Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Journal of Economic Dynamics & Control 32 (2008) 3148–3165

JOURNAL OF Economic Dynamics & Control

www.elsevier.com/locate/jedc

Learning, monetary policy rules, and macroeconomic stability

Fabio Milani

Department of Economics, University of California, Irvine, 3151 Social Science Plaza, Irvine, CA 92697-5100, USA

Received 9 September 2005; accepted 12 December 2007
Available online 18 January 2008

Abstract

Several papers have documented a regime switch in US monetary policy from 'passive' and destabilizing in the pre-1979 period to 'active' and stabilizing afterwards. These studies typically work with DSGE models with rational expectations.

This paper relaxes the assumption of rational expectations and allows for learning instead. Economic agents form expectations from simple models and update the parameters through constant-gain learning. In this setting, the paper aims to test whether monetary policy may have been a source of macroeconomic instability in the 1970s by inducing unstable learning dynamics.

The model is estimated by Bayesian methods. The constant-gain coefficient is jointly estimated with the structural and policy parameters in the system.

The results show that monetary policy was respecting the Taylor principle also in the pre-1979 period and, therefore, did not trigger macroeconomic instability.

JEL classification: C11; D84; E30; E50; E52; E58

Keywords: Monetary policy; Learnability; Constant-gain learning; Expectations; Bayesian estimation; Macroeconomic instability

1. Introduction

A large literature has studied the evolution of US monetary policy over the post-war period and the effects of monetary policy on macroeconomic stability. Influential papers

E-mail address: fmilani@uci.edu
URL: http://www.socsci.uci.edu/~fmilani

0165-1889/$ - see front matter
doi:10.1016/j.jedc.2007.12.004

by Clarida et al. (CGG, 2000) and Lubik and Schorfheide (LS, 2004) have argued that policy was substantially different in the pre-1979 period compared with the following two decades. CGG estimate a single equation – a forward-looking Taylor rule – by GMM, while LS use full-information methods to estimate a New Keynesian model with rational expectations. Both papers conclude that the estimated monetary policy rule in the pre-1979 sample fails to satisfy the so-called 'Taylor principle' and may have been a source of macroeconomic instability by allowing the existence of 'sunspot equilibria'.

This paper avoids imposing rational expectations and it introduces, instead, learning by economic agents. A recent literature[1] highlights, in fact, the strong informational requirements of economic agents under rational expectations and proposes to relax this assumption in favor of agents that form expectations from simple economic models and need to learn the true model details over time.

In a model with learning, a failure to satisfy the Taylor principle, as implied by CGG and LS's estimates, would still be a source of endogenous macroeconomic instability. But it would produce instability for a different reason. In a model with learning, in fact, a monetary policy rule that fails to satisfy the Taylor principle would prevent the learning dynamics from converging to the rational expectations equilibrium (REE) of the economy. The system may, therefore, be characterized by *unstable learning dynamics*.

This paper aims to estimate a model with learning to evaluate whether unstable learning dynamics indeed existed in the pre-1979 period. In the model I will present, the Taylor principle represents a necessary and sufficient condition for learnability of the true rational expectations solution. Therefore, the paper will focus on checking whether monetary policy satisfied the Taylor principle to derive evidence on unstable learning dynamics in the 1960s and 1970s.

Similarly to LS, I adopt full-information Bayesian methods in the estimation. But differently from them, I relax the assumption of rational expectations and allow for near-rational expectations and learning.

Under rational expectations, the papers that estimate structural models 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. This approach rules out by construction estimates of the monetary policy rule that do not respect the Taylor principle. LS are the first to provide the tools to extend the likelihood function to the indeterminacy region, thus allowing for determinacy and indeterminacy in the estimation under rational expectations. But such complications are not needed under subjective expectations and learning. My framework is, therefore, particularly suited to study the evolution of US monetary policy over time and to investigate the determinacy, indeterminacy, and learnability properties of estimated monetary policy rules, both in the pre- and post-1979 samples.

I find that monetary policy has satisfied the Taylor principle also in the pre-1979 period. The results, therefore, indicate that monetary policy was not a source of instability in the pre-Volcker sample.[2] The estimates imply that, in the case of a decreasing gain, there

[1]See Sargent (1993, 1999), Evans and Honkapohja (2001), and Bullard (2007).

[2]The results seem consistent with the evidence from time-varying coefficients VARs. Sims and Zha (2006), Primiceri (2005), and Canova and Gambetti (2006), 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.

should have been convergence in the limit to the REE. Since the model assumes a constant, rather than decreasing gain, and the data favor a rather small constant gain, it can be concluded that the learning dynamics should have led the economy to fluctuate in the limit around the REE, but without departing too far from it.

The results are, therefore, suggestive that evidence on the instability of monetary policy in the 1970s may be dependent on the assumptions about expectations. The results are consistent with the findings in Orphanides (2001, 2004), who replicate the exercise of CGG, but exploiting the forecasts that were available to the fed in real time. Using forecast from the Greenbook, instead of rational expectations, Orphanides likewise does not detect evidence of failure of the Taylor principle in the pre-1979 sample.[3]

2. The model

I assume that the aggregate dynamics of the economy can be summarized by the following New Keynesian model, which can be derived from the optimizing behavior of households and firms (see Woodford, 2003 for details):

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

$$x_t = \hat{\mathrm{E}}_t x_{t+1} - \sigma(i_t - \hat{\mathrm{E}}_t \pi_{t+1} - r_t^n), \tag{2}$$

$$i_t = \chi_t^\pi \hat{\mathrm{E}}_t \pi_{t+1} + \chi_t^x \hat{\mathrm{E}}_t x_{t+1} + \varepsilon_t, \tag{3}$$

where π_t is inflation, x_t is the output gap, and i_t is the nominal interest rate.[4] Eq. (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. $\hat{\mathrm{E}}_t$ indicates subjective (possibly non-rational) expectations.[5]

Eq. (2) represents the log-linearized intertemporal Euler equation that derives from the households' optimal choice of consumption. The output gap depends on expected one-period ahead output gap and on the ex-ante real interest rate. The coefficient $\sigma>0$ represents the intertemporal elasticity of substitution of consumption. u_t denotes cost-push shocks and r_t^{n} the natural real interest rate. They evolve according to univariate AR(1) processes as $u_t = \rho_u u_{t-1} + v_t^u$ and $r_t^{\mathrm{n}} = \rho_r r_{t-1}^{\mathrm{n}} + v_t^{\mathrm{r}}$, where $v_t^u \sim iid(0, \sigma_u^2)$ and $v_t^r \sim iid(0, \sigma_r^2)$. Eq. (3) describes monetary policy. The central bank follows a rule by adjusting its policy instrument, a short-term nominal interest rate, in response to deviations of expected inflation and expected output gap from their targets. 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 CGG, that a regime switch in policy occurs in 1979, when Paul Volcker

[3]Orphanides exploits real-time forecasts from the Greenbook. The similarity of results should not be surprising, since the approach described in this paper may be seen as a simple way to model the formation of those forecasts.

