reconcile the evidence from DSGE models with what obtained by time-varying VAR studies, which typically find only modest changes in policy coefficients over the post-war sample.

Keywords: monetary policy, new Keynesian DSGE model, constant-gain learning, expectations, Bayesian estimation, macroeconomic instability.

JEL classification: C11, D84, E30, E50, E52, E58.

1. Introduction

A large literature has studied the evolution of U.S. monetary policy over the post-war period and the effects of monetary policy on macroeconomic stability. Clarida, Gali', and Gertler (2000) show that policy was substantially different in the pre-1979 period compared with the following two decades. They find that policy switched from 'passive' in the 1970s, i.e. one that did not raise real interest rates in response to an increase in inflation, to 'active' in the 1980s and 1990s. Their estimated monetary policy rule in the pre-1979 sample implies instability if inserted in a

Date: First version: April, 2005. This version: July 14, 2005.

This paper constitutes the third chapter of my dissertation at Princeton University.

Address for correspondence: Princeton University, Department of Economics, Fisher Hall, Princeton, NJ 08544-1021. *E-mail:* fmilani@princeton.edu. *Homepage:* http://www.princeton.edu/~fmilani.

simple new Keynesian macroeconomic model. It does so, in particular, by leaving the economy subject to fluctuations induced by self-fulfilling expectations. During the Volcker-Greenspan period, instead, the estimated policy rule is characterized by a feedback to inflation larger than one and it is therefore stabilizing. A switch in regime towards improved monetary policy is often used as a possible explanation of the decline in inflation and output volatility in the U.S. starting from the 1980s, also known as "the great moderation".

Other papers provide similar evidence. Ireland (2001), Boivin and Giannoni (2004), and Lubik and Schorfheide (2004) estimate small DSGE models and find that monetary policy has become more aggressive since the early 1980s. Their results are also consistent with monetary policy as a source of instability in the 1970s.

But the view of a regime switch in monetary policy is challenged by a number of papers that obtain contrasting results. Sims and Zha (2004) and Canova and Gambetti (2004), for example, find little evidence in support of substantial changes in policy. Sims and Zha find that the best-fitting specification has time-variation in the variances of the shocks, but not in the coefficients. According to this view, bad luck rather than bad policy is responsible for the macroeconomic outcomes in the Seventies. Also Primiceri (2003) and Sargent, Williams, and Zha (2004) obtain changes in policy that are far from drastic. A split seems to exist in the literature. The papers that estimate simple theoretical models typically find evidence of a regime switch in monetary policy. Those that impose only a minimal set of theoretical restrictions, instead, estimating structural time-varying VARs, support a relatively stable monetary policy rule over the post-war sample.

The papers that reveal a change in monetary policy practice typically work with DSGE models under the conventional assumption of *rational expectations*. But a recent literature highlights the strong informational requirements of economic agents under rational expectations and proposes to relax the assumption in favor of agents that form expectations from simple economic models and learn the true model details over time.[1]

This paper aims to verify the evidence on changes in post-war U.S. monetary policy without imposing rational expectations and introducing learning instead. I estimate a simple macroeconomic model that can be derived from the optimizing decisions of households and firms. The preference parameters are assumed stable over time; the monetary policy coefficients are instead allowed to vary. In particular, I suppose that a shift in monetary policy takes place after the start of Volcker's term as Chairman of the Federal Reserve. Several studies have in fact shown that the Volcker-Greenspan period has been characterized by a more aggressive policy stance compared to the previous decades.

Individuals and firms in the model form expectations from correctly-specified economic models that correspond to the Minimum State Variable solution of the system under rational expectations. But they do not know the values of the model coefficients. Therefore, they use historical data to estimate the model and learn the relevant parameters over time, updating their beliefs through constant-gain learning. The use of constant-gain learning, according to which recent observations are weighted more than distant observations, rather than recursive least squares (RLS) learning, aims at modeling the learning process of individuals concerned about potential structural breaks at unknown dates. Rational expectations are nested in the model as a special limiting case.

I adopt Bayesian methods in the estimation. Several recent papers employ a Bayesian approach to estimate DSGE models. Smets and Wouters (2003, 2004a), for example, estimate a rich structural model that is found to fit the data as well as unrestricted VARs. Empirical studies typically work with DSGE models with rational expectations. This paper instead uses Bayesian methods to estimate a DSGE model with non-fully rational expectations and learning. A potential criticism of models with adaptive learning, discussed for example in Marcet and Nicolini (2003),

[1]Milani (2004b) shows that learning improves the fit of monetary DSGE models.

stresses the arbitrary choices often available to the researcher, which render the model hardly falsifiable. Milani (2004a), in fact, shows how his estimates strongly vary over the range of possible chosen gain coefficients. In the present paper, the gain coefficient is therefore jointly estimated with the rest of the parameters. The paper's joint estimation is hence similar to Milani (2004b).

Notice also that the papers that estimate full models under rational expectations by likelihood methods typically impose restrictions in the estimation to guarantee that the parameters fall in the determinacy region, so that the models can be solved by standard procedures. Those approaches rule out by construction regime switches from passive to active monetary policy. Lubik and Schorfheide (2004) are the first to provide the tools to extend the likelihood function to the indeterminacy region, finally allowing for determinacy and indeterminacy in the estimation under rational expectations. But such complications are not even needed under subjective expectations and learning. My framework is therefore particularly suitable to study the evolution of U.S. monetary policy over time and to investigate the determinacy and indeterminacy properties of estimated monetary policy rules.

I find that monetary policy has been active and stabilizing also in the pre-1979 period. Although some time variation in policy exists over the sample, the feedback coefficient to expected inflation is well above 1 also in the 1960s and 1970s. The evidence also indicates a decline in the Fed's response to movements in the output gap. The evidence of a regime switch from passive to active policy that is obtained under rational expectations would therefore be due a misspecification of the agents' expectations. Under learning, the private sector and the central bank are found, in fact, to consistently underestimate inflation in their forecasts during the 1970s.[2] Therefore, even though the policymaker responds to inflation with a coefficient above 1 when the true forecasts are used, it appears ex-post that the policy reaction has been weak if one assumes rational expectations. The persistent forecast errors derive from the economic agents' real-time perceptions of a low autocorrelation in

[2]Croushore (1998), in fact, find that inflation forecasts were systematically biased in the 1970s.

inflation and a large elasticity of inflation to changes in real activity in the 1960s and 1970s compared to the full-sample estimates.[3]

In the context of a structural macroeconomic model, Orphanides (2001, 2004) also provides results that contrast with the described prevalent view from DSGE models. He estimates a policy rule, substantially replicating the exercise of Clarida, Gali', and Gertler (2000), but exploiting the forecasts that were available to the Fed in real time. Using forecasts from the Greenbook, instead of rational expectations, Orphanides does not detect a passive monetary policy in the pre-1979 sample. What appears as an attenuated response to shocks is instead a stabilizing response to macroeconomic conditions in real time. My results therefore differ from those of Clarida, Gali', and Gertler and are consistent with those of Orphanides. As mentioned, he exploits real time forecasts from the Greenbook. But, in fact, this paper's approach represents a way to model such forecasts' formation. In this way, the model with learning may be more suitable than the model with rational expectations to study the interaction between policy and the economy over time.

My results are also consistent with the cited evidence from time-varying coefficients VARs. This similarity supports the view that a misspecification of standard simple DSGE models with rational expectations might be responsible for the impression of dramatic changes in policy.

The paper's contribution to the literature is therefore threefold. First, as described, the paper adds to the vast literature on the evolution of U.S. monetary policy, by obtaining results that contrast with common evidence from DSGE models and can be reconciled instead with the findings of better fitting structural VARs. Then, the paper contributes to the literature on adaptive learning, evoking how assuming learning rather than rational expectations might lead to considerably different empirical results. Finally, the paper offers a methodological contribution by showing how to estimate a DSGE model with subjective expectations and learning,

[3] The low autocorrelation of inflation in the 1960s and a large recession therefore attenuated the inflation forecasts in the first part of the 1970s.

and recommending the estimation of the learning parameters jointly with the rest of the system.

Furthermore, this paper, together with Milani (2004b), is the first to estimate the value of the constant gain coefficient jointly with the other deep parameters of the model. Estimating the gain from actual data is crucial, since one's results can heavily depend on the otherwise arbitrarily chosen gain (Milani 2004a shows how results vary over a large range of gain coefficients).

A limitation of my model lies in assuming a single regime switch taking place in 1979. In fact, even though the private agents learn the structure of the economy and its shifts over time, the true model has only a single structural break. This choice mirrors most work on monetary policy stability in rational expectations DSGE models and aims to clarify the different outcomes. The paper tries to suggest one direction in which the typical monetary DSGE model may fail to match the data, namely in the assumed expectations formation mechanism, and it illustrates how relaxing that assumption might overturn previous beliefs about monetary policy.

The paper is organized as follows. Section 2 presents the aggregate dynamics of the model with learning and describes the expectations formation mechanism. Section 3 describes the Bayesian techniques employed for estimation. Section 4 presents the empirical results, discussing the evidence about a regime switch in monetary policy and its effects on macroeconomic stability. It also checks the evidence regarding changes in the volatility of the shocks and the learning speed of the economic agents. Section 5 evaluates the robustness of the results to different specifications, while section 6 compares the model specifications computing the posterior marginal data densities. Section 7 concludes and discusses possible directions for future research.

2. The Model

I employ a simple DSGE model that has become popular for the analysis of the monetary transmission mechanism. The model can be derived from the optimizing behavior of households and firms (see, for example, Woodford 2003 for details). Clarida, Gali, and Gertler (2000) work with a similar model and find that U.S. monetary policy has changed in their sample, contributing to macroeconomic instability in the 1970s.

