WHY INFLATION ROSE AND FELL: POLICYMAKERS' BELIEFS
AND US POSTWAR STABILIZATION POLICY

Giorgio E. Primiceri

Why Inflation Rose and Fell: Policymakers' Beliefs and US Postwar Stabilization Policy
Giorgio E. Primiceri
NBER Working Paper No. 11147
February 2005
JEL No. E31, E32, E5

ABSTRACT

This paper provides an explanation for the run-up of U.S. inflation in the 1960s and 1970s and the sharp disinflation in the early 1980s, which standard macroeconomic models have difficulties in addressing. I present a model in which rational policymakers learn about the behavior of the economy in real time and set stabilization policy optimally, conditional on their current beliefs. The steady state associated with the self-confirming equilibrium of the model is characterized by low inflation. However, prolonged episodes of high inflation ending with rapid disinflations can occur when policymakers underestimate both the natural rate of unemployment and the persistence of inflation in the Phillips curve. I estimate the model using likelihood methods. The estimation results show that the model accounts remarkably well for the evolution of policymakers' beliefs, stabilization policy and the postwar behavior of inflation and unemployment in the United States.

Giorgio E. Primiceri
Northwestern University
Department of Economics
2001 Sheridan Road
3218 Andersen Hall
Evanston, Il. 60208-2600

1   Introduction

This paper aims to explain the behavior of inflation and unemployment in the United States. Figure 1 presents a plot of the annualized quarterly growth rate of the GDP deflator and the total civilian unemployment rate over the postwar period. The striking feature of the graph is the long and pronounced run-up of inflation, which occurred in the 1960s and 1970s. This episode, known as the Great Inflation, is not just "America's only peacetime inflation" (DeLong 1997), but has also been called "the greatest failure of American macroeconomic policy in the postwar period" (Mayer 1999).

At least four stylized facts characterize the Great Inflation.

- *Dimension.* Between 1963 and 1981 the inflation rate in the United States rose by more than 9 percentage points. If we exclude the peak in 1974 (which is due to the effect of the first oil price shock), the rate of increase was approximately constant.

- *Duration.* The episode of high inflation lasted for more than 20 years. Inflation started to increase around 1963 and came back under control, at a level of about 2 percent, only around 1985.

- *Asymmetry.* The episode of high inflation was asymmetric. In the early 1980s, the duration of the so called "Volcker disinflation" was much shorter than the phase of rising inflation.

- *Unemployment lagged inflation.* Unemployment lagging behind inflation is a general characteristic of the business cycle. However, this feature of the data was particularly evident in the period of high inflation, with unemployment peaking always a few quarters after inflation.

This paper puts forward a theory of the behavior of inflation and unemployment, which fits the U.S. data well and, in particular, explains all four of the stylized facts above. This theory is based on the evolution of policymakers' beliefs about the structure of the economy.

Previous attempts to explain the Great Inflation fall apart in three categories, which I label the "bad luck," the "lack of commitment" and the "policy mistakes" views. I briefly discuss each of these branches of literature below.

The "bad luck" view   The first type of explanations is based on bad luck, in view of the fact that it has been well documented that the volatility of the exogenous, non-policy shocks

was higher in the 1960s and 1970s than in the last two decades of the century (see, for instance, Cogley and Sargent 2003, Kim and Nelson 1999a, McConnell and Perez-Quiros 2000, Sims and Zha 2004, Stock and Watson 2002). However, although non-policy shocks definitely played an important role, it is hard to reconcile the existing estimates with the exceptional dimension and duration of the Great Inflation.

The "lack of commitment" view The second class of explanations is what Christiano and Fitzgerald (2003) have called the "institution vision of inflation". According to this view, inflation was high in the 1960s and 1970s because policymakers did not have any incentive to keep inflation low. The motivation for this relies on the time-consistency problem of optimal policy, first emphasized by Kydland and Prescott (1977) and Barro and Gordon (1983). The importance of this line of research has been recently emphasized by Chari, Christiano, and Eichenbaum (1998), Christiano and Gust (2000) and Christiano and Fitzgerald (2003).

However, the inflation bias generated by the time-consistency problem seems to be quantitatively too small to explain the high inflation of the 1960s and 1970s (see, for example, Reis 2003). Ireland (1999) formally tests the inflation bias hypothesis on U.S. data. While he is not able to reject it, his estimates suggest the presence of an inflationary bias of small magnitude. Moreover, it is hard to reconcile the time-consistency view with the rapid Volcker disinflation. In fact, it is not clear what exactly changed between the pre and post 1980s period from the institutional point of view.[1] The final difficulty with the "lack of commitment" approach is the fact that it would predict unemployment leading, rather than lagging inflation. This is due to the fact that the advantages of inflationary surprises depend on the level of unemployment. As mentioned above, this is clearly at odds with the data.

The "policy mistakes" view This approach focuses on policy mistakes and stresses that in the 1960s and 1970s monetary policymakers were not as good as the ones of the last two decades. For example, many authors have argued that U.S. monetary policy was less responsive to inflationary pressures under the Fed chairmanship of Arthur Burns than under Paul Volcker and Alan Greenspan (among others, see Boivin and Giannoni 2002, Clarida, Gali, and Gertler

---

[1] Recent work has made some progress in this direction. Sargent (1999), for example, explains the disinflation as escape dynamycs from the inflation biased equilibrium. Rogoff (2003) argues that Central Banks' lower incentive to inflate is related to globalization and the consequent increase in world competition. Albanesi, Chari, and Christiano (2003), instead, analyze the lack of commitment problem in an optimizing agents model and show the existence of multiple equilibria, which can potentially explain the disinflation.

2000, Cogley and Sargent 2001, Judd and Rudebusch 1998, Lubik and Schorfheide 2004).[2]

In this respect, the line of research started by Orphanides represents an attempt to rationalize why the policy authorities behaved so differently in the pre and post 1980s period. Orphanides (2000 and 2002) has argued that policymakers in the 1970s overlooked a break in potential output. They overestimated potential output leading to overexpansionary policies, which ultimately resulted in high inflation. Among others, this explanation has also been proposed by Cukierman and Lippi (2002), Lansing (2002), Bullard and Eusepi (2003), Reis (2003) and Tambalotti (2003). While this strand of literature represents a step forward, the dimension of the high inflation episodes explained by such models is usually much lower than what we observe in the data, unless the model is augmented with additional propagation mechanisms like, for instance, private sector learning (as in Orphanides and Williams 2003). Furthermore, the explanations based on the misperception of potential output fail to address the Volcker disinflation, unless an exogenous shift in policymakers' preferences is specified (see, for instance, Bullard and Eusepi 2003).

While there is clearly some truth in all of these theories, they also seem to have difficulties in addressing at least some of the stylized facts of the hump-shaped behavior of inflation and unemployment. This paper proposes instead an explanation of the Great Inflation that matches all these stylized facts.

I present a model, in which rational policymakers form their beliefs about the behavior of the economy in real time and set stabilization policy optimally, conditional on the information available to them. Although the equilibrium of the model is characterized by low inflation, episodes of high inflation and unemployment can occur when policymakers simultaneously underestimate *both* the natural rate of unemployment *and* the persistence of inflation in the Phillips curve. Such initial conditions result in peculiar dynamics of policymakers' beliefs, ultimately affecting also their perception of the slope of the Phillips curve and of the cost of the inflation-unemployment trade-off.

Intuitively, if real-time policymakers underestimate the natural rate of unemployment, this results in overexpansionary policies and higher inflation. Moreover, if policymakers' estimate of the persistence of inflation in the Phillips curve is also downward biased, they rationally choose not to react strongly to inflation, amplifying the initial effect. The reason is that the

---

[2] This view is controversial. Other studies have in fact found either little evidence of changes in the systematic part of monetary policy (for example, Bernanke and Mihov 1998, Hanson 2003, Leeper and Zha 2002, Primiceri 2005) or no evidence of unidirectional drifts in policy toward a more active behavior (Sims 2001b, Sims and Zha 2004).

more stationary inflation is perceived to be, the sooner it is expected to revert to its mean and the less urgent is the need for anti-inflationary action. This period of "overoptimism" ends when inflation reaches a level that concerns policymakers. However, when policymakers start reacting to inflation, pushing unemployment above the *perceived* natural rate does not seem to reduce inflation. This is because they still have a downward biased estimate of the natural rate of unemployment. In this period of "overpessimism", they temporarily and mistakenly perceive a very costly inflation-unemployment trade-off, which explains why anti-inflationary policy is postponed even further. The disinflation occurs only when the perceived inflation-unemployment trade-off becomes favorable, relative to the level of inflation.

Among others, Orphanides (2000), DeLong (1997) and Romer and Romer (2002) have argued in favor of the policy misperception of potential output and the natural rate of unemployment in the 1960s and 1970s. Policymakers' misperception of the persistence of inflation in the Phillips curve in the 1960s is also no longer controversial. For example, Blanchard and Fischer (1989) and Mayer (1999) have noted that, at least until the early 1970s, most of the econometric studies underestimated inflation persistence. In relation to the overpessimism phase, DeLong (1997), Romer and Romer (2002) and Cogley and Sargent (2004) have emphasized that policy was cautious in the 1970s because the cost of lowering inflation seemed too high.

From a quantitative and statistical standpoint, I show that the evolution of policymakers' beliefs about the coefficients of the Phillips curve is very important to explain the behavior of inflation and unemployment. I estimate the model using likelihood methods. The estimated version of the model accounts remarkably well for the evolution of policymakers' beliefs, stabilization policy and the postwar behavior of inflation and unemployment in the United States.

The importance of policymakers' learning dynamics has been recognized by many authors. In the context of the "natural rate" literature, policymakers' learning has been introduced by Sims (1988). Theoretical advances include Sargent (1999), Cho, Williams, and Sargent (2002) and Williams (2003). Empirical studies include Chung (1990), Sargent (1999), Cogley and Sargent (2004) and Sargent, Williams, and Zha (2004). The main insight of this literature is that policymakers' learning introduces temporary deviations from the model's equilibrium, which is characterized by an inflation bias. These temporary deviations are in the direction of the optimal, low inflation outcome. Unlike these studies, in this paper the equilibrium outcome is a low inflation regime. Nevertheless, the model explains the run-up of U.S. inflation in the 1960s and 1970s and the sharp disinflation in the early 1980s. Although the explanation roughly belongs to the "policy mistakes" category, in this paper policymakers are assumed to be

rational and optimizing. As a consequence, I find that the mismeasurement of the natural rate of unemployment alone is not sufficient to generate fluctuations of the inflation rate comparable to what we observe in the data. This differs importantly from Orphanides (2000) and other similar approaches.

The paper is organized as follows. Section 2 presents the theoretical model of the economy and policymakers' behavior. Section 3 offers a model-based interpretation of the Great Inflation. Section 4 focuses instead on statistical evidence, i.e. estimation, fit and quantitative simulation results. Sections 5 and 6 demonstrate the robustness of the results to two modifications of the baseline framework. In particular, section 5 introduces private agents' forward looking behavior in the model and section 6 allows for stochastic volatility of the exogenous innovations. Section 7 makes an attempt to uncover the deeper reasons of the Great Inflation, i.e. why policymakers underestimated the persistence of inflation in the Phillips curve in the 1960s. Section 8 concludes with some final remarks.

2 Imperfect Information and Inflation-Unemployment Dynamics

In this section I present a simple model of inflation-unemployment dynamics when policymakers have imperfect information. The source of imperfect information is the fact that policymakers do not know the exact model of the economy. In particular, they are uncertain about the value of the model's parameters. Therefore, policymakers update their beliefs about the model's unknowns in every period and implement optimal policy, conditional on their current beliefs. In turn, the policy variable affects the behavior of inflation and unemployment because it enters the model describing the true evolution of key macroeconomic variables.

2.1 The Model Economy

As a "true" model of the economy, I consider a simple rational expectations model that can be rewritten as a backward looking one. Even if conceptually similar to modern New-Keynesian specifications, the benchmark model is more in the spirit of the empirical literature following along the lines of King, Stock, and Watson (1995) and, more recently, Gordon (1997 and 1998), Rudebusch and Svensson (1999) and Staiger, Stock, and Watson (1997 and 2001).

This framework, not only is tractable and convenient for estimation, but is also simple and transparent, providing a clear intuition for the role played by policymakers' learning dynamics

in the behavior of inflation and unemployment. Integrating learning dynamics in a forward looking specification is instead computationally more expensive and this case in analyzed in section 5.

The private sector part of the model is described by the following equations:

$$\pi_t = \pi_t^e - \tilde{\theta}(L)(u_{t-1} - u_{t-1}^N) + \varepsilon_t. \tag{1}$$

$$(u_t - u_t^N) = \rho(L)(u_{t-1} - u_{t-1}^N) + V_{t-1} + \eta_t, \tag{2}$$

$$u_t^N = (1 - \gamma)u^* + \gamma u_{t-1}^N + \tau_t. \tag{3}$$

Equation (1) represents a standard expectation augmented Phillips curve, where π_t is the inflation rate, π_t^e is the agents' expected inflation rate, u_t is the unemployment rate and u_t^N is the time varying natural rate of unemployment. $\tilde{\theta}(L)$ is a lag polynomial and ε_t is a random innovation, assumed to be i.i.d. $N(0, \sigma_\varepsilon^2)$.[3] I assume that some of the agents are fully rational, while the rest of them form their expectations adaptively, so that

$$\pi_t^e = (1 - \tilde{\alpha}(1))E_{t-1}\pi_t + \tilde{\alpha}(L)\pi_{t-1}. \tag{4}$$

$\tilde{\alpha}(L)$ is a lag polynomial and the combination of (1) and (4) leads to the following familiar reduced form Phillips curve:

$$\pi_t = \alpha(L)\pi_{t-1} - \theta(L)(u_{t-1} - u_{t-1}^N) + \varepsilon_t, \tag{5}$$

where $\alpha(L) = \tilde{\alpha}(L)/\tilde{\alpha}(1)$ and $\theta(L) = \tilde{\theta}(L)/\tilde{\alpha}(1)$. Note that, no matter what the exact fraction of agents with adaptive expectations is, $\alpha(1) = 1$, implying the absence of a long run trade-off between unemployment and inflation, which is consistent with the natural rate hypothesis. The interpretation of (5) is straightforward: the inflation rate changes either because of a random "cost push" term or because unemployment is not in line with the natural rate.

Equation (2) is a very simple aggregate demand equation, where $\rho(L)$ is a lag polynomial, η_t is an i.i.d. $N(0, \sigma_\eta^2)$ random innovation and V_t is a variable controlled by policymakers. In other words, unemployment deviates from the natural rate either because of a random shock or because of policymakers' decisions about stabilization policy.

Although a natural interpretation of the policy variable V_t is of real rate of interest, more generally V_t can be thought as capturing the joint effect of monetary and fiscal policy. In particular, this modeling strategy avoids complications related to the specification of two aspects

---

[3] The case of heteroskedastic innovations is particularly interesting in the context of learning models and is analyzed in section 6.

of the policy process: the relative importance of the monetary and fiscal policy actions on real activity and the particular channels through which monetary and fiscal policy affect real activity. This is in the spirit of the recent renovated interest on the effect of fiscal policy (see, among others, Blanchard and Perotti 2002, Mountford and Uhlig 2002) and the direct role of monetary aggregates in macro models (Favara and Giordani 2002, Leeper and Roush 2003).

Equation (3) describes the exogenous stochastic process for the natural rate of unemployment, which is assumed to evolve as an AR(1), where τ_t is *i.i.d.* $N(0, \sigma_\tau^2)$. u^* represents the unconditional expectation of u_t^N. This assumption on the evolution of u_t^N is standard in the literature (see, for instance, Staiger, Stock, and Watson 2001, although they set $\gamma = 1$ in their empirical specification).

2.2 Optimal Policy under Imperfect Information

The value of the policy variable V is chosen in every period by policymakers. They base their decision on the available information and on current beliefs about the state of the economy. I assume that policymakers know the structure of the true model of the economy (given by equations (5) and (2)), but they are uncertain about the value of the unobservable variables (the natural rate) and the coefficients. Policymakers estimate the model's parameters in every period and use these estimates as true values, neglecting both estimates uncertainty and the possibility of future updates.[4]

Policymakers' beliefs about unobservables and the model's constant coefficients are denoted by hats. All these beliefs are formed at time t, but the subscript is omitted for simplicity. In particular, $\hat{u}_{t-1}^N$ stands for $\hat{u}_{t-1|t}^N$ and indicates the estimate at time t of the value of the natural rate at time $t-1$. Policymakers determine the optimal value of the policy variable V by solving the following optimization problem:

$$\min_{\{V_t\}} L = \hat{E} \sum_{t=s}^{\infty} \delta^{t-s} \left[(\pi_t - \pi^*)^2 + \lambda \left(u_t - k\hat{u}_t^N \right)^2 + \phi \left(V_t - V_{t-1} \right)^2 \right], \tag{6}$$

$$\text{s.t.} \quad \pi_t = \hat{c}_\pi + \hat{\alpha}(L)\pi_{t-1} - \hat{\theta}(L)(u_{t-1} - \hat{u}_{t-1}^N) + \hat{\varepsilon}_t, \tag{7}$$

$$(u_t - \hat{u}_t^N) = \hat{c}_u + \hat{\rho}(L)(u_{t-1} - \hat{u}_{t-1}^N) + V_{t-1} + \hat{\eta}_t. \tag{8}$$

L represents the familiar quadratic loss function, which depends on deviations of inflation and unemployment from the respective targets. λ represents the weight on the unemployment objective. Notice that, like in Barro and Gordon (1983), the target for the unemployment

---

[4] These assumptions are standard in the adaptive learning literature, although the resulting policymakers' behavior is suboptimal. For alternative approaches, see Beck and Wieland (2002) and Wieland (2000a and 2000b).

rate is given by $k\hat{u}_t^N$. When $k = 0$ the unemployment target is equal to zero and policymakers' preferences resemble Kydland and Prescott's (1977). This would be the case in which the policy time-consistency problem is most pronounced. On the other hand, when $k = 1$, the target is the natural rate and the time-consistency and the related inflation bias completely disappear from the policy problem. Blinder (1998), among others, has argued in favor of these kind of policy preferences. The loss function has also a "smoothing" component, which penalizes big shifts of the policy variable. From an empirical perspective, this term is crucial in order to match the actual policy behavior because it helps to account for the strong autocorrelation shown by the instruments of economic policy (Dennis 2001, Favero and Rovelli 2003 and Soderstrom, Soderlind, and Vredin 2003). See also Woodford (2003) for an overview of the theoretical desirability of the smoothing term in the monetary policy context. Moreover, it is easy to think to models in which instrument smoothing is desirable also for fiscal policy (for example, Barro 1979).

Policymakers minimize their loss function subject to two constraints, (7) and (8). These constraints are the estimated counterparts of the true Phillips curve and aggregate demand equations. Observe that in the policymakers' model $\hat{\alpha}(1)$ is not constrained to be equal to one and $\hat{c}_\pi$ in (7) controls their beliefs about the level of average inflation. $\hat{c}_u$ in (8) instead controls their beliefs about the effect of setting V equal to zero. In other words, $-\hat{c}_u$ represents the "natural" level of policy, i.e. the level of V that does not affect unemployment.[5] Notice that, without further assumptions, $\hat{c}_\pi$, $\hat{c}_u$ and $\hat{u}_t^N$ would not be separately identified in the policy econometric model. A discussion about this issue is postponed until the next subsection.