[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 obtained under rational expectations and introducing learning from that point. For a different approach, see Preston (2005, 2006), who introduces learning directly from the primitive assumptions of multi-period decision problems.

begins his term as Chairman of the Federal Reserve (August 1979). Duffy and Engle-Warnick (2006), using nonparametric methods, similarly identify a switch in policy exactly in the third quarter of 1979. The policy coefficients, therefore, evolve as follows:

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

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 (see Evans and Honkapohja, 2001 for an introduction) in assuming that agents have the same knowledge an econometrician would have about the economy and need to learn the relevant parameters over time.

2.1. Expectations formation and constant-gain learning

Economic agents in the model need to form forecasts of future macroeconomic conditions and they are assumed to use simple linear economic models to form those forecasts.

Notice that expectations appear also in the monetary policy rule (3). Such a rule is consistent with two different assumptions: (1) the central bank reacts to observed private sector expectations; (2) the central bank reacts to internal forecasts, which are formed in the same way as private sector forecasts.[6] Both interpretations are possible here and do not affect the results on learnability. They may affect, however, the interpretation of the results as I will discuss later.

Agents estimate the linear specification

$$Z_t = a_t + b_t u_t + c_t r_t^{\mathrm{n}} + \varepsilon_t \tag{4}$$

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 a_t, b_t, and c_t are coefficient vectors of appropriate dimensions).[7] Expression (4) represents the *Perceived Law of Motion* or *PLM* of the agents. The PLM hence assumes that economic agents and the central bank condition their learning on variables that would be unobservable in reality. This clearly represents a weakness of the estimated model. One could enrich the model so that the MSV solution, which agents are assumed to adopt as their PLM, includes only observables, or both observable and unobservable variables.

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

$$\hat{\phi}_t = \hat{\phi}_{t-1} + \bar{\mathbf{g}} R_{t-1}^{-1} X_t (Z_t - X_t' \hat{\phi}_{t-1}), \tag{5}$$

$$R_t = R_{t-1} + \bar{\mathbf{g}} (X_{t-1} X_{t-1}' - R_{t-1}), \tag{6}$$

[6] I thank a referee for helping clarify this point.

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

where $\hat{\phi}_t = (a_t', vec(b_t, c_t)')'$ describes the updating of the learning rule coefficients, and R_t the updating of the matrix of second moments of the stacked regressors $X_t \equiv \{1, u_t, r_t^n\}_0^{t-1}$. $\bar{\mathbf{g}}$ denotes the constant-gain coefficient. Under recursive least squares learning, the gain would equal $1/t$; here the gain is instead constant, providing a simple way to model learning of economic agents concerned about future unknown structural breaks.[8]

Although the true constants in the REE of the model equal zero, I do not endow agents with this information. The agents also need to learn about the constant terms in the model.[9]

In its baseline specification, the PLM is very close to rational expectations and, therefore, it can be similarly unrealistic. Even minor deviations from rational expectations; however, are often found to lead to substantially different results.

Economic agents use (4), the observed value of the shocks,[10] and the updated parameter estimates to form expectations for $t+1$:

$$\hat{\mathrm{E}}_t \begin{bmatrix} \pi_{t+1} \\ x_{t+1} \end{bmatrix} = \begin{bmatrix} a_{1,t} \\ a_{2,t} \end{bmatrix} + \begin{bmatrix} b_{1,t} & c_{1,t} \\ b_{2,t} & c_{2,t} \end{bmatrix} \begin{bmatrix} \rho_u u_t \\ \rho_x r_t^n \end{bmatrix}. \tag{7}$$

Learning in this paper is, therefore, modeled as adaptive as in Sargent (1999) and Evans and Honkapohja (2001): in each period t, agents and the central bank update their parameter estimates, but they act as if those estimates will remain fixed for all future periods. This is sub-optimal. An alternative (optimal) approach to model learning is to assume agents engage in Bayesian, or rational, learning, as studied in Wieland (2000, 2006). In this case, economic agents and the central bank update their parameter estimates, but they also take their likely future updating into account when solving current optimization problems. Rational learning would lead to an optimal degree of experimentation to reduce future uncertainty.

2.2. Equilibrium

2.2.1. Determinacy

The monetary DSGE model described by Eqs. (1)–(3) has a unique REE if the following necessary and sufficient conditions (corresponding to Proposition 4 in Bullard and Mitra, 2002) are satisfied:

$$\chi^x < \sigma^{-1}(1+\beta^{-1}), \tag{8}$$

$$\kappa(\chi^\pi - 1) + (1+\beta)\chi^x < 2\sigma^{-1}(1+\beta), \tag{9}$$

[8]Constant-gain learning has been used in Sargent (1999), Orphanides and Williams (2005, 2007), Williams (2003), and Primiceri (2006), among others.

[9]Federal Reserve's policy is assumed consistent with a steady-state inflation rate of $\hat{\pi}^*$ both in the pre- and post-1979 periods (hence Eq. (3) also has a zero constant), even though the feedback coefficients in the Taylor rule are allowed to vary. But the economic agents in the model do not know, at any point in time, whether the Fed's policy is indeed consistent with $\hat{\pi}^*$ or not, and, therefore, they also need to estimate intercepts a_t in their PLMs. Those estimates change over time and provide an indication of the agents' perceived average inflation rate.