The aggregate dynamics of the economy can be summarized by the following three equations

$$\pi_t = \beta \widehat{E}_t \pi_{t+1} + \kappa x_t + u_t \quad (2.1)$$

$$x_t = \widehat{E}_t x_{t+1} - \sigma \left(i_t - \widehat{E}_t \pi_{t+1} - r_t^n \right) \quad (2.2)$$

$$i_t = \chi_t^\pi \widehat{E}_t \pi_{t+1} + \chi_t^x \widehat{E}_t x_{t+1} + \varepsilon_t \quad (2.3)$$

where π_t is inflation, x_t is the output gap, and i_t is the nominal interest rate (the policy instrument).[4] Equation (2.1) is the forward-looking New Keynesian Phillips curve that can be derived from the optimizing behavior of monopolistically competitive firms under Calvo price setting or quadratic adjustment costs in nominal prices. Inflation depends on expected inflation in $t+1$ and on current output gap. The parameter $0 < \beta < 1$ represents the households' discount factor, while κ denotes the slope of the Phillips curve. $\widehat{E}_t$ indicates subjective (possibly non-rational) expectations.[5]

Equation (2.2) represents the log-linearized intertemporal Euler equation that derives from the households' optimal choice of consumption. This simple model has no investment or government spending and therefore output equals consumption. Output depends on the expected one-period ahead output and on the ex-ante real interest rate. The coefficient $\sigma > 0$ represents the intertemporal elasticity of

[4] The empirical proxies of these variables will be demeaned prior to estimation.

[5] Notice that I follow the vast majority of papers in the adaptive learning literature in starting from the linearized equations under rational expectations and introducing learning from that point.

For a different approach, see Preston (2005), who introduces learning directly from the primitive assumptions of multi-period decision problems.

substitution of consumption. u_t and r_t^n denote exogenous shocks to firms' marginal costs of production and to the natural real interest rate.[6] They evolve according to univariate AR(1) processes

$$u_t = \rho_u u_{t-1} + \nu_t^u \tag{2.4}$$

$$r_t^n = \rho_r r_{t-1}^n + \nu_t^r \tag{2.5}$$

where $\nu_t^u \sim iid\left(0, \sigma_u^2\right)$ and $\nu_t^r \sim iid\left(0, \sigma_r^2\right)$. Equation (2.3) describes monetary policy. The central bank follows a rule by adjusting its policy instrument, a short-term nominal interest rate, to deviations of expected inflation and expected output gap from their targets. ε_t represents deviations from the systematic part of monetary policy and it has mean 0 and variance σ_ε^2. The coefficients χ_t^π and χ_t^x are the policy feedback coefficients and are time-varying. I assume, for simplicity and to be consistent with Clarida, Gali', and Gertler (2000), that a regime switch in policy occurs in 1979, when Paul Volcker begins his term as Chairman of the Federal Reserve (August 1979). Duffy and Engle-Warnick (2005), using nonparametric methods, similarly identify a switch in policy exactly in the third quarter of 1979. The policy coefficients therefore evolve as follows

$$\chi_t^\pi = \begin{cases} \chi_{pre-79}^\pi & t < 1979:03 \\ \chi_{post-79}^\pi & t \geq 1979:03 \end{cases} \quad \text{and } \chi_t^x = \begin{cases} \chi_{pre-79}^x & t < 1979:03 \\ \chi_{post-79}^x & t \geq 1979:03 \end{cases}$$

The post-1979 sample therefore refers to the Volcker-Greenspan tenures as Chairman of the Fed.

The monetary DSGE model described by Equations (2.1) to (2.3) has a unique stable solution if monetary policy is 'active', that is if the central bank raises the real interest rate in response to inflation, respecting the so-called "Taylor principle"

[6]Since x_t represents the theoretical output gap, i.e. the deviation of actual output from the natural rate of output that would prevail under flexible price, any deviation of the empirical output gap from its theoretical counterpart will materialize in u_t.

($\chi^{\pi} > 1$).[7] If monetary policy does not respect the Taylor principle it is instead defined as 'passive'.[8]

As already discussed, in a similar model under rational expectations, Clarida, Gali, and Gertler (2000) find that the estimated policy rules in the U.S. are passive and destabilizing before 1979 and active and stabilizing afterwards. Here, I re-estimate the monetary policy rule to assess if those results can be confirmed when learning replaces the strong informational assumptions required by rational expectations.

Differently from Clarida, Gali, and Gertler (2000), I jointly estimate the full model using likelihood-based Bayesian methods, whereas they estimate a single equation by GMM. My full system estimation is instead similar to Lubik and Schorfheide (2004), who nonetheless confirm Clarida, Gali, and Gertler's results. Full system estimation is more efficient. Moreover, as Sims and Zha (2004) discuss, a single equation estimation by instrumental variables relies on the claim that the instruments influence monetary policy only through their effects on expected variables. But in reality, they continue, it is unlikely to think that the central bank responds in the same way to the same future inflation rate, no matter what the recent level of inflation has been. The downside of the multivariate approach is that misspecifications in any part of the model will also bias the policy coefficients I intend to examine. Another difference in the model lies in the monetary policy rule: the presented model does not include an interest-rate smoothing term. I have found that the choice to exclude the interest-rate smoothing term simplifies the estimation of the learning model.[9] But later in the paper I examine the robustness of the results to this choice. The central bank responds to one-period ahead forecasts

[7] This is true under the assumption that the response to the output gap coefficient χ^{x} does not fall outside a specific range. The commonly estimated values for this coefficient, however, do not pose problems for determinacy.

[8] Notice that the correspondence between respecting the Taylor principle and stability depends on the common assumption of ignoring fiscal policy. The joint consideration of fiscal and monetary policy leads in fact to different stability conditions.

[9] In particular, what simplifies the estimation is not having the lagged interest rate term in the minimum state variable solution used by economic agents in their learning rule.

of inflation and output gap. A response to one-period ahead forecasts is, in fact, optimal in this kind of model under learning (see Evans and Honkapohja 2003).

I assume constant preference parameters over the whole sample. Therefore, even if there is a switch in policy after 1979, the β, σ, and κ coefficients do not vary. The only time variation in the model lies in the monetary policy rule. Having relaxed the assumption of rational expectations, I need to specify a model of expectations formation for the agents. I follow the recent adaptive learning literature (Evans and Honkapohja 2001 for an introduction) in assuming that agents have the same knowledge an econometrician would have and learn the relevant parameters over time.

2.1. **Expectations Formation and Constant-Gain Learning.** Consumers, firms, and the monetary authority in the model need to form forecasts of future macroeconomic conditions. I assume that these economic agents behave as econometricians, employing an economic model and forming expectations from that model. For simplicity, all economic agents share the same model.

Agents estimate the linear specification

$$Z_t = b_t u_t + c_t r_t^n + \epsilon_t \tag{2.6}$$

using variables that appear in the Minimum State Variable (MSV) solution of the system under rational expectations (where I define $Z_t \equiv [\pi_t, x_t, i_t]'$ and where b_t and c_t are coefficient vectors of appropriate dimensions)[10]. Expression (2.6) represents the "**Perceived Law of Motion**" or **PLM** of the agents. Agents use a correctly-specified model of the economy, i.e. they use the regressors that appear in the true solution of the system, but they do not know the relevant model parameters. They use historical data to learn those parameters over time. As additional data become available in subsequent periods, they update their estimates of the

[10]See McCallum (1999) for a presentation of the MSV criterion to choose the "fundamental" solution of linear rational expectations systems.

coefficients (a_t, b_t) according to the constant-gain learning formula

$$\widehat{\boldsymbol{\phi}}_t = \widehat{\boldsymbol{\phi}}_{t-1} + \overline{\mathbf{g}} R_{t-1}^{-1} X_t (Z_t - X_t' \widehat{\boldsymbol{\phi}}_{t-1}) \tag{2.7}$$

$$R_t = R_{t-1} + \overline{\mathbf{g}} (X_{t-1} X_{t-1}' - R_{t-1}) \tag{2.8}$$

where $\widehat{\boldsymbol{\phi}}_t = (vec(b_t, c_t)')'$ describes the updating of the learning rule coefficients, and R_t the updating of the matrix of second moments of the stacked regressors $X_t \equiv \{u_t, r_t^n\}_0^{t-1}$. $\overline{\mathbf{g}}$ denotes the constant gain coefficient. Under recursive least squares learning, the gain equals $1/t$. Here the gain is instead constant, providing a simple way to model learning of economic agents concerned about future unknown structural breaks.[11] Constant-gain learning weighs more heavily recent observations (it is a variant of what is more generally called bounded-memory learning). Differently from recursive least squares learning, the economy will not converge to the Rational Expectations Equilibrium, but might only converge to an ergodic distribution around it. A larger value of the gain coefficient $\overline{\mathbf{g}}$ implies faster learning of shifts, but it also induces a larger variance around the steady state.

Economic agents form expectations for $t+1$ using (2.6) and the updated parameter estimates

$$\widehat{E}_t Z_{t+1} = \begin{bmatrix} b_{1,t-1} & c_{1,t-1} \\ b_{2,t-1} & c_{2,t-1} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \rho_u u_t \\ \rho_x r_t^n \end{bmatrix}. \tag{2.9}$$

The model informational assumptions are as follows: in period t, agents observe the values of the shocks and they use the estimated parameters in $t-1$ (which they have derived from a regression they have run in the previous period), to form expectations for the future period $t+1$[12].

The learning framework I have presented nests the standard rational expectations specification as a special limiting case, i.e. the case for $\overline{\mathbf{g}} \to 0$. The model also nests a wide range of possible learning cases indexed by the single parameter $\overline{\mathbf{g}}$, according to which agents are more or less willing to update their expectations based on new information.

[11] Constant-gain learning has been used in various recent empirical studies, such as Orphanides and Williams (2003a,b,c, 2004), Williams (2003), Primiceri (2003), and Milani (2004a,b), among others. Orphanides and Williams also refer to it as perpetual learning.

[12] The information assumptions are similar to Evans and Honkapohja (2003a,b), Preston (2003), and Milani (2004b).

Furthermore, as discussed by Orphanides and Williams (2003), modeling agents' learning allows researchers to accommodate the Lucas critique, since private expectations endogenously adjust to changes in policy.

To summarize, the model economy is represented by the inflation dynamics equations (2.1), output equation (2.2), time-varying monetary policy rule (2.3), shock processes (2.4), (2.5), and the expectations formation mechanism expressed by (2.7), (2.8) and (2.9).