The optimal rule for fixing V_t is given by

$$V_t = g(\hat{\beta})S_t, \tag{9}$$

where $\hat{\beta}$ represents the vector of values for the model's parameters that policymakers treat as certainty-equivalents; $g(\hat{\beta})$ is the standard solution of a linear-quadratic problem, obtained solving the corresponding Riccati equation. S_t meanwhile denotes the set of relevant state variables and beliefs about unobservable states of the economy. To be more concrete, assuming that all the lag polynomials have order one, $\hat{\beta}$ would be given by the vector $\left[\hat{c}_\pi; \hat{\alpha}_1; \hat{\alpha}_2; \hat{\theta}_1; \hat{\theta}_2; \hat{c}_u; \hat{\rho}_1; \hat{\rho}_2; \hat{u}_t^N\right]$ and S_t by the vector $\left[1; \pi_t; \pi_{t-1}; u_t - \hat{u}_t^N; u_{t-1} - \hat{u}_{t-1}^N; V_{t-1}\right]$.

---

[5] In the true model the "natural" level of policy is normalized to zero.

9

2.3 Learning

To implement the optimal rule and fix the value of the policy variable V_t, policymakers must estimate the parameters of interest, which are the unobservables and the coefficients. While I relax this assumption in Primiceri (2004), here I assume that policymakers form their beliefs about the natural unemployment rate using univariate methods, i.e. they extract information about the natural rate, only looking at the behavior of unemployment. Observe that this is suboptimal as, conditional on the true model of the economy, better estimates could be obtained by exploiting the information contained not only in the unemployment rate, but also in the inflation rate. However, there are several reasons motivating this choice.

First, historical narrative evidence (see, for instance, DeLong 1997, Romer and Romer 2002) suggests that this is a realistic assumption for the behavior of past policymakers. Even now, univariate algorithms are commonly used to define the potential of the economy, especially in the output gap and monetary policy literature (see, for instance, Orphanides and Van Norden 2001, Lansing 2002, Taylor 1999). Second, Staiger, Stock, and Watson (2001) show that the natural rate estimated using formally the Phillips curve approach is basically indistinguishable from the univariate trend in unemployment. The last reason is substantial. In fact, as mentioned above, $\hat{c}_\pi$, $\hat{c}_u$ and $\hat{u}_t^N$ are clearly not separately identified in equations (7) and (8). Therefore, I assume that policymakers solve this identification problem by imposing the prior belief that on average, unemployment is equal to its natural rate. This assumption provides a very natural way of estimating the natural rate, which is to use univariate algorithms on the series of unemployment, in order to isolate the low frequency component. Furthermore, observe that this assumption is not contradictory and respects the coherence of the policymakers' model because, as will be shown in section 2.4 and appendix A, it correspond to a self-confirming equilibrium.

Conditional on their estimate of the natural rate of unemployment, policymakers can estimate the model's coefficients using standard regression methods. Following a large part of the most recent literature (see, among others, Sargent 1999, Williams 2003) in the baseline specification of the model I assume that policymakers update their beliefs using constant gain algorithms (CG). These algorithms allow to update beliefs discounting the past and giving more weight to recent data. Recent data are considered more informative possibly because of the suspicion of drift in the parameters.

As mentioned above, the estimation works in two steps. In the first step policymakers obtain

an estimate of the current value of the natural rate using the following updating formula:

$$\hat{u}_{t|t}^{N} = \hat{u}_{t-1|t-1}^{N} + g_N R_{N,t-1}^{-1} \left(u_t - \hat{u}_{t-1|t-1}^{N} \right), \tag{10}$$

$$R_{N,t} = R_{N,t-1} + g_N \left(1 - R_{N,t-1} \right). \tag{11}$$

Equation (10) states that the current estimate of the natural rate is obtained by updating the previous estimate according to the current realization of the unemployment rate. The weight given to the last observation depends on the gain (g_N) and the inverse of the variance of the regressor (the constant 1 in this case), which, in turn, is updated in equation (11).

In the second step, policymakers use their estimate of the natural rate to update their beliefs about the Phillips curve and aggregate demand coefficients:

$$\hat{\beta}_{t}^{i} = \hat{\beta}_{t-1}^{i} + g R_{i,t-1}^{-1} x_t^i \left(y_t^i - x_t^{i\prime} \hat{\beta}_{t-1}^{i} \right), \tag{12}$$

$$R_{i,t} = R_{i,t-1} + g \left(x_t^i x_t^{i\prime} - R_{i,t-1} \right), \qquad i = \{\pi, u\}, \tag{13}$$

where $\hat{\beta}_t^{\pi} = [\hat{c}_\pi; \hat{\alpha}_1; \hat{\alpha}_2; \hat{\theta}_1; \hat{\theta}_2]'$; $y_t^\pi = \pi_t$; $x_t^\pi = [1; \pi_{t-1}; \pi_{t-2}; u_{t-1} - \hat{u}_{t|t}^N; u_{t-2} - \hat{u}_{t|t}^N]$; $\hat{\beta}_t^{u} = [\hat{c}_u; \hat{\rho}_1; \hat{\rho}_2]'$; $y_t^u = u_t - \hat{u}_{t|t}^N - V_{t-1}$; $x_t^u = [1; u_{t-1} - \hat{u}_{t|t}^N; u_{t-2} - \hat{u}_{t|t}^N]$. Notice from the expressions of the vectors x_t^π and x_t^u that policymakers approximate $\hat{u}_{t-1|t}^N$ and $\hat{u}_{t-2|t}^N$ with their last estimate of the current level of the natural rate, $\hat{u}_{t|t}^N$. Equations (12) and (13) update beliefs with a mechanism similar to the one illustrated for equations (10) and (11).

Observe that I allow for the possibility of different gain parameters in the algorithms for the estimation of the natural rate and the coefficients (respectively g_N and g). The gain parameters control the rate at which new information affects beliefs. If g and g_N were decreasing and equal to $\frac{1}{t-1}$, equations (10), (11), (12) and (13) would be recursive representations of ordinary least squares estimates (if properly initialized).

As robustness checks, I also consider the cases in which policymakers form estimates of the natural rate using a moving average and estimates of the coefficients by ordinary least squares (OLS) or discounted least squares (DLS), a weighted least square method with weight Δ^{t-s} to time s observation (where t is the time period of the most recent data and $\Delta < 1$ is a discount factor).

2.4   Equilibrium and Steady State

As is standard in most of the recent literature on learning, I focus on the concept of self-confirming equilibria.[6] In period t, policymakers form beliefs about the model's parameters.

[6] For a formal treatment of the issue, see Sargent (1999), Evans and Honkapohja (2001), Williams (2003).

Their beliefs imply an optimal value for the policy variable V_t, which, in turn, affects the stochastic data generating process and, ultimately, next periods beliefs. This defines a map from today's beliefs to tomorrow's beliefs. A fixed point of this map is a self-confirming equilibrium. In other words, a self-confirming equilibrium is a situation in which policymakers' beliefs are expected not to change with the new vintage of data. Appendix A gives a formal definition and derives the model's self-confirming equilibrium for the baseline case in which beliefs are formed using the CG algorithm.

I define the model's steady state as the unconditional mean of the stationary stochastic process for the vector $\left[\pi_t, u_t, V_t, u_t^N\right]$ in a self-confirming equilibrium. Observe that, in steady state, (5) implies $u_t = u^*$. Consequently, $V_t = 0$ follows from (2). Finally (9) implicitly defines the steady state inflation as a function of the equilibrium beliefs.

As an example and for simplicity, consider the intuitive case of a constant natural unemployment rate.[7] Notice that, if $k = 1$, the steady state inflation is just the inflation target π^*. In other words, if policymakers do not wish to push unemployment below the natural rate, the outcome is the optimal one in which unemployment is at the natural level and inflation at the target. If $0 \leq k < 1$ instead, the steady state inflation will be higher. As it will be clear later, the data favor a model with a limited amount of inflation bias, due to the fact that postwar policymakers did not seem to have an excessively low unemployment target.

3   Interpreting the Great Inflation

In section 2 I have presented the baseline model of the paper and discussed the related technical issues. I now introduce a simplified version of the model and use it to interpret and explain the postwar behavior of inflation and unemployment in the United States. Therefore, while the rest of the paper focuses on estimation and simulations, the objective of this section is to provide the intuition and the main ideas necessary to interpret the quantitative results.

---

[7] In the case of a constant natural rate of unemployment it is easy to check that, in the self-confirming equilibrium, beliefs about the model's coefficients coincide with the true values. As shown in the appendix, when $Var\left(u_t^N\right)$ is bigger than zero equilibrium beliefs about the model's parameters do not necessarily coincide with the true parameters of the model, but can be arbitrarily close.

3.1   A Special Case

As an illustrative example, consider the special case of the previous model given by the following simplified Phillips curve and aggregate demand equations:

$$\pi_t = \pi_{t-1} - \theta(u_{t-1} - u^N) + \varepsilon_t \tag{14}$$

$$(u_t - u^N) = V_t, \tag{15}$$

where policymakers determine V by solving the following problem:

$$\min_{\{V_t\}} L = \hat{E}\sum_{t=s}^{\infty}\delta^{t-s}\left[\pi_t^2 + \left(u_t - k\hat{u}^N\right)^2\right],$$

$$\text{s.t.} \quad \pi_t = \hat{c}_\pi + \hat{\alpha}\pi_{t-1} - \hat{\theta}(u_{t-1} - \hat{u}^N) + \widehat{\varepsilon}_t, \tag{16}$$

$$(u_t - \hat{u}^N) = \hat{c}_u + V_t.$$

Observe that in order to obtain a closed form solution I have modified the timing of V in (15). The set of estimated parameters is given by $\hat{\beta} = [\hat{c}_\pi; \hat{c}_u; \hat{\alpha}; \hat{\theta}; \hat{u}^N]$ and the states of the policy optimization problem are collected in $S_t = [1; \pi_t]$. The solution of the policy problem is given by the following linear control rule:

$$V_t = g(\hat{\beta})S_t = -\hat{c}_u + A(\hat{\beta}) + B(\hat{\beta})\pi_t, \tag{17}$$

where

$$A(\hat{\beta}) = -(1-k)\hat{u}^N + \frac{\left(1 + B(\hat{\beta})\frac{\hat{\alpha}}{\hat{\theta}}\right)\left(\hat{c}_\pi + \theta(1-k)\hat{u}^N\right)}{\hat{\theta}\left(1 + B(\hat{\beta})\frac{\hat{\alpha}}{\hat{\theta}}\right) - \left(\frac{\hat{\alpha}}{\hat{\theta}} - \frac{1}{\delta\hat{\theta}}\right)}$$

and

$$B(\hat{\beta}) = \frac{-\left(\frac{1}{\delta\hat{\theta}} + \hat{\theta} - \frac{\hat{\alpha}^2}{\hat{\theta}}\right) + \sqrt{\left(\frac{1}{\delta\hat{\theta}} + \hat{\theta} - \frac{\hat{\alpha}^2}{\hat{\theta}}\right)^2 + 4\hat{\alpha}^2}}{2\hat{\alpha}}.$$

Observe that $B(\hat{\beta})$ is always positive (for positive values of $\hat{\alpha}$). This implies that policy reacts to inflation by pushing the unemployment rate upwards. Substituting (17) into (15) and the resulting equation into (14) we obtain the following equation for the inflation rate:

$$\pi_t = \theta\left(u^N - \hat{u}^N\right) - \theta A(\hat{\beta}) + \left(1 - \theta B(\hat{\beta})\right)\pi_{t-1} + \varepsilon_t. \tag{18}$$

This equation can be interpreted as a local approximation of inflation dynamics for given policy-makers' beliefs about the state of the economy. Notice that the mismeasurement of the natural rate of unemployment shifts the mean of the inflation process. Instead, the strength of the policy reaction to inflation affects the persistence and therefore both the mean and the variance

of the inflation process. In other words, the stronger policy reacts to inflation in the feedback rule, the more stationary and less volatile is inflation. In particular, it is easy to show that $B(\cdot)$ is a positive function of $\hat{\alpha}$ (the estimated persistence) for all positive values of $\hat{\alpha}$ and $\hat{\theta}$ (the estimated slope of the Phillips curve). This implies that the lower the estimated persistence of inflation in the Phillips curve, the lower the reaction to inflation and, given (18), the higher the actual persistence of the univariate, reduced form process for inflation.

Similar perverse dynamics can occur for wrong estimates of θ. In fact, $B(\cdot)$ is a positive function of $\hat{\theta}$ for values of $\hat{\theta} < \sqrt{\frac{1}{\delta} - \hat{\alpha}^2}$, which corresponds to estimated set of possible values of $\hat{\theta}$.[8] This means that the lower the estimated slope of the Phillips curve, the lower the reaction to inflation and, given (18), the higher the actual inflation persistence. That is, if policymakers perceive a costly inflation-unemployment trade-off, they will not be willing to accept higher unemployment for a limited relief from inflation. Therefore, they will react to inflation less strongly, ultimately leading to a less stationary inflation.

The extreme case of a unit root in the reduced form, univariate inflation process (18) occurs if policy does not respond to inflation at all. For example, this can happen if the perceived slope of the Phillips curve is zero. Figure 2 gives an idea of the shape of the function $B(\cdot)$. Figure 2a shows B as a function of $\hat{\alpha}$, when $\hat{\theta}$ is fixed to three different values. Besides the positive slope in all cases, notice the pronounced nonlinearity of B as a function of $\hat{\alpha}$, especially for high values of $\hat{\alpha}$. Figure 2b shows instead the opposite case in which $\hat{\alpha}$ is fixed to three different values and $\hat{\theta}$ is allowed to change. Interestingly, the effect of increasing $\hat{\theta}$ on B, heavily depends on the value of $\hat{\alpha}$: the higher the estimate of inflation persistence, the higher the effect of the estimate of the slope of the Phillips curve on the strength of the reaction to inflation. This will be crucial to explain the sharpness of the disinflation.

It is important to realize that the case of weak policy reaction to inflation leads to inflation close to a unit root process. This is particularly dangerous for the stability of this economy. The reason is that if inflation is stationary, a mistake in the estimation of the natural rate shifts the mean of the process. If inflation exhibits a unit root instead, a mistake in the estimate of the natural rate creates a time trend in the inflation process.

It is easy to check that for this simple economy, the self-confirming equilibrium corresponds to $\hat{\beta} = \beta = \left[0; \alpha; \theta; u^N\right]$. It follows that the associated stationary stochastic process for the random vector $[\pi_t, u_t, V_t]$ is a simple VAR(1), whose coefficients are omitted for brevity. The steady state is $[\pi_t, u_t, V_t] = \left[\frac{(1-k)u^N}{\theta + B(\beta) + 1 - 1/\delta}; u^N; 0\right]$. Notice that if $k = 1$ the steady state inflation

---

[8] Overall $B(\cdot)$ is a hump-shaped function of $\hat{\theta}$. See Rudebusch (2001) for an example of parameter values for which, locally, $B(\cdot)$ is a negative function of $\hat{\theta}$.

is equal to zero. If $0 \leq k < 1$ instead, we have a positive inflation bias. For values of k close to one, the inflation bias is rather small.

3.2 A Simple Story for the Great Inflation

The Great Inflation refers to the high inflation and unemployment episode of the 1960s, 1970s and early 1980s (figure 1). The model of the previous section provides a useful and powerful tool for the interpretation of this long and important episode of U.S. recent economic history.

I begin the imaginary simulation in 1960. Figure 3 plots real-time estimates of the natural rate of unemployment, inflation persistence in the Phillips curve and the slope of the Phillips curve, starting in 1960 and formed using data from 1948. These represent measures of real-time policymakers' beliefs affecting the choice of the policy variable V. Inflation persistence is measured by the sum of coefficients on lagged inflation, i.e. $\hat{\alpha}(1)$ in (7). The slope of the Phillips curve is measured as the sum of coefficients on unemployment deviations from the natural rate, i.e. $-\hat{\theta}(1)$ in (7). Finally, these estimates are constructed using the baseline, constant gain learning algorithm of section 2.3, but their qualitative behavior is very robust to the alternative specifications of the learning algorithm.

3.2.1 The period of *overoptimism*

To start, notice that in the first part of the sample policymakers' estimates of the natural rate of unemployment were between 4 percent and 5 percent. These are low numbers, compared to our current estimates of the level of the natural rate in the 1960s and the first part of the 1970s.[9]

This erroneous belief that the natural rate was so low led to overexpansionary monetary and fiscal policies. However, while this can explain why inflation started rising in the early 1960s, it is not sufficient to explain why rational policymakers let inflation increase so much and for such a long period of time. What is key to rationalize policymakers' behavior in the 1960s and 1970s, is realizing that they were uncertain not only about the value of the natural rate, but also about the value of all remaining parameters of their model. In particular, observe that the real-time estimate of inflation persistence ($\alpha(1)$) in the early 1960s was approximately equal to 0.5 (figure 3b).[10]

---

[9] For example, compare figure 3a to the model-based, smoothed estimates of the natural rate of unemployment plotted in figure 4.

[10] Remember that if the sum of coefficients on past inflation in the Phillips curve is less than one, this implies the existence of a long-run trade-off between inflation and unemployment.

According to the model, a low estimate of $\hat{\alpha}(1)$ implies a low reaction to inflation, due to the wrong belief that inflation is rather stationary around some mean. This explains the passive behavior of policymakers in the 1960s. I will refer to this period as the *overoptimism* period, during which inflation was perceived as stationary, the natural rate of unemployment as low and the optimal policy was keeping unemployment close to the estimated natural rate, without much concern for the accelerating inflation.

By now, it is not controversial that policymakers' real-time estimates of the natural rate, inflation persistence and the inflation-unemployment long run trade-off were too optimistic in that period (see, for example, Orphanides and Williams 2002, DeLong 1997, Mayer 1999 and Romer and Romer 2002).

3.2.2 The period of *overpessimism*

Slowly something changed. In fact, as suggested by the model and confirmed by figure 3b the estimates of $\alpha(1)$ should be slowly revised upwards, towards the true value of one.

However, while this would imply a reinforcement of the policy reaction to inflation, policy reaction remained low because of the following perverse mechanism: policymakers noticed that pushing unemployment above the *underestimated* natural rate did not provide any relief from inflation. As a consequence they revised toward zero their beliefs about the slope of the Phillips curve ($\hat{\theta}$ decreases in the early 1970s, as shown in figure 3c). As implied by the model, this reduced the strength of policy reaction to inflation. In other words, even after the overoptimism period, policymakers kept reacting weakly to inflation, this time because they perceived a very costly inflation-unemployment trade-off. Okun (1978) provides clear evidence that this was actually the case. In fact, he surveys a number of papers written in the 1970s by important economists and concludes that "the average estimate of the cost of 1 point reduction in the basic inflation rate is 10 percent of a year's GNP" (Okun 1978, p. 348). Another example that the perceived sacrifice ratio was very high is the following statement by the Economic Report of the President (EROP):

> When inflation failed to respond significantly to macroeconomic policy, a 90-day wage and price freeze was announced on August 15, 1971; it was followed by a period of mandatory wage and price controls. (EROP 1979, pp. 54-55.)

Commenting on the current economic conditions, the 1972 EROP further referred to a:

> tendency to an unsatisfactorily high rate of inflation which persists over a long

period of time and is impervious to variations in the rate of unemployment, so that the tendency cannot be eradicated by any feasible acceptance of unemployment. (EROP 1972, p. 113.)

Even in 1979, the EROP wrote:

We will not try to wring inflation out of our economic system by pursuing policies designed to bring about a recession. That course of action ... would be ineffective. Twice in the past decade inflation has accelerated and a recession has followed, but each recession brought only limited relief from inflation. (EROP 1979, p. 7.)