[10]The information assumptions are similar to Evans and Honkapohja (2006), Preston (2005), and Bullard and Mitra (2002).

$$\chi^{\pi} + \left(\frac{1-\beta}{\kappa}\right)\chi^{x} > 1. \tag{10}$$

Condition (10) represents the so-called 'Taylor principle', also discussed in Woodford (2003). When the response to the output gap is small, $\chi^{\pi} > 1$ guarantees determinacy. But a sufficiently large response to the output gap leads to determinacy even if $\chi^{\pi} < 1$. Notice also that a too large response to inflation can lead to violation of condition (9) and, therefore, to indeterminacy.

2.2.2. *Learnability*

The necessary and sufficient condition for the MSV solution to be learnable (see Proposition 5 in Bullard and Mitra, 2002), or 'E-stable', is

$$\chi^{\pi} + \left(\frac{1-\beta}{\kappa}\right)\chi^{x} > 1, \tag{11}$$

i.e. the Taylor principle should be satisfied. Notice that learnability in this model does not imply determinacy.[11]

2.3. *Pre-Volcker monetary policy and the Taylor principle*

In a similar model under rational expectations, CGG and LS find that the estimated policy rule in the US violates the Taylor principle in the pre-1979 period. Conditions (8) and (9) are, instead, satisfied both in the pre-1979 and in the post-1982 samples for typical parameter estimates.

Here, I re-estimate the monetary policy rule to assess if their results can be confirmed when learning replaces the strong informational assumptions required by rational expectations. The above-cited papers conclude that the failure of policy to satisfy the Taylor principle in the pre-Volcker sample may have been conducive to macroeconomic instability, by allowing the existence of sunspot equilibria. In a model with learning, a monetary policy that fails to satisfy the Taylor principle would also be a source of endogenous instability. In a model with learning, however, instability would arise from the failure of learning dynamics to converge to the REE. A policy that does not satisfy the Taylor principle would, therefore, trigger unstable learning dynamics.

In the paper, I aim to test whether the Taylor principle was satisfied in the pre-1979 period, to check for the possibility of unstable learning dynamics in the 1970s (as described, the Taylor principle is the necessary and sufficient condition for learnability).

Differently from CGG, 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 akin to LS.[12] A difference in the model lies in the monetary policy rule: the

[11]In some cases, it is possible, as Evans and McGough (2005) show, that if the determinacy conditions are not satisfied, agents may be able to learn sunspot solutions.

[12]Full system estimation is more efficient. Moreover, as Sims and Zha (2006) 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.

presented model does not include an interest-rate smoothing term. But later in the paper I examine the robustness of the results to this choice. As in CGG, the central bank is assumed to respond to forecasts of inflation and output gap.

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 are truly structural and do not vary. The only time-variation in the model lies in the monetary policy rule.

To summarize, the model economy is represented by the inflation dynamics equations (1), output equation (2), time-varying monetary policy rule (3), and the expectations formation mechanism expressed by (5), (6), and (7).

3. Econometric approach

I estimate the model by likelihood-based Bayesian methods. A recent literature in macroeconomics (e.g. An and Schorfheide, 2007; Schorfheide, 2003; Smets and Wouters, 2003, 2004, 2005) employs a similar Bayesian approach to estimate DSGE models under the standard hypothesis of rational expectations, whereas here the approach is extended to deal with for non-fully rational expectations and learning. The econometric procedure allows me to estimate the main learning parameters, the constant-gain coefficient here, *jointly* with the parameters describing preferences and monetary policy in the economy (a similar estimation strategy was used in Milani, 2007).[13] The initial beliefs of the agents are, instead, estimated from pre-sample data.[14]

The parameters are collected in the vector θ

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

where Q is the variance–covariance matrix collecting the variances of the supply, demand, and policy disturbances. The model can be written in state-space form and its likelihood computed at each iteration through the Kalman filter. I then generate draws from the posterior distribution using the Metropolis–Hastings algorithm (see Appendix A). The model is fitted to US data on inflation, output gap, and nominal interest rates. The data are quarterly for the period 1960:I–2004:II.[15]

3.1. Specifying the prior distribution

Table 1 presents information about the priors, which are assumed independent. I fix β equal to 0.99. I assume a Gamma distribution with mean 1 for σ. The slope coefficient of the Phillips curve, κ, follows a Normal distribution with mean 0.1 and standard deviation

[13] Estimating also the constant-gain coefficient is crucial, since one's results might heavily depend on the chosen gain. For instance, in Milani (2004) I have shown how the estimated backward-lookingness in inflation varies over the possible gain values.

[14] An earlier version of the paper tested the robustness of results to also estimating the initial beliefs jointly with the other parameters.

[15] The series were obtained from FRED, the database of the Federal Reserve Bank of Saint Louis. 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 the federal funds rate is used as the policy instrument. All the variables have been demeaned prior to estimation. I run 300,000 draws for the Markov Chain, discarding the first 50% as burn-in and saving one draw every 100.

Table 1
Prior distributions

Parameters	Range	Prior distr.	Prior mean	Prior std.	95% Prior prob. int.
β	0.99	–	0.99	–	–
κ	$\mathbb{R}$	Normal	0.1	0.05	[0.002, 0.198]
σ	$\mathbb{R}^+$	Gamma	1	0.7	[0.12, 2.78]
$\chi_{\pi,\text{pre-79}}$	$\mathbb{R}$	Normal	1.5	0.5	[0.52, 2.48]
$\chi_{x,\text{pre-79}}$	$\mathbb{R}$	Normal	0.5	0.25	[0.01, .99]
$\chi_{\pi,\text{post-79}}$	$\mathbb{R}$	Normal	1.5	0.5	[0.52, 2.48]
$\chi_{x,\text{post-79}}$	$\mathbb{R}$	Normal	0.5	0.25	[0.01, 0.99]
ρ_r	[0, 0.97]	Uniform	0.485	0.28	[0.024, 0.946]
ρ_u	[0, 0.97]	Uniform	0.485	0.28	[0.024, 0.946]
σ_ε	$\mathbb{R}^+$	InvGamma	1	0.5	[0.34, 2.81]
σ_r	$\mathbb{R}^+$	InvGamma	1	0.5	[0.34, 2.81]
σ_u	$\mathbb{R}^+$	InvGamma	1	0.5	[0.34, 2.81]
$\bar{\mathbf{g}}$	$\mathbb{R}^+$	Gamma	0.031	0.022	[0.004, 0.086]

0.05. The monetary policy rule coefficients also follow Normal distributions with mean 1.5 and standard deviation 0.5 for the inflation feedback coefficients, and mean 0.5 and standard deviation 0.25 for the output feedback coefficients. The 95% prior probability interval, therefore, includes values of the response to inflation below 1, which are likely to lead to macroeconomic instability (not satisfying the Taylor principle). I choose inverse gamma distributions for the standard deviations of the shocks and uniform distributions for the autoregressive coefficients. Finally, I assume a Gamma distribution for the constant-gain coefficient: the gain has prior mean 0.031 and prior standard deviation 0.022.