3. Econometric Approach

I estimate the model by likelihood-based Bayesian methods. A recent literature in macroeconomics employs a similar Bayesian approach to estimate DSGE models. The papers in this literature, however, typically estimate the models under the standard hypothesis of rational expectations, whereas here the novelty lies in estimating a simple DSGE model with non-fully rational expectations and learning. The econometric procedure allows one to estimate the learning parameters jointly with the structural parameters of the economy. In particular, here I estimate the constant gain coefficient *jointly* with the parameters describing preferences and the monetary policy rule parameters. A similar estimation strategy is used in Milani (2004b). Estimating also the constant gain coefficient is crucial, since one's results might heavily depend on the chosen gain. For instance, Milani (2004a) shows how the estimated backward-lookingness in inflation varies over the possible gain values.

I collect the parameters in the vector θ

$$\theta = \left\{\beta, \kappa, \sigma, \chi^{\pi}_{pre-79}, \chi^{x}_{pre-79}, \chi^{\pi}_{post-79}, \chi^{x}_{post-79}, \rho_u, \rho_r, vec(Q), \overline{g}\right\}$$

where Q is the variance-covariance matrix collecting the variances of the supply, demand, and policy disturbances. I fix β equal to 0.99, σ equal to 0.2,[13] and the autoregressive coefficient ρ_u and ρ_r equal to 0.9. I initialize the learning algorithm using data from the pre-sample (1951-1959) period. The model can be simply written in state-space form and its likelihood computed at each iteration through

[13] I have found that fixing σ considerably improve the properties of the Metropolis algorithm.

the Kalman Filter. I then generate draws from the posterior distribution using the Metropolis-Hastings algorithm.

The information about the parameters is summarized by the posterior distribution, obtained by Bayes Theorem

$$p\left(\theta \mid Y^T\right) = \frac{p\left(Y^T \mid \theta\right) p\left(\theta\right)}{p\left(Y^T\right)} \tag{3.1}$$

where $p\left(Y^T \mid \theta\right)$ is the likelihood function, $p\left(\theta\right)$ the prior for the parameters, and $Y^T = [y_1, ..., y_T]'$ collects the data histories.

The model is fitted to U.S. data on inflation, output gap, and nominal interest rates. The data are quarterly for the period 1960:I to 2004:II. Inflation is defined as the annualized quarterly rate of change of the GDP Implicit Price Deflator, output gap as the log difference between GDP and Potential GDP (CBO estimate), and I use the federal funds rate as the policy instrument.[14] All the variables have been demeaned prior to estimation. I run $100,000$ draws for the Markov Chain, discarding the first 50% as burn-in and saving one draw every 100.[15]

3.1. **Specifying the Prior Distribution.** First, I need to specify a prior distribution over the structural parameters collected in θ. Table 1 presents information about the priors.

I assume that the parameters are a priori independent. I fix β, σ, ρ_r, and ρ_u: they can be thought to have dogmatic priors with zero variance. The slope coefficient of the Phillips curve, κ, follows a Normal distribution with mean .1 and standard deviation .05. The monetary policy rule coefficients are also assumed to follow Normal distributions with mean 1.5 and standard deviation .5 for the inflation feedback coefficients, and mean .5 and standard deviation .25 for the output feedback coefficients. The 95% prior probability interval therefore includes values of the response to inflation below 1, which would lead to macroeconomic instability (not satisfying the Taylor principle). I choose inverse gamma distributions for the

[14]The series were obtained from FRED, the database of the Federal Reserve Bank of Saint Louis.

[15]Thinning the chain leads to inefficient estimates, but it reduces the autocorrelation of the draws and it considerably reduces the computational and storage requirements of the MCMC procedure.

standard deviations of the shocks with mean 1 and standard deviations .5. Finally, I assume a Gamma distribution for the constant gain coefficient to assure that it remains positive. The gain has prior mean .031 and prior standard deviation .022.

3.2. **Metropolis-Hastings Algorithm.** To generate draws from the posterior distribution of θ using the Metropolis-Hastings algorithm, I need to evaluate the likelihood function $p\left(Y^T \mid \theta\right)$ at each iteration. The model can be expressed as a linear Gaussian system, so that I can easily compute the likelihood recursively with the Kalman Filter. Appendix A illustrates the details of the procedure.

I obtain a Markov Chain $\{\theta_1, \theta_2, ..., \theta_D\}$ of parameter values, where D represents the total number of draws ($D = 100,000$ in the paper). Given this Markov Chain and a function of interest $g(\cdot)$, it is possible to prove that $\widehat{\mu} = \frac{1}{D}\sum_{j=1}^{D} g(\theta_j)$ converges almost surely to $\mu = E[g(\theta)]$ as $D \to \infty$. Application of a Central Limit Theorem implies $\sqrt{D}(\widehat{\mu} - \mu) \xrightarrow{D} N\left(0, \sigma_\mu^2\right)$.

3.3. **Convergence.** To assess convergence of the MCMC (Markov Chain Monte Carlo) simulation, I performed various checks. First, a simple initial check consists of looking at the trace plots of the draws. More formally, I have considered a number of convergence tests as the ones proposed by Geweke (1992), and Raftery and Lewis (1995). Raftery and Lewis (1995)'s diagnostics suggests a minimum number of total draws, a thinning parameter, and a minimum burn-in, by computing the autocorrelation of the draws. Geweke's test instead compares the partial means $\widehat{\mu}_1 = \frac{1}{D_1}\sum_{j=1}^{D_1} g(\theta_j)$ and $\widehat{\mu}_2 = \frac{1}{D_2}\sum_{j=D_1+1}^{D_2} g(\theta_j)$, obtained from the first D_1 and last D_2 simulation draws. The null hypothesis of equal means between the two samples of draws can be tested knowing that for $D \to \infty$ the quantity $(\widehat{\mu}_1 - \widehat{\mu}_2) / \left(\frac{\widehat{S}_g^1(0)}{D_1} + \frac{\widehat{S}_g^2(0)}{D_2}\right)^{1/2} \Longrightarrow N(0,1)$. I also look at the plots derived from the test proposed by Yu and Mykland (1994), based on CUMSUM plots of the draws[16]. Finally, I ascertain convergence by looking at the recursive mean plots

[16] They propose the statistics $CS_t = \left(\frac{1}{t}\sum_{d=1}^{t} \theta^d - \mu_\theta\right)/\sigma_\theta$, where μ_θ and σ_θ are the empirical mean and standard deviations of the D draws of the Markov Chain. The plot of CS_t converges to 0 as t increases.

and bivariate scatter plots among the parameters to evaluate the mixing of the chain.[17]

4. Estimating the Model with Learning: Empirical Results

4.1. **Has Monetary Policy Changed?** I first estimate the baseline model described in (2.1) to (2.3) for the 1960-2004 sample. Agents form expectations from economic models and learn the relevant parameters over time. The model allows for different response coefficients to expected inflation and output gap in the monetary policy rule in the pre- and post-1979 samples.

It is well known that in the simple monetary model of the kind studied in this paper, the equilibrium is undetermined if monetary policy is 'passive', i.e. if the central bank does not raise nominal interest rates more than one to one in response to increasing inflation. The empirical literature estimating DSGE models under rational expectations typically imposes the restriction that the parameters lay within the determinacy region. Lubik and Schorfheide (2004), instead, are the first to provide the econometric tools to estimate a DSGE model allowing for indeterminacies and sunspot fluctuations. They find that pre-Volcker policy is consistent with indeterminacy. The estimation under indeterminacy is simpler under learning. With learning, in fact, there is no need to use the techniques required under rational expectations to solve the model and therefore it is possible to estimate parameter values that would lead to indeterminacy without the complications needed in Lubik and Schorfheide (2004) to extend the likelihood function to the indeterminacy region.

Table 2 presents the estimation results.

The results suggest some time variation in the monetary policy rule coefficients. But the feedback coefficient to inflation appears well above 1 also in the pre-Volcker sample. I therefore find that monetary policy was 'active' also in the 1960s and

[17] Details on the convergence checks will be reported on a technical appendix that will be available on my website.

1970s and it was not a source of macroeconomic instability. The feedback coefficients to inflation and output are estimated equal to 1.688 and .166 in the pre-1979 sample and equal to 2.708 and $-.005$ in the Volcker-Greenspan sample. The reported 95% posterior probability interval also lays well above 1 in both samples.

Figure 1 overlaps the posterior distributions of the pre- and post-1979 policy feedback coefficients. From the figure, it is evident that the posterior distribution of the inflation response coefficients has moved rightward, whereas the response to output fluctuations has instead attenuated.

A central coefficient in my estimation is represented by the constant gain $\overline{\mathbf{g}}$. The posterior mean estimate for the gain equals 0.0234. The estimated value is close to values used in learning studies that conditioned their analysis on a given chosen gain (often working with gains in the interval $0.015-0.04$), and to the value found by Orphanides and Williams (2003, 2004), using data on expectations from the Survey of Professional Forecasters[18]. To facilitate intuition, the gain can be interpreted as an indication of how many past observations economic agents use to form their expectations. A gain equal to 0.0234 therefore indicates that agents roughly employ 10 or 11 years of data (42.735 quarters).

Figure 2 plots the evolution of agents' beliefs over the sample together with a 95% posterior probability interval. Those beliefs display some variation over time; a larger uncertainty seems to characterize those estimates at the end of the sample.

4.2. **Time-Varying Shocks and Time-Varying Policy.** Sims and Zha (2004) find little evidence of time-variation in the policy coefficients over the post-war sample. They find instead that their best-fitting specification has time-variation in the variances of the innovations only. In the context of structural DSGE models, therefore, the omitted consideration of regime changes in the variances might lead researchers to spuriously find ampler time-variation in the coefficients. Notice, however, that this paper finds no evidence of dramatic changes in the coefficients,

[18]They found values of $\overline{\mathbf{g}}$ in the range $0.01-0.04$ to perform better and they adopted $\overline{\mathbf{g}}=0.02$ in their baseline model.

i.e. a switch from passive to active policy, even omitting time variation in the variances of the disturbances, when learning is considered.

Here, I re-estimate the model allowing for a structural change in the volatilities of the shocks as well. I check if this modification leads to changes in the estimates of the policy coefficients and the gain coefficient.

The results in table 3 show that the variances of the supply and demand innovations decreased in the 1980s and 1990s, consistently with what expected. Surprising is instead the increased variance of the policy disturbance. Certainly, as Sims and Zha (2004) discuss, a single shift in variance hardly captures the evolution of the policy shock variance. In particular, in my estimation, the larger post-1979 variance originates from the 1979-1982 years of non-borrowed reserves targeting that led to a strong volatility of the federal funds rate.