I will refer to this period as the *overpessimism* period, during which policy did not fight inflation because policymakers "did not believe it would work at an acceptable cost" (DeLong 1997, p. 264). The *overpessimism* period is successive to the *overoptimism* one and accounts for the long duration of the hump-shaped episode.

3.2.3   The disinflation

In the meantime, policymakers' estimate of inflation persistence in the Phillips curve had been updated toward the true value of one. In this situation, even small changes of the policy variable are perceived to have long-lasting consequences on the inflation rate. Hence, as shown in figure 2b, the model predicts that small updates of the estimate of the slope of the Phillips curve have large effects on the strength of the policy reaction. Therefore, the episode of high inflation ended after a proper revision of the estimate of the slope of the Phillips curve. This happened because of a sequence of new exogenous shocks, which caused updates of $\hat{\theta}$ toward the self-confirming equilibrium. When the bias of $\hat{\theta}$ decreased and the perceived inflation-unemployment trade-off improved, policymakers reacted strongly to high inflation, because they finally had a model of the economy that was approximately correct. Consequently, unemployment was pushed quickly way above the estimated natural rate. The sharp disinflation was the result of this prompt and strong action. Policy maintained a high unemployment rate until inflation came back under control. At that point unemployment slowly returned to levels close to the natural rate.

Notice that the model's predictions match very well the stylized facts. In fact, not only is the model able to account for the dimension and the duration of the episode, but it is also able to explain why the disinflation period was shorter than the run-up period and why unemployment increased and decreased during the 1970s and early 1980s, lagging behind inflation.

It is important to stress that this paper offers an explanation of the Volcker disinflation, which is not based on a sudden change of policymakers' preferences in the late 1970s. Here, instead, the disinflation occurs when the perceived inflation-unemployment trade-off becomes favorable, relative to the level of inflation.

3.2.4 The cost of disinflation

The fact that the disinflation is delayed due to the perceived high sacrifice ratio represents an important difference between this paper and Sargent (1999). In Sargent (1999), policymakers believe that policy is able to affect real activity only by directly controlling inflation. Therefore, when the correlation between inflation and unemployment approaches zero (like in the early 1970s), for Sargent's (1999) policymakers this represents a unique opportunity to disinflate, since it can be done at no cost in terms of real activity. This does not seem to correspond to what we observe in the data.

Like this paper, also Cogley and Sargent (2004) interpret the rise and fall of inflation as the result of a learning process of U.S. policimakers. However, the policymakers in their paper consider the prescriptions of three policy models and the delayed disinflation is explained by their concern for robustness. Although I do not deny that this might be an important part of the story, here I move in another direction. In fact, I show that a perfect explanation of the behavior of inflation can be obtained when policy is based simply on one model of the economy, if only one is willing to assume that policymakers in the 1960s and 1970s regarded real activity and not inflation under their control. Although they do not provide conclusive evidence, the quotes of section (3.2.2) suggest that this might have actually been the case. Building a model with this feature constitute one important innovation of this paper.

Another crucial difference between this paper and Cogley and Sargent (2004) (as well as Sargent 1999 or Sargent, Williams, and Zha 2004) is the fact that, as a "true" structure of the economy, I use a reduced form new-Keynesian model, which allows me to explain the behavior of unemployment during the Great Inflation. The use of a "true" model of the economy similar to Lucas (1972) and Sargent (1973) (together with their assumption about the public's expectations formation) prevents Cogley and Sargent (2004), Sargent (1999) and Sargent, Williams, and Zha (2004) from offering any explanation for the hump-shape behavior of unemployment.

4 Empirical Evidence

This section presents empirical evidence supporting the model of section 2 and the dynamics illustrated in section 3.

I estimate the model by maximum likelihood methods, over a sample period running from 1960:I to 2002:IV, using quarterly data on inflation and unemployment.[11] I let policymakers estimate the approximate model and, consequently, choose V in every period. Putting the model in state space form, the likelihood can be computed using the Kalman filter (see appendix B). The likelihood is maximized with respect to nine parameters, $\left[\alpha_1; \theta_1; \theta_2; \rho_1; \rho_2; k; \phi; \sigma_\varepsilon^2; \sigma_\eta^2\right]$, while I fix $\delta = 0.99$, $\pi^* = 2$, $\lambda = 1$, $u^* = 6$, $\gamma = 0.99$, $\sigma_\tau^2 = 0.0199$. Fixing δ is standard practice. I fix $\pi^* = 2$ because this number is close to the estimates of the inflation target level for the post-Volcker era (see, for instance, Bullard and Eusepi 2003, Schorfheide 2003, Favero and Rovelli 2003). However, differently from these previous studies, there is no exogenous switch in the level of the inflation target in my model. λ is set to 1 to be consistent with some previous studies (for example, Sims 1988 and Sargent 1999).[12] Finally, for the coefficients of the exogenous process driving the natural rate, $u^* = 6$ is chosen to match the average of unemployment during the sample period. γ is fixed at 0.99 and σ_τ^2 at 0.0199 to impose the prior belief that the natural rate is smoother than unemployment itself. Notice that the value of γ and σ_τ^2 imply an unconditional variance of one for the natural rate. Furthermore, in the case of the constant gain learning, g is set to 0.015 and g_N to 0.03. Observe that the constant gain for the estimation of the natural rate of unemployment is higher than the one for the estimation of the parameters. This captures the fact that policymakers expect the natural rate to drift more than the model's coefficients. I calibrate the initial beliefs in the most natural way, that is using the actual data from 1948 to 1960.[13] This procedure results in the following set of initial beliefs in 1960: $\left[\hat{c}_\pi; \hat{\alpha}_1; \hat{\alpha}_2; \hat{\theta}_1; \hat{\theta}_2; \hat{c}_u; \hat{\rho}_1; \hat{\rho}_2; \hat{u}^N\right] = [1.156; 0.330; 0.131; -0.914; 0.885; 0.012; 1.536; -0.717; 4.701]$.

4.1 Statistical Evidence

The point estimates and the standard errors of the free coefficients are reported in table 1 for the three specifications of the learning algorithm. Four results stand out. First, the estimates

---

[11] As in the rest of the paper, inflation is measured by the annualized quarterly growth rate of the GDP deflator and unemployment by quarterly averages of the monthly civilian unemployment rate.

[12] Interestingly, if λ is treated as a free parameter in the estimation procedure, the point estimate is 0.99, with a standard error of 0.36.

[13] I used the DLS algorithm with discount rate $\Delta = 1 - 1/120$.

obtained using different assumptions for the policymakers' learning process are very similar to each other. Second, the estimate of ϕ, the smoothing coefficient in the policy loss function, might appear as too high. However, once movements in the policy variable V are interpreted in terms of deviations of unemployment from the natural rate, the size of ϕ is not a puzzle anymore.[14] Third, all the other coefficients have reasonable sign and size. Fourth, the estimate of k, the parameter that affects the unemployment target in the loss function (6), is close to unity, casting doubt on the inflationary bias story that I mentioned in the introduction.

Figure 4 plots the smoothed estimate of the natural rate of unemployment, which resembles previous estimates in the literature (Staiger, Stock, and Watson, 1997 and 2001, Gordon, 1997).

Figure 5 plots the evolution over time of the model's policy variable V. A measure of the real rate of interest (rescaled)[15] is reported for comparison. Notice the similarities between the two series, especially in the second part of the sample. This is remarkable, if we consider that the time series for V has been obtained without any information related to the interest rate.

4.2 Model's fit

In order to evaluate the fit of the model, I compare it to the fit of unrestricted multivariate linear models, i.e. bivariate vector autoregressions (VAR) with inflation and unemployment, similar to the ones used in King and Watson (1994). I consider the three alternatives of a VAR(2), a VAR(3) and a VAR(4), where the number in parenthesis indicates the number of lags included in the VAR. Table 2 reports three measures of fit for the learning models and the reference ones. The first column of table 2 reports the value of the log-likelihood, evaluated at the peak. The log-likelihood is useful, but certainly not very informative in a case in which the candidate models have such a different number of free parameters. A VAR(2) has 13 free parameters, a VAR(3) has 17 and a VAR(4) has 21, as opposed to the 9 free parameters of my model. The difference in the number of free parameters is so large that, even though the models are non-nested, it is reasonable to expect a lower log-likelihood for the learning models with respect to the reference ones.

To solve this problem, the second column of table 2 reports the Bayesian Information Crite-

---

[14] For example, consider a permanent unitary increase in the level of V_t. This would create a permanent deviation of unemployment from the natural rate equal to $\frac{1}{1-\rho(1)}$. Therefore the term ϕV_t^2 in the loss function could be expressed in terms of deviations of unemployment from the target as $\phi \left(1 - \rho(1)\right)^2 \left(u_t - \hat{u}_t^N\right)^2$, where the weight becomes $\phi \left(1 - \rho(1)\right)^2$. If unemployment is very persistent, $\rho(1)$ will be close to one and $\phi \left(1 - \rho(1)\right)^2$ will be a small number, despite a very high ϕ.

[15] The real rate of interest is computed as federal funds rate minus the quarterly inflation rate averaged over the last four quarters.

rion (BIC). BIC is an asymptotic approximation of the marginal likelihood and automatically penalizes models with higher number of parameters.

The third column of table 2 is in principle the most reliable measure of fit. The last column in fact reports the logarithm of the marginal likelihood itself, which, is proportional to the posterior probability of the model, under flat prior on the models' space.[16]

From table 2 it is clear that, as expected, the log-likelihood of the learning models is lower than the unrestricted alternatives. However, once the different number of parameters is taken into account, it seems that the learning models fit the data better than the reference VARs. In particular, the marginal likelihood favors the learning models over the alternative ones. Observe that the results of the model comparison exercise are robust to different policymakers' learning schemes.

4.3   Simulations

This subsection considers quantitative simulations of the model in the case of the benchmark (CG) specification. The purpose of these simulations is twofold. First, they show that the model produces a pronounced, prolonged and asymmetric hump-shaped behavior of inflation and unemployment. In addition, these simulations highlight that the mismeasurement of the natural rate of unemployment alone is not sufficient to reproduce this kind of hump-shaped behavior.

In order to do so, I conduct the following exercise. I start in 1960:I, assuming that data on inflation and unemployment are the actual data from 1948:I to 1959:IV. In 1960:I, policymakers optimize their objective function, on the base of the current estimates and fix the value of V for the current period. Unemployment and inflation in 1960:II are determined through the true Phillips curve and aggregate demand equations, (5) and (2). In the next period, policymakers reestimate their approximate model and choose the new value for the policy variable. With this mechanism I simulate 42 years of quarterly data, up to 2002. In the simulations I assume that the true value of the model's parameters are the values estimated in the previous subsection and reported in the first column of table 1. Also as "true" natural rate of unemployment I use the smoothed estimate plotted in figure 4. Finally, I perform the simulations generating sequences

---

[16] The marginal likelihood is computed using the Laplace method, based on a second order approximation of the posterior around the peak. For the VAR models instead, it is possible to compute the exact marginal likelihood. Since the computation of the marginal likelihood requires the use of a proper prior, to remain as agnostic as possible, for the model comparison I use training sample priors. Therefore, while the original sample starts in 1960, I extend it back to 1953. The (properly rescaled) likelihood of the training sample is then interpreted as a prior density for the original sample. Further details can be found in Sims (2002) and Sims (2003).

of *i.i.d.* random innovations with mean zero and variance corresponding to the estimated value reported in table 1.

I perform $3,000$ simulations using identical initial conditions for policymakers' beliefs in 1960 (the ones reported in section 4), but different series of exogenous shocks. The main results are summarized in figure 6. Figures 6a-d graph the empirical distribution of four objects: the maximum level reached by the inflation rate between 1960 and 2002 ($\max(\pi)$); the time period in which inflation reaches its maximum level (t_π^*); the maximum level reached by the unemployment rate between 1960 and 2002 ($\max(u)$); the difference (expressed in years) between the time period in which unemployment peaks and the time period in which inflation peaks ($t_u^* - t_\pi^*$). The scatter plots of figures 6e and 6f are meant to illustrate two bivariate relations: the relation between peak time and the peak level of inflation; the relation between the peak level of unemployment and inflation. Overall, figure 6 makes clear that the model reproduces remarkably well the main characteristics of the Great Inflation. In particular, the model fully captures the dimension, the duration of the high inflation episode and the fact that unemployment peaks after inflation.

Showing that, on average, the simulated time paths exhibit a rapid disinflation is less straightforward, because simply averaging the simulated time paths would not preserve the typical shape. To solve this problem I compute the average path, but only after rescaling the horizontal and vertical axis in every simulation, to make them peak at the same time and at the same level.[17] The result is plotted in figure 7, where the peak time of the average inflation path is normalized to zero and the peak value to the peak level of actual inflation in 1981. Actual inflation is also reported for comparison. The striking feature of figure 7 is that the average of the simulated paths captures perfectly the so called Volcker disinflation, that many macro models have difficulties in addressing.

Figure 8a plots a simulation of the behavior of inflation and unemployment, when the standard deviation of the shocks to the Phillips curve and the aggregate demand are chosen as small as possible and are fixed respectively to 0 and 0.005.[18] This is clearly improper, since the model is nonlinear and the size of the shocks might affect the speed of the learning process. Nevertheless, if we keep this in mind, this simulation shows that the model can capture perfectly

---

[17] In order to determine the peak time of inflation in a robust way, I compute a five years moving average for every simulated path and select the point in time in which the resulting smooth series reaches the maximum.

[18] Two observations are necessary at this point. First, the standard deviation of the shocks to the aggregate demand cannot be set to zero because some random variation in the regressors of the Phillips curve is necessary for a meaningful learning dynamics and the convergence to the self-confirming equilibrium. Second, the simulations with low or zero variance shocks look all exactly the same except for the time period in which inflation (and, of course, unemployment) peaks.

well the low frequency behavior of inflation and unemployment, even with small exogenous shocks. This exercise stresses further the role of initial beliefs, which play a more prominent role than the exogenous shocks in the generation of these peculiar dynamics. This differs from Sargent, Williams, and Zha (2004).

Figure 8a is also very interesting for the behavior of the simulated path of inflation in the last part of the sample. In the 1990s, the "true" series of the natural rate of unemployment exhibits a sharp decline (corresponding to the advent of the so called "New Economy"), which is recognized only gradually by policymakers in the simulation. Nevertheless, since in the 1990s policymakers have approximately learnt the true values of the long and short run slopes of the Phillips curve, they are able to prevent inflation from falling to undesirably low levels. This matches very well the actual behavior of inflation in the last decade.

To stress even more the importance of the misperception about the coefficients of the Phillips curve in the 1960s and 1970s, figures 8b and 8c plot the results of counterfactual simulation exercises. Here I use the same sequence of random shocks used to generate figure 8a, but I change some of the values of the parameters of the model. In the simulation of figure 8b I assume that policymakers know the exact value of the slope of the Phillips curve. However they are uncertain about the natural rate and inflation persistence in the Phillips curve and they have to estimate them. It is evident that the high inflation episode would have lasted less long and that the rapid disinflation would have disappeared. In figure 8c instead, I assume that policymakers have the correct estimate of the parameters of the Phillips curve, except for the natural rate, which is estimated in the usual way. It is clear that, if this had been the case, inflation would have increased much less than it did. Indeed, this graph does not exhibit any sizable low frequency behavior.

In order to shed some light on the model's implications for the future path of inflation, I performed 1,000 simulations for the next 100 years. In these simulations over the period from 2003 to 2102, policymakers' estimates of inflation persistence never reach the low levels of the 1960s, which means that normal supply and demand shocks do not seem to have a major effect on policymakers' estimates of the Phillips curve coefficients when enough data are available for the estimation. As a consequence of this, in all these forward simulations, if and when inflation goes up, it remains high only for a short period of time because, differently from the Great Inflation, the propagation mechanism through learning is absent.

4.4 The fit of alternative explanations

As shown in the previous section, the learning model is able to reproduce remarkably well the postwar behavior of inflation and unemployment in the United States. The purpose of this section is to compare the fit of the baseline learning model to the fit of some alternative explanations of the Great Inflation.

As I mentioned in the introduction, most of the alternative theories of the Great Inflation adopt an exogenous shift in the preferences of policymakers to explain the sharp disinflation and the low level of inflation since the early 1980s. On the other hand, the great advantage of the approach taken in this paper is the fact that the time variation in the conduct of stabilization policy is fully endogenized. To address this issue and assess the importance of the exogenous and unexplainable channels of time variation in policy, I estimate a version of the learning model where also the parameters of the loss function are allowed to vary over time. In particular, in this subsection, I will assume the following form for the loss function of policymakers:

$$L = \hat{E} \sum_{s=t}^{\infty} \delta^{s-t} \left[(\pi_t - \pi_t^*)^2 + \lambda_t \left(u_t - k_t \hat{u}_t^N \right)^2 + \phi \left(V_t - V_{t-1} \right)^2 \right],$$

which differs from (6) because three of the parameters representing the preferences of policymakers (k, π^* and λ) are allowed to change over time. Following many previous studies (for example, Clarida, Gali, and Gertler 2000), I model the time variation in k, π^* and λ in the simplest possible way, i.e. allowing them to differ between the pre and post-Volcker period:

$$[\pi_t^*, k_t, \lambda_t] = \begin{cases} [\pi_1^*, k_1, \lambda_1] & \text{for } t < \bar{t} \\ [\pi_2^*, k_2, \lambda_2] & \text{for } t \geq \bar{t}, \end{cases}$$

where $\bar{t}$ is set to the fourth quarter of 1979.

The estimates of the model with a break in the preferences of policymakers in addition to learning are reported in the first column of table 4, together with the value of the log-likelihood and the BIC.[19] Notice that the parameter controlling the size of the inflationary bias (k) does not change between the pre and post-Volcker periods. Instead, as expected, both the target for inflation and the weight on the unemployment objective seem to decrease after 1979, although the estimates are quite imprecise. In particular, the p-values for the tests with null hypothesis of no change in π^* and λ are respectively 0.5 and 0.34, implying a failure to reject the null of no change in both cases. This is confirmed by the relatively small improvement in the log-likelihood and the substantial deterioration of the BIC.

---

[19] The marginal likelihood for this model is not computed, since a prior for the preference parameters of the post-Volcker period cannot be obtained using a training sample prior (and, therefore, the results would not be easily comparable to the last column of table 2).

As a further check, I consider another alternative, which is a model without learning on the parameters of the Phillips Curve (except for the natural rate of unemployment), but with shifts in the policymakers' preferences. The estimates of this model are reported in the second column of table 4. Notice that in this case k is estimated to be zero, both in the pre and post-Volcker periods. On the other hand, both π^* and λ, seem to decrease in the second part of the sample. The changes in the values of π^* and λ before and after 1979 are even larger than in the previous case, but they are also even less precisely estimated. In fact, the p-values for the tests with null hypothesis of no change in π^* and λ are respectively 0.52 and 0.5, implying again a failure to reject the null of no change in both cases. The log-likelihood and the BIC improve over the previous specification, but the BIC is still inferior with respect to the BIC of the baseline learning model.

The overall impression is that the fit of the baseline model is superior to the alternatives presented in this subsection. In other words, once policymakers are assumed to learn over time about the structural relations of the economy, the time variation in the policy preferences seems redundant.