4. Estimating the model with learning: empirical results

4.1. Has monetary policy changed?

I first estimate the baseline model described in (1)–(3) for the 1960–2004 sample.

In the simple monetary model studied in this paper, the equilibrium is indeterminate and not learnable if monetary policy does not respect the Taylor principle. As already discussed, the empirical literature estimating DSGE models under rational expectations typically imposes the restriction that the parameters lay within the determinacy region. LS, 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 result in a failure of the Taylor principle without the complications needed in LS to extend the likelihood function to the indeterminacy region.

Table 2
Posterior estimates model with time-varying policy

Description	Parameters	Post. mean	Post. std.	95% Post. prob. int.
Discount rate	β	0.99	–	–
Phillips curve slope	κ	0.052	0.04	[−0.03, 0.13]
IES	σ	0.03	0.02	[0.003, 0.08]
Feedback infl. (pre-79)	$\chi_{\pi,\text{pre-79}}$	1.77	0.24	[1.29, 2.28]
Feedback gap (pre-79)	$\chi_{x,\text{pre-79}}$	0.34	0.09	[0.17, 0.53]
Feedback infl. (post-79)	$\chi_{\pi,\text{post-79}}$	2.62	0.25	[2.14, 3.13]
Feedback gap (post-79)	$\chi_{x,\text{post-79}}$	0.25	0.1	[0.06, 0.45]
Autoregr. dem shock	ϕ_r	0.91	0.02	[0.86, 0.95]
Autoregr. sup shock	ϕ_u	0.79	0.04	[0.72, 0.87]
MP shock	σ_ε	2.13	0.12	[1.91, 2.37]
Demand shock	σ_r	1.22	0.14	[0.97, 1.52]
Supply shock	σ_u	1.13	0.09	[0.96, 1.34]
Gain coeff.	$\bar{\mathbf{g}}$	0.020	0.003	[0.014, 0.027]

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 consistent with the Taylor principle also in the 1960s and 1970s and it was not a source of macroeconomic instability. The posterior means for the inflation and output feedback coefficients equal 1.77 and 0.34 in the pre-1979 sample and equal to 2.62 and 0.25 in the Volcker–Greenspan sample. The reported 95% posterior probability interval for χ_π also remains well above 1 in both samples.

Fig. 1 overlaps the posterior distributions of the pre- and post-1979 policy feedback coefficients: the posterior distribution of the inflation response coefficients has moved rightward, whereas the response to output fluctuations has, instead, attenuated.

A central coefficient in the estimation is represented by the constant gain $\bar{\mathbf{g}}$. The posterior mean estimate for the gain equals 0.020. 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 values found by Orphanides (2004) and Orphanides and Williams (2005), using data from the Survey of Professional Forecasters, by Branch and Evans (2006), and by Milani (2007), who estimates $\bar{\mathbf{g}} = 0.0183$ by Bayesian methods.

Fig. 2 shows the evolution of inflation and output expectations compared with realized values. Agents have constantly underestimated inflation during the late 1960s and most of the 1970s, while they have overestimated inflation in the first half of the 1960s and in the 1990s. This pattern is consistent with inflation forecasts from surveys, as analyzed by Croushore (1998).

Therefore, bad policy in the pre-1979 period arises in the model not from a low reaction coefficient to inflation, but from a reaction to inflation forecasts that persistently underestimated actual inflation. Under the interpretation of policy rule (3) that favors the central bank's reaction to its own forecasts, the model would attribute the high inflation in the 1970s to the policymaker's slowness in learning about the economy. The alternative

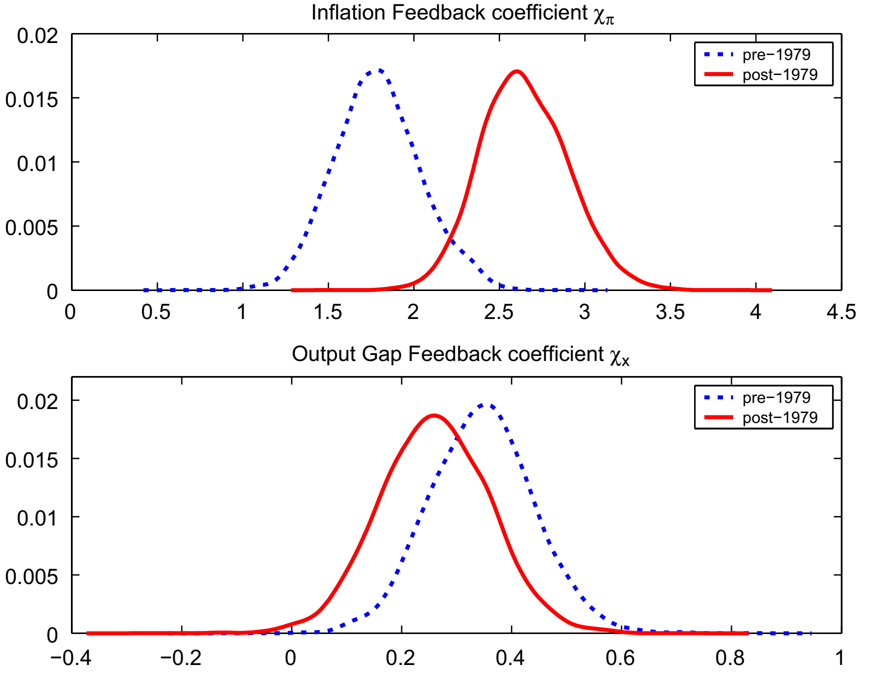

Fig. 1. Policy coefficients: Posterior distributions.