The policy rule coefficients, instead, are substantially unchanged. The gain estimate is now a little lower: 0.01957.

4.3. **Has Learning Speed Changed over Time?** I now allow for different learning speeds in the two sub-samples. I re-estimate the model with time-varying monetary policy, now also allowing for a one-time break in the constant-gain coefficient that occurs in 1979. Obviously, a one-time discrete switch in the speed of learning is not entirely realistic but it can be a useful first approximation, at least to sense how learning might differ in environments characterized by different policies or volatility levels.

The posterior estimates of the reaction function and private sector parameters are substantially unchanged. I detect a change in learning speed between the two samples. I estimate, in fact, a gain that decreases from the estimated value of 0.03 until 1979, to 0.0152 afterwards. It seems that the data support slower learning in the 1980s and 1990s. The evidence seems also robust to starting the Volcker-Greenspan period only in 1983. The 95% posterior probability intervals in the table, however, are similar.

Figure 3 plots the histograms of the gain coefficients' distributions in the two samples. The posterior distribution of the gain has clearly switched after 1979 and indicates values closer to zero in the last two decades.

5. Robustness

5.1. **Forecasts k Periods Ahead.** In the previous section, the monetary policy rule posited a response to one-period ahead expectations. I now assume that the monetary authority reacts to 4-quarters ahead forecasts of inflation and 2-quarters ahead forecasts of the output gap. The timing assumption becomes similar to Clarida, Gali', and Gertler (2000).

I report the results in table 5. The reaction of monetary policy to changes in inflation is now stronger. The feedback coefficient to inflation equals 2.04 in the 1960s-1970s and it increases to 3.20 in the 1980s-1990s. Again, although policy has become more aggressive in the Volcker-Greenspan era, policy has been sufficiently active over the whole post-war period to guarantee determinacy of the equilibrium. Again, I find a reduction in the response to output movements. The gain is now estimated to be 0.0305.

5.2. **Interest-Rate Smoothing.** I now verify the robustness of the results to introducing an interest-rate smoothing term in the monetary policy rule, which becomes

$$i_t = \rho_t^i i_{t-1} + (1-\rho_t^i)\left[\chi_t^\pi \widehat{E}_t \pi_{t+4} + \chi_t^x \widehat{E}_t x_{t+2}\right] + \varepsilon_t \tag{5.1}$$

where ρ_t^i is the smoothing coefficient; ρ_t^i, χ_t^π, and χ_t^x are time-varying. The policy rule is now similar to the one in Clarida, Gali', and Gertler (2000) and Lubik and Schorfheide (2004). Table 6 shows the posterior estimates.

Figure 4 shows that the evidence of a drastically more aggressive policy response to inflation in the post-1979 sample is even weaker than found in the previous section. It seems instead that policy has become more inertial, since the posterior distribution of the lagged interest-rate term is shifting towards 1. The posterior mean estimate of ρ_t^i in fact goes from .825 to .903. When the interest-rate smoothing term is allowed, the policy reaction to output movements is larger (χ_x equals.701 and .57).

5.3. **VAR(1) as Learning Rule.** The abandonment of rational expectations opens a wide range of possibilities on how to best model the expectations formation mechanism of economic agents. With the uncertainty surrounding the modeling of expectations, it is necessary to check that the empirical results are robust to alternative forecasting rules. For example, I now suppose that agents employ the following PLM:

$$Z_t = \Phi Z_{t-1} + \epsilon_t \tag{5.2}$$

where $Z_t \equiv [\pi_t, x_t]'$. They therefore estimate a small VAR(1) and use it to form forecasts of future output and inflation. Table 7 presents the new estimation results. The evidence about the switch in monetary policy coefficients is similar. A strong difference lies instead in the estimated gain coefficient: under the VAR(1) learning rule the estimate drop to .0011, very close to lower bound of 0.

Figure 5 instead shows the evolution of inflation and output one-period ahead expectations derived from the agents' PLM and compared with the actual ex-post values. It is apparent that agents have constantly underestimated inflation during most of the 1970s, while they have overestimated inflation in the first half of the 1960s and in the second half of the 1990s. In particular, the underestimation of inflation in the 1970s originates from an underestimation of its persistence and an overestimation of its sensitivity to movements in real activity (the recession in the 1970s was therefore expected to mitigate future inflation pressures). Regarding output, it appears instead that agents' forecast errors have been less dramatic and less serially correlated.[19]

5.4. **Joint Estimation of Initial Beliefs.** So far, I have estimated the initial beliefs of the agents on pre-sample data. Now, I estimate the initial beliefs jointly with the other parameters of the system exploiting the whole sample. The joint

[19] Assuming, however, that agents need to learn also the level of potential output in real time would certainly worsen their forecasting performance.

estimation further reduces the arbitrariness implicit in the choice of the initial values. The parameter vector to be estimated becomes

$$\theta = \left\{ \beta, \kappa, \sigma, \rho_{pre-79}, \chi^{\pi}_{pre-79}, \chi^{x}_{pre-79}, \rho_{post-79}, \chi^{\pi}_{post-79}, \chi^{x}_{post-79}, \rho_u, \rho_r, vec(Q), \phi_{1|0}, \overline{g} \right\}$$

where $\phi_{1|0}$ are the initial beliefs used to initialize the learning algorithm. Notice that I need to estimate only the initial values: agents' beliefs are then updated over the sample through the learning formula (2.7). As in the previous paragraph, agents use a VAR(1) as their perceived law of motion of the economy. The results are shown in table 8. Monetary policy seems to respond with similar coefficients to inflation in both subsamples now. From the estimation of the initial values, notice that agents start with the belief that inflation is very volatile (the estimated autoregressive coefficient $b_{1,1|0}$ equals .005); they perceive output as a more persistent process instead ($c_{2,1|0} = .54$). The elasticity of inflation to movements in the output gap $c_{1,1|0}$ is perceived by the agents to equal 0.135 at the beginning of the sample.

5.5. **Recursive Least Squares (RLS) Learning.** I have assumed that agents update their beliefs through constant-gain learning. In this way, they discount more heavily observations in the distant past than recent observations. Another popular approach to model learning assumes instead that agents equally weight all observations: this case is known as recursive least squares (RLS) learning. The updating formulas become

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + t^{-1} R^{-1}_{t-1} X_t (Z_t - X'_t \widehat{\phi}_{t-1}) \qquad (5.3)$$

$$R_t = R_{t-1} + t^{-1} (X_{t-1} X'_{t-1} - R_{t-1}) \qquad (5.4)$$

Under RLS learning, the gain equals t^{-1} and it decreases over time toward 0. I re-estimate the system under the new learning specification and I present the results in table 9. Monetary policy was active even before 1979, but it became more aggressive towards inflation afterwards. As before, monetary policy seems to have responded less to real activity and it has become more inertial after 1979. Assuming a different learning mechanism therefore does not alter the results.

6. Model Comparison

In this section, I compare the log marginal likelihoods of the various estimated specifications. In this way, I let the data decide about the plausibility of the different assumptions. The marginal data densities $p\left(Y^T \mid \mathcal{M}_i\right)$ are defined as

$$p\left(Y^T \mid \mathcal{M}_i\right) = \int L\left(\theta_j \mid Y^T, \mathcal{M}_i\right) p\left(\theta_j \mid \mathcal{M}_i\right) d\theta_j \tag{6.1}$$

where $\mathcal{M}_i$ represent the possible models. This measure of model fit is appealing even because it automatically penalizes for model dimensionality. I compute the marginal likelihood using Geweke's (1999) Modified Harmonic Mean approximation.[20]

From the table, it is clear that the models that incorporate interest-rate smoothing in the monetary policy rule fit much better than those assuming a simple Taylor rule with full adjustment. Among the cases without smoothing, the introduction of time-varying shocks together with time-varying policy significantly improves the fit of the model. Allowing for time variation in the speed at which economic agents are learning does not instead lead to big alterations in the model's marginal likelihood. Considering k-periods ahead forecasts in the monetary policy rule (as in Clarida, Gali', and Gertler 2000) worsens the performance of the model. In the robustness section, I have considered several possible assumptions about the learning process of the agents. If agents use the VAR(1) in output and inflation as their perceived model of the economy, the model's fit worsens (the log marginal likelihood falls from -767 to -787). The joint estimation of initial beliefs improves the fit. The paper's assumption of constant-gain learning, under which agents discount more heavily past observations, leads to a much better fit than the alternative recursive least squares learning.

[20]See appendix for details.

7. Conclusions and Future Directions

Several papers find that monetary policy has considerably changed over the post-war sample. Clarida, Gali', and Gertler (2000), and Lubik and Schorfheide (2004) conclude that policy was 'passive' in the pre-1979 sample and became 'active' in the Volcker-Greenspan period. The policy rule they estimate leads to instability if inserted in a simple DSGE model with rational expectations commonly used to study monetary policy transmission.

This paper revisits the evidence on the evolution of monetary policy by departing from the assumption of rational expectations. I estimate a model where agents form expectations from econometric models and learn the relevant parameters over time.

In the model with learning, I find that some time variation in monetary policy emerges. But, in contrast to Clarida, Gali', and Gertler (2000) and Lubik and Schorfheide (2004), I find that the response coefficient to inflation was well above 1 also in the pre-1979 period. Therefore my results suggest that monetary policy was 'active' also in the 1960s-1970s and, conditional on the proposed model of the economy, it was not contributing to macroeconomic instability.

I have estimated the model by likelihood-based Bayesian methods, which have allowed me to estimate the constant gain coefficient jointly with the rest of the system. Such exercise tries to close a gap in the literature, which lacks estimates of the gain. The different estimates in the paper lie around 0.02 (except in the case of VAR(1) forecasting rules), and are close to Milani (2004b). Milani (2004a) and Branch and Evans (2005), instead, do not directly estimate the gain, but provide evidence on the values that would provide the best in-sample fit of a standard Phillips curve (Milani 2004a), and the best out-of-sample forecasting performance (Branch and Evans 2005). I find also some evidence of a change in the learning speed of the public after 1979. It would be interesting to study how changes in learning speed are affected by and affect different policies and different levels of macroeconomic volatility. Branch, Carlson, Evans, and McGough (2004), for example, study the interaction between optimal monetary policy and the anchoring of inflation expectations to explain the decline in macroeconomic volatility after the 1980s. An interesting direction is, in fact, to investigate the role of agents' learning on the "great moderation".