5 Policymakers' Learning and Forward Looking Agents

In the model of section 2 I used the Phillips curve (1) that can be rewritten as the backward looking equation (5), which is similar to the one estimated by policymakers. In other words I have assumed that policymakers have the correct model of the economy in hand, but are uncertain about the value of the model's parameters. This section tests the robustness of the results to alternative popular assumptions about the price setting equation. In particular, here I will concentrate on the case in which price setters are forward looking (for example, as in Calvo 1983). This assumption generates a forward looking Phillips curve, in which today's prices and inflation depend on expectations of future prices and inflation. For symmetry and in the spirit of the recent new-Keynesian literature, I will also assume that the aggregate demand equation has potentially a forward looking component. To be concrete, in this subsection I will replace the Phillips curve and the aggregate demand, (5) and (2), with the following forward looking versions:

$$\pi_t = \frac{1}{1+d}\pi_{t-1} + \frac{d}{1+d}E_{t-1}\pi_{t+1} - \theta E_{t-1}\left(u_t - u_t^N\right) + \varepsilon_t, \tag{19}$$

$$u_t - u_t^N = \rho_b\left(u_{t-1} - u_{t-1}^N\right) + \rho_f E_{t-1}\left(u_{t+1} - u_{t+1}^N\right) + V_{t-1} + \eta_t, \tag{20}$$

where d is the private sector's discount factor. The first equation is a forward looking Phillips curve and, similarly to Christiano, Eichenbaum, and Evans (2005), can be derived from the firms' profits maximization problem. As in Calvo (1983), firms are assumed to be able to reoptimize and set their prices with a certain probability in every period. If they do not reoptimize, they are assumed to set their prices with an indexation mechanism to past inflation. This simple indexation rule in the price setters behavior explains the backward looking component in (19). As pointed out in many studies, this component helps to account for the high degree of inflation persistence observed in the data (see, for instance, Fuhrer and Moore 1995, Gali and Gertler 1999, Christiano, Eichenbaum, and Evans 2005). Observe that (19) is a vertical Phillips curve in the long run, implying absence of long run trade-off between inflation and unemployment.

The second equation is a forward looking aggregate demand equation. This expression can be derived from the household maximization problem, allowing for external habit formation, as in Smets and Wouters (2003), when $\rho_b + \rho_f = 1$ and $\rho_b < 0.5$. Habit formation explains the backward looking component, which helps to account for the persistence observed in the data. In my estimation and simulations I will relax the previous equality and inequality restrictions on ρ_b and ρ_f because those restrictions seem to be at odds with the data. Observe that the value of V is chosen by the policymakers in every period in the same way of section 2. In other words, policymakers estimate the backward looking model given by (7) and (8) and fix V solving the linear quadratic problem based on their current beliefs. Differently from the benchmark case, policymakers not only are using estimated parameters instead of true ones, but they are also using a model structure which is fundamentally different from the true model of the economy.

In order to solve the model and possibly estimate it, we need to make assumptions on the private sector's expectation formation process. As in Sargent (1999), I assume that the private sector knows exactly the way policy is made. In other words, people know that policymakers estimate (7) and (8) and fix V by solving the linear quadratic problem (6). However, there are many possible assumptions for the way agents form their expectations on future policy. For completeness, I analyze two alternative (and completely opposite) cases. As a starting point, following most of the adaptive learning literature, I assume that the private sector thinks that policymakers will not revise their estimates of (7) and (8) in the future and will continue to implement policy based on their latest estimates of (7) and (8). For simplicity, I will denote this behavior of the private sector as "partially rational". As an alternative and complication of the baseline hypothesis, I also analyze the case in which the private sector is "fully rational"

and takes into account the fact that policymakers will revise their estimates of (7) and (8) on the base of future data.

5.1 Partially Rational Private Agents

The solution of this first specification of the model is reasonably standard. In fact, under the assumptions that the coefficients $g(\hat{\beta})$ in (9) remain constant in the infinite future, equations (19), (20), (9) and (3) form a *linear* rational expectations system of equations (RESE), which can be solved using standard methods (like Sims 2001a). In reality the coefficients $g(\hat{\beta})$ change, therefore the RESE must be solved in every period of the sample to find the time varying data generating process. Given a random sequence of exogenous, non-policy shocks, this method can be used to construct simulated paths of the variables of interest. Given the actual data, this method can be used to compute the likelihood and maximize it with respect to the model's parameters. I fix the private sector discount factor (d) to 0.99 and maximize the likelihood with respect to seven parameters $\left[\theta; \rho_b; \rho_f; k; \phi; \sigma_\varepsilon^2; \sigma_\eta^2\right]$.

The log-likelihood and BIC for the forward looking model are reported in table 5. Although the data seem to favor the backward looking specification, it is worth focusing on the estimates of the forward looking model. The point estimates and the standard errors of the parameters are reported in table 5.[20] First notice that the point estimate of θ has the correct sign, but it is very small. However this must be expected. The reason is that, if the private sector thinks that the policy rule parameters will not change in the future, a higher θ would imply an immediate enormous effect on inflation's expectations and, consequently on inflation itself. We do not observe this in the data. The second thing to notice is that the weight on the forward looking component of the aggregate demand equation is estimated to be zero. The statistical evidence for this result is very strong. In fact, when I reestimate the model imposing the restriction $\rho_f > 0.5$, the log-likelihood drops dramatically. This makes equation (20) hard to reconcile with the household optimization behavior and supports the skepticism shown, among others, by Estrella and Fuhrer (2002) and Fuhrer and Rudebusch (2002).

As in the previous section, I fix the parameters to their estimated value and perform $3,000$ simulations of the pattern of inflation and unemployment by generating random sequences of exogenous shocks. The second column of table 3 reports medians, lower and upper quartiles of $\max(\pi)$, $\max(u)$, t_π^* and $t_u^* - t_\pi^*$. The median peak level of inflation is even higher than for the baseline model, unemployment seems to peak a bit later and there is more uncertainty related

---

[20] As in the benchmark case, here I assumed that policymakers learn using a constant gain algorithm.

to the peak time of inflation. However, overall, the summary statistics are in line with the actual behavior of inflation and unemployment during the Great Inflation. The asymmetry of the high inflation episodes generated by the forward looking model is demonstrated in figures 9a and 9b, which are constructed in the same way as figures 8a and 7.

Summarizing, the results obtained with the baseline model are very robust to the use of the forward looking Phillips curve. We observe both inflation increasing and decreasing faster, as well as unemployment peaking after inflation.

5.2 Fully Rational Private Agents

The solution of the model with the full rationality assumption is non-standard. The reason is that the system given by (19), (20), (3), (9), (10), (11), (12) and (13) is a *nonlinear* RESE, because the coefficients $g(\hat{\beta})$ are nonlinear functions of the endogenous variables. To solve the model I use numerical methods. In particular I apply the method that Fackler and Miranda (2001) propose for a general nonlinear RESE. It consists in approximating unknown functions (which, in our case, are the expectations $E_{t-1}\pi_{t+1}$ and $E_{t-1}u_{t+1}$) with projection techniques. In other words, the response function is approximated by the function Φs_{t-1}, where Φ is a matrix of coefficients and s_{t-1} is a collection of basis functions of the state variables. The method consists of replacing $E_{t-1}\pi_{t+1}$ and $E_{t-1}u_{t+1}$ with the approximations and finding the matrix of coefficients Φ^* that solve the RESE. The difficulty here is the high dimension of the problem and of the state vector. The details of the solution method are given in appendix C.

The solution of the model can take hours, therefore estimation, though theoretically easy, is not practical.[21] For this reason I perform only a set of simulations, calibrating the model in the following way: as before, d is fixed to 0.99; the parameters $\left[\rho_b; \rho_f; k; \phi; \sigma_\varepsilon^2; \sigma_\eta^2\right]$ are fixed to the point estimates in the model's specification of section 5.1; θ is fixed to 0.01, which is in the range of the values that have been used by the literature. Notice that θ is fixed to be higher than the point estimate in the previous specification. The reason is that, under this new assumption on the expectation formation, the private sector recognizes that policy mistakes will be slowly corrected in the future. Consequently, observing policy mistakes today does not cause a huge effect on expected inflation and inflation itself.

I perform $3,000$ simulations and report medians, lower and upper quartiles of $\max(\pi)$, $\max(u)$, t_π^* and $t_u^* - t_\pi^*$ in the last column of table 3. Observe that inflation and unemployment continue to peak at high levels, which are comparable to what we observe in the data. However,

---

[21] Since it would require several likelihood evaluation and, therefore, several model's solutions for different values of the parameters.

on average, inflation peaks earlier and difference between the peak time of unemployment and inflation is higher than in the simulations with the other models. This is confirmed by figure 9c and 9d.

Generally, two points stand out from the analysis of table 3 and figures 9c and 9d: first, the simulations of this model continue to exhibit the pronounced and prolonged hump-shape behavior of inflation and unemployment; second, the typical simulated path of inflation peaks earlier and does not show a disinflation as sharp as the one observed in the data and reproduced by the baseline model and the forward looking model of section 5.1. However, this is not at all surprising, since the private sector is assumed to be forward looking and to take fully into account future policy changes. If the private sector knows the way policy is made, it will predict that policy will become very active against inflation as soon as policymakers' estimates of the Phillips curve will be revised. Inflation expectations will be automatically adjusted to take this into account. Consequently, inflation will reflect these adjustments in expected inflation.

This section suggests two conclusions. First, that even in the forward looking Phillips curve case, introducing policymakers' learning is important to explain the dimension and duration of high inflation episodes. Second, that the assumption of fully rational agents, who form their expectations taking into account the future evolution of policymakers' beliefs about the economy, is probably too strong and at odds with the data on the disinflation episode of 1981.

6 Stochastic Volatility and the Role of Non-Policy Shocks

In this section I extend the baseline framework to account for stochastic volatility. The fact that exogenous shocks have exhibited high heteroskedasticity over the last forty years has been highlighted in many studies (see, for instance, Cogley and Sargent 2003, Primiceri 2005, Sims and Zha 2004, Stock and Watson 2002, Stock and Watson 2003). Furthermore, many authors have argued that one of the main causes of the American inflation in the 1970s was exactly a particularly bad sequence of exogenous, non-policy shocks (see, for instance, Blinder 1979). On the other hand, this paper has so far provided an explanation of the high inflation episode based on the policy behavior. Therefore, it seems natural to extend and reestimate the model of section 2 allowing for heteroskedastic innovations. In this way, I leave it up to the data to determine whether the channel of the policymakers' behavior remains important once the possibility of time varying variances is taken into account.

There is also another reason why it is potentially very important to take heteroskedasticity into account. This is because, in a learning model, the size of the exogenous shocks affects

the speed of the learning process. In very general terms, higher volatility of an equation's disturbances would slow down the learning process. On the other hand, higher volatility of the regressors would speed up the learning process.

In order to allow for heteroskedasticity, I make the assumption that the standard deviations of the exogenous innovations of equations (5) and (2) follow a geometric random walk process:

$$\log \sigma_{\varepsilon,t} \quad = \quad \log \sigma_{\varepsilon,t-1} + \nu_{\varepsilon,t}, \tag{21}$$

$$\log \sigma_{\eta,t} \quad = \quad \log \sigma_{\eta,t-1} + \nu_{\eta,t}. \tag{22}$$

This class of models is known as stochastic volatility models and constitutes an alternative to ARCH and GARCH models. The crucial difference is that the variances generated by (21) and (22) are unobservable components.

The estimation of the model augmented with stochastic volatility is considerably more involved than the benchmark case. I adopt Markov chain Monte Carlo (MCMC) methods for the posterior numerical evaluation of the parameters of interest. MCMC deals efficiently with the nonlinearities of the model, allowing to draw from lower dimensional and standard distributions as opposed to the high dimensional joint posterior of the whole parameters set. Appendix D gives the details of the estimation algorithm.

The fourth column of table 1 reports the posterior median and the posterior standard deviations of the parameters of interest. They are similar to the benchmark case of the previous section. Figure 10 plots the posterior median (and the 68 percent error band) of the time varying standard deviations of the innovations to the Phillips curve and the aggregate demand equation. Observe that the time path is consistent with the past literature (see, for instance, Cogley and Sargent 2003, Primiceri 2005, Sims and Zha 2004, Stock and Watson 2002). The variance of the innovations to the Phillips curve is low in the 1960s, high in the 1970s and early 1980s and low again since 1985 to the end of the sample. The variance of the demand shocks in the aggregate demand equation follows approximately the same time pattern. Notice the sharp decrease in the standard deviation of the shocks to the unemployment rate that occurred in the early 1980s, a phenomenon which is known as the Great Moderation.

I evaluate the role of the exogenous shocks in the high inflation episode of the 1970s by performing counterfactual simulation exercises. The methodology I adopt is straightforward. I fix the model's coefficients to the estimated posterior medians and reconstruct an estimate of the sequence of exogenous shocks $\{\varepsilon_t\}_{t=1}^T$ and $\{\eta_t\}_{t=1}^T$. Starting from 1960:I, these shocks can be used to simulate counterfactual data, constructed using different values of the parameters. These new series can be interpreted as the realization of the data that would have been observed,

had the parameters of the model been the ones used to generate the series. In this context, the interesting experiment consists of replaying history assuming that the standard deviations of the exogenous shocks were lower that the estimated standard deviations of the 1960s and 1970s. For comparison, figure 11a plots actual inflation and unemployment. Figure 11b plots the counterfactual paths of inflation and unemployment when the time varying standard deviations of supply and demand shocks are both replaced by their value in 1995, which is one of the least volatile periods. It is clear that the differences from the actual behavior of inflation and unemployment are minor, except, of course, for the reduced volatility. Figure 11c and 11d plot the counterfactual paths of inflation and unemployment when the time varying standard deviations of respectively supply and demand shocks are replaced by their value in 1995. Figure 11d shows that the worst scenario for inflation would have occurred in the case of reduced volatility of the shocks to unemployment. The reason is that it would have slowed down the updating process toward the self-confirming value for the beliefs about the slope of the Phillips curve parameter. For the same reason, the best scenario would have been the one of figure 11c, i.e. the case of reduced volatility of the supply shocks and the estimated historical volatility of the demand shocks. Finally notice that, although the behavior of inflation differs across the four plots, even in the most favorable scenario of figure 11c, inflation would still have peaked at the high level of about 9 percent.

Overall, this section suggests two conclusions. First, allowing for heteroskedasticity of the exogenous shocks does not undermine the results that the low frequency behavior of inflation and unemployment is largely explained by the evolution of policymakers' beliefs. Second, in the context of this model, allowing for time varying variances does not generate important differences in the speed of the policymakers' learning process and the timing of the disinflation. Instead, the dynamics seem to be mainly driven by the convergence process of policymakers' beliefs to the self-confirming equilibrium, starting from the initial conditions. This finding also facilitates the interpretation of figures 8a, 9a and 9c.

7   Why Did Inflation Rise?

In the previous sections I have argued that policymakers' mistakes in the estimation of the natural rate of unemployment can only explain why inflation started to increase in the early 1960s. Observe that mistakes in the estimation of the natural rate in real time are, to some extent, unavoidable (see, for example, Orphanides and Van Norden 2001). This is because the natural rate is intrinsically a time varying entity, due to a number of factors like demographics,

changes in productivity growth or labor market conditions (for an overview, see Ball and Mankiw 2002).

On the other hand, the optimistic view about the natural rate is not enough to understand why rational policymakers let inflation rise for more than fifteen years. In fact, I showed that in the 1960s policymakers did not fight inflation strongly enough because they underestimated the persistence of inflation in the Phillips curve. This induced the peculiar dynamics described in section III.

Of course, a fundamental question is why policymakers started out with such downward biased beliefs about the degree of inflation persistence in the Phillips curve ($\alpha(1)$) in the 1960s? The answer to this question is very simple and natural: because these beliefs were perfectly consistent with the available data prior to 1960. Observe that this is true not only if we look at data between 1948 and 1960 (like this paper does to calibrate initial beliefs), but also if we look further back in the past. For example, Christiano and Fitzgerald (2003) provide unambiguous evidence that all data between 1900 and 1960 appeared consistent with a stable long run trade-off between inflation and unemployment. Similarly, Barsky (1987) uses data from 1890 to formally show that strong inflation persistence emerged only around 1960. In other words, there was basically no way for adaptive policymakers learning from the past to anticipate the high degree of persistence of the last decades.

There are obviously many possible reasons why the properties of the inflation-unemployment process might have changed so drastically in the postwar period. One conjecture is that the different nature of the data is related to a change in monetary regimes like, for example, the movement away from the Gold Standard and Bretton Woods (for an overview, see Bordo 1993). This does not necessarily imply that inflation rose because policymakers abandoned the commitment technology provided by fixed exchange rate regimes (like, for instance, in Bordo and Kydland 1995). The interpretation provided in this paper is instead that inflation rose because policymakers were simply slow to learn how to conduct policy under a new regime.

8   Concluding Remarks

This paper presents a simple model of inflation-unemployment dynamics when policymakers have imperfect information. The source of imperfect information is the fact that policymakers do not know the exact model of the economy. Therefore, they update their beliefs about the model's unknowns in every period and implement optimal policy, conditional on their current beliefs.

The model's self-confirming equilibrium is characterized by low inflation. Nevertheless, the model can generate prolonged and asymmetric episodes of high inflation, which closely resemble the run-up of US inflation in the 1960s and 1970s and the sharp disinflation in the early 1980s. In particular, these episodes occur when policymakers simultaneously underestimate *both* the natural rate of unemployment *and* the persistence of inflation in the Phillips curve. Starting from this situation of overoptimistic beliefs, the model endogenously generates periods in which policy is overpessimistic. During these periods policy mistakenly perceives the inflation-unemployment trade-off as too costly. I show that it is not optimal to create a recession to stop inflation, either if beliefs are overoptimistic or if they are overpessimistic. This is why the policy action against inflation is delayed until the moment in which the perceived trade-off improves. When this happens, inflation is already very high. Therefore, the policy action against inflation is strong and decisive.

Unlike most of the existing literature, the model matches many recognized stylized facts of the American Great Inflation of the 1960s, 1970s and 1980s. I also formally estimate the model by likelihood methods and find that it fits the behavior of US inflation and unemployment remarkably well.

Given the empirical support, the model can be used to evaluate the possibility that episodes similar to the Great Inflation could happen again in the future. In this respect, the conclusion of the paper is more optimistic than alternative theories. In this model, in fact, high inflation episodes do not occur as the consequence of the mismeasurement of the natural rate of unemployment only (Orphanides 2000) or as a result of policymakers' lack of commitment to low inflation (for example, Sargent 1999). High inflation is the consequence of the unlikely combination of two factors. As mentioned above, one of these two factors is the mismeasurement of the natural rate, which, to some extent, is unavoidable. However, the serious underestimation of the persistence of inflation in the Phillips curve seems to be much more uncommon and appears more related to special circumstances, such as structural breaks in the true model of the economy.

A Self-confirming equilibrium

Appendix A formally defines a self-confirming equilibrium and derives the self-confirming equilibrium of the model of section 2, in the case of constant gain learning. Let y_t^{π}, y_t^u, x_t^{π}, x_t^u, $\hat{\beta}^{\pi}$ and $\hat{\beta}^u$ be the same objects defined in section 2.3. I follow Sargent (1999) in defining a self-confirming equilibrium in this framework:

Definition 1 *A self-confirming equilibrium is a set of policymakers' beliefs about the models' parameters $\hat{\beta} \equiv [\hat{\beta}^{\pi}, \hat{\beta}^u, \hat{u}^N]$, a fixed optimal policy rule $g(\hat{\beta})$ and an associated stationary stochastic process for the vector $\left[\pi_t, u_t, V_t, u_t^N\right]$ such that: (a) $\hat{u}^N$, $\hat{\beta}^{\pi}$ and $\hat{\beta}^u$ satisfy*

$$E\left[u_t - \hat{u}^N\right] \;=\; 0 \tag{23}$$

$$E\left[x_t^i\left(y_t^i - x_t^{i\prime}\hat{\beta}^i\right)\right] \;=\; 0, \qquad i = \{\pi, u\} \tag{24}$$

where the expectations are taken with respect to the probability distribution generated by (5), (2), (3) and (9); (b) the vector $\left[\pi_t, u_t, V_t, u_t^N\right]$ is generated by the stationary stochastic process implied by (5), (2), (3) and (9).

It is straightforward to verify that the set of beliefs $\hat{u}^N = u^*$, $\hat{\beta}^{\pi} = [0; \alpha_1; \alpha_2; \theta_1; \theta_2]$ and $\hat{\beta}^u = [0; \rho_1; \rho_2]$ satisfy (23) and (24) and therefore represents a self-confirming equilibrium, in the case of $\sigma_{\tau}^2 = 0$. When $\sigma_{\tau}^2 > 0$, finding the self-confirming equilibrium is more involved and requires a numerical solution of the system of equations given by (23) and (24). The procedure works as follows: any given fixed value of $\hat{u}^N$, $\hat{\beta}^{\pi}$ and $\hat{\beta}^u$ implies a linear stochastic process for $\left[\pi_t, u_t, V_t, u_t^N\right]$ via equations (5), (2), (3) and (9). The linear process can be rewritten as a first order system of the form $z_t = C + A z_{t-1} + B\nu_t$. Thus $E(z_t) = (I - A)^{-1}C$ and $Var(z_t)$ can be found by solving the Lyapunov equation $Var(z_t) = A Var(z_t) A' + B Var(\nu_t) B'$. The elements of $E(z_t)$ and $Var(z_t)$ can be used to compute $E\left[u_t - \hat{u}^N\right]$ and $E\left[x_t^i(y_t^i - x_t^{i\prime}\hat{\beta}^i)\right]$, for $i = \{\pi, u\}$, which, in general, will not be equal to zero. A simple equation solver can be used to solve for the set of beliefs $\hat{u}^N$, $\hat{\beta}^{\pi}$ and $\hat{\beta}^u$, which satisfy (23) and (24). Of course the solution will depend on the value of the true parameters of the model. As an illustrative example I consider the case in which the true parameters of the model are the point estimates of the baseline specification, presented in the first column of table 1. The self-confirming equilibrium in this case corresponds to $\hat{u}^N = 6$, $\hat{\beta}^{\pi} = [0.0394; 0.7203; 0.2623; -0.8409; 0.7637]$ and $\hat{\beta}^u = [0; 1.5703; -0.6269]$. Furthermore, the eigenvalues of the Jacobian of the expressions contained in (23) and (24), evaluated at the self-confirming equilibrium, have negative real parts. This guarantees the stability of the equilibrium.

As $Var\left(u_t^N\right)$ increases the mistakes associated with the estimation of the current level of the natural rate of unemployment will bias toward zero the estimate the slope of the Phillips curve and of the persistence of unemployment deviations from the natural rate in the aggregate demand equation. This leads to self-confirming equilibria whose distance from the true values of the parameters is increasing in $Var\left(u_t^N\right)$. Finally, when $Var\left(u_t^N\right)$ is large with respect to σ_ε^2 and σ_η^2, the model does not admit a self confirming equilibrium anymore. Figure 12 plots the Euclidean distance between the true parameters $([0; \alpha_1; \alpha_2; \theta_1; \theta_2; 0; \rho_1; \rho_2; u^*])$ and the set of beliefs about these parameters corresponding to the self-confirming equilibria. These self-confirming equilibria are computed for different values of $Var\left(u_t^N\right) \geq 0$. This graph confirms the intuition that the distance between equilibrium beliefs and true parameters increases with $Var\left(u_t^N\right)$. The line is truncated at the value 4.63, because for $Var\left(u_t^N\right)$ larger than this value a self confirming equilibrium cannot be found. All the found self-confirming equilibria are stable.

B State space form for model's estimation

Appendix B gives the details of the state space form representation of the model for the estimation with the Kalman filter.

The canonical state space form is given by:

$$y_t = AZ_t + BX_t + Re_t, \tag{25}$$

$$X_t = C + GX_{t-1} + Qs_t, \tag{26}$$

$$\begin{pmatrix} e_t \\ s_t \end{pmatrix} \sim \ i.i.d. \ N(0, I). \tag{27}$$

In our case, $y_t = [\pi_t; u_t]'$; $Z_t = [\pi_{t-1}; \pi_{t-2}; u_{t-1}; u_{t-2}; V_{t-1}]'$; $X_t = \left[u_t^N; u_{t-1}^N; u_{t-2}^N\right]'$;

$$A = \begin{bmatrix} \alpha_1 & \alpha_2 & -\theta_1 & -\theta_2 & 0 \\ 0 & 0 & \rho_1 & \rho_2 & 1 \end{bmatrix}; B = \begin{bmatrix} 0 & \theta_1 & \theta_2 \\ 1 & -\rho_1 & \rho_2 \end{bmatrix}; C = \begin{bmatrix} (1-\gamma)u^* \\ 0 \\ 0 \end{bmatrix}; G = \begin{bmatrix} \gamma & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix};$$

$$R = \begin{bmatrix} \sigma_\varepsilon & 0 \\ 0 & \sigma_\eta \end{bmatrix}; Q = \begin{bmatrix} \sigma_\tau & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The standard Kalman filter recursion formulas can be found in Hamilton (1994). To start the recursion it is necessary to specify $E\left(X_0|\Omega_0\right)$ and $Var\left(X_0|\Omega_0\right)$, where Ω_0 represents the information set available at time 0. Following a common practice, I set $E\left(X_0|\Omega_0\right)$ and $Var\left(X_0|\Omega_0\right)$ to the unconditional values implied by the transition equation. In particular, this results in

$E\left(X_0|\Omega_0\right) = [6; 6; 6]'$, which corresponds approximately to the estimate of the natural rate of unemployment of Staiger, Stock, and Watson (2001) in 1960 (which is the initial date of our sample).

C   The solution method for the forward looking model

Appendix C illustrates in more detail the method used to solve the forward looking model with fully rational agents. As mentioned in section 5.2, the model is hard to solve because it is a nonlinear system of rational expectation equations. The source of nonlinearity is the learning behavior of policymakers. I will rely on numerical methods. The adopted solution method is based on Fackler and Miranda (2001). A similar method is in Fernandez-Villaverde and Rubio (2002).

Consider the system of rational expectation equations given by (19) and (20). To simplify the analysis and only for the purposes of this section I will assume that u_t^N is a deterministic function, known by the private sector, but, as usual, unknown by policymakers. I will set u_t^N to be equal to the smoothed estimate of u_t^N obtained in the estimation of the forward looking model with partially rational agents (section 5.1). Thus, let $\tilde{u}_t \equiv u_t - u_t^N$. Equations (19) and (20), the only ones involving expectations, can be rearranged and rewritten in the following compact form:

$$y_t = AE_{t-1}y_{t+1} + BX_{t-1} + v_t, \tag{28}$$

where $y_t \equiv [\pi_t, \tilde{u}_t]'$ is the vector of observed endogenous variables; $X_{t-1} \equiv [\pi_{t-1}, \tilde{u}_{t-1}, V_{t-1}]$ is the vector of observed predetermined variables; $v_t \equiv [\varepsilon_t, \eta_t]'$ is the vector of unobservable shocks; A and B are matrices of coefficients, omitted for brevity. (28) is linear, but the complete system, given by (28), (9), (10), (11), (12) and (13) is nonlinear. The solution of the model is the unknown response function $E_{t-1}y_{t+1} = \Psi(\Omega_{t-1})$, where Ω_t represents the information available at time t and $\Psi(\cdot)$ satisfies

$$\Psi(\Omega_{t-1}) = AE_{t-1}\Psi(\Omega_t) + BE_{t-1}X_t.$$

When the model is nonlinear in general there is not a closed form expression for $\Psi(\cdot)$ and it must be approximated numerically by projection methods. The basic idea of Fackler and Miranda (2001) is approximating $\Psi(\Omega_t)$ with a linear combination of basis functions of the state variables. This is given by Φs_t, where s_t is an $m \times 1$ vector of basis functions and Φ is a $2 \times m$ matrix of coefficients. In the numerical procedure, also the expectation operator must be approximated

using quadrature methods. Therefore, the expectation of a generic function $f\left(\cdot\right)$ of the model's source of randomness, v_t, is approximated by a discrete version of the integral, given by

$$Ef(v_t) \approx \sum_{j=1}^{k} \omega_j f\left(v_t^j\right).$$

For a given value of the innovations v_t^j,

$$y_t^j = A\Phi s_{t-1} + BX_{t-1} + v_t^j$$

and

$$y_{t+1}^j = A\Phi s_t^j + BX_t + v_{t+1},$$

where the superscript j for s indicates that the value of s at time t depends on the realization of the shocks at time t. Now we can compute $E_{t-1}y_{t+1} = E_{t-1}E_t y_{t+1} \approx \sum_{j=1}^{k} \omega_j \left(A\Phi s_t^j + BX_t\right) \equiv z_{t-1}$. Let $S \equiv [s_{1,t-1}, ..., s_{n,t-1}]$ be a collection of $n \geq m$ values of s_{t-1} and $Z \equiv [z_{1,t-1}, ..., z_{n,t-1}]$ the collection of the corresponding n values of z_{t-1}. The solution consists in the Φ^* which solves $\Phi^* S = Z$, or $\Phi^* SS' = ZS'$ in the case in which $n > m$. It can be done using standard equation solvers.

In this application, to approximate the integrals and expectation operators, I use a Gauss-Hermite quadrature with 3 nodes (see Judd 1998). As mentioned above, the dimension of the state vector is high. Thus, the use of tensor product bases or complete polynomial bases is unfeasible for any polynomial degree bigger or equal to 2. For this reason I chose the following ad hoc collection of 21 basis functions of the states, which turned out to work well:

$$
\begin{aligned}
s_t \;=\; & \left[1; \pi_t; \tilde{u}_t; \pi_{t-1}; \tilde{u}_{t-1}; V_t; \hat{c}_{\pi,t}; \hat{\alpha}_{1,t}; \hat{\alpha}_{2,t}; \hat{\theta}_{1,t}; \hat{\theta}_{2,t}; \hat{c}_{u,t}; \hat{\rho}_{1,t}; \hat{\rho}_{2,t}; \hat{u}_t^N; \right. \\
& \left. \pi_t^2; \tilde{u}_t^2; \hat{\alpha}_{1,t}\pi_t; \hat{\alpha}_{2,t}\pi_{t-1}; \hat{\theta}_{1,t}\left(u_t - \hat{u}_t^N\right); \hat{\theta}_{1,t}\left(u_{t-1} - \hat{u}_{t-1}^N\right)\right].
\end{aligned}
$$

Notice that the choice of s_t includes all linear terms in the state variables of the problem and some potentially relevant second order terms. The dimension of s_t is so large that the choice of S based on standard grid methods is unfeasible, even specifying only two values for any state variable. To solve this problem I chose a collection of $n = 86$ s_t's, corresponding to the actual values of s_t observed in the data, every 2 quarters, from 1959:IV to 2002:IV. The results are only marginally affected by a different choice of values for S, like for example the observed data, every 2 quarters, 1960:I to 2002:III.

D   The MCMC algorithm for the stochastic volatility model

Appendix D illustrates the details of the MCMC algorithm used in section 6 for the estimation of the model with stochastic volatility. The parameters of interest are the coefficients $\Psi_1 \equiv [\alpha_1; \theta_1; \theta_2; \rho_1; \rho_2]$, $\Psi_2 \equiv [k; \phi]$, $\Psi_3 \equiv \left[\sigma^2_{\nu_\varepsilon}; \sigma^2_{\nu_\eta}\right]$ and the unobservable states $u^N \equiv \left\{u^N_t\right\}^T_{t=1}$, $\sigma_\varepsilon \equiv \{\sigma_{\varepsilon,t}\}^T_{t=1}$ and $\sigma_\eta \equiv \{\sigma_{\eta,t}\}^T_{t=1}$. The estimation consists of the simulation of the posterior of the parameters of interest, conditional on the observed data. MCMC allows to simulate lower dimensional conditional posteriors instead of the high dimensional unconditional one.

Notice that the model can be rewritten like in (25), (26) and (27), with the difference that now the elements of R are time varying and follow the processes (21) and (22). The algorithm works in 5 steps.

D.1   Step 1: drawing u^N

Conditional on σ_ε, σ_η, Ψ_1 and Ψ_2, the observation equation (25) is linear and has Gaussian innovations with known variance. Therefore, the vector u^N can be drawn using standard simulation smoothers, like, for instance, Carter and Kohn (1994). Details of this procedure can be also found in Kim and Nelson (1999b).

D.2   Step 2: drawing σ_ε and σ_η

Consider now the system of equations

$$y_t - AZ_t - BX_t = y^*_t = R_t e_t \tag{29}$$

where, taking u^N, Ψ_1 and Ψ_2 as given, y^*_t is observable. This is a system of nonlinear measurement equations, but can be easily converted in a linear one, by squaring and taking logarithms of every element of (29), which leads to the following approximating state space form:

$$y^{**}_t = 2h_t + e^{**}_t \tag{30}$$

$$h_t = h_{t-1} + \omega_t. \tag{31}$$

$y^{**}_{it} = \log[(y^*_{it})^2 + \bar{c}]$; $\bar{c}$ is an offset constant (set to 0.001); $e^{**}_{it} = \log(e^2_{it})$; $h_t = \log(\text{diag}(R_t))$. Observe that the e^{**}'s and the ω's are not correlated. The system in this form has a linear, but non-Gaussian state space form, because the innovations in the measurement equations are distributed as a $\log \chi^2(1)$. In order to further transform the system in a Gaussian one, a mixture of normals approximation of the $\log \chi^2$ distribution is used, as described in Kim, Shephard,

and Chib (1998). Observe that the variance covariance matrix of the e's is the identity matrix. This implies that the variance covariance matrix of the e^{**}'s is also diagonal, allowing to use the same (independent) mixture of normals approximation for any element of e^{**}. Kim, Shephard and Chib (1998) select a mixture of 7 normal densities with component probabilities q_i, means m_i, and variances v_i^2, $i = 1, ..., 7$. The constants $\{q_i, m_i, v_i^2\}$ are chosen to match a number of moments of the $\log \chi^2(1)$ distribution. The constants $\{q_i, m_i, v_i^2\}$ can be found in Kim, Shephard, and Chib (1998).

For the innovation to the variable y_{jt}, define as $s_j^T = [s_{j1}, ..., s_{jT}]'$ the vector of indicator variables selecting at every point in time which member of the mixture of normal approximation has to be used. Conditional on u^N, Ψ and s^T (which denotes the collection of s_j^T), the system has an approximate linear and Gaussian state space form. Again, exactly like in the previous step of the sampler, this procedure allows to draw every h_t using a simulation smoother.

Conditional on the data and the new series of h_t's, it is possible to sample the new s_j^T vectors, to be used in the next iteration. This is easily done (separately for every j) by sampling from the discrete densities defined by

$$\Pr(s_{jt} = i \mid y_{jt}^{**}, h_{jt}) \propto q_i f_N(y_{jt}^{**} \mid 2h_{jt} + m_i, v_i^2), \quad i = 1, ..., 7.$$

Further details can be found in Kim, Shephard, and Chib (1998) or Primiceri (2005).

D.3 Step 3: drawing Ψ_1

Conditional on Ψ_2, σ_ε, σ_η and u^N, the objects Z_t, X_t and R_t are observable. Therefore, the elements of Ψ_1 (which correspond to the elements of A and B) can be easily drawn from the posterior of the coefficients of a regression with known variance. This posterior is normally distributed with mean equal to the OLS coefficients and variance equal to the variance of the OLS coefficients.

D.4 Step 4: drawing Ψ_2

Ψ_2 enters the model non-linearly. Therefore, in order to draw from the conditional posterior of Ψ_2, I use a Metropolis step, nested in the Gibbs sampler. The procedure works as follows: I draw a candidate value Ψ_2^* from a proposal distribution $\varphi\left(\Psi_2^* | \Psi_2^{i-1}\right)$, where Ψ_2^{i-1} is the previous draw of the chain. At this point I compute the value of the posterior associated to the draw, $p\left(\Psi_2^* | \Psi_1, \sigma_\varepsilon, \sigma_\eta, u^N, \{y_t\}_{t=1}^T\right)$, which, under flat prior, is proportional to the value of the

likelihood. The new draw is accepted with probability

$$a = \min\left\{\frac{p\left(\Psi_2^*\right)/\varphi\left(\Psi_2^*\right)}{p\left(\Psi_2^{i-1}\right)/\varphi\left(\Psi_2^{i-1}\right)}, 1\right\}.$$

If the proposal value is rejected, the next element of the chain is set to be Ψ_2^{i-1}. In order to satisfy the constraints $\phi \geq 0$ and m, $0 \leq k \leq 1$, I chose the proposal distribution to be normal in $f(\Psi_2^*)$, where $f(a,b) = \left[\log(a), \log\left(\frac{b}{1-b}\right)\right]$. The mean is chosen to be $f(\Psi_2^{i-1})$, while I fix the variance to a diagonal matrix with elements 0.001 and 0.005 on the main diagonal.

D.5 Step 5: drawing Ψ_3

Conditional on σ_ε, σ_η, each element of Ψ_3 has an inverse-Gamma posterior distribution, independent of the other element. Conditional on σ_ε, σ_η, it is easy to draw from these inverse-Gamma posteriors because the innovations are observable.[22]

---

[22] See Gelman, Carlin, Stern, and Rubin (1995) for a description of the sampling procedure from an inverse-Gamma or inverse-Wishart distributions.

References

ALBANESI, S., V. V. CHARI, AND L. J. CHRISTIANO (2003): "Expectations Traps and Monetary Policy," mimeo, Northwestern University.

BALL, L., AND N. G. MANKIW (2002): "The NAIRU in Theory and Practice," mimeo, Johns Hopkins University.

BARRO, R. J. (1979): "On the Determination of Public Debt," *Journal of Political Economy*, 87, 940–971.

BARRO, R. J., AND D. B. GORDON (1983): "A Positive Theory of Monetary Policy in a Natural Rate Model," *Journal of Political Economy*, 91(4), 589–610.

BARSKY, R. B. (1987): "The Fisher Hypothesis and the Forecastibility and Persistence of Inflation," *Journal of Monetary Economics*, 19(1), 3–24.

BECK, G. W., AND V. WIELAND (2002): "Learning and Control in a Changing Economic Enviroment," *Journal of Economic Dynamics and Control*, 26, 1359–1377.

BERNANKE, B. S., AND I. MIHOV (1998): "Measuring Monetary Policy," *The Quarterly Journal of Economics*, 113, 869–902.

BLANCHARD, O. J., AND S. FISCHER (1989): *Lectures on Macroeconomics*. MIT Press, Cambridge, Massachusetts.

BLANCHARD, O. J., AND R. PEROTTI (2002): "An Empirical Investigation of the Dynamic Effect of Shocks to Government Spending and Taxes on Output," *The Quarterly Journal of Economics*, 117-4, 1329–1368.

BLINDER, A. S. (1979): *Economic Policy and the Great Stagflation*. Academic Press, New York.

——— (1998): *Central Banking in Theory and Practice*. MIT Press, Camdridge, Massachusetts.

BOIVIN, J., AND M. P. GIANNONI (2002): "Assessing Changes in the Monetary Trasmission Mechanism: A VAR Approach," *Federal Reserve Bank of New York Economic Policy Review*, 8(1), 97–111.

BORDO, M. D. (1993): "The Gold Standard, Bretton Woods and Other Monetary Regimes: An Historical Appraisal," *Federal Reserve Bank of St. Louis Review*, 75(2), 123–191.

BORDO, M. D., AND F. E. KYDLAND (1995): "The Gold Standard as a Rule," *Explorations in Economic History*, 32, 423–464.

BULLARD, J., AND S. EUSEPI (2003): "Did the Great Inflation Occur Despite Policymaker Committment to a Taylor Rule?," mimeo, Federal Reserve Bank of St. Louis.

CALVO, G. (1983): "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 12(3), 383–98.

CARTER, C. K., AND R. KOHN (1994): "On Gibbs Sampling for State Space Models," *Biometrika*, 81(3), 541–553.

CHARI, V. V., L. J. CHRISTIANO, AND M. EICHENBAUM (1998): "Expectations Trap and Discretion," mimeo, Northwestern University.

CHO, I.-K., N. WILLIAMS, AND T. J. SARGENT (2002): "Escaping Nash Inflation," *The Review of Economic Studies*, 69, 1–40.

CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (2005): "Nominal Rigidities and the Dynamic Effect of a Shock to Monetary Policy," *The Journal of Political Economy*, forthcoming.

CHRISTIANO, L. J., AND T. J. FITZGERALD (2003): "Inflation and Monetary Policy in the 20th Century," *Federal Reserve Bank of Chicago Economic Perspectives*, 27, 22–45.

CHRISTIANO, L. J., AND C. J. GUST (2000): "The Expectations Trap Hypothesis," *Federal Reserve Bank of Chicago Economic Perspectives*, 2000(2), 21–39.

CHUNG, H. (1990): "Did Policy Makers Really Believe in the Phillips Curve? An Econometric Test," Ph.D. thesis, University of Minnesota.

CLARIDA, R., J. GALI, AND M. GERTLER (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *The Quarterly Journal of Economics*, 115(1), 147–180.

COGLEY, T., AND T. J. SARGENT (2001): "Evolving Post-World War II U.S. Inflation Dynamics," in *NBER Macroeconomic Annual*, pp. 331–373, Cambridge, Massachusetts. MIT Press.

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

———— (2004): "The Conquest of U.S. Inflation: Learning, Model Uncertainty, and Robustness," mimeo, University of California, Davis.

CUKIERMAN, A., AND F. LIPPI (2002): "Endogenous Monetary Policy with Unobserved Potential Output," mimeo, Tel Aviv University.

DELONG, B. (1997): "America's Only Peacetime Inflation: The 1970s," in *Reducing Inflation*, ed. by Christina, and D. Romer. University of Chicago Press, Chicago.

DENNIS, R. (2001): "The Policy Preferences of the U.S. Federal Reserve," Federal Reserve Bank of San Francisco, Working Paper No. 2001-08.

ESTRELLA, A., AND J. C. FUHRER (2002): "Dynamic Inconsistencies: Counterfactual Implications of a Class of Rational-Expectations Models," *American Economic Review*, 92-4, 1013–1028.

EVANS, G. W., AND S. HONKAPOHJA (2001): *Learning and Expectations in Macroeconomics*. Princeton University Press, Princeton, NJ.

FACKLER, P. L., AND M. J. MIRANDA (2001): "Solving Non-Linear Rational Expectations Models," mimeo, North Carolina State University.

FAVARA, G., AND P. GIORDANI (2002): "Reconsidering the Role of Money for Output, Prices and Interest Rates," SSE/EFI Working Paper Series in Economics and Finance No. 514.

FAVERO, C. A., AND R. ROVELLI (2003): "Macroeconomic Stability and the Preferences of the Fed. A Formal Analysis," *Journal of Money, Credit and Banking*, 35(4), 545–556.

FERNANDEZ-VILLAVERDE, J., AND J. RUBIO (2002): "Estimating Nonlinear Dynamic Equilibrium Economies: A Likelihood Approach," mimeo, University of Pennsylvania.

FUHRER, J., AND G. MOORE (1995): "Inflation Persistence," *The Quarterly Journal of Economics*, 110(1), 127–159.

FUHRER, J. C., AND G. D. RUDEBUSCH (2002): "Estimating the Euler Equation for Output," mimeo, Federal Reserve Bank of Boston.

GALI, J., AND M. GERTLER (1999): "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature*, 37(4), 1661–1707.

GELMAN, A., J. B. CARLIN, H. S. STERN, AND D. B. RUBIN (1995): *Bayesian Data Analysis.* Chapman and Hall, London.

GORDON, R. J. (1997): "The Time-Varying NAIRU and its Implications for Economic Policy," *The Journal of Economic Perspectives*, 11(1), 11–32.

————— (1998): "Foundations of the Goldilocks Economy: Supply Shocks and the Time-Varying NAIRU," *Brooking Papers on Economic Activity*, 1998(2), 297–346.

HAMILTON, J. D. (1994): *Time Series Analysis.* Princeton University Press, Princeton, New Jersey.

HANSON, M. (2003): "Varying Monetary Policy Regimes: A Vector Autoregressive Investigation," mimeo, Wesleyan University.

IRELAND, P. N. (1999): "Does the Time-Consistency Problem Explain the Behavior of Inflation in the United States?," *Journal of Monetary Economics*, 44(2), 279–91.

JUDD, J. P., AND G. D. RUDEBUSCH (1998): "Taylor's Rule and the Fed: 1970-1997," *Federal Reserve Bank of San Francisco Economic Review*, 98(3), 3–16.

JUDD, K. L. (1998): *Numerical Methods in Economics.* The MIT Press, Cambridge, MA.

KIM, C.-J., AND C. R. NELSON (1999a): "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle," *The Review of Economics and Statistics*, 81, 608–616.

————— (1999b): *State-Space Models with Regime Switching: Classical and Gibbs Sampling Approaches with Applications.* M.I.T. Press, Cambridge, Massachussetts.

KIM, S., N. SHEPHARD, AND S. CHIB (1998): "Sthochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *The Review of Economic Studies*, 65(3), 361–393.

KING, R. G., J. H. STOCK, AND M. W. WATSON (1995): "Temporal Instability of the Unemployment-Inflation Relationship," *Economic Perspectives of the Federal Reserve Bank of Chicago*, pp. 2–12.

KING, R. G., AND M. W. WATSON (1994): "The Postwar US Phillips Curve: A Revisionist Econometric History," *Carnegie-Rochester Conference Series on Public Policy*, 41, 157–219.

KYDLAND, F. E., AND E. C. PRESCOTT (1977): "Rules Rather Than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, 85(3), 473–492.

LANSING, K. J. (2002): "Learning About a Shift in Trend Output: Implications for Monetary Policy and Inflation," Federal Reserve Bank of San Francisco Working Paper 00-16.

LEEPER, E., AND T. ZHA (2002): "Modest Policy Interventions," NBER working paper no. 9192.

LEEPER, E. M., AND J. E. ROUSH (2003): "Putting 'M' Back in Monetary Policy," mimeo, University of Indiana.

LUBIK, T. A., AND F. SCHORFHEIDE (2004): "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *America Economic Review*, 94(1), 190–217.

LUCAS, R. E. (1972): "Expectations and the Neutrality of Money," *Journal of Economic Theory*, 4, 103–124.

MAYER, T. (1999): *Monetary Policy and the Great Inflation in the United State: The Federal Reserve and the Failure of Macroeconomic Policy, 1965-1979.* Edward Elgar, Cheltenam, UK.

MCCONNELL, M. M., AND G. PEREZ-QUIROS (2000): "Output Fluctuations in the United States: What Has Changed Since the Early 1980's," *American Economic Review*, 90(5), 1464–1476.

MOUNTFORD, A., AND H. UHLIG (2002): "What are the Effects of Fiscal Policy Shocks?," Tilburg University, Center for Economic Research Discussion Paper 2002-31.

OKUN, A. M. (1978): "Efficient Disinflationary Policies," *American Economic Review*, 68(2), 348–352.

ORPHANIDES, A. (2000): "The Quest for Prosperity Without Inflation," European Central Bank Working Paper No. 15.

————— (2002): "Monetary Policy Rules and the Great Inflation," *American Economic Review*, 92(2), 115–120.

ORPHANIDES, A., AND S. VAN NORDEN (2001): "The Unreliability of Output Gap Estimates in Real Time," CIRANO working paper 2001s-57.

ORPHANIDES, A., AND J. C. WILLIAMS (2002): "Robust Monetary Policy Rules with Unknown Natural Rates," *Brooking Papers on Economic Activity*, 2002(2), 63–145.

————— (2003): "The Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning, and Expectations," Federal Reserve Bank of San Francisco Working Paper 03-24.

PRIMICERI, G. E. (2004): "The Effect of Stabilization Policy on US Postwar Business Cycle Fluctuations," Ph.D. thesis, Princeton University.

————— (2005): "Time Varying Structural Vector Autoregressions and Monetary Policy," *The Review of Economic Studies*, forthcoming.

REIS, R. (2003): "Where is the Natural Rate? Rational Policy Mistakes and Persistent Deviations of Inflation from Target," *Advances in Macroeconomics*, 3(1), Article 1.

ROGOFF, K. (2003): "Globalization and Global Disinflation," mimeo, International Monetary Fund.

ROMER, C., AND D. ROMER (2002): "The Evolution of Economic Understanding and Postwar Stabilization Policy," in *Rethinking Stabilization Policy*. Federal Reserve Bank of Kansas City.

RUDEBUSCH, G. D. (2001): "Is the Fed Too Timid? Monetary Policy in an Uncertain World," *The Review of Economics and Statistics*, 83(2), 203–217.

RUDEBUSCH, G. D., AND L. E. O. SVENSSON (1999): "Policy Rules for Inflation Targeting," in *Monetary Policy Rules*, ed. by J. B. Taylor, pp. 203–246. University of Chicago Press.

SARGENT, T. J. (1973): "Rational Expectations, the Real Rate of Interest, and the Natural Rate of Unemployment," in *Brooking Papers on Economic Activity*, pp. 429–472.

――― (1999): *The Conquest of American Inflation*. Princeton University Press, Princeton, New Jersey.

SARGENT, T. J., N. WILLIAMS, AND T. ZHA (2004): "Schock and Government Beliefs: The Rise and Fall of American Inflation," mimeo, New York University.

SCHORFHEIDE, F. (2003): "Learning and Monetary Policy Shifts," mimeo, University of Pennsylvania.

SIMS, C. A. (1988): "Projecting Policy Effects with Statistical Models," *Revista de Analisis Economico*, 1988(3), 3–20.

――― (2001a): "Solving Linear Rational Expectations Models," *Journal of Computational Economics*, 20(1-2), 1–20.

――― (2001b): "Stability and Instability in US Monetary Policy Behavior," mimeo, Princeton University.

――― (2002): "Testing Restrictions and Comparing Models," lecture notes, Princeton University.

SIMS, C. A. (2003): "Probability Models for Monetary Policy Decisions," mimeo, Princeton University.

SIMS, C. A., AND T. ZHA (2004): "Were There Regime Switches in US Monetary Policy?," mimeo, Princeton University.

SMETS, F., AND R. WOUTERS (2003): "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 1(5), 1123–1175.

SODERLIND, P., U. SODERSTROM, AND A. VREDIN (2002): "New-Keynesian Models and Monetary Policy: A Reexamination of the Stylized Facts," Sveriges Riksbank, Working Paper No. 140.

STAIGER, D., J. H. STOCK, AND M. W. WATSON (1997): "The NAIRU, Unemployment and Monetary Policy," *The Journal of Economic Perspectives*, 11(1), 33–49.

――― (2001): "Prices, Wages and the U.S. NAIRU in the 1990s," mimeo, Princeton University.

STOCK, J. H., AND M. W. WATSON (2002): "Has the Business Cycle Changed and Why?," NBER working paper no. 9127.

――― (2003): "Understanding Changes in International Business Cycle Dynamics," mimeo, Princeton University.

TAMBALOTTI, A. (2003): "Inflation, Productivity and Monetary Policy: From the Great Stagflation to the New Economy," mimeo, Federal Reserve Bank of New York.

TAYLOR, J. B. (1999): "A Historical Analysis of Monetary Policy Rules," in *Monetary Policy Rules*, ed. by J. B. Taylor, pp. 319–341. University of Chicago Press.

WIELAND, V. (2000a): "Learning by Doing and the Value of Optimal Experimentation," *Journal of Economic Dynamics and Control*, 24(4), 501–534.

$\qquad$ (2000b): "Monetary Policy, Parameter Uncertainty and Optimal Learning," *Journal of Monetary Economics*, 46(1), 199–228.

WILLIAMS, N. (2003): "Escape Dynamics in Learning Models," mimeo, Princeton University.

WOODFORD, M. (2003): "Optimal Interest-Rate Smoothing," *The Review of Economic Studies*, 70(4), 861–886.

Coefficients	CG	OLS	DLS	SV
α_1	0.707	0.711	0.707	0.663
	(0.074)	(0.074)	(0.073)	(0.076)
θ_1	-1.053	-1.021	-1.057	$-.718$
	(0.292)	(0.290)	(0.291)	(0.244)
θ_2	0.928	0.903	0.943	0.63
	(0.289)	(0.287)	(0.289)	(0.241)
ρ_1	1.661	1.756	1.640	1.376
	(0.057)	(0.065)	(0.056)	(0.083)
ρ_2	-0.737	-0.779	-0.719	-0.449
	(0.057)	(0.060)	(0.056)	(0.082)
σ_ε^2	1.033	1.041	1.033	$-$
	(0.113)	(0.113)	(0.109)	
σ_η^2	0.036	0.036	0.035	$-$
	(0.006)	(0.006)	(0.006)	
ϕ	2131	475.5	1902	2763
	(1570)	(273.9)	(1119)	(2375)
k	0.872	0.960	0.809	0.869
	(0.026)	(0.014)	(0.032)	(0.026)

Table 1: Maximum likelihood estimates of the model's parameters. CG: model with constant gain learning; DLS: model with discounted least squares learning; OLS: model with ordinary least squares learning; SV: model with stochastic volatility. Standard errors in parentheses.

Models	log-Likelihood	BIC	log-Marginal Likelihood
CG	−250.16	−273.30	−273.66
OLS	−252.02	−275.16	−268.22
DLS	−248.99	−272.13	−263.68
VAR(2)	−242.74	−276.17	−278.70
VAR(3)	−234.52	−278.22	−279.79
VAR(4)	−229.90	−283.89	−279.98

Table 2: Measures of fit of different models. CG: model with constant gain learning; DLS: model with discounted least squares learning; OLS: model with ordinary least squares learning; VAR: vector autoregressions with 2, 3 of 4 lags.

Parameters	CG	FL_1	FL_2
$\max(\pi)$	11.95	14.71	13.90
	[10.52; 13.44]	[12.87; 16.45]	[12.60; 15.27]
$\max(u)$	10.60	11.16	10.16
	[9.31; 12.25]	[9.04; 13.42]	[8.89; 11.86]
t_π^*	1975:III	1975:II	1968
	[1971; 1981]	[1970:I; 1985:I]	[1966; 1975:II]
$t_u^* - t_\pi^*$	2:III	4:II	6:I
	[2; 6:I]	[2:II; 8:I]	[4:I; 9:III]

Table 3: Summary statistics of the simulations (medians, lower and upper quartiles). CG: baseline model, with constant gain learning; FL_1: forward looking model with partially rational agents; FL_2: forward looking model with fully rational agents (t_π^* and t_u^* stand for the time periods in which inflation and unemployment peak).

Coefficients	SP	SP (no learning)
α_1	0.709	0.718
	(0.073)	(0.074)
θ_1	-1.035	-1.102
	(0.295)	(0.285)
θ_2	0.910	1.013
	(0.291)	(0.278)
ρ_1	1.635	1.568
	(0.062)	(0.066)
ρ_2	-0.724	-0.699
	(0.055)	(0.062)
σ_ε^2	1.037	1.039
	(0.111)	(0.114)
σ_η^2	0.034	0.030
	(0.006)	(0.006)
ϕ	950.6	2764
	(583.4)	(16871)
k_1	1.000	0.001
	(0.000)	(0.001)
k_2	1.000	0.000
	(0.007)	(0.054)
π_1^*	4.182	3.598
	(0.739)	(2.020)
π_2^*	3.535	2.297
	(0.345)	(1.138)
λ_1	0.847	1.706
	(0.566)	(2.205)
λ_2	0.420	0.324
	(0.270)	(1.109)
log-Likelihood	-245.99	-241.94
BIC	-281.98	-277.94

Table 4: Maximum likelihood estimates of the model's parameters. SP: model with constant gain learning and shift in the parameters representing the policy preferences; SP (no learning): model with shift in the parameters representing the policy preferences and without learning on the persistence of inflation and the slope of the Phillips curve. Standard errors in parentheses. See section 4.4 for details.

θ	ρ_b	ρ_f	σ_ε^2	σ_η^2	ϕ	k	log-lh	BIC
-0.00095	0.936	0.0000	1.18	0.0833	1972	0.880	-307.53	-325.52
(0.00081)	(0.0237)	(0.0025)	(0.1739)	(0.0152)	(1396)	(0.034)		

Table 5: Maximum likelihood estimates, log-likelihood and BIC for the forward looking model with partially rational agents. Standard errors in parentheses.
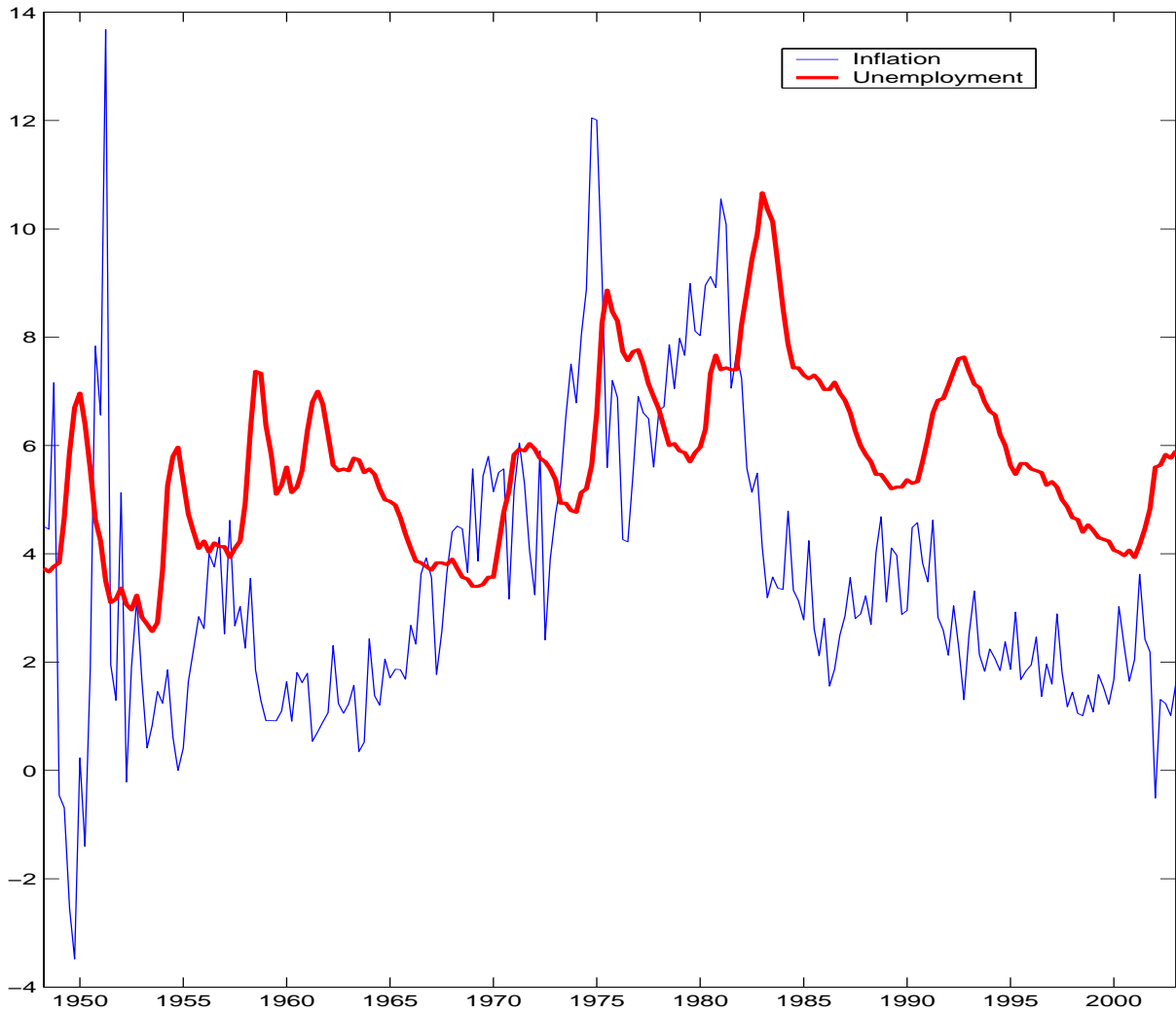
Figure 1: US inflation and unemployment.
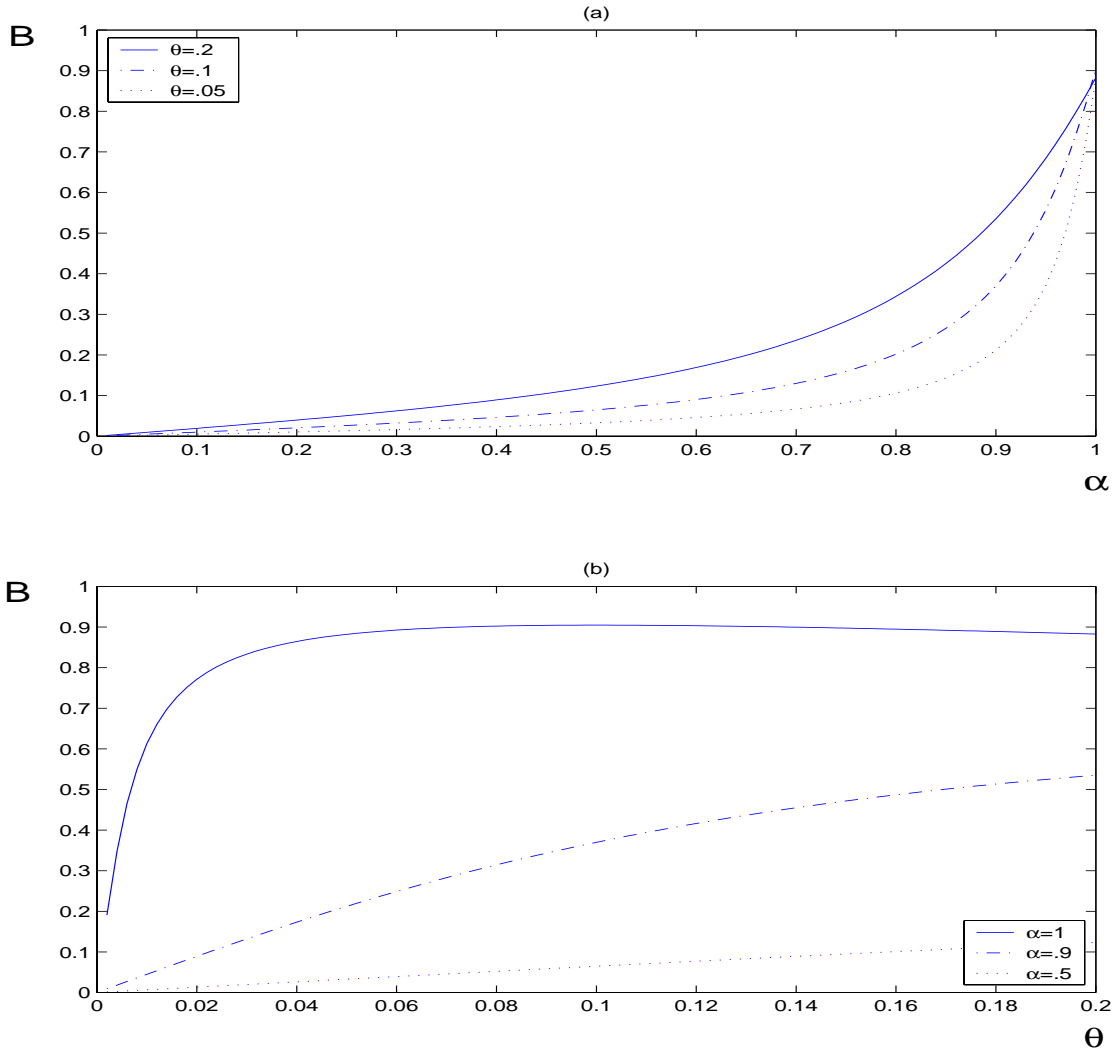
Figure 2: Strength of the policy reaction to inflation as a function of (a) estimated persistence of inflation in the Phillips curve and (b) estimated slope of the Phillips curve.
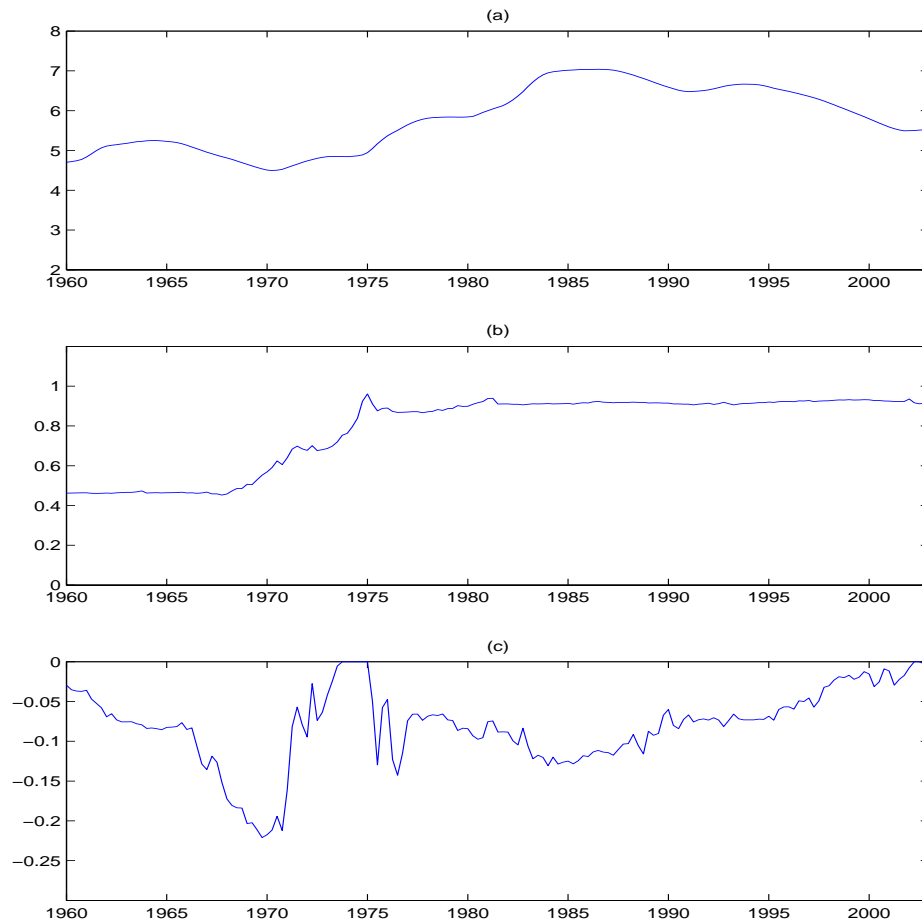
Figure 3: Evolution of policymakers' beliefs about: (a) the natural rate of unemployment; (b) the persistence of inflation in the Phillips curve; (c) the slope of the Phillips curve.
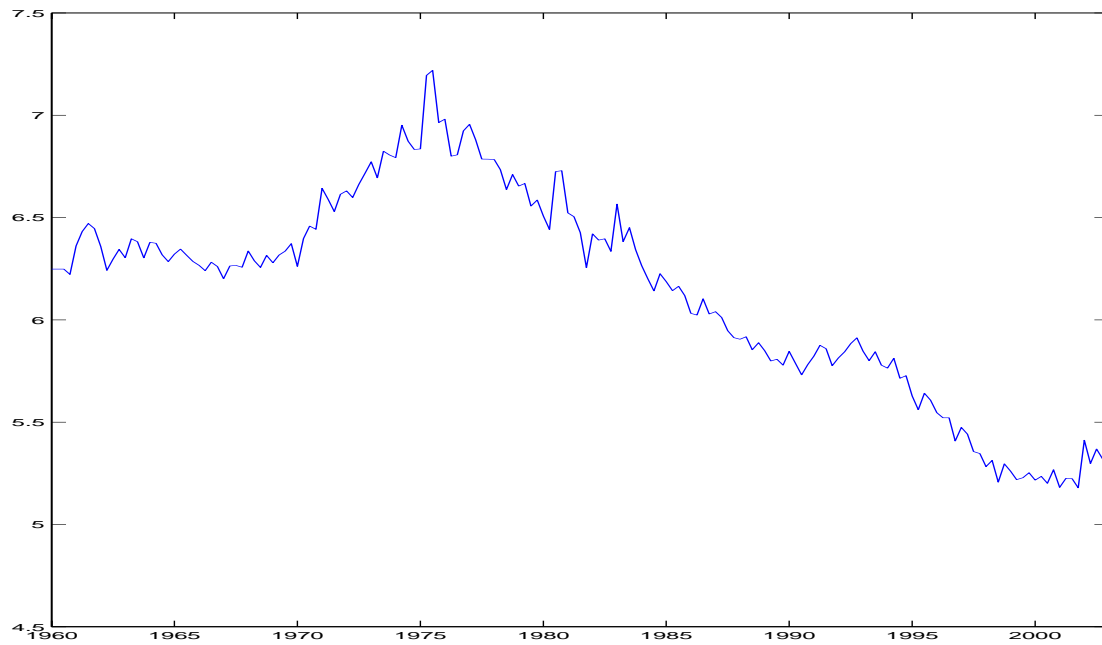
Figure 4: Smoothed estimates of the natural rate of unemployment.
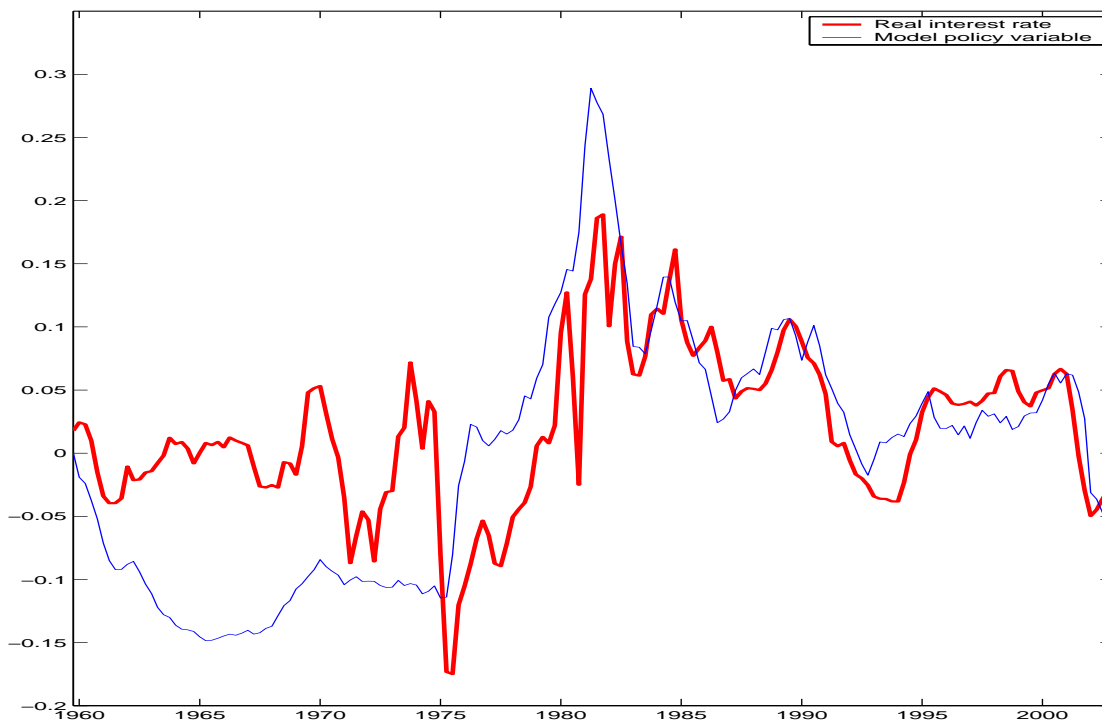
Figure 5: Models' policy variable and real rate of interest (rescaled).

53
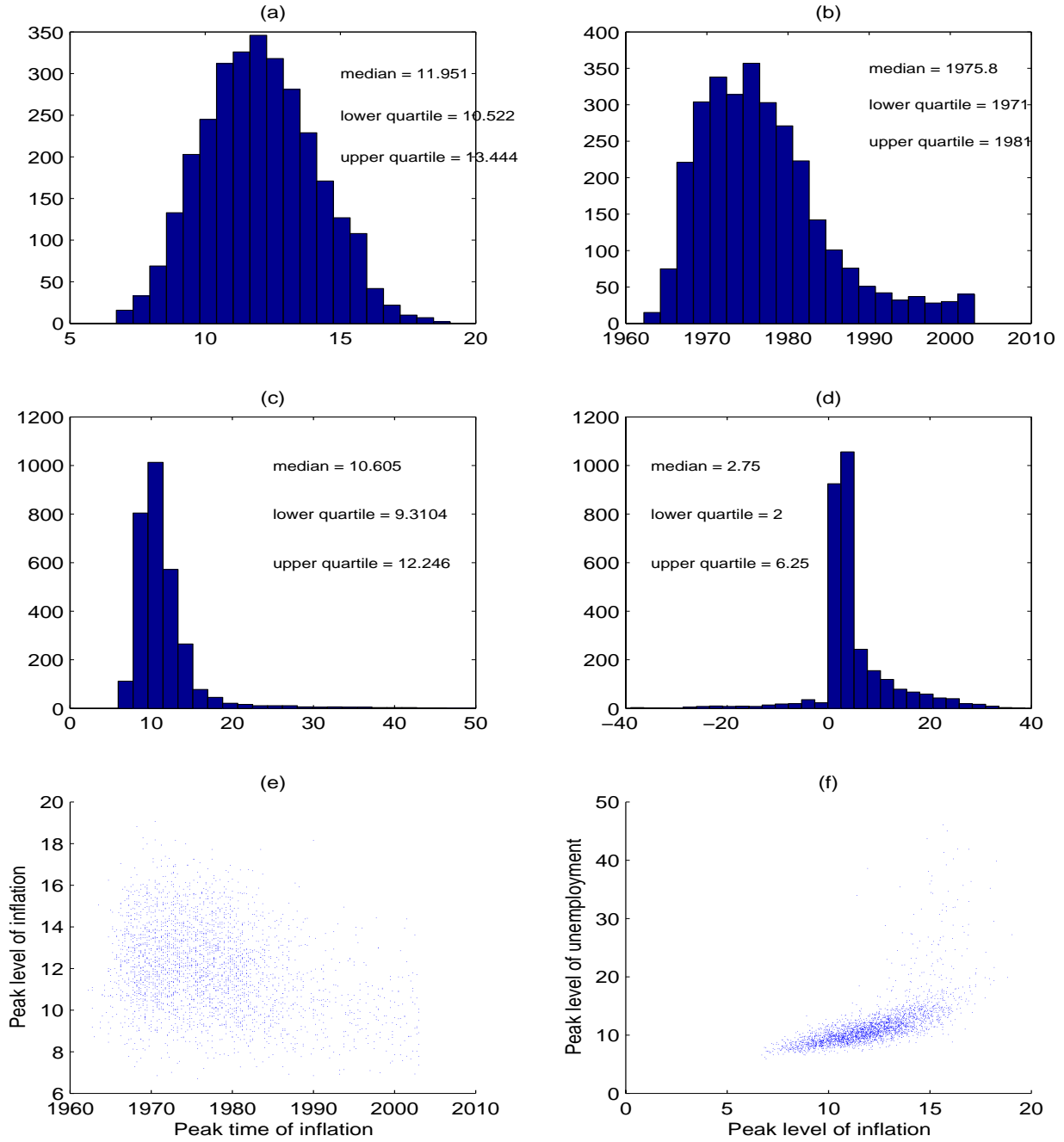
Figure 6: Simulation results for the benchmark model. Empirical distribution of: (a) $\max(\pi)$, (b) t_π^*, (c) $\max(u)$ and (d) $t_u^* - t_\pi^*$. Scatter plot of the relation between (e) t_π^* and $\max(\pi)$ and (f) $\max(\pi)$ and $\max(u)$. ($\max(\pi)$ and t_π^* stand respectively for the peak level and the peak time of inflation, while $\max(u)$ and t_u^* stand for the peak level and the peak time of unemployment).
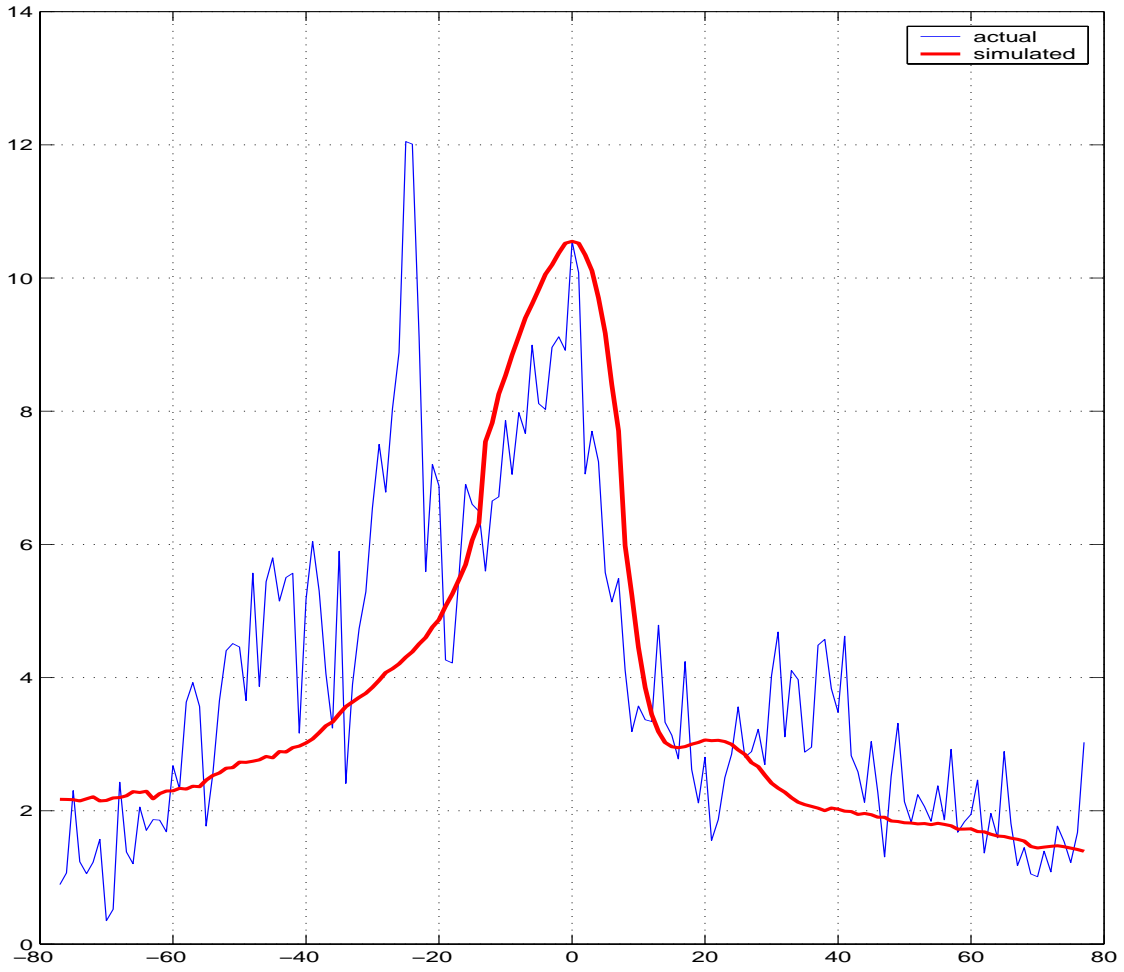
Figure 7: Actual data and average of the simulated inflation paths around the peak time (time expressed in quarters on the horizontal axis).
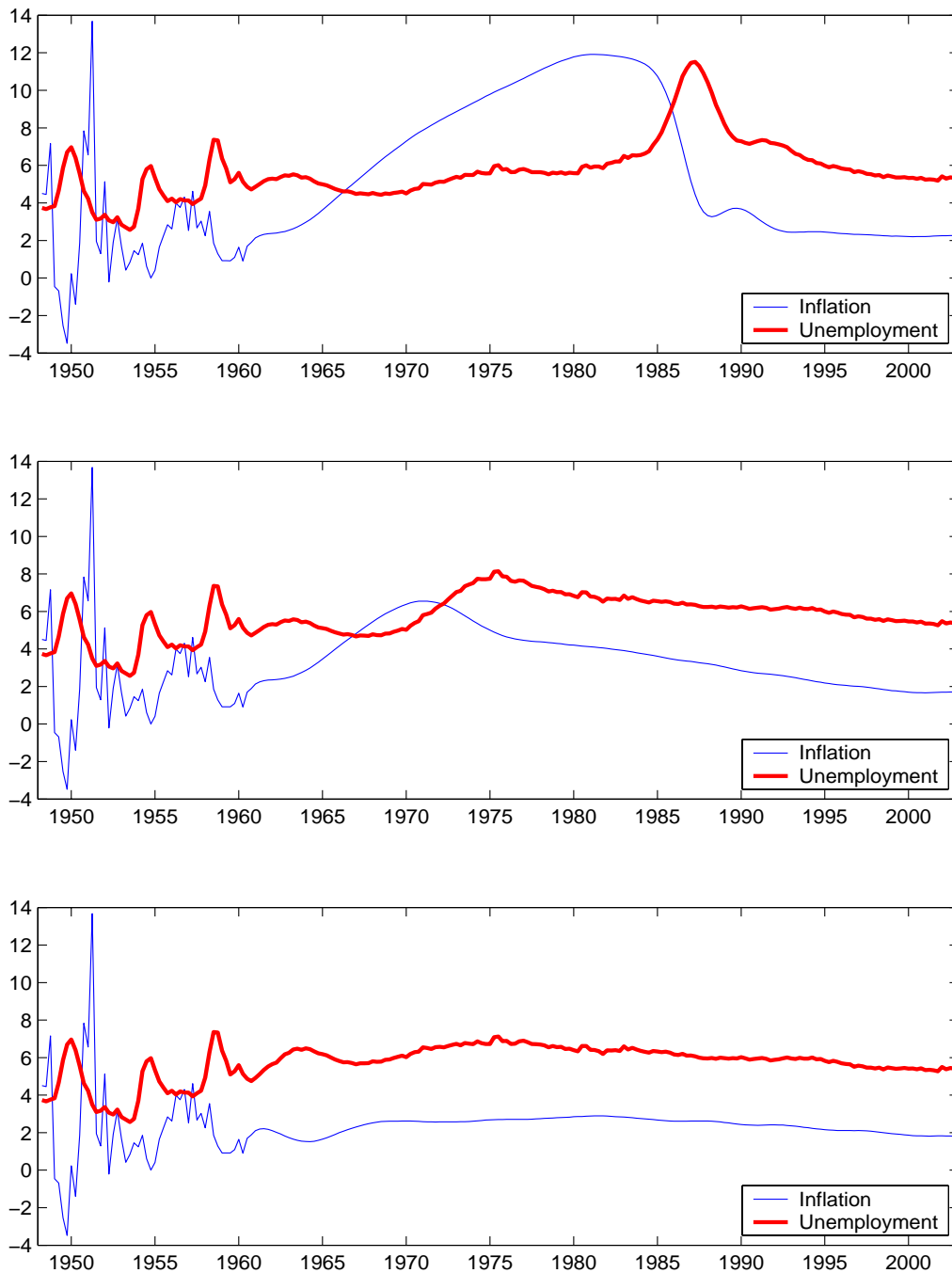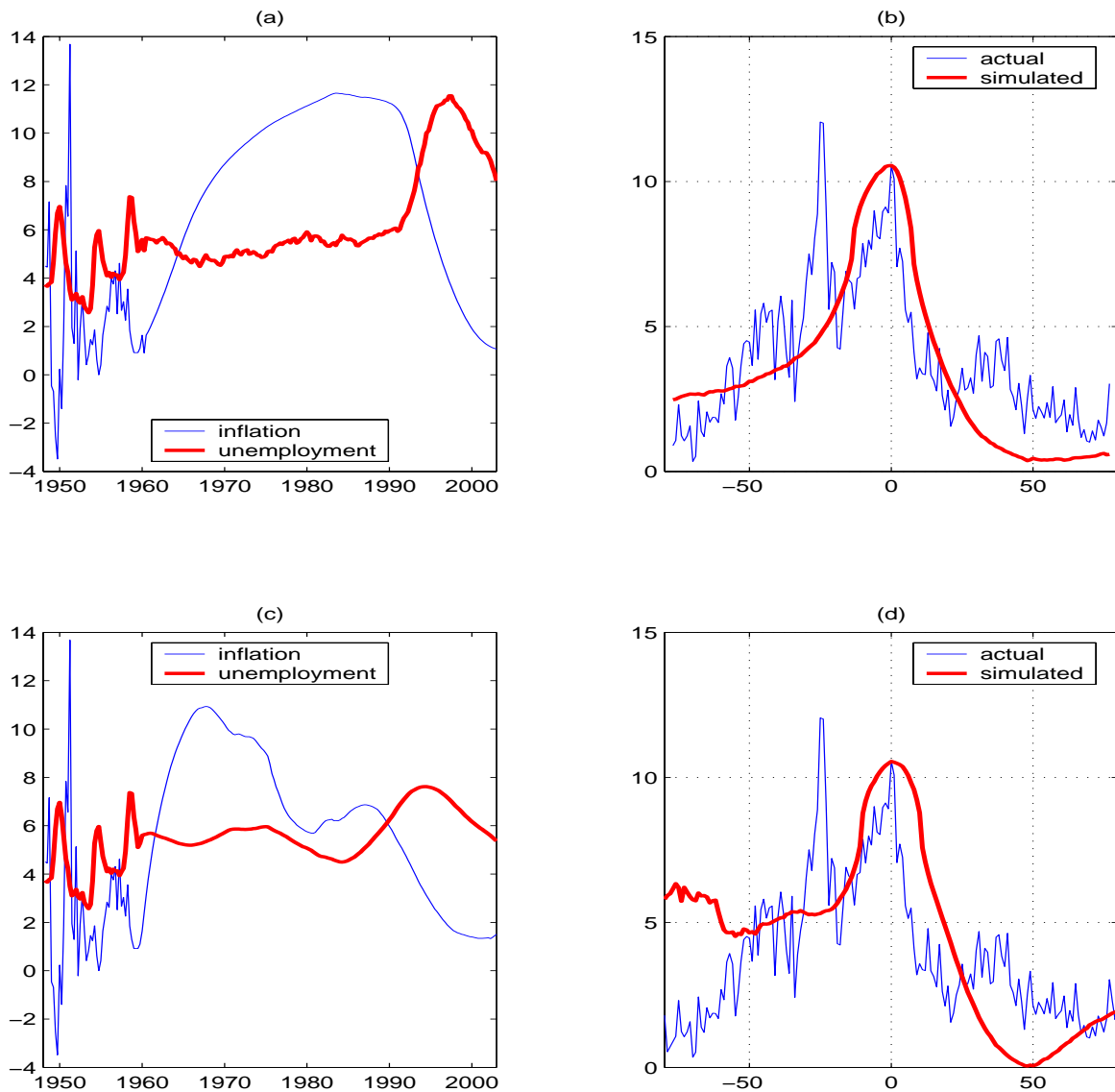
Figure 8: Simulation of the inflation and unemployment behavior under a scenario of low volatility of the exogenous disturbances. (a) Simulation for the baseline constant gain learning model; (b) counterfactual simulation under the assumption that policymakers know the slope of the Phillips curve; (c) counterfactual simulation under the assumption that policymakers know all the parameters of the Phillips curve except for the natural rate of unemployment.

Figure 9: Simulation of inflation and unemployment under a scenario of low volatility of the exogenous disturbances in (a) the partially rational agents model of section 5.1 and (c) the fully rational agents model of section 5.2. Average of all simulated inflation paths around the peak under realistic volatility of the exogenous disturbances in (b) the partially rational agents model and (d) the fully rational agents model.
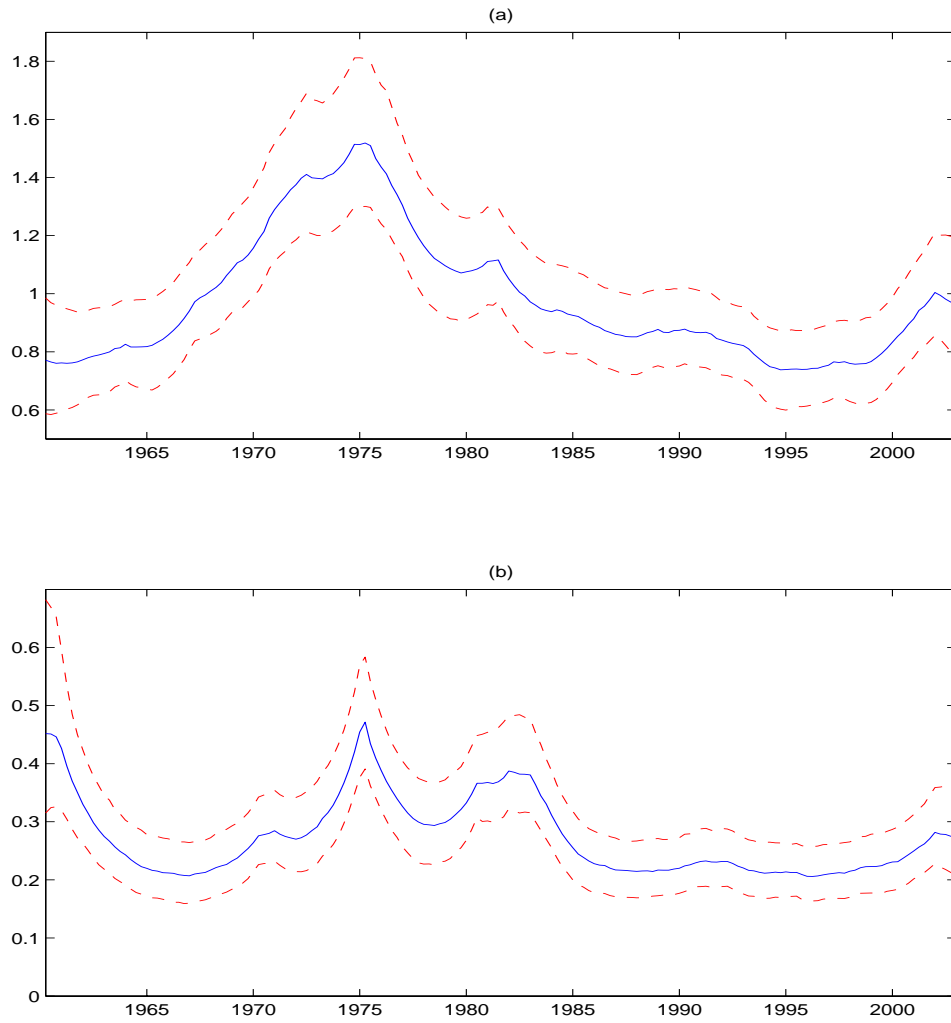
Figure 10: Posterior median and 68% error bands for the time varying standard deviations of (a) shocks to the Phillips curve and (b) shocks to the aggregate demand.
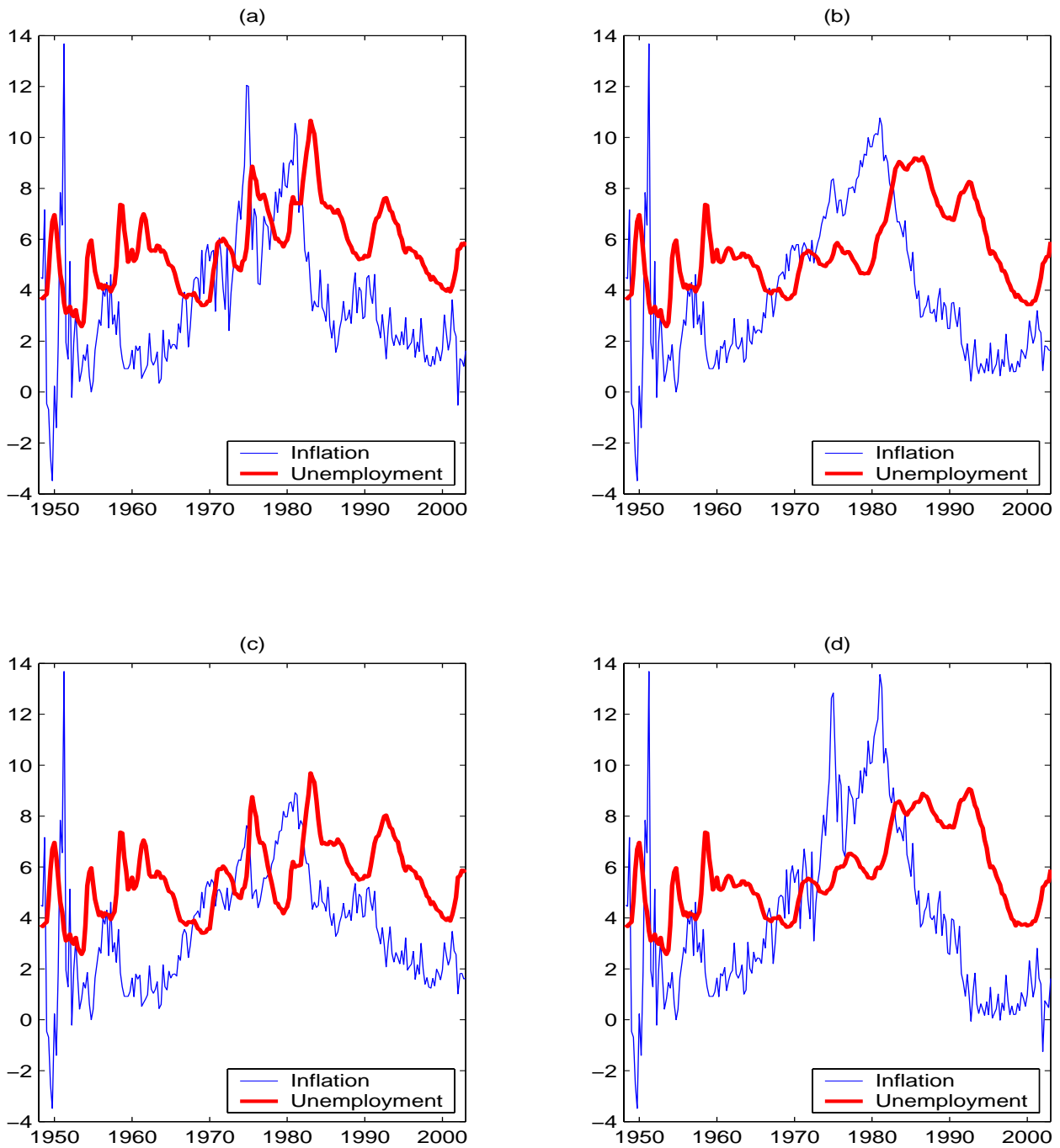
Figure 11: (a) Actual data for inflation and unemployment and counterfactual simulations when (b) the standard deviations of the shocks to the Phillips curve and the aggregate demand are fixed to their 1995 value; (c) only the standard deviation of the shocks to the Phillips curve is fixed to its 1995 value; (d) only the standard deviation of the shocks to the aggrgate demand is fixed to its 1995 value.
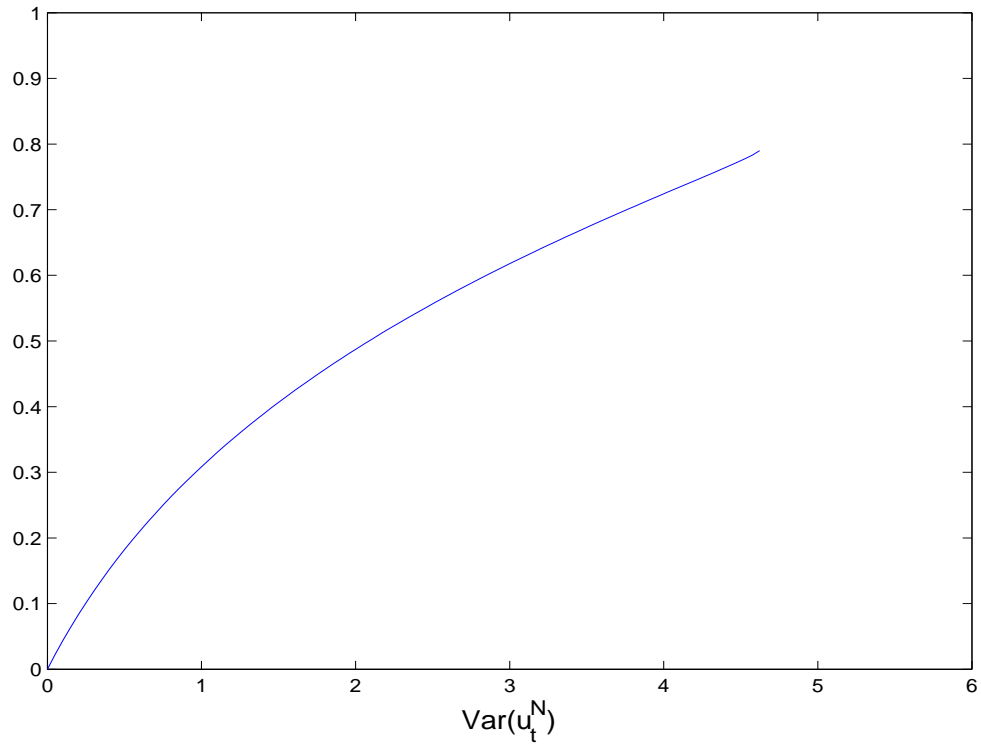
Figure 12: Euclidean norm of the distance between the true value of the coefficients and the value of policymakers beliefs in a self-confirming equilibrium (as a function of the variance of the non-inflationary rate of unemployment).