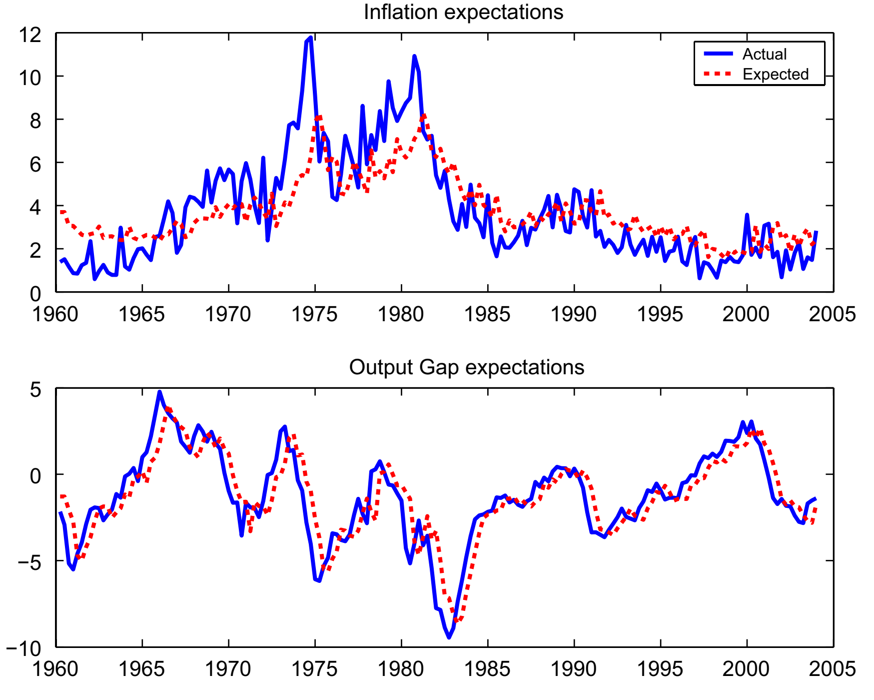

Fig. 2. Inflation and output gap expectations and actual data.

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

Description	Parameters	Post. mean	Post. std.	95% Post. prob. int.
Discount rate	β	0.99	–	–
Phillips curve slope	κ	0.092	0.04	[0.02, 0.17]
IES	σ	0.076	0.06	[0.005, 0.22]
Feedback infl. (pre-79)	$\chi_{\pi,\text{pre-79}}$	1.62	0.23	[1.16, 2.11]
Feedback gap (pre-79)	$\chi_{x,\text{pre-79}}$	0.37	0.08	[0.21, 0.54]
Feedback infl. (post-79)	$\chi_{\pi,\text{post-79}}$	2.32	0.24	[1.88, 2.81]
Feedback gap (post-79)	$\chi_{x,\text{post-79}}$	0.23	0.11	[0.02, 0.45]
Autoregr. dem shock	ϕ_r	0.92	0.02	[0.88, 0.96]
Autoregr. sup shock	ϕ_u	0.82	0.04	[0.74, 0.9]
MP shock (pre-79)	$\sigma_{\varepsilon,\text{post-79}}$	1.87	0.16	[1.6, 2.23]
Demand shock (pre-79)	$\sigma_{r,\text{pre-79}}$	1.2	0.15	[0.94, 1.5]
Supply shock (pre-79)	$\sigma_{u,\text{pre-79}}$	1.37	0.14	[1.12, 1.67]
MP shock (post-79)	$\sigma_{\varepsilon,\text{pre-79}}$	2.34	0.18	[2.03, 2.73]
Demand shock (post-79)	$\sigma_{r,\text{post-79}}$	0.91	0.18	[.63, 1.32]
Supply shock (post-79)	$\sigma_{u,\text{post-79}}$	0.59	0.07	[0.46, 0.75]
Gain coeff.	$\bar{\mathbf{g}}$	0.025	0.005	[0.017, 0.037]

interpretation, instead, shows the dangers of responding to private sector expectations when those are largely misaligned.

4.2. *Robustness*

4.2.1. *Time-varying shocks and time-varying policy*

Sims and Zha (2006), using a structural VAR with regime switches, 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.[16]

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 reported in Table 3 and Fig. 3 show that the standard deviations of the supply and demand innovations decreased in the 1980s and 1990s, consistently with what expected (particularly significant is the drop in the volatility of the supply shock).[17]

The estimates of the policy rule and gain coefficients remain in the ballpark of the previous case.

[16]Notice, however, that this paper finds no evidence of dramatic changes in the coefficients, i.e. a switch from passive to active policy, even omitting time-variation in the variances of the disturbances when learning is considered.

[17]It might be, instead, surprising to see that the variance of the policy disturbance increases. Certainly, as Sims and Zha (2006) discuss, a single shift in variance hardly captures the evolution of the policy shock variance. In my estimation, the larger post-1979 variance basically originates from the 1979 to 1982 years of non-borrowed reserves targeting that led to a strong volatility of the federal funds rate.

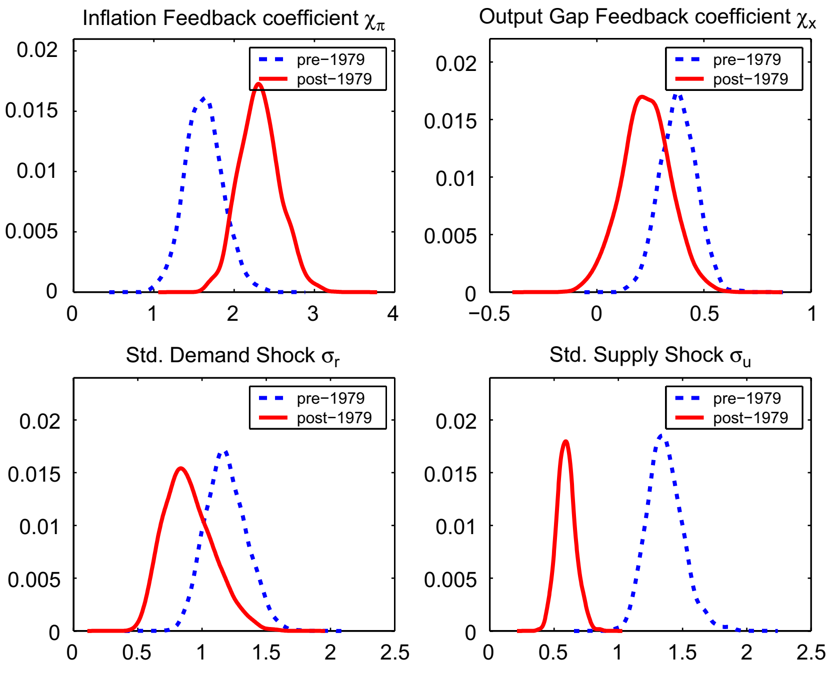

Fig. 3. Policy coefficients and standard deviations of the shocks: Posterior distributions.