I have shown that under learning, the evidence of a regime switch of U.S. monetary policy from passive to active is weak. The paper has not proved, however, that the model with learning provides a better explanation of the data compared with an alternative rational expectations model in which policy indeed changes. I leave for future research to investigate the relative fit of the presented model with learning and a corresponding model with rational expectations that allows for instability and can be estimated with the techniques proposed by Lubik and Schorfheide (2004). But the evidence coming from well-fitting structural VARs seems to indicate a relatively stable policy, hence more in line with the results under learning.

The same literature also emphasizes the importance to correctly model time-varying volatilities. A useful extension would therefore consist of embedding stochastic volatility in DSGE models under both learning and rational expectations, to investigate the changes in the results.

Finally, it is important to remind that the evidence on instability is model specific and, therefore, considering a richer model or introducing fiscal policy together with monetary policy might importantly alter the results.

References

[1] Boivin, J., and M. Giannoni, (2003). "Has Monetary Policy Become More Effective?", NBER Working Paper no. 9459, January.

[2] Branch, W.A., Carlson, J., Evans, G.W., and B. McGough, (2004). "Monetary Policy, Endogenous Inattention, and the Volatility Trade-off", mimeo.

[3] Branch, W.A., and Evans, G.W., (2005). "A Recursive Forecasting Model", mimeo, University of California, Irvine, and University of Oregon.

[4] Canova, F., and L. Gambetti, (2004). "Structural Changes in the US Economy: Bad Luck or Bad Policy?", mimeo.

[5] Clarida, R., Gali', J., and M. Gertler, (2000). "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory", *Quarterly Journal of Economics*, vol. CXV, issue 1, pp. 147-180.

[6] Croushore, D., (1998). "Evaluating inflation forecasts", WP 98-14, Federal Reserve Bank of Philadelphia.

[7] Duffy, J., and J. Engle-Warnick, (2005). "Multiple Regimes in U.S. Monetary Policy? A Nonparametric Approach", forthcoming, *Journal of Money, Credit, and Banking.*

[8] Evans, G.W., and S. Honkapohja, (2001). *Learning and Expectations in Macroeconomic*s, Princeton University Press.

[9] ——— (2003). "Monetary Policy, Expectations and Commitment", mimeo, University of Oregon and University of Cambridge.

[10] Geweke, J.F., (1992). "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments", in J.O. Berger, J.M. Bernardo, A.P. Dawid, and A.F.M. Smith (eds.), *Proceedings of the Fourth Valencia International Meeting on Bayesian Statistics*, pp. 169-194, Oxford University Press.

[11] Ireland, P. N., (2001). "Sticky-Price Models of the Business Cycle: Specification and Stability," *Journal of Monetary Economics*, 47, 3-18.

[12] Justiniano, A., and B. Preston, (2004). "Small Open Economy DSGE Models: Specification, Estimation and Model Fit", mimeo, IMF and Columbia University.

[13] Lubik, T., and F. Schorfheide, (2004). "Testing for Indeterminacy: an Application to U.S. Monetary Policy", *American Economic Review*, 94(1), pp. 190-217.

[14] Marcet, A., and Nicolini J.P., (2003). "Recurrent Hyperinflations and Learning", *American Economic Review*, vol. 93, no. 5, pp. 1476-1498.

[15] McCallum, B.T., (1999). "Role of the Minimal State Variable Criterion in Rational Expectations Models", *International Tax and Public Finance*, Kluwer Academic Publishers, vol. 6(4), pages 621-39, November.

[16] Milani, F., (2004a). "Adaptive Learning and Inflation Persistence", mimeo, Princeton University.

[17] ——— (2004b). “Expectations, Learning and Macroeconomic Persistence”, mimeo, Princeton University.

[18] Orphanides, A., (2001). “Monetary Policy Rules Based on Real-Time Data”, *American Economic Review*, vol. 91(4), pages 964-985.

[19] ——— (2004). “Monetary Policy Rules, Macroeconomic Stability, and Inflation: A View from the Trenches”, *Journal of Money, Credit, and Banking*, vol. 36 (2), pp. 151-176.

[20] Orphanides, A., and J. Williams, (2003). “The Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning, and Expectations”, Federal Reserve Bank of San Francisco Working Paper 03-24.

[21] ——— (2004). “Robust Monetary Policy with Imperfect Knowledge”, paper presented at the Conference on *Monetary Policy and Imperfect Knowledge*, Wurzburg, October 14-15.

[22] Preston, B., (2003). “Learning About Monetary Policy Rules When Long-Horizon Expectations Matter”, *International Journal of Central Banking*, forthcoming.

[23] ——— (2004a). “Adaptive Learning and the Use of Forecasts in Monetary Policy”, *Journal of Monetary Economics*, forthcoming.

[24] ——— (2005). “Adaptive Learning in Infinite Horizon Decisions Problems”, mimeo, Columbia University.

[25] Primiceri, G., (2003). “Why Inflation Rose and Fell: Policymakers’ Beliefs and Postwar Stabilization Policy”, mimeo, Princeton University.

[26] ——— (2005). “Time Varying Structural Vector Autoregressions and Monetary Policy”, forthcoming, *Review of Economic Studies*.

[27] Raftery, A.E., and S.M. Lewis, (1995). “The Number of Iterations, Convergence Diagnostics, and Generic Metropolis algorithms”, *Practical Markov Chain Monte Carlo*, W.R. Gilks, D.J. Spiegelhalter, and S. Richardson, eds., Chapman and Hall, London.

[28] Sargent, T.J., (1993). *Bounded Rationality in Macroeconomics*. Oxford University Press.

[29] ——— (1999). *The Conquest of American Inflation*. Princeton University Press.

[30] Sargent, T.J., Williams, N., and T. Zha, (2004). “Shocks and Government Beliefs: The Rise and Fall of American Inflation”, NBER Working Paper No. 10764.

[31] Schorfheide, F., (2003). “Learning and Monetary Policy Shifts”, mimeo, University of Pennsylvania.

[32] Sims, C.A., and T. Zha, (2004). “Were There Regime Switches in US Monetary Policy?”, *American Economic Review*, forthcoming.

[33] Smets, F., and R. Wouters, (2003). “Monetary Policy in an Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area”, *Journal of the European Economic Association*, v. 1, iss. 5, pp. 1123-75.

[34] ——— (2004a). “Comparing Shocks and Frictions in US and Euro Business Cycles: A Bayesian DSGE Approach”, forthcoming, *Journal of Applied Econometrics*.

[35] ——— (2004b). “Forecasting with a Bayesian DSGE Model: An Application to the Euro Area”, *Journal of Common Market Studies,* 42 (4), 841-867.

[36] Williams, N., (2003). "Adaptive Learning and Business Cycles", mimeo, Princeton University.

[37] Woodford, M., (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press.

[38] Yu, B., and P. Mykland (1994). "Looking at Markov Samplers through CUMSUM path plots: a simple diagnostic idea", Technical Report 413, Dept. of Statistics, University of California at Berkeley.

Appendix A. Econometric Procedure

A.1. **Kalman Filter.** To generate draws from the posterior distribution of θ using the Metropolis algorithm, I need to evaluate the likelihood function $p\left(Y^T \mid \theta\right)$ at each iteration.

Substituting private expectations into (2.1), (2.2), and (2.3) yields the state-space form

$$\begin{aligned} \xi_t &= F_t \xi_{t-1} + G_t w_t \\ Z_t &= H \xi_t \end{aligned} \tag{A.1}$$

where $\xi_t = [x_t, \pi_t, i_t, u_t, r_t^n]$, $w_t \sim N(0, Q)$, H is a matrix of zeros and ones just selecting variables from ξ_t[21], and F_t, G_t are time-varying matrices of coefficients, which are convolutions of preference parameters and agents beliefs. Expression $(A.1)$ also represents the implied "**Actual Law of Motion**", or **ALM**, of the economy. Having expressed the model as a linear Gaussian system, I can easily compute the likelihood recursively with the Kalman Filter.

I run the Kalman filter recursion with standard steps to compute first and second moments of the following conditional distributions:

(1) Start from initial values:

$$p\left(\xi_t \mid Y^t, \theta\right) \tag{A.2}$$

(2) Prediction:

$$p\left(\xi_{t+1} \mid Y^t, \theta\right) = \int p\left(\xi_{t+1} \mid \xi_t, \theta\right) p\left(\xi_t \mid Y^t, \theta\right) d\xi_t \tag{A.3}$$

$$p\left(Y_{t+1} \mid Y^t, \theta\right) = \int p\left(Y_{t+1} \mid \xi_{t+1}, Y^t, \theta\right) p\left(\xi_{t+1} \mid Y^t, \theta\right) d\xi_{t+1} \tag{A.4}$$

(3) Updating:

$$p\left(\xi_{t+1} \mid Y^{t+1}, \theta\right) = \frac{p\left(Y_{t+1} \mid \xi_{t+1}, Y^t, \theta\right) p\left(\xi_{t+1} \mid Y^t, \theta\right)}{p\left(Y_{t+1} \mid Y^t, \theta\right)} \tag{A.5}$$

[21] Because of the well-known stochastic singularity of standard RE systems, where there are more endogenous variables than shocks (5 vs. 3 here), I follow the common approach of computing the likelihood only on a subset of the variables (x_t, π_t, i_t in this case).

(4) Evaluate likelihood function:

$$p\left(Y^T \mid \theta\right)=\prod_{t=1}^{T} p\left(Y_t \mid Y^{t-1}, \theta\right)$$

A.2. **Metropolis-Hastings Algorithm.** To generate draws from the posterior distribution $p\left(\theta \mid Y^T\right)$, I use the Metropolis algorithm. The procedure works as follows.

1. Start from an arbitrary value for the parameter vector θ_0. Set $j=1$.
2. Evaluate $p\left(Y^T \mid \theta_0\right) p\left(\theta_0\right)$
3. Generate $\theta_j^*=\theta_{j-1}+\varepsilon$, where θ_j^* is the proposal draw and $\varepsilon \sim N(0, c\Sigma_\varepsilon)$. c is a scale factor that is usually adjusted to keep the acceptance ratio of the MH algorithm at an optimal rate (25%-50%, see Geweke (1999)). I set c at 0.1 in the estimation that guarantees an acceptance rate slightly larger than 40%.
4. Generate u from $Uniform[0,1]$
5. Set $\begin{cases} \theta_j=\theta_j^* & \text{if } \; u \leq \alpha\left(\theta_{j-1}, \theta_j^*\right)=\min\left\{\frac{p(Y^T|\theta_j^*)p(\theta_j^*)}{p(Y^T|\theta_{j-1})p(\theta_{j-1})}, 1\right\} \\ \theta_j=\theta_{j-1} & \text{if } \; u>\alpha\left(\theta_{j-1}, \theta_j^*\right) \end{cases}$
6. Repeat for $j+1$ from 2. until $j=D$ ($D=$ total number of draws).

A.3. **Model Comparison.** I follow Geweke (1999)'s modified harmonic mean approximation to compute the marginal likelihood. The procedure works as follows. Let $\widehat{\theta}_D=\frac{1}{D}\sum_{d=1}^{D}\theta_d$ and $\widehat{\Sigma}_D=\frac{1}{D}\sum_{d=1}^{D}\left(\theta_d-\widehat{\theta}\right)\left(\theta_d-\widehat{\theta}\right)'$ be the estimates of $E(\theta \mid Y, \mathcal{M}_i)$ and $var(\theta \mid Y, \mathcal{M}_i)$ obtained from the output of the posterior simulator. Then, for a given $p \in (0,1)$ define the support of $f(\theta)$ as:

$$\widehat{\Theta}_D=\left\{\theta:\left(\theta_d-\widehat{\theta}\right)\widehat{\Sigma}_D^{-1}\left(\theta_d-\widehat{\theta}\right)' \leq \chi_{1-p}^2(k)\right\} \tag{A.6}$$

where $\chi_{1-p}^2(k)$ is the $(1-p)$th percentile of the Chi-squared distribution with degrees of freedom equal to the number of parameters k. Geweke (1999) recommends using the following Multivariate Normal density truncated to the region $\widehat{\Theta}_D$ as $f(\theta)$:

$$f(\theta)=\frac{1}{p\left(2\pi\right)^{k/2}}\left|\widehat{\Sigma}_D\right|^{-1/2}\exp\left[-\frac{1}{2}\left(\theta_d-\widehat{\theta}\right)\widehat{\Sigma}_D^{-1}\left(\theta_d-\widehat{\theta}\right)'\right]1\left(\theta \in \widehat{\Theta}_D\right) \tag{A.7}$$

where $\mathbf{1}(\cdot)$ is an indicator function.

Description	Parameters	Range	Prior Distr.	Prior Mean	Prior Std	95% Prior Prob. Int.
Discount rate	β	.99	–	.99	–	–
Phillips curve slope	κ	$\mathbb{R}$	$Normal$	.1	.05	$[.002, .198]$
IES	σ	.2	–	.2	–	–
Feedback Infl. (pre-79)	$\chi_{\pi,pre79}$	$\mathbb{R}$	$Normal$	1.5	.5	$[.52, 2.48]$
Feedback Gap (pre-79)	$\chi_{x,pre79}$	$\mathbb{R}$	$Normal$	.5	.25	$[.01, .99]$
Feedback Infl. (post-79)	$\chi_{\pi,post79}$	$\mathbb{R}$	$Normal$	1.5	.5	$[.52, 2.48]$
Feedback Gap (post-79)	$\chi_{x,post79}$	$\mathbb{R}$	$Normal$	.5	.25	$[.01, .99]$
Autoregr. Dem shock	ρ_r	0.9	–	.9	–	–
Autoregr. Sup shock	ρ_u	0.9	–	.9	–	–
MP shock	σ_ε	$\mathbb{R}^+$	$InvGamma$	1	.5	$[.34, 2.81]$
Demand shock	σ_r	$\mathbb{R}^+$	$InvGamma$	1	.5	$[.34, 2.81]$
Supply shock	σ_u	$\mathbb{R}^+$	$InvGamma$	1	.5	$[.34, 2.81]$
Gain Coeff.	$\bar{\mathbf{g}}$	$\mathbb{R}^+$	$Gamma$	.031	.022	$[.004, .086]$

Table 1 - Prior distributions.

Description	Parameters	Posterior Mean	Posterior Std.	95% Post. Prob. Interval
Discount rate	β	.99	–	–
Phillips curve slope	κ	.104	.045	$[.014, .19]$
IES	σ	.2	–	–
Feedback Infl. (pre-79)	$\chi_{\pi,pre79}$	1.688	.246	$[1.23, 2.21]$
Feedback Gap (pre-79)	$\chi_{x,pre79}$	.166	.112	$[-.061, .378]$
Feedback Infl. (post-79)	$\chi_{\pi,post79}$	2.708	.252	$[2.20, 3.25]$
Feedback Gap (post-79)	$\chi_{x,post79}$	$-.005$	.134	$[-.271, .246]$
Autoregr. Dem shock	ϕ_r	.9	–	–
Autoregr. Sup shock	ϕ_u	.9	–	–
MP shock	σ_ε	2.302	.130	$[2.05, 2.55]$
Demand shock	σ_r	.972	.064	$[.86, 1.11]$
Supply shock	σ_u	.815	.052	$[.72, .92]$
Gain Coeff.	$\overline{\mathbf{g}}$	.0234	.0178	$[0.0012, 0.06]$

Table 2 - Posterior estimates model with time-varying policy.

Description	Parameters	Posterior Mean	Posterior Std.	95% Post.Prob.Interval
Discount rate	β	.99	–	–
Phillips curve slope	κ	0.11045	0.043617	$[0.030739, 0.1933]$
IES	σ	.2	–	–
Feedback Infl. (pre-79)	$\chi_{\pi,pre79}$	1.6393	0.19089	$[1.2417, 2]$
Feedback Gap (pre-79)	$\chi_{x,pre79}$	0.21087	0.084768	$[0.050686, 0.37303]$
Feedback Infl. (post-79)	$\chi_{\pi,post79}$	2.5883	0.24729	$[2.1268, 3.0708]$
Feedback Gap (post-79)	$\chi_{x,post79}$	0.0027344	0.13442	$[-0.27952, 0.25214]$
Autoregr. Dem shock	ϕ_r	.9	–	–
Autoregr. Sup shock	ϕ_u	.9	–	–
MP shock (pre-79)	$\sigma_{\varepsilon,pre79}$	1.8704	0.15672	$[1.5943, 2.2232]$
Demand shock (pre-79)	$\sigma_{r,pre79}$	1.1393	0.09718	$[0.97582, 1.3515]$
Supply shock (pre-79)	$\sigma_{u,pre79}$	0.98827	0.082506	$[0.84048, 1.177]$
MP shock (post-79)	$\sigma_{\varepsilon,post79}$	2.5838	0.22926	$[2.2185, 3.1349]$
Demand shock (post-79)	$\sigma_{r,post79}$	0.80779	0.062761	$[0.69894, 0.94691]$
Supply shock (post-79)	$\sigma_{u,post79}$	0.64158	0.04759	$[0.55909, 0.74233]$
Gain Coeff.	$\overline{\mathbf{g}}$	0.01957	0.011233	$[0.0027297, 0.046046]$

Table 3 - Posterior estimates: model with time-varying policy and variances.

Description	Parameters	Posterior Mean	Posterior Std.	95% Post. Prob. Interval
Discount rate	β	.99	–	–
Phillips curve slope	κ	.093	.0428	$[.013, .173]$
IES	σ	.2	–	–
Feedback Infl. (pre-79)	$\chi_{\pi,pre79}$	1.649	.215	$[1.24, 2.10]$
Feedback Gap (pre-79)	$\chi_{x,pre79}$	.153	.108	$[-.071, .36]$
Feedback Infl. (post-79)	$\chi_{\pi,post79}$	2.66	.227	$[2.17, 3.09]$
Feedback Gap (post-79)	$\chi_{x,post79}$	$-.043$	.130	$[-0.29, 0.21]$
Autoregr. Dem shock	ϕ_r	.9	–	–
Autoregr. Sup shock	ϕ_u	.9	–	–
MP shock	σ_ε	2.3	.128	$[2.09, 2.58]$
Demand shock	σ_r	.96	.057	$[0.86, 1.07]$
Supply shock	σ_u	.80	.046	$[0.72, 0.905]$
Gain Coeff. (pre-79)	$\overline{\mathbf{g}}_{pre79}$	.03	.014	$[0.006, 0.056]$
Gain Coeff. (post-79)	$\overline{\mathbf{g}}_{post79}$	.0152	.0135	$[0.001, 0.053]$

Table 4 - Posterior estimates: model with time-varying policy and changing learning speed.

Description	Parameters	Posterior Mean	Posterior Std.	95% Post. Prob. Interval
Discount rate	β	.99	–	–
Phillips curve slope	κ	.092	.041	$[.011, .17]$
IES	σ	.2	–	–
Feedback Infl. (pre-79)	$\chi_{\pi,pre79}$	2.04	.29	$[1.47, 2.57]$
Feedback Gap (pre-79)	$\chi_{x,pre79}$	.17	.12	$[-.06, .41]$
Feedback Infl. (post-79)	$\chi_{\pi,post79}$	3.20	.29	$[2.59, 3.74]$
Feedback Gap (post-79)	$\chi_{x,post79}$	$-.03$	.14	$[-.30, .22]$
Autoregr. Dem shock	ϕ_r	.9	–	–
Autoregr. Sup shock	ϕ_u	.9	–	–
MP shock	σ_ε	2.33	.12	$[2.11, 2.58]$
Demand shock	σ_r	.95	.06	$[.84, 1.08]$
Supply shock	σ_u	.80	.05	$[.72, .89]$
Gain Coeff.	$\bar{\mathbf{g}}$	.0305	.015	$[.0049, .061]$

Table 5 - Posterior estimates: model with κ-period ahead forecasts

Description	Parameters	Posterior Mean	Posterior Std.	95% Post. Prob. Interval
Discount rate	β	.99	–	–
Phillips curve slope	κ	.016	.04	$[-.06, .10]$
IES	σ	.2	–	–
Interest-Rate Smooth (pre-79)	$\rho_{i,pre79}$	.825	.046	$[.73, .91]$
Feedback Infl. (pre-79)	$\chi_{\pi,pre79}$	1.851	.408	$[1.04, 2.68]$
Feedback Gap (pre-79)	$\chi_{x,pre79}$	.701	.197	$[.32, 1.10]$
Interest-Rate Smooth (post-79)	$\rho_{i,post79}$	.903	.031	$[.84, .96]$
Feedback Infl. (post-79)	$\chi_{\pi,post79}$	2.07	.51	$[1.09, 3.04]$
Feedback Gap (post-79)	$\chi_{x,post79}$	.57	.23	$[.14, 1]$
Autoregr. Dem shock	ϕ_r	.9	–	–
Autoregr. Sup shock	ϕ_u	.9	–	–
MP shock	σ_ε	.958	.050	$[.86, 1.06]$
Demand shock	σ_r	.997	.069	$[.88, 1.14]$
Supply shock	σ_u	.811	.050	$[.72, .92]$
Gain Coeff.	$\overline{\mathbf{g}}$	.0258	.021	$[.0008, .068]$

Table 6 - Posterior estimates: model with interest-rate smoothing.

Description	Parameters	Posterior Mean	Posterior Std.	95% Post. Prob. Interval
Discount rate	β	.99	–	–
Phillips curve slope	κ	.071	.045	$[-.02, .16]$
IES	σ	.2	–	–
Interest-Rate Smooth (pre-79)	$\rho_{i,pre79}$	.85	.05	$[.74, .95]$
Feedback Infl. (pre-79)	$\chi_{\pi,pre79}$	1.556	.446	$[.65, 2.41]$
Feedback Gap (pre-79)	$\chi_{x,pre79}$	.72	.21	$[.31, 1.11]$
Interest-Rate Smooth (post-79)	$\rho_{i,post79}$	.91	.03	$[.85, .96]$
Feedback Infl. (post-79)	$\chi_{\pi,post79}$	2.016	.50	$[1.06, 2.99]$
Feedback Gap (post-79)	$\chi_{x,post79}$	.489	.23	$[.06, .97]$
Autoregr. Dem shock	ϕ_r	.9	–	–
Autoregr. Sup shock	ϕ_u	.9	–	–
MP shock	σ_ε	.976	.053	$[.88, 1.09]$
Demand shock	σ_r	4.61	.22	$[4.17, 5.02]$
Supply shock	σ_u	1.45	.080	$[1.3, 1.6]$
Gain Coeff.	$\overline{\mathbf{g}}$	.0011	.0008	$[.0001, .0031]$

Table 7 - Posterior estimates: VAR(1) as learning rule.

Description	Parameters	Posterior Mean	Posterior Std.	95% Post. Prob. Interval
Discount rate	β	.99	–	–
Phillips curve slope	κ	.067	.047	$[-.03, .16]$
IES	σ	.2	–	–
Interest-Rate Smooth (pre-79)	$\rho_{i,pre79}$	.89	.04	$[.79, .96]$
Feedback Infl. (pre-79)	$\chi_{\pi,pre79}$	1.665	.49	$[.70, 2.64]$
Feedback Gap (pre-79)	$\chi_{x,pre79}$	.62	.26	$[.10, 1.11]$
Interest-Rate Smooth (post-79)	$\rho_{i,post79}$	.94	.02	$[.90, .97]$
Feedback Infl. (post-79)	$\chi_{\pi,post79}$	1.53	.51	$[.58, 2.58]$
Feedback Gap (post-79)	$\chi_{x,post79}$	.50	.24	$[.04, 1.01]$
Autoregr. Dem shock	ϕ_r	.9	–	–
Autoregr. Sup shock	ϕ_u	.9	–	–
MP shock	σ_ε	.99	.05	$[.90, 1.11]$
Demand shock	σ_r	4.48	.24	$[4.03, 4.90]$
Supply shock	σ_u	1.31	.07	$[1.17, 1.47]$
Initial beliefs	$b_{1,1\|0}$	.005	.011	$[-.23, .20]$
"	$c_{1,1\|0}$	.135	.091	$[-.04, .31]$
"	$b_{2,1\|0}$	.09	.06	$[-.04, .21]$
"	$c_{2,1\|0}$	.54	.07	$[.39, .68]$
Gain Coeff.	$\overline{\mathbf{g}}$	.0031	.0020	$[.0005, .0082]$

Table 8 - Posterior estimates: VAR(1) and joint estimation of initial beliefs.

Description	Parameters	Posterior Mean	Posterior Std.	95% Post. Prob. Interval
Discount rate	β	.99	–	–
Phillips curve slope	κ	$-.039$	.04	$[-.12, .05]$
IES	σ	.2	–	–
Interest-Rate Smooth (pre-79)	$\rho_{i,pre79}$	.83	.04	$[.74, .95]$
Feedback Infl. (pre-79)	$\chi_{\pi,pre79}$	1.65	.38	$[.88, 2.40]$
Feedback Gap (pre-79)	$\chi_{x,pre79}$	.57	.20	$[.16, .99]$
Interest-Rate Smooth (post-79)	$\rho_{i,post79}$	.88	.03	$[.85, .96]$
Feedback Infl. (post-79)	$\chi_{\pi,post79}$	2.37	.44	$[1.48, 3.21]$
Feedback Gap (post-79)	$\chi_{x,post79}$	.44	.23	$[-.01, .91]$
Autoregr. Dem shock	ϕ_r	.9	–	–
Autoregr. Sup shock	ϕ_u	.9	–	–
MP shock	σ_ε	.94	.05	$[.84, 1.3]$
Demand shock	σ_r	1.02	.06	$[.92, 1.14]$
Supply shock	σ_u	.799	.04	$[.72, .89]$
Gain Coeff.	$\overline{\mathbf{g}} = 1/t$	–	–	–

Table 9 - Posterior estimates: RLS learning.

Model Specifications	**Log Marginal Likelihood** (**Stand. Dev.**)
Time-Varying Mon. Policy	−919.02 (1.33)
Time-Varying MP and Variances	−908.21 (1.37)
Time-Varying MP and Learning Speed	−919.27 (1.54)
k-periods ahead Forecasts in MP rule	−924.90 (1.36)
MP rule with Interest-rate Smoothing	−767.11 (1.33)
VAR(1) Learning rule + IRS	−787.27 (1.33)
VAR(1) Learning rule + Est. Initial Beliefs + Int-rate Smooth.	−767.98 (1.37)
RLS Learning + Interest-rate Smoothing	−801.76 (1.47)

Table 10 - Model Comparison: Log Marginal Likelihoods and Standard Deviations.

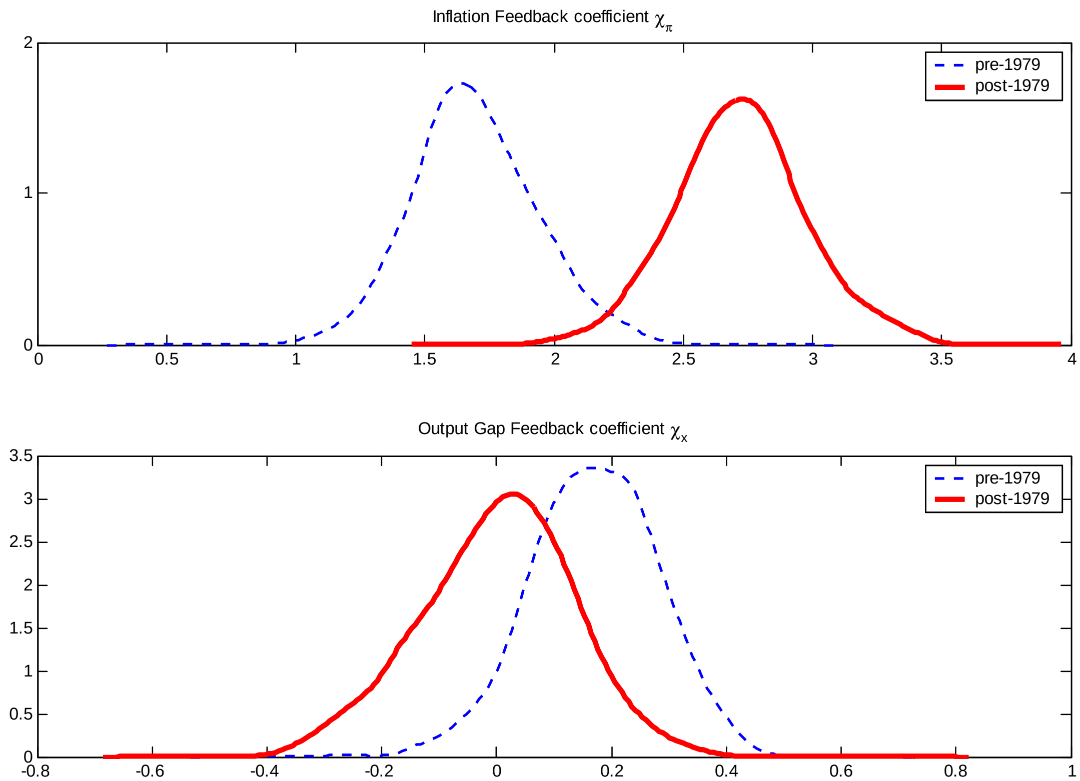

Figure 1 - Policy Coefficients Posterior Distributions.

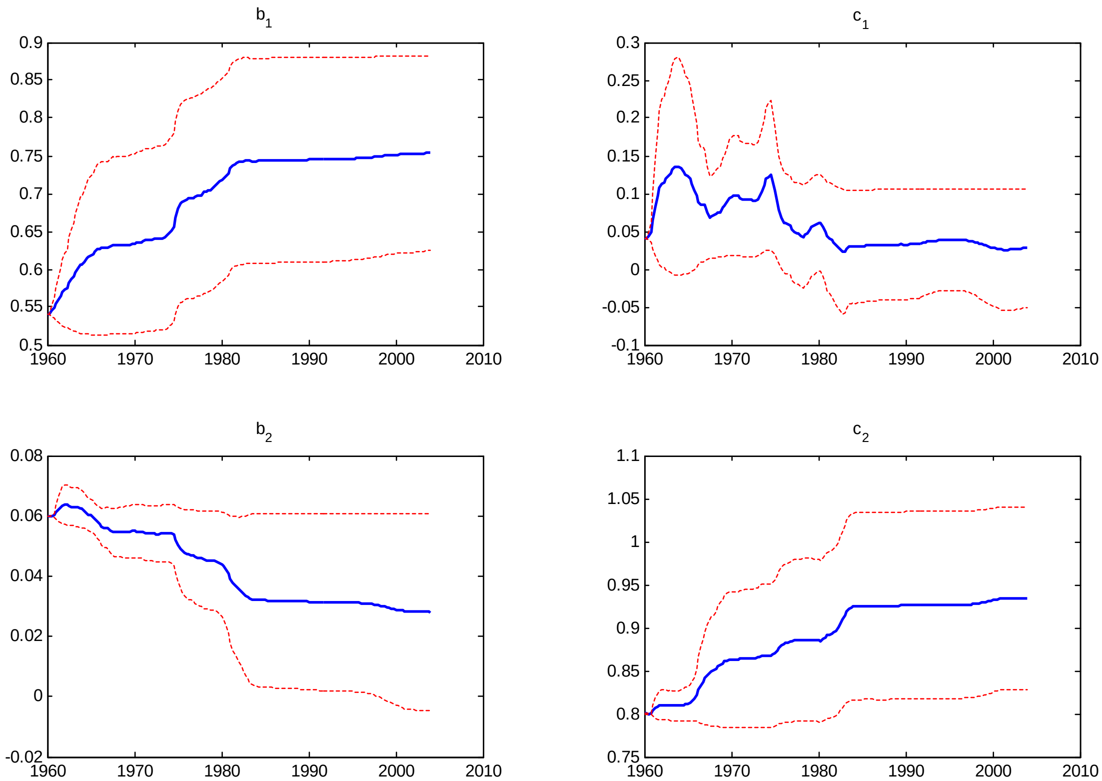

Figure 2 - Evolution of economic agents' beliefs.

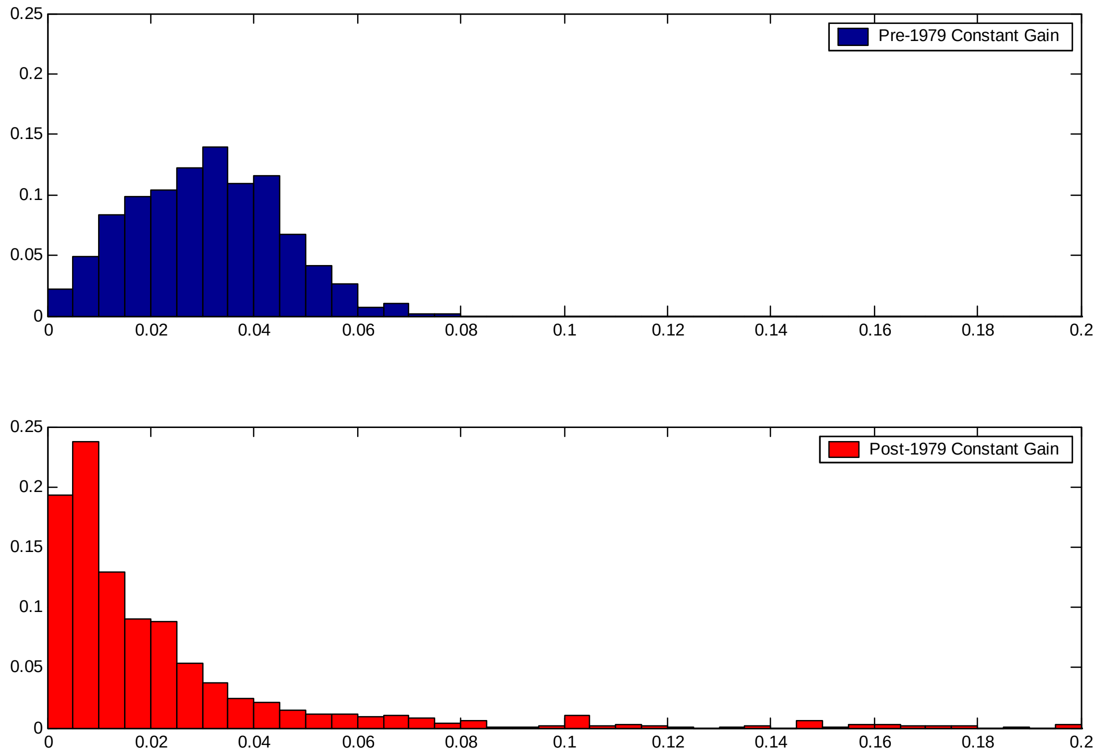

Figure 3 - Histograms about the constant gain coefficient distribution in the pre-1979 and post-1979 samples.

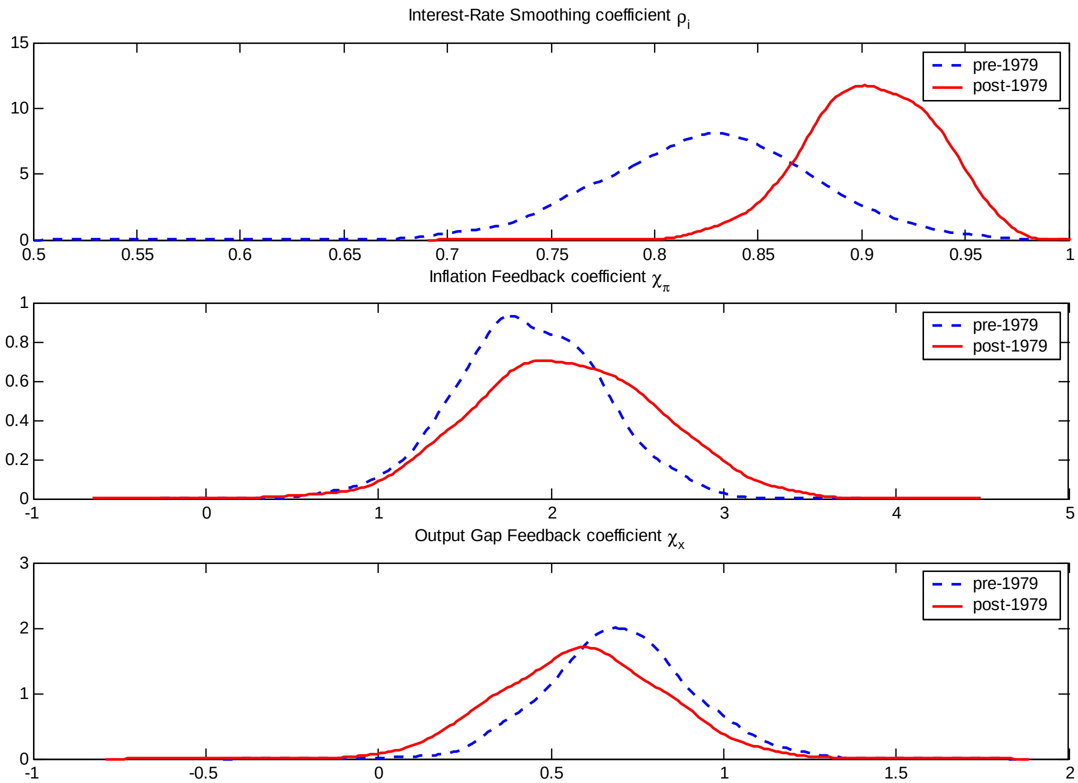

Figure 4 - Policy Coefficients Posterior Distributions.

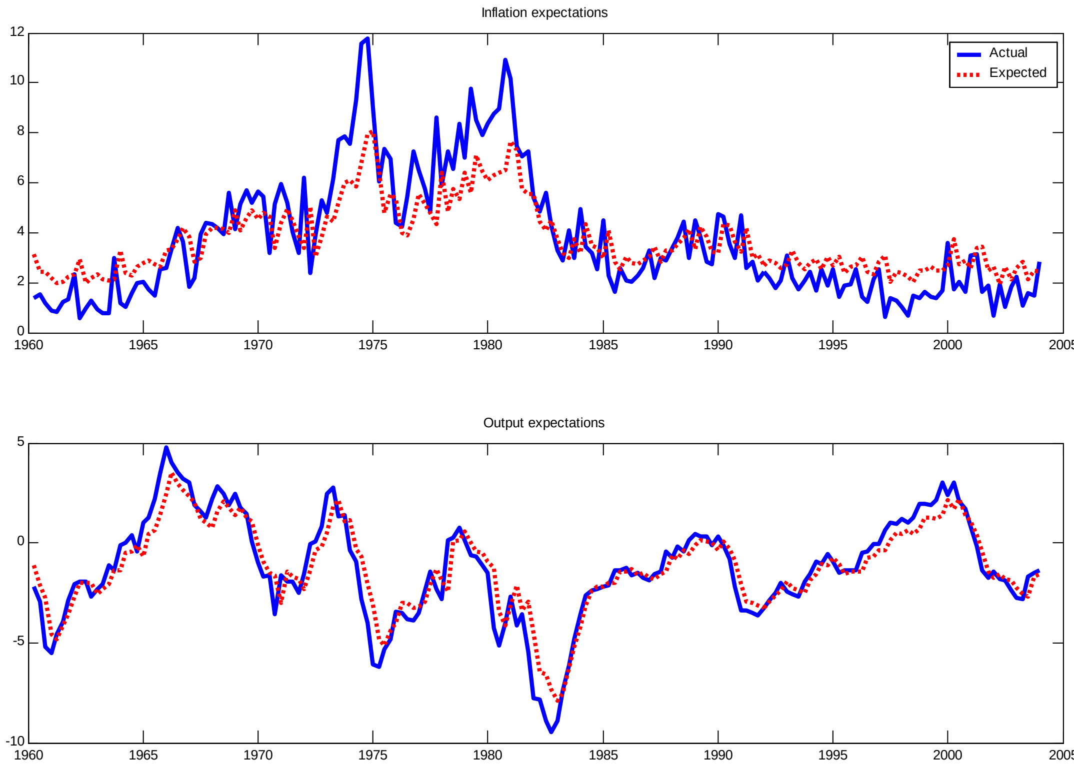

Figure 5 - Inflation and output gap expectations and actual data.

Fabio Milani, Princeton University, Department of Economics, Fisher Hall, Princeton, NJ 08544-1021.

E-mail address: `fmilani@princeton.edu`

URL: `http://www.princeton.edu/~fmilani`