

NBER WORKING PAPER SERIES

SHOCKS AND GOVERNMENT BELIEFS:
THE RISE AND FALL OF AMERICAN INFLATION

Thomas Sargent
Noah Williams
Tao Zha

Working Paper 10764
<http://www.nber.org/papers/w10764>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2004

We thank seminar participants at the NBER Summer Institute, Eric Leeper, James Nason, Ricardo Reis, William Roberds, Frank Schorfheide, Christopher Sims, and Lars Svensson for helpful discussions and Jordi Mondria for excellent research assistance. The views expressed herein do not necessarily reflect those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Sargent and Williams thank the National Science Foundation for support. The views expressed herein are those of the author(s) and not necessarily those of the National Bureau of Economic Research.

©2004 by Thomas Sargent, Noah Williams, and Tao Zha. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Shocks and Government Beliefs: The Rise and Fall of American Inflation
Thomas Sargent, Noah Williams, and Tao Zha
NBER Working Paper No. 10764
September 2004
JEL No. E5

ABSTRACT

We use a Bayesian Markov Chain Monte Carlo algorithm to estimate a model that allows temporary gaps between a true expectational Phillips curve and the monetary authority's approximating non-expectational Phillips curve. A dynamic programming problem implies that the monetary authority's inflation target evolves as its estimated Phillips curve moves. Our estimates attribute the rise and fall of post WWII inflation in the US to an intricate interaction between the monetary authority's beliefs and economic shocks. Shocks in the 1970s altered the monetary authority's estimates and made it misperceive the tradeoff between inflation and unemployment. That caused a sharp rise in inflation in the 1970s. Our estimates say that policymakers updated their beliefs continuously. By the 1980s, their beliefs about the Phillips curve had changed enough to account for Volcker's conquest of US inflation in the early 1980s.

Thomas Sargent
Department of Economics
New York University
269 Mercer Street, 7th Floor
New York, NY 10003
and NBER
thomas.sargent@nyu.edu

Noah Williams
Department of Economics
Princeton University
001 Fisher Hall
Princeton, NJ 08544-1021
and NBER
noahw@princeton.edu

Tao Zha
Research Department
Federal Reserve Bank of Atlanta
1000 Peachtree Street N.E.
Atlanta, GA 30309-4470
tao.zha@atl.frb.org

I. INTRODUCTION

Today, many statesmen and macroeconomists believe that inflation can largely be determined by a government monetary authority. Then why did the Federal Reserve Board preside over high US inflation during the late 1960s and the 1970s? And why, under Paul Volcker, did it rapidly bring inflation down during the early 1980s? This paper answers these questions by estimating a model that features a particular process that makes a procession of economic shocks induce the monetary authority to alter its model of inflation-unemployment dynamics, the Phillips curve. At each date t , the monetary authority updates its beliefs about the Phillips curve and then recomputes a first-period action recommended by a “Phelps problem”, a discounted dynamic programming problem that minimizes the expected value of a discounted quadratic loss function of inflation and unemployment.¹ The monetary authority pursues the same objectives at each date, using the same structural model, with only its estimates of that model changing over time.² This model of the systematic part of inflation puts the monetary authority’s beliefs about the Phillips curve front and center.³

We assume that the monetary authority’s model of the Phillips curve deviates in two subtle but important ways from what it would be in a rational expectations model (e.g., Kydland and Prescott (1977)). The first deviation is that, while the true Phillips curve is like Kydland and Prescott’s, we assume that the monetary authority omits the public’s expected rate of inflation from its Phillips curve. By itself, this omission need not prevent the outcomes of our model from coinciding with those of Kydland and Prescott’s, nor need it imply that the government’s model is wrong in a way that could be detected from even

¹Sargent (1999) contributed the nomenclature Phelps problem.

²There is some debate about whether policy objectives or the structural models used by policymakers have evolved over time. However introducing such an evolution of understanding into formal models is difficult without arbitrarily imposing exogenous changes. We need no such exogenous shifts.

³As does Kydland and Prescott’s (1977) model of time-consistent suboptimal inflation.

an infinite sample. Whether the monetary authority's model is wrong in a statistically detectable way depends on how we allow the monetary authority to reestimate the parameters of its model. In particular, if the monetary authority were to believe that the coefficients of its Phillips curve are constant over time, then its estimates would converge to ones that support a self-confirming equilibrium (SCE). After convergence, its estimated Phillips curve would correctly describe occurrences along the SCE path for inflation and unemployment. Such an after-convergence version of our model has little hope of explaining the rise and fall of US inflation: that model would have inflation fluctuating randomly around a constant SCE level that coincides with Kydland and Prescott's time consistent suboptimal (i.e., excessive) level.⁴

This outcome motivates our second subtle deviation from a rational expectations equilibrium. Instead of thinking that the regression coefficients in its Phillips curve are time invariant (which they indeed *are* in an SCE), our monetary authority believes that they form a vector random walk with innovation covariance matrix V . Given that model, the monetary authority updates its beliefs using Bayes' rule. The covariance matrix V and the initial condition for the regression parameters in the monetary authority's Phillips curve become the hyperparameters of a model that shapes evolution of the monetary authority's beliefs.⁵ After calibrating the initial condition and imposing that the systematic part of inflation is determined by the time t solution of the Phelps problem, we estimate V along with parameters of the *true* expectational Phillips curve that, unbeknownst to the monetary authority, truly describes inflation-unemployment dynamics. Using a Bayesian Markov Chain Monte Carlo (MCMC) algorithm, we report statistics that describe the posterior distribution of these parameters of our model. We obtain a much better explanation of the

⁴Michael Parkin (1993) and Peter Ireland (1999) advocate the hypothesis that the post WWII US inflation data can be accounted for by well understood medium term movements in the natural rate of unemployment, stable government preferences, and a steady adherence to the time-consistent suboptimal equilibrium of Kydland and Prescott (1977).

⁵As would be true in a rational expectations model, the monetary authority's beliefs are outcomes, not free parameters.

monetary authority's inflation choices than earlier efforts to estimate similar models had achieved.

With very particular a priori settings of the parameter innovation covariance matrix V , Sims (1988), Chung (1990), Sargent (1999), and Cho, Williams, and Sargent (2002) all studied versions of our model.⁶ When Chung and Sargent estimated their a-priori-fixed- V versions of our model, they obtained discouraging results. They did not come close to explaining the rise and fall of US inflation in terms of a process of the monetary authority's learning about its Phillips curve.⁷

This paper estimates settings for V that attain substantial improvements in the model's ability to rationalize the choices made by the US monetary authority. The MCMC algorithm finds values for V that allow the model to reverse engineer from the data a sequence of government beliefs about the Phillips curve that, through the intermediation of the Phelps problem, capture both the acceleration of US inflation in the 1970s and its rapid decline in the early 1980s. Our MCMC method estimates a V that accommodates an avenue by which economic shocks impinge on the monetary authority's beliefs, via its use of Bayes' rule, and its decisions, via successive solutions of its Phelps problem. The monetary authority's views about parameter drift and its application of Bayes' rule add a source of history dependence to its procession of decisions that is absent, for example, in either Sargent's SCE or the Markov perfect equilibrium of Kydland and Prescott's model. The resulting interactions of shocks and monetary beliefs forms the basis for our explanation of the rise and fall of US inflation.

The rest of the paper is organized as follows. In Section II, we lay out the model and discuss the theoretical characterizations of it. In Section III, we develop a method for the

⁶Sargent and Williams (2003) is an extensive theoretical study of a version of our model that focuses on the impact of different settings of V on rates of convergence to, escapes from, and cycles around an SCE.

⁷Previous failures to match the data with a model like ours seem to be widely recognized and helped to promote a literature that makes the "stickiness" (or persistence) of inflation exogenous.

estimation and inference of the model. Section IV presents our empirical results, examines the performance of the model, and explores implications. Section V conducts some counterfactual exercises. Section VI relates our findings to some other literature. Finally, Section VII concludes. Six appendices describe the data, provide technical details about the setting of our prior distribution for estimation and the sampling scheme for inference, compare key forecasting features of our learning model to those of alternative statistical models, and report statistical inference on the estimates of the government's inflation policy and perceived tradeoffs.

II. THE MODEL

The model is an extension of Sargent and Williams (2003), which is composed of a Lucas-Sargent natural-rate version of the Phillips curve and a true inflation process:

$$u_t - u^* = \theta_0(\pi_t - E_{t-1}\pi_t) + \theta_1(\pi_{t-1} - E_{t-2}\pi_{t-1}) + \tau_1(u_{t-1} - u^*) + \sigma_1 w_{1t}, \quad (1)$$

$$\pi_t = x_{t-1} + \sigma_2 w_{2t}, \quad (2)$$

where u_t is the unemployment rate, π_t is inflation, x_t is the part of inflation controllable by the government given the information up to time t , and w_{1t} and w_{2t} are i.i.d. uncorrelated standard normal random variables. Equation (1) is an expectations-augmented Phillips curve. If $\text{abs}(\theta_0) > \text{abs}(\theta_1)$, (1) becomes a version of a natural-rate Phillips curve that allows a serially correlated disturbance (Sargent 1999). Equation (2) states that the government sets policy to influence inflation up to a random shock. The government's policy is to minimize a loss function that dislikes both inflation and unemployment. The decision x_{t-1} solves the ‘‘Phelps problem’’:

$$\min_{x_{t-1}} \hat{E} \sum_{t=1}^{\infty} \delta^t ((\pi_t - \pi^*)^2 + \lambda(u_t - u^{**})^2) \quad (3)$$

subject to (2) and

$$u_t = \hat{\alpha}'_{t|t-1} \Phi_t + \sigma w_t, \quad (4)$$

where π^* and u^{**} are the targeted levels of inflation and unemployment, both $\hat{\alpha}_{t|t-1}$ and Φ_t are $r \times 1$ vectors, w_t is an i.i.d. standard normal random variable, and where (4) is the core of the monetary authority's model of inflation-unemployment dynamics. The vector Φ_t of regressors consists of lags of unemployment and inflation. By comparing (4) with (1), we see that the government fails to account explicitly for the role of expectations in determining the unemployment rate. Here \hat{E} represents expectations with respect to the government's subjective model, and the subscript $t - 1$ means that the government updates $\hat{\alpha}_{t|t-1}$ and at each t computes x_{t-1} by solving the time t Phelps problem before observing π_t and u_t . Thus, the government sets policy based on its misspecified Phillips curve (4), not the true Phillips curve (1). A self-confirming equilibrium (SCE) is a vector of government beliefs $\bar{\alpha}$ that is consistent with what it observes in the sense of satisfying the population least squares orthogonality condition:

$$E [\Phi_t(u_t - \Phi_t' \bar{\alpha})] = 0, \quad (5)$$

where the expectation is taken with respect to the probability distribution of u_t , π_t , and x_{t-1} that satisfies (1), (2), and the decision rule when the government optimizes based on $\bar{\alpha}$.

Self-confirming equilibrium outcomes agree with the time-consistent Nash equilibrium outcomes in which policymakers set inflation at a higher level than the socially optimal Ramsey level (see Sargent 1999).⁸ Nash inflation is

$$\pi^{\text{Nash}} = \pi^* - \lambda(u^* - u^{**}) [(1 + \delta\tau_1)\theta_0 + \delta\theta_1]. \quad (6)$$

The larger are $u^* - u^{**}$, θ_0 , and θ_1 in absolute value, the higher is the Nash inflation rate compared to the Ramsey rate π^* .

A self-confirming equilibrium is a population concept that restricts beliefs to be time-invariant and that is a benchmark – and as it sometimes turns out, a limit point – for our model. In our model, unlike an SCE, at any point in history the government updates its beliefs as it learns. In particular, the government bases $\hat{\alpha}_{t|t-1}$, its mean estimate of the

⁸As explained by Sargent (1999, chapter 3), the gap between the Ramsey and Nash or SCE outcomes for inflation reflects the benefit to the government of being able to commit to a policy.

drifting parameter vector α_t , on the observations up to and including time $t - 1$ from the following (misspecified) econometric model:

$$u_t = \alpha_t' \Phi_t + \sigma w_t, \quad (7)$$

$$\alpha_t = \alpha_{t-1} + \Lambda_t, \quad (8)$$

where Λ_t , uncorrelated with w_t , is an i.i.d. Gaussian random vector with mean 0 and covariance matrix V . Thus, the government believes that the true economy drifts over time. That is why it continually adapts its parameter estimates. The innovation covariance matrix V governs the perceived volatility of increments to the parameters, and is a key component of the model. The mean estimate of α_t for the econometric model (7)-(8) is

$$\hat{\alpha}_{t|t-1} \equiv \hat{E}(\alpha_t | \mathcal{I}_{t-1}),$$

$$\mathcal{I}_t \equiv \{u_1, \pi_1, \dots, u_t, \pi_t\}.$$

Let

$$P_{t|t-1} \equiv \widehat{Var}(\alpha_t | \mathcal{I}_{t-1}).$$

Given the government's model, the mean estimates are optimally updated via the special case of Bayes rule known as the Kalman filter. Given $\hat{\alpha}_{1|0}$ and $P_{1|0}$, the Kalman filter algorithm updates $\hat{\alpha}_{t|t-1}$ with the following formula:

$$\hat{\alpha}_{t+1|t} = \hat{\alpha}_{t|t-1} + \frac{P_{t|t-1} \Phi_t (u_t - \Phi_t' \hat{\alpha}_{t|t-1})}{\sigma^2 + \Phi_t' P_{t|t-1} \Phi_t}, \quad (9)$$

$$P_{t+1|t} = P_{t|t-1} - \frac{P_{t|t-1} \Phi_t \Phi_t' P_{t|t-1}}{\sigma^2 + \Phi_t' P_{t|t-1} \Phi_t} + V. \quad (10)$$

Instead of using the Kalman filter, many models of adaptation assume a learning rule known as recursive least squares (RLS), which is closely related to the Kalman filter.

An important issue about any learning rule is whether the learning process will converge to a self-confirming equilibrium. To summarize what we know about this, we scale the innovation covariance matrix as $V = \varepsilon^2 \hat{V}$, for $\varepsilon > 0$. Key analytical results from Sargent and

Williams (2003) that underscore the role of the government's learning from misspecified models are:

- (1) In this model, inflation converges much faster to the SCE under Kalman filtering learning than under RLS. In effect, the Kalman filter learning rule with drifting coefficients seems to discount the past data more rapidly than the constant gain RLS learning rule.
- (2) As the government's prior belief parameter $\varepsilon \rightarrow 0$ (at the 0 limit there is no time variation in the parameters), inflation converges to the self-confirming equilibrium (SCE) and the mean escape time becomes arbitrarily long.
- (3) As the government's prior belief parameter $\sigma \rightarrow 0$ (in the 0 limit, there is no variation in the government's regression error or arbitrarily large time variation in the drifting parameters), large escapes from an SCE can happen arbitrarily often and nonconvergence is possible.
- (4) The covariance matrix V in the government's prior belief about the volatility of the drifting parameters affects the speed of escape. The covariance matrix V combined with the prior belief parameter ε , affects the speed of convergence to the SCE from a low inflation level.

III. ESTIMATION AND INFERENCE

The theoretical results indicate how very different outcomes can emerge from different government beliefs. The overriding task of this paper is to fit the model to the data and thereby to estimate and quantify the uncertainty about the parameters, σ^2 and V , jointly with the model's other structural parameters, including those governing the "true" expectational Phillips curve (1). Before estimation, we fix the values of δ , λ , π^* , u^{**} , and $\hat{\alpha}_{1|0}$. Group all other free structural parameters as

$$\phi = \{v^*, \theta_0, \theta_1, \tau_1, \zeta_1, \zeta_2, u(C_P), u(C_V)\},$$

where $v^* = u^*(1 - \tau_1)$, C_P and C_V are upper triangular such that $P_{1|0} = C_P' C_P$ and $V = C_V' C_V$, and $\zeta_1 = 1/\sigma_1^2$ and $\zeta_2 = 1/\sigma_2^2$ represent the precisions of the corresponding innovations. The notation $u(C_P)$ or $u(C_V)$ means that only the upper triangular part of C_P or C_V are among the free parameters.

The structural parameter $\zeta = 1/\sigma^2$ is not a free parameter. It is clear from (9), (10), and (B1) that if we scale V and $P_{1|0}$ by κ and ζ by $1/\kappa$, the likelihood value remains the same. There would exist a continuum of maximum likelihood estimates (MLEs) if ζ were not restricted (i.e., the model is unidentified). Some normalization is necessary. Following Sargent and Williams, we impose the restriction $\zeta = \zeta_1$. This normalization implies that the variation that policymakers observe in the unemployment rate is correctly decomposed into variation in the regressors and variation due to exogenous shocks. This assumption has an advantage because it makes limiting results easier to derive.⁹

As we've noted, Sargent and Williams (2003) show that whether monetary policy stays close to a path associated with a self-confirming equilibrium, and if not, how it evolves over time are both sensitive to the model's parameters (especially the government's belief about the covariance matrix for the drifting coefficients). No doubt this sensitivity contributes to the outcome that we are able to estimate the key structural parameters sharply.

To take into account the parameter uncertainty explicitly, we employ the Bayesian method and develop a Monte Carlo Markov Chain (MCMC) algorithm that breaks ϕ into three separate blocks: θ , $\{\zeta_1, \zeta_2\}$, and φ where

$$\theta = \begin{bmatrix} v^* \\ \theta_0 \\ \theta_1 \\ \tau_1 \end{bmatrix},$$

⁹Note that an SCE requires the orthogonality conditions, not necessarily the equality restriction $\zeta = \zeta_1$. Indeed, the examples of Sims (1988) allow $\zeta \neq \zeta_1$.

and $\varphi = \{u(C_P), u(C_V)\}$. The prior distributions of both θ and φ take a Gaussian form:

$$p(\theta) = \text{Normal}(\bar{\theta}, \bar{\Sigma}_\theta); \quad (11)$$

$$p(\varphi) = \text{Normal}(\bar{\varphi}, \bar{\Sigma}_\varphi). \quad (12)$$

The prior probability density for the precision parameters ζ_1 and ζ_2 is a Gamma distribution:

$$p(\zeta_1, \zeta_2) = \text{Gamma}(\bar{\alpha}, \bar{\beta}) = \prod_{i=1}^2 \frac{1}{\Gamma(\bar{\alpha}) \bar{\beta}^{\bar{\alpha}}} \zeta_i^{\bar{\alpha}-1} e^{-\frac{\zeta_i}{\bar{\beta}}}. \quad (13)$$

The posterior distribution of ϕ can be simulated by alternately sampling from the conditional posterior distributions (a Gibbs sampler):

$$p(\theta \mid \mathcal{I}_T, \zeta_1, \zeta_2, \varphi),$$

$$p(\zeta_1, \zeta_2 \mid \mathcal{I}_T, \theta, \varphi),$$

$$p(\varphi \mid \mathcal{I}_T, \theta, \zeta_1, \zeta_2).$$

The following two propositions state that the first two conditional posterior distributions are Gaussian. Appendix B offers a proof.

Proposition 1.

$$p(\theta \mid \mathcal{I}_T, \zeta_1, \zeta_2, \varphi) = \text{Normal}(\tilde{\theta}, \tilde{\Sigma}_\theta), \quad (14)$$

where

$$\tilde{\Sigma}_\theta^{-1} = \zeta_1 \sum_{t=1}^T (y_t y_t') + \bar{\Sigma}_\theta^{-1},$$

$$\tilde{\theta} = \tilde{\Sigma}_\theta \left(\zeta_1 \sum_{t=1}^T (u_t y_t) + \bar{\Sigma}_\theta^{-1} \bar{\theta} \right),$$

$$y_t = \begin{bmatrix} 1 & z_{2t} & z_{2t-1} & u_{t-1} \end{bmatrix}',$$

$$z_{1t} = u_t - u^* - \theta_0(\pi_t - E_{t-1}\pi_t) - \theta_1(\pi_{t-1} - E_{t-2}\pi_{t-1}) - \tau_1(u_{t-1} - u^*),$$

$$z_{2t} = \pi_t - x_{t-1}.$$

Proposition 2.

$$p(\zeta_1, \zeta_2 \mid \mathcal{I}_T, \theta, \varphi) = \text{Gamma}(\tilde{\alpha}_{\zeta_1}, \tilde{\beta}_{\zeta_1}) \text{Gamma}(\tilde{\alpha}_{\zeta_2}, \tilde{\beta}_{\zeta_2}), \quad (15)$$

where

$$\begin{aligned} \tilde{\alpha}_{\zeta_1} &= \tilde{\alpha}_{\zeta_2} = \frac{T}{2} + \bar{\alpha}, \\ \tilde{\beta}_{\zeta_i} &= \frac{1}{0.5 \sum_{t=1}^T z_{it}^2 + \bar{\beta}^{-1}}, \quad \forall i \in \{1, 2\}. \end{aligned}$$

The government's optimization problem renders the conditional posterior pdf

$$p(\varphi \mid \mathcal{I}_T, \theta, \zeta_1, \zeta_2)$$

one of nonstandard form. To draw from this distribution, therefore, we use the following Metropolis algorithm.

Metropolis Algorithm. We employ four steps to simulate φ from its conditional posterior distribution.

- (1) Given the value φ^{last} , compute the proposal draw

$$\varphi^{\text{prop}} = \varphi^{\text{last}} + \xi,$$

where ξ is randomly drawn from the normal distribution with mean zero and covariance $\epsilon \tilde{\Sigma}_\varphi$ specified in (D1). The scale factor ϵ will be adjusted to keep the acceptance ratio optimal (around 25% – 40%).

- (2) Compute

$$q = \min \left\{ \frac{p(\varphi^{\text{prop}} \mid \mathcal{I}_T, \theta, \zeta_1, \zeta_2)}{p(\varphi^{\text{last}} \mid \mathcal{I}_T, \theta, \zeta_1, \zeta_2)}, 1 \right\}.$$

- (3) Randomly draw v from the uniform distribution $U(0, 1)$.

- (4) If $v \leq q$, accept φ^{prop} as the value of the current draw; otherwise, keep φ^{last} as the value of the current draw.

It follows from Propositions 1 and 2 and the properties of the Metropolis algorithm that a large number of MCMC samples alternately drawn from these conditional posterior distributions will eventually form an empirical distribution of ϕ that emulates the posterior distribution.¹⁰

IV. EMPIRICAL EVIDENCE

In this section, we present our results. Using the monthly US data described in Appendix A and the prior specified in Appendix C, we estimate ϕ by maximizing the posterior density function. We obtained similar results using maximum likelihood, but the prior is crucial for small sample inference. In estimation, we set $\delta = 0.9936$, $\lambda = 1$, $\pi^* = 2$, and $u^{**} = 1$. The value of u^{**} is set at a low value to allow Nash inflation to be higher than Ramsey inflation.¹¹ Setting the unemployment target closer to the natural rate has no effect on our main results.¹²

The initial belief $\hat{\alpha}_{1|0}$ is set at the regression estimate using the presample data from January 1948 to December 1959.¹³ We tried to fix $P_{1|0}$ at the value that scales up and down the presample regression estimate $\hat{\sigma}^2(\Phi'\Phi)^{-1}$, but the fit was extremely poor. Similarly, letting V be fixed at the presample-estimated covariance matrix with different scales does

¹⁰For each draw of ϕ , ζ is normalized to be equal to ζ_1 before the government's inflation policy is solved. This normalization is consistent with Wald normalization discussed in Hamilton, Waggoner, and Zha (2003).

¹¹Alan Blinder (1998) emphasizes that the source of time inconsistency in Kydland and Prescott's (1977) Phillips curve example is their setting $u^{**} \neq u^*$ in the monetary authority's preferences (3). Citing his experience as Vice Chairman of the Federal Reserve, Blinder questions whether there was ever much of a gap between u^{**} and u^* .

¹²Narrowing the gap between u^* and u^{**} slightly increases the magnitude of the estimated Phillips curve slope parameters θ_0 and θ_1 . But this change is small enough as to be essentially inconsequential for our results.

¹³In an earlier draft, we followed Chung (1990) to let this belief be estimated using the sample data. Since it is influenced by the updated beliefs in the sample, the value estimated this way is as difficult to interpret as that in Chung.

not improve the poor fit.¹⁴ Departing from Sargent (1999), therefore, we estimate the government’s prior beliefs $P_{1|0}$ and V within the sample. Our MCMC or maximum likelihood algorithm is “reverse engineering” the empirical Phillips curve at each date that, in conjunction with the Phelps problem, is needed to rationalize that date’s inflation rate. Such flexibility is crucial to allow this reverse engineering to back out correct correlations in $P_{1|0}$ and V . Moreover, this flexibility is arguably reasonable. The presample data may be informative of the government’s subjective point estimates (which we fix), but less so of its subjective uncertainty (which we estimate). Thus, we use the presample data to pin down the mean of the government’s estimate of the empirical Phillips curve, but *not* to estimate the belief innovation covariance matrix V .

IV.1. Model Fit. The posterior estimate of ϕ is reported in Table 1, along with the 68% and 90% probability intervals around the estimate.¹⁵ In our estimation and inference, the regressor vector in the government’s (subtly) misspecified Phillips regression is:

$$\Phi_t = \left[\pi_t \quad \pi_{t-1} \quad u_{t-1} \quad \pi_{t-2} \quad u_{t-2} \quad 1 \right]'$$

As mentioned, among the parameters that we estimate are those of the expectational Phillips curve (1) that we assume truly governs the data. As can be seen in Table 1, the natural rate of unemployment u^* in equation (1) is estimated to be 6.1 and its probability intervals are wide, consistent with the confidence interval in the statistical model of Staiger, Stock, and Watson (1997). Responses of unemployment to inflation surprises (θ_0 and θ_1) are extremely weak and are statistically insignificant by the probability intervals. This is an important finding for us, partly because it implies from (6) that Nash inflation is close to π^* despite the large difference between u^* and u^{**} . In the long run, therefore, the mean dynamics of inflation are always close to the Ramsey outcomes. Unemployment is by itself a persistent series and the persistence is tightly estimated.

¹⁴This point is illustrated further in Section V.2.

¹⁵All probability intervals are derived from the empirical joint posterior distribution generated from a sequence of 50,000 MCMC draws.

It can be seen from the estimates and probability intervals of ζ_1 and ζ_2 that their posterior distribution is tight but skewed downward, especially for ζ_1 whose posterior estimate is outside the 90% interval.

The estimated $P_{1|0}$ shows strong correlations (at least above 0.95) among all the elements. The relatively large variance for the drifting coefficient on π_{t-2} (the 4th element) implies that the government is quite uncertain about this coefficient, which affects the uncertainty about other coefficients even though their marginal variances are relatively small.

The estimated V shows strong correlations among the coefficients on current and lagged inflation variables. The variances on the inflation variables are large in the sense of Sargent and Williams (2003). The constant term has a very large variance and affects the coefficients on the lagged unemployment variables because of high correlations, but it has a small influence on the inflation coefficients in the government's model. Because V is not small, the government's beliefs are likely to drift significantly and inflation is likely to escape to the near-Ramsey region. Our estimates of the true expectational Phillips curve (1) imply that there is a negligible difference between the SCE and π^* . But as we show later, even when we artificially alter the parameters of (1) to allow the SCE inflation rate to be considerably higher than the Ramsey rate, this large V permits frequent escapes to low inflation rates.

The one-step forecasts of inflation are plotted against the actual path in Figure 1, and one-step forecasts of unemployment are plotted against the actual path in Figure 2. It is evident from these figures that the model fits the data well; so well, in fact, that it is difficult to discern the difference between the series. By this fit criterion, our reverse engineering exercise is a success, especially compared to those carried out by Chung (1990) and Sargent (1999). Figures 3 and 4 plot the one-step forecast errors for inflation and unemployment, showing that for most of the sample, the forecasts are within one half a percentage point of the realized value.¹⁶ In Appendix E we provide a more formal comparison of our model

¹⁶These forecast errors are comparable to those from the random walk model or an AR(1) model. These statistical models treat an inflation process to be exogenous.

TABLE 1. Posterior estimates of model parameters

Log value of posterior kernel at its peak: 81.56096

Estimates of coefficients in true Phillips curve and inflation process
with 68% and 90% probability intervals in parentheses

u^* : 6.1104 (5.2500, 7.1579) (4.2238, 9.0586)

θ_0 : -0.0008 (-0.0237, 0.0475) (-0.0458, 0.0719)

θ_1 : -0.0122 (-0.0375, 0.0297) (-0.0589, 0.0526)

τ_1 : 0.9892 (0.9852, 0.9960) (0.9817, 0.9996)

ζ_1 : 35.6538 (28.7565, 32.4947) (27.6017, 33.7890)

ζ_2 : 18.97671 (15.6565, 18.2557) (14.7008, 19.1196)

Estimate of $P_{1|0}$:

10.8705	14.3236	2.2518	-25.4037	-0.9279	-10.1548
14.3236	19.3721	2.9624	-33.9832	-1.1883	-13.5923
2.2518	2.9624	0.4690	-5.2629	-0.1928	-2.1050
-25.4037	-33.9832	-5.2629	59.8997	2.1339	23.9551
-0.9279	-1.1883	-0.1928	2.1339	0.0816	0.8526
-10.1548	-13.5923	-2.1050	23.9551	0.8526	9.5810

Estimate of V :

8.2323	-7.7781	0.9208	4.9782	-0.8136	-41.414
-7.7781	8.1400	0.0303	-5.089	1.9353	68.591
0.9208	0.0303	2.9854	0.1187	3.7012	72.067
4.9782	-5.089	0.1187	3.2032	-1.0548	-39.963
-0.8136	1.9353	3.7012	-1.0548	5.1362	100.6400
-41.414	68.591	72.067	-39.963	100.6400	2588.3000

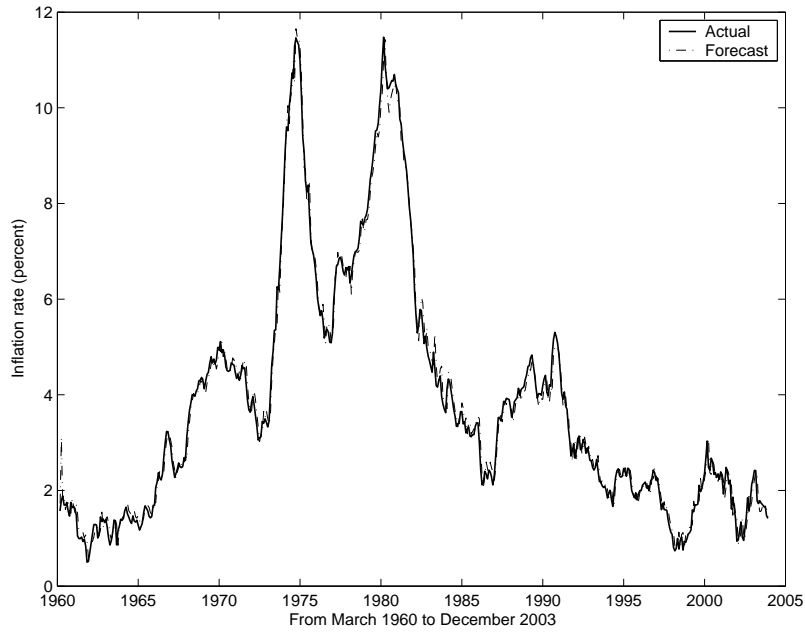


FIGURE 1. Inflation: actual vs one-step forecast (i.e, government controlled inflation)

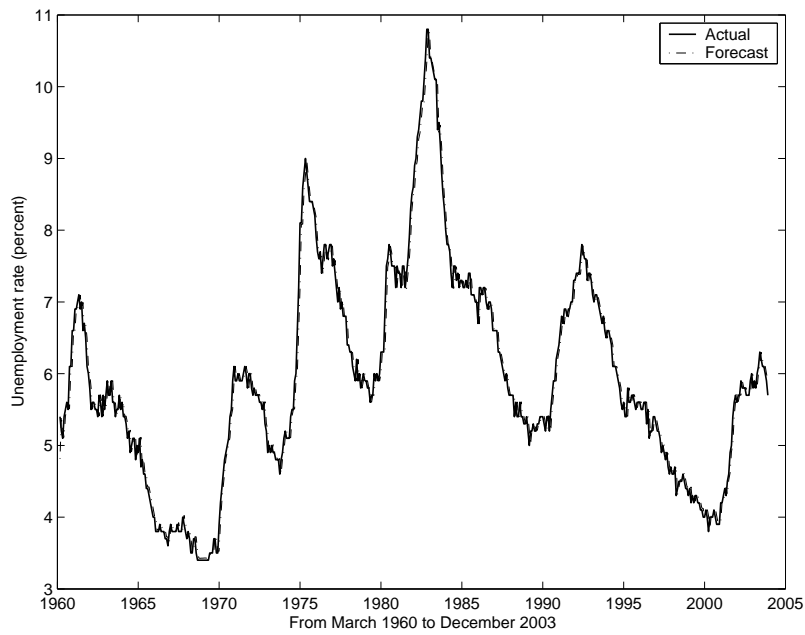


FIGURE 2. Unemployment rate: actual vs one-step forecast

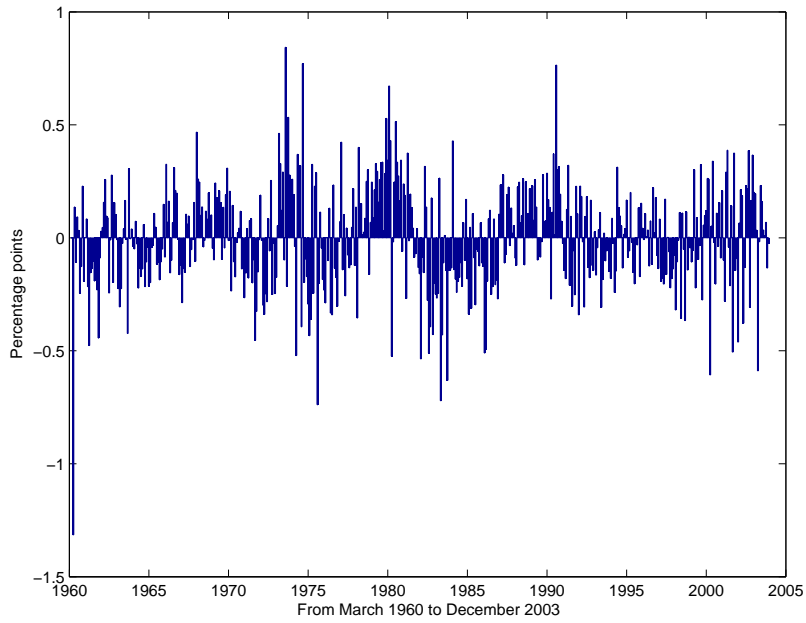


FIGURE 3. Differences between actual values and one-step forecasts of inflation

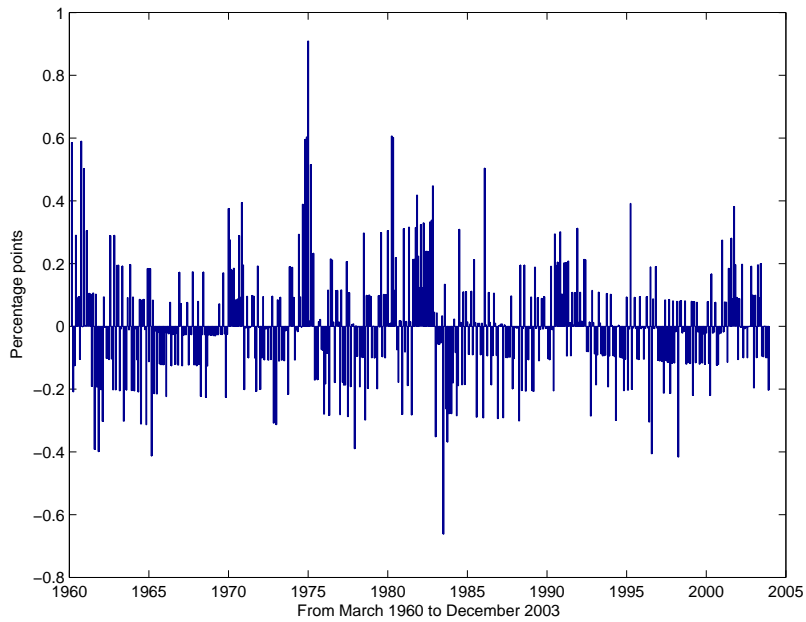


FIGURE 4. Differences between actual values and one-step forecasts of unemployment

to an atheoretical VAR. There we show that the VAR outperforms our model based on a marginal likelihood criterion, although this may be sensitive to the priors we chose in estimating our model. In terms of forecasting the rise and fall of inflation, our model performs as well as or better than the VAR. Without any assumption about exogenous components of the persistence of inflation, the government's inflation policy explains, almost entirely, the rise and fall of post-war American inflation (Figure 1). This kind of result has not been achieved in previous work (e.g., Sims 1988, Chung 1990, and Sargent 1999).

IV.2. Short-Run Dynamics: Shocks and Beliefs. The rise and fall of inflation in our model is driven by the Phelps problem in conjunction with government's belief in an exploitable tradeoff between inflation and unemployment, which leads to a high inflation rate in the early 70s. But then occasional sequences of stochastic shocks lead the government temporarily to believe that it can cut inflation with no rise in unemployment, which leads to rapid disinflations in the early 80s. During these episodes, the government learns a version of the natural rate theory in which the sum of the coefficients on inflation in the government's model is nearly zero, reflecting a vertical long-run Phillips curve.

Such an evolution of the government's updated beliefs are displayed in Figure 5. The sum of the coefficients on inflation becomes very negative in the early 70s and stay quite negative until the late 70s. In the 1980s, although the sum of the inflation coefficients is still negative, it is small enough for policymakers to decide cutting inflation without worrying much about costs in unemployment.

Figure 6 displays the suspected covariations in the drift of some key parameters in the government's Phillips curve, derived from our estimated V reported in Table 1. These key parameters are the sum of the coefficients on current and lagged inflation variables ($\alpha_1 + \alpha_2 + \alpha_4$), the sum of the coefficients on current and lagged unemployment variables ($1 - \alpha_3 - \alpha_5$), and the coefficient on the constant term (α_6). As shown by the symbol "*" in the first row of graphs of Figure 6, the estimated constant coefficient has a large, positive value while the sum of the estimated inflation coefficients is quite negative. This

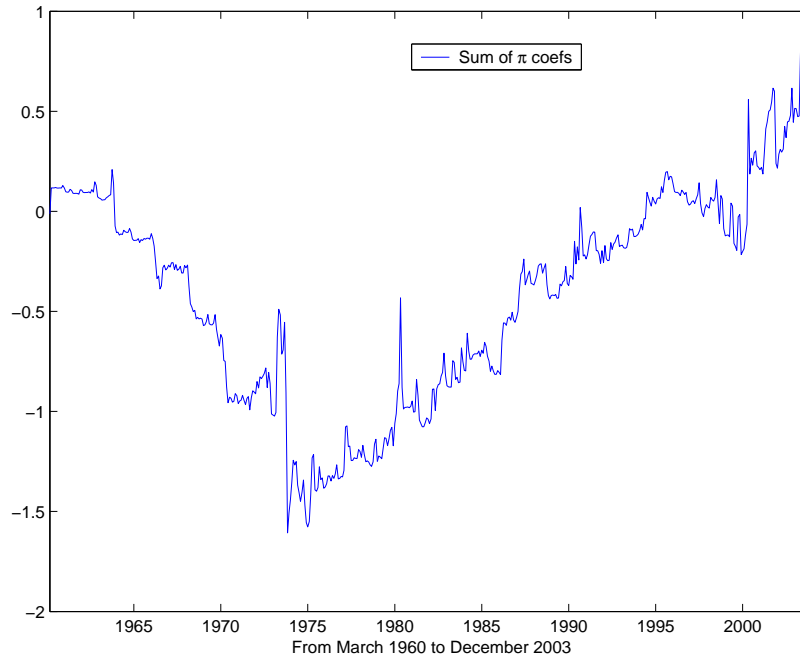


FIGURE 5. Evolution of the government's beliefs

combination leads to a perceived high tradeoff between unemployment and inflation in December 1973.

In contrast, at the point associated with the SCE (indicated by the symbol “o” in the second row of Figure 6), the estimated constant coefficient is small and the sum of the inflation coefficients is near zero, giving the government no incentive to inflate in pursuit of lower unemployment.

The probability ellipses shown in Figure 6 are quite large along the dimension of the constant coefficient. The large variation implies that a tradeoff between inflation and unemployment can be severe if there is a high probability of the constant coefficient and the sum of the inflation coefficients falling far into the north-west quadrant, as in the case of the upper-left graph. The bottom-left graph shows the historical estimates of these two belief parameters, induced by the particular sequence of shocks throughout our post-war sample. The area in which the sum of the inflation coefficients is less than -1 and the constant coefficient is greater than 15 covers most of the estimates for the 70s.

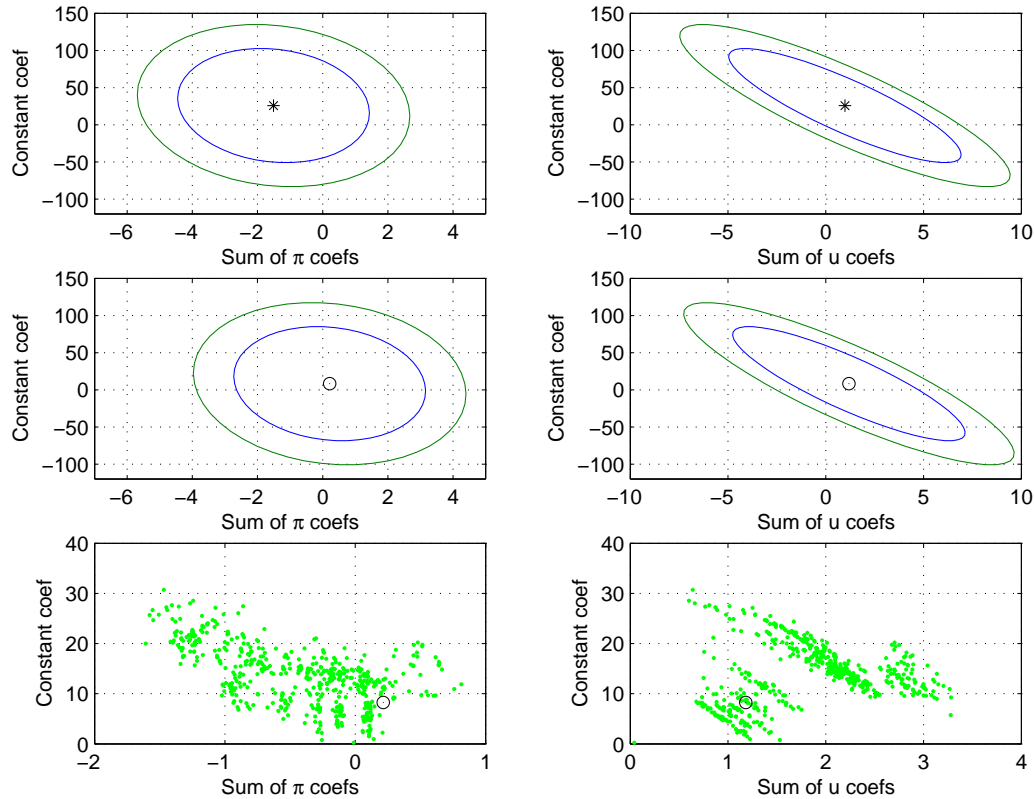


FIGURE 6. 68% and 90% probability ellipses about key parameters in the government's Phillips curve. The first row is based on the observation at 73:12; the second row is based on a limiting case associated with an SCE; the third row displays scattered plots of the estimates throughout our 60:02-03:12 sample. The asterisk symbol * in the first row depicts the government's estimates at 73:12. The circle symbol \circ in the second and third rows depicts SCE values, which also equal limiting estimates from the mean dynamics.

The constant and the sum of the unemployment coefficients are highly but negatively correlated, as shown in the first two graphs in the second column of Figure 6. Later we will see that in the transition to the SCE the economy may go through periods of very volatile inflation. The negative correlation between the constant and the unemployment coefficients is the most likely cause of these volatile inflation paths, if these two parameters

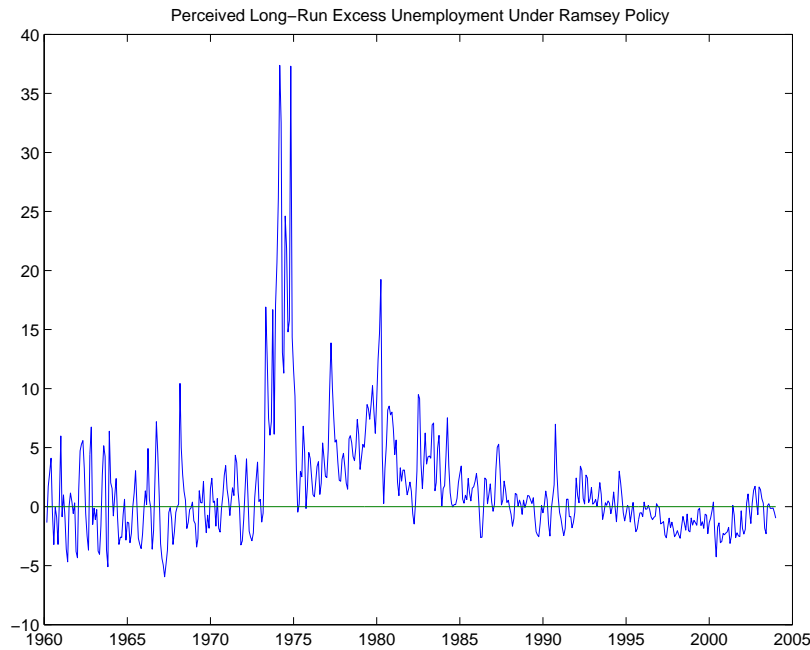


FIGURE 7. Perceived long-run excess unemployment under the Ramsey policy of 2% inflation according to the government's model.

fall in the south-east and north-west quadrants. Fortunately for US inflation outcomes, our historical estimates have been concentrated around the north-east quadrant, as shown in the bottom-right graph. It is only in out of sample simulations that we enter the more volatile regions.

The three belief parameters discussed above are key inputs to the government's perceived sacrifice ratio. The government's estimate of the long-run unemployment above the natural rate u^* under the government's Ramsey inflation policy is

$$\frac{\pi^*(\alpha_1 + \alpha_2 + \alpha_4) + \alpha_6}{1 - \alpha_3 - \alpha_5} - u^*. \quad (16)$$

Under any inflation policy in history that differs from Ramsey, the tradeoff will be a simple scaled version of (16) in proportion to that difference.¹⁷ Figure 7 plots the government's

¹⁷Note that our measure of the sacrifice ratio differs from the more conventional usage, which gives the cost of disinflating from a current inflation rate. Instead, ours is a long-run measure, independent of current inflation.

perceived long run excess unemployment under the Ramsey policy. Here we see that, throughout the 1970s, the government's model implied that substantial increases in unemployment would result from a low inflation policy.¹⁸ It wasn't until the early 1980s that this ratio fell to near zero, at which time the disinflation commenced. This point will be reinforced below when we present medium term forecasts around that time.

IV.3. Two Peaks and an Enduring Decline. We now analyze how the model forecasts the two peaks of inflation in the 1970s and the sharp decline in the early 1980s. We use Monte Carlo simulations to assess the distribution of forecasts going forward over four year horizons from different initial conditions. In each case, we take the estimated beliefs at the starting date and draw 5000 simulations of 50 periods each.¹⁹ The figures then plot the actual experienced inflation and the estimated government controlled inflation (x_{t-1}), along with 68% and 90% probability bands. In each plot, the initial condition is shown as date zero, from which we look forward 50 periods.

Figure 8 reports the forecasts. The upper left panel starts in January 1973 when inflation was at a very low level (3.3%). This is also near the time that the government most overestimated the tradeoff between inflation and unemployment (see Figure 5). According to the model, the government exploited the tradeoff and pushed up inflation to lower unemployment. The model predicts a steadily rising inflation path as high as 10% towards the end of the 4-year horizon (the upper 90% band), and gives little probability to a lower inflation rate in the medium run.

Due to a sequence of shocks, the inflation path actually experienced reached its peak earlier than the model predicts. But this is a treacherous period in which to predict, and

¹⁸A temporary drop in this sacrifice ratio around 1976 led to a temporary decline in inflation around that time. See Cogley and Sargent (2004) for a story in which the government was deterred from stabilizing in the mid 1970s because it attached a small positive probability to a model that assigned high unemployment costs to a rapid deflation.

¹⁹Adding uncertainty in the parameters would certainly widen our forecast bands.

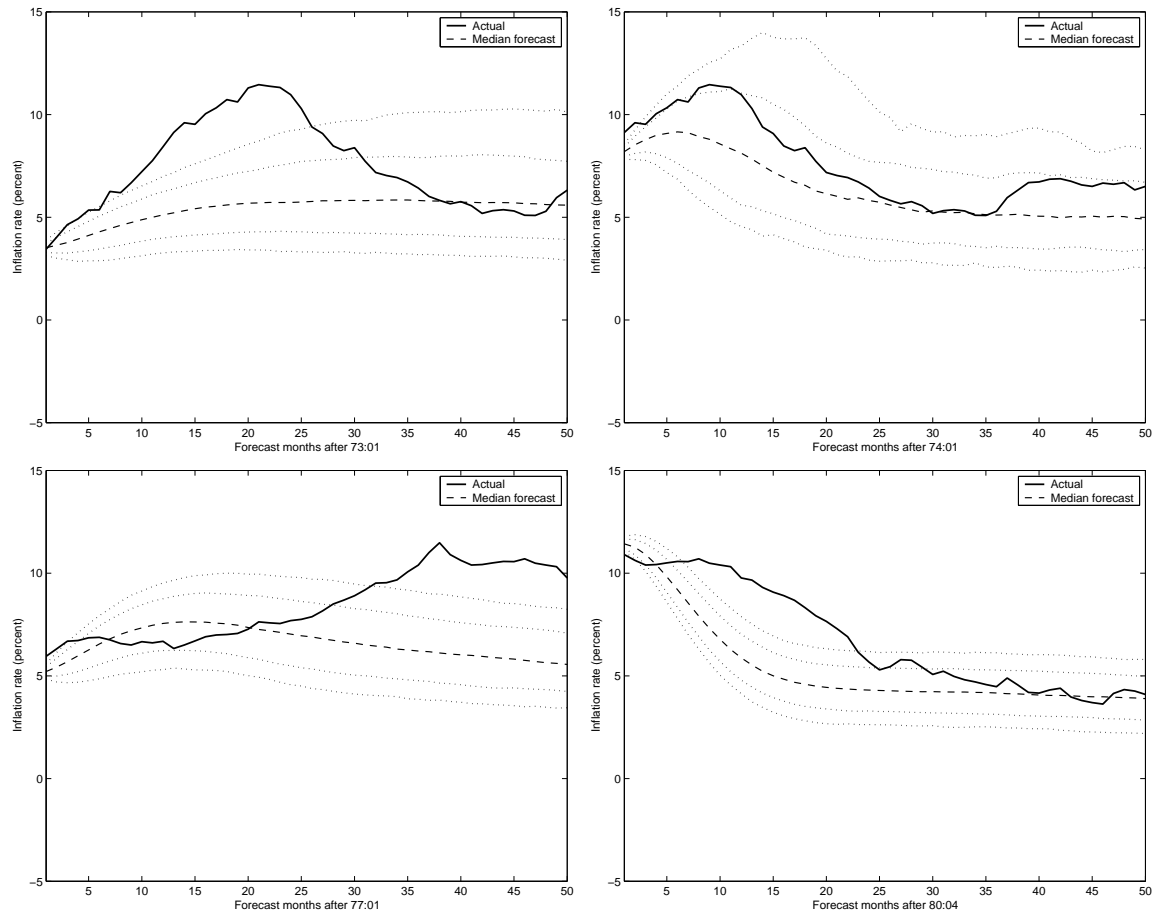


FIGURE 8. Dynamic forecasts of inflation with 68% and 90% error bands, using as initial estimated conditions at 73:01, 74:01, 77:01, and 80:04.

the prediction of rising inflation by the model compares favorably to alternative statistical models (see Appendix E).

A year later in January 1974, which is shown in the upper right panel of Figure 8, inflation had continued upward, now reaching 8.4%. Here we see that the model tracks the actual inflation path quite well, predicting a further increase in inflation prior to a return to lower levels.

January 1977 was another difficult time to predict inflation because inflation was at its trough and a second run-up was about to begin. Although actual inflation reached its peak

at the later date, the model assigns an overwhelming probability to higher inflation and the upper 90% reaches as high as 10% (the lower left panel).

The disinflation episode in the early 1980s is often interpreted as reflecting the intellectual triumph of the rational expectations version of the natural rate theory. What does our learning model say about this period? Would the government continue to pursue a higher inflation policy? After all, from the vantage point of April 1980 when inflation reached its second peak, most forecasting models either predict that inflation was very likely to go higher than it actually did, or they fail to predict the fall of inflation. The lower right panel of Figure 8 displays the forecast from our learning model. While actual inflation declines at somewhat a slower speed than what the model predicts in 1980 and 1981, the forecast of a fast decline in inflation is remarkable. The model's prediction is especially good further out in the forecasting period. Unlike many forecasting models, our model gives almost no probability to rising inflation in the medium horizon, because the tradeoff between inflation and unemployment by then is not high enough for the government to pursue higher inflation.

IV.4. Long-Run Dynamics. So far we have shown that our model is able to explain and predict the rise and fall of US inflation experienced. It is equally important to examine the long run properties of the model to see if the government's adapting beliefs will eventually lead to good inflation outcomes. We first discuss the convergence of our baseline model to a limit distribution. Then we analyze the small variation limits as in Sargent and Williams (2003).

IV.4.1. Long-Run Convergence. Figure 9 shows the inflation dynamics over 30,000 months starting at our estimated initial conditions: 1960:03 (the beginning of the sample), 1973:12 (the date where the sum of the π coefficients is most negative in the sample), and 2003:12 (the end of the sample). Clearly, they all converge to a limiting distribution around the Ramsey outcomes. This convergence occurs from the estimated initial conditions at any

date. The fluctuations at the beginning of the simulation reflect the rise and fall of American inflation that was temporarily off the equilibrium. As shown in the lower left panel, we are likely to see some high inflation in the near future but such high inflation is caused purely by exogenous random shocks to inflation, as the government will continue to see no tradeoff between inflation and unemployment (see the lower right panel of Figure 9). The government's beliefs are volatile for a while but eventually the sum of coefficients on inflation converges to near zero.²⁰ Consequently, the mean dynamics suggest that inflation converges to around 2%. These long run properties foster a view of US monetary history as a process of continual learning before inflation becomes stable around the Ramsey outcomes.

IV.4.2. *Small Variation Limits.* In the previous section, we saw evidence from simulations that the economy converges to a limit distribution. In order to obtain more explicit analytic results, we consider small variation limits. While it is difficult to obtain explicit results for any arbitrary setting of V , for smaller V the beliefs drift at a slower rate, allowing us to approximate their evolution with a differential equation. In particular, as in Sargent and Williams (2003), we let $V = \varepsilon^2 \hat{V}$ and study limits as $\varepsilon \rightarrow 0$. However, $P_{t|t} \rightarrow 0$ as $\varepsilon \rightarrow 0$, so we define a scaled matrix $\hat{P}_{t|t} = P_{t|t}/\varepsilon$ that does not vanish. Sargent and Williams show that as $\varepsilon \rightarrow 0$, the sequence $\{\alpha_{t|t}, \hat{P}_{t|t}\}$ generated by (9)-(10) converges weakly to the solution

²⁰ In those volatile periods, the constant coefficient in the government's estimated Phillips curve is often very large (on the order of 100) and the sum of the unemployment coefficients tends to be negative. Thus, these two government Phillips curve parameters fall in the north-west quadrant of the graph discussed in Section IV.2. If the sum of the inflation coefficients is negative, one can see from (16) that the government's dynamic programming problem implies a large increase in inflation to restrain adverse fluctuations in unemployment. Similarly, if this sum is positive, the government tends to generate a large rate of deflation. Such values for the government parameters in our simulations are far outside of the range attained by the historical estimates, as shown in the third row of graphs in Figure 6. When by chance we draw a sequence of shocks that keeps these government Phillips curve parameters within their historical range, convergence to a stable inflation path occurs without large swings of inflation.