4.2.2. *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:

$$i_t = \rho_t^i i_{t-1} + (1 - \rho_t^i)[\chi_t^\pi \hat{\mathrm{E}}_t \pi_{t+1} + \chi_t^x \hat{\mathrm{E}}_t x_{t+1}] + \varepsilon_t, \tag{12}$$

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 CGG and LS. The relevant determinacy and E-stability conditions become those derived in Bullard and Mitra (2007). Table 4 shows the posterior estimates, which are again similar.

4.2.3. *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, the results may be sensitive to the specific PLM that is chosen. A serious comparison of alternative learning models would be worthwhile, but it is beyond the scope of this paper. Here, however, I check whether the empirical results are robust to agents using a VAR(1) as their PLM:

$$Z_t = \Phi Z_{t-1} + \varepsilon_t, \tag{13}$$

where $Z_t \equiv [\pi_t, x_t, i_t]'$. Table 5 presents the new estimation results. The evidence about the switch in monetary policy coefficients is similar. The estimated gain coefficient is now lower and equals 0.010.

Table 4
Posterior estimates: model with interest-rate smoothing

Description	Parameters	Post. mean	Post. std.	95% Post. prob. int.
Discount rate	β	0.99	–	–
Phillips curve slope	κ	0.048	0.04	$[-0.02, 0.13]$
IES	σ	0.026	0.02	$[0.003, 0.08]$
IRS (pre-79)	$\rho_{i,\text{pre-79}}$	0.81	0.13	$[0.5, 0.98]$
Feedback infl. (pre-79)	$\chi_{\pi,\text{pre-79}}$	1.8	0.25	$[1.34, 2.30]$
Feedback gap (pre-79)	$\chi_{x,\text{pre-79}}$	0.34	0.09	$[0.17, 0.49]$
IRS (post-79)	$\rho_{i,\text{post-79}}$	0.82	0.14	$[0.46, 0.99]$
Feedback infl. (post-79)	$\chi_{\pi,\text{post-79}}$	2.65	0.24	$[2.18, 3.19]$
Feedback gap (post-79)	$\chi_{x,\text{post-79}}$	0.26	0.1	$[0.06, 0.45]$
Autoregr. dem shock	ϕ_r	0.91	0.02	$[0.86, 0.96]$
Autoregr. sup shock	ϕ_u	0.8	0.04	$[0.71, 0.88]$
MP shock	σ_ε	2.13	0.11	$[1.94, 2.35]$
Demand shock	σ_r	1.21	0.14	$[0.96, 1.52]$
Supply shock	σ_u	1.12	0.09	$[0.95, 1.32]$
Gain coeff.	$\bar{\mathbf{g}}$	0.0196	0.003	$[0.015, 0.026]$

Table 5
Posterior estimates: VAR(1) as learning rule

Description	Parameters	Post. mean	Post. std.	95% Post. prob. int.
Discount rate	β	0.99	–	–
Phillips curve slope	κ	0.065	0.05	$[-0.02, 0.16]$
IES	σ	0.029	0.02	$[0.003, 0.07]$
IRS (pre-79)	$\rho_{i,\text{pre-79}}$	0.85	0.05	$[0.76, 0.94]$
Feedback infl. (pre-79)	$\chi_{\pi,\text{pre-79}}$	1.51	0.45	$[0.61, 2.46]$
Feedback gap (pre-79)	$\chi_{x,\text{pre-79}}$	0.63	0.23	$[0.19, 1.08]$
IRS (post-79)	$\rho_{i,\text{post-79}}$	0.91	0.03	$[0.84, 0.96]$
Feedback infl. (post-79)	$\chi_{\pi,\text{post-79}}$	2.03	0.36	$[1.23, 2.68]$
Feedback gap (post-79)	$\chi_{x,\text{post-79}}$	0.40	0.21	$[0.03, 0.9]$
Autoregr. dem shock	ϕ_r	0.55	0.06	$[0.43, 0.67]$
Autoregr. sup shock	ϕ_u	0.78	0.05	$[0.69, 0.87]$
MP shock	σ_ε	0.97	0.06	$[0.87, 1.1]$
Demand shock	σ_r	1.56	0.19	$[1.24, 1.97]$
Supply shock	σ_u	1.57	0.13	$[1.34, 1.84]$
Gain coeff.	$\bar{\mathbf{g}}$	0.010	0.002	$[0.007, 0.013]$

Fig. 4 examines the robustness of the constant-gain estimates across alternative model specifications. When the agents adopt the MSV solution as their PLM, the estimates fall fairly robustly around a value of 0.02. The posterior distribution is, instead, more concentrated around 0.01 when the PLM is given by the VAR(1).[18]

[18] Additional robustness checks (Taylor rule with k-periods ahead forecasts, RLS learning, estimated initial conditions) were performed in an earlier version of the paper without large changes in the results.

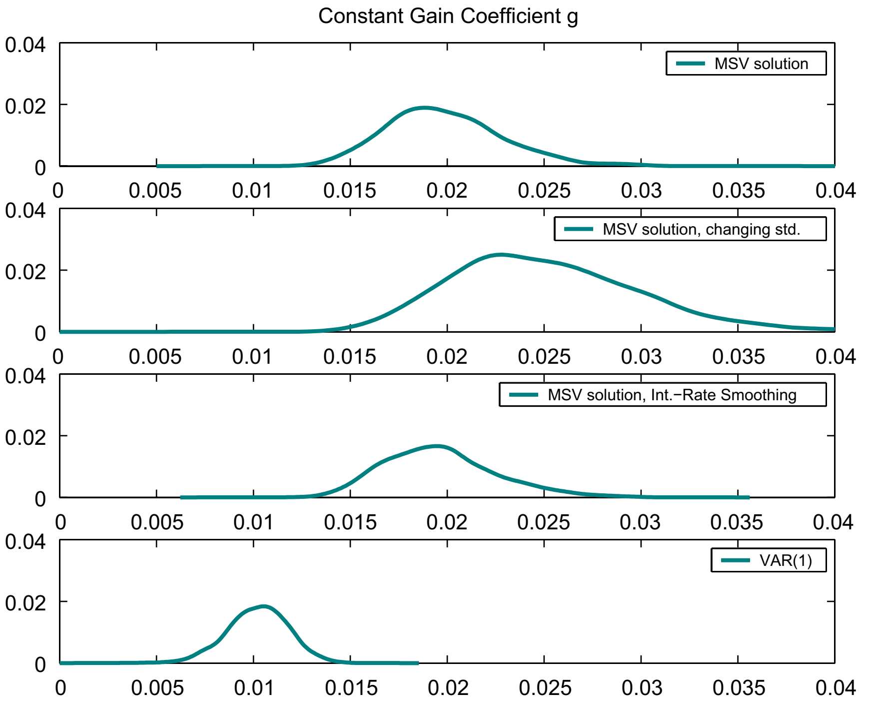

Fig. 4. Constant-gain coefficients: Posterior distributions across model specifications.

5. Conclusions and future directions

Several papers find that monetary policy has considerably changed over the post-war sample. CGG and LS conclude that policy was ‘passive’ in the pre-1979 sample and became ‘active’ in the Volcker–Greenspan period. The policy rule they estimate fails to satisfy the Taylor principle in the pre-1979 period and it leads to instability if considered within a simple monetary DSGE model with rational expectations.

This paper revisits the evidence on the evolution of monetary policy by departing from the assumption of rational expectations. I estimate a model in which agents form expectations from econometric models and learn the relevant parameters over time. In the model with learning, the failure to satisfy the Taylor principle would lead to instability, by preventing agents' learning from converging to the REE of the economy. By estimating a model that allows for time-variation in the monetary policy rule, I can check whether Fed's policy may have been a source of unstable learning dynamics in the 1970s.

I find some time-variation in monetary policy. But the paper shows that the estimated policy rule satisfied the Taylor principle also in the pre-1979 period. Therefore the results suggest that also in the 1960s–1970s, conditional on the proposed model of the economy, monetary policy was not contributing to macroeconomic instability.

I have shown that under learning the evidence of a regime switch of US monetary policy from passive to active is weak. Therefore, the results on the instability of policy during the Great Inflation may be dependent on the assumed expectations formation mechanism. As shown, small deviations from rational expectations may lead to very different results.

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. Moreover, the New Keynesian model I have used is likely to be misspecified: future research should use Del Negro and Schorfheide (2004)'s tools to determine the degree of misspecification of the model by comparing its restrictions under learning with unrestricted VARs (and by comparing the results with what one would find using the model under rational expectations).

Acknowledgments

I would like to thank the editor Volker Wieland, Athanasios Orphanides, and two anonymous referees for helpful comments. I am grateful to Michael Woodford, Chris Sims, Ricardo Reis, and seminar participants at the Federal Reserve of San Francisco for comments and discussions. This paper is a revised version of the third chapter of my dissertation at Princeton University and it was written while I was visiting the Department of Economics at Columbia University, which I would like to thank for hospitality.

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 $p(Y^{\mathrm{T}}|\theta)$ at each iteration.

Substituting private expectations into (1), (2), and (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 $\xi_t = [x_t, \pi_t, i_t, u_t, r_t^n]$, $w_t \sim \mathrm{N}(0, Q)$, H is a matrix of zeros and ones just selecting variables from ξ_t,[19] 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 by standard steps.

A.2. Metropolis–Hastings algorithm

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

$$p(\theta|Y^{\mathrm{T}}) = \frac{p(Y^{\mathrm{T}}|\theta)p(\theta)}{p(Y^{\mathrm{T}})}, \tag{A.2}$$

where $p(Y^{\mathrm{T}}|\theta)$ is the likelihood function, $p(\theta)$ the prior for the parameters, and $Y^{\mathrm{T}} = [y_1, \ldots, y_T]'$ collects the data histories.

[19]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).

To generate draws from the posterior distribution $p(\theta|Y^{\mathrm{T}})$, 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 $p(Y^{\mathrm{T}}|\theta_0)p(\theta_0)$.
3. Generate $\theta_j^*=\theta_{j-1}+\varepsilon$, where θ_j^* is the proposal draw and $\varepsilon\sim\mathrm{N}(0,c\Sigma_\varepsilon)$. 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 1992). I set c at 0.1 in the estimation that guarantees an acceptance rate slightly larger than 40%.
4. Generate u from a Uniform[0, 1]
5. Set $\begin{cases} \theta_j=\theta_j^* & \text{if } u\leqslant\alpha(\theta_{j-1},\theta_j^*)=\min\left\{\dfrac{p(Y^{\mathrm{T}}|\theta_j^*)p(\theta_j^*)}{p(Y^{\mathrm{T}}|\theta_{j-1})p(\theta_{j-1})},1\right\}, \\ \theta_j=\theta_{j-1} & \text{if } u>\alpha(\theta_{j-1},\theta_j^*). \end{cases}$
6. Repeat for $j+1$ from step 2 until $j=D$ ($D=$ total number of draws).

A.3. Convergence

To assess convergence of the MCMC (Markov Chain Monte Carlo) simulation, I performed various checks, besides looking at the trace plots of the draws. I have considered the convergence tests proposed by Geweke (1992), and Raftery and Lewis (1995). Raftery and Lewis's (1995) 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 $\hat{\mu}_1=\frac{1}{D_1}\sum_{j=1}^{D_1}g(\theta_j)$ and $\hat{\mu}_2=\frac{1}{D_2}\sum_{j=D_1+1}^{D_2}g(\theta_j)$, 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 $(\hat{\mu}_1-\hat{\mu}_2)/(\frac{\hat{S}_g^1(0)}{D_1}+\frac{\hat{S}_g^2(0)}{D_2})^{1/2}\Longrightarrow\mathrm{N}(0,1)$. I also look at the plots derived from the test proposed by Yu and Mykland (1994), based on CUMSUM plots of the draws.[20] Finally, I ascertain convergence by looking at the recursive mean plots and bivariate scatter plots among the parameters to evaluate the mixing of the chain.[21]

References

An, S., Schorfheide, F., 2007. Bayesian analysis of DSGE models. Econometric Reviews 26 (2–4), 113–172.

Branch, W.A., Evans, G.W., 2006. A recursive forecasting model. Economics Letters 91 (2), 158–166.

Bullard, J., 2007. The learnability criterion and monetary policy. Federal Reserve Bank of St. Louis Review 88 (3), 203–217.

Bullard, J., Mitra, K., 2002. Learning about monetary policy rules. Journal of Monetary Economics 49 (6), 1105–1129.

Bullard, J., Mitra, K., 2007. Determinacy, learnability, and monetary policy inertia. Journal of Money, Credit, and Banking 39 (5), 1177–1212.

Canova, F., Gambetti, L., 2006. Structural changes in the US economy: bad luck or bad policy? CEPR Discussion Papers No. 5457.

[20]They propose the statistics $CS_t=(\frac{1}{t}\sum_{d=1}^{t}\theta^d-\mu_\theta)/\sigma_\theta$, 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.

[21]Details on the convergence checks are available upon request.

Clarida, R., Galí, J., Gertler, M., 2000. Monetary policy rules and macroeconomic stability: evidence and some theory. Quarterly Journal of Economics CXV (1), 147–180.

Croushore, D., 1998. Evaluating inflation forecasts. WP 98-14, Federal Reserve Bank of Philadelphia.

Del Negro, M., Schorfheide, F., 2004. Priors from general equilibrium models for VARs. International Economic Review 45(2) (1), 643–673.

Duffy, J., Engle-Warnick, J., 2006. Multiple regimes in U.S. monetary policy? A nonparametric approach. Journal of Money, Credit, and Banking 38 (5), 1363–1378.

Evans, G.W., Honkapohja, S., 2001. Learning and Expectations in Macroeconomics. Princeton University Press, Princeton.

Evans, G.W., Honkapohja, S., 2006. Monetary policy, expectations and commitment. Scandinavian Journal of Economics 108 (1), 15–38.

Evans, G.W., McGough, B., 2005. Monetary policy, indeterminacy and learning. Journal of Economic Dynamics and Control 29, 1809–1840.

Geweke, J.F., 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Berger, J.O., Bernardo, J.M., Dawid, A.P., Smith, A.F.M. (Eds.), Proceedings of the Fourth Valencia International Meeting on Bayesian Statistics. Oxford University Press, pp. 169–194.

Lubik, T., Schorfheide, F., 2004. Testing for indeterminacy: an application to U.S. monetary policy. American Economic Review 94 (1), 190–217.

McCallum, B.T., 1999. Role of the minimal state variable criterion in rational expectations models. International Tax and Public Finance 6 (4), 621–639.

Milani, F., 2004. Adaptive learning and inflation persistence. Mimeo.

Milani, F., 2007. Expectations, learning, and macroeconomic persistence. Journal of Monetary Economics 54 (7), 2065–2082.

Orphanides, A., 2001. Monetary policy rules based on real-time data. American Economic Review 91 (4), 964–985.

Orphanides, A., 2004. Monetary policy rules, macroeconomic stability, and inflation: a view from the trenches. Journal of Money, Credit, and Banking 36 (2), 151–176.

Orphanides, A., Williams, J., 2005. The decline of activist stabilization policy: natural rate misperceptions, learning, and expectations. Journal of Economic Dynamics and Control 29 (11), 1927–1950.

Orphanides, A., Williams, J., 2007. Robust monetary policy with imperfect knowledge. Journal of Monetary Economics 54 (5), 1406–1435.

Preston, B., 2005. Learning about monetary policy rules when long-horizon expectations matter. International Journal of Central Banking, (2).

Preston, B., 2006. Adaptive learning, forecast-based instrument rules and monetary policy. Journal of Monetary Economics 53 (3), 507–535.

Primiceri, G., 2005. Time varying structural vector autoregressions and monetary policy. Review of Economic Studies 72, 821–852.

Primiceri, G., 2006. Why inflation rose and fell: policymakers' beliefs and postwar stabilization policy. Quarterly Journal of Economics 121, 867–901.

Raftery, A.E., Lewis, S.M., 1995. The number of iterations, convergence, diagnostics, and generic metropolis algorithms. In: Gilks, W.R., Spiegelhalter, D.J., Richardson, S. (Eds.), Practical Markov Chain Monte Carlo. Chapman and Hall, London.

Sargent, T.J., 1993. Bounded Rationality in Macroeconomics. Oxford University Press, Oxford.

Sargent, T.J., 1999. The Conquest of American Inflation. Princeton University Press, Princeton.

Schorfheide, F., 2003. Learning and monetary policy shifts. Review of Economic Dynamics 8 (2), 392–419.

Sims, C.A., Zha, T., 2006. Were there regime switches in US monetary policy? American Economic Review 96 (1), 54–81.

Smets, F., Wouters, R., 2003. Monetary policy in an estimated stochastic dynamic general equilibrium model of the Euro area. Journal of the European Economic Association 1 (5), 1123–1175.

Smets, F., Wouters, R., 2004. Forecasting with a Bayesian DSGE model: an application to the Euro area. Journal of Common Market Studies 42 (4), 841–867.

Smets, F., Wouters, R., 2005. Comparing shocks and frictions in US and Euro business cycles: a Bayesian DSGE approach. Journal of Applied Econometrics 20 (2), 161–183.

Wieland, V., 2000. Monetary policy, parameter uncertainty and optimal learning. Journal of Monetary Economics 46 (1), 199–228.

Wieland, V., 2006. Monetary policy and uncertainty about the natural unemployment rate: Brainard-style conservatism versus experimental activism. Advances in Macroeconomics 6 (1) (Article 1).
Williams, N., 2003. Adaptive learning and business cycles. Mimeo, Princeton University.
Woodford, M., 2003. Interest and Prices: Foundations of a Theory of Monetary Policy. Princeton University Press, Princeton.
Yu, B., Mykland, P., 1994. Looking at Markov samplers through CUMSUM path plots: a simple diagnostic idea. Technical Report 413, Department of Statistics, University of California at Berkeley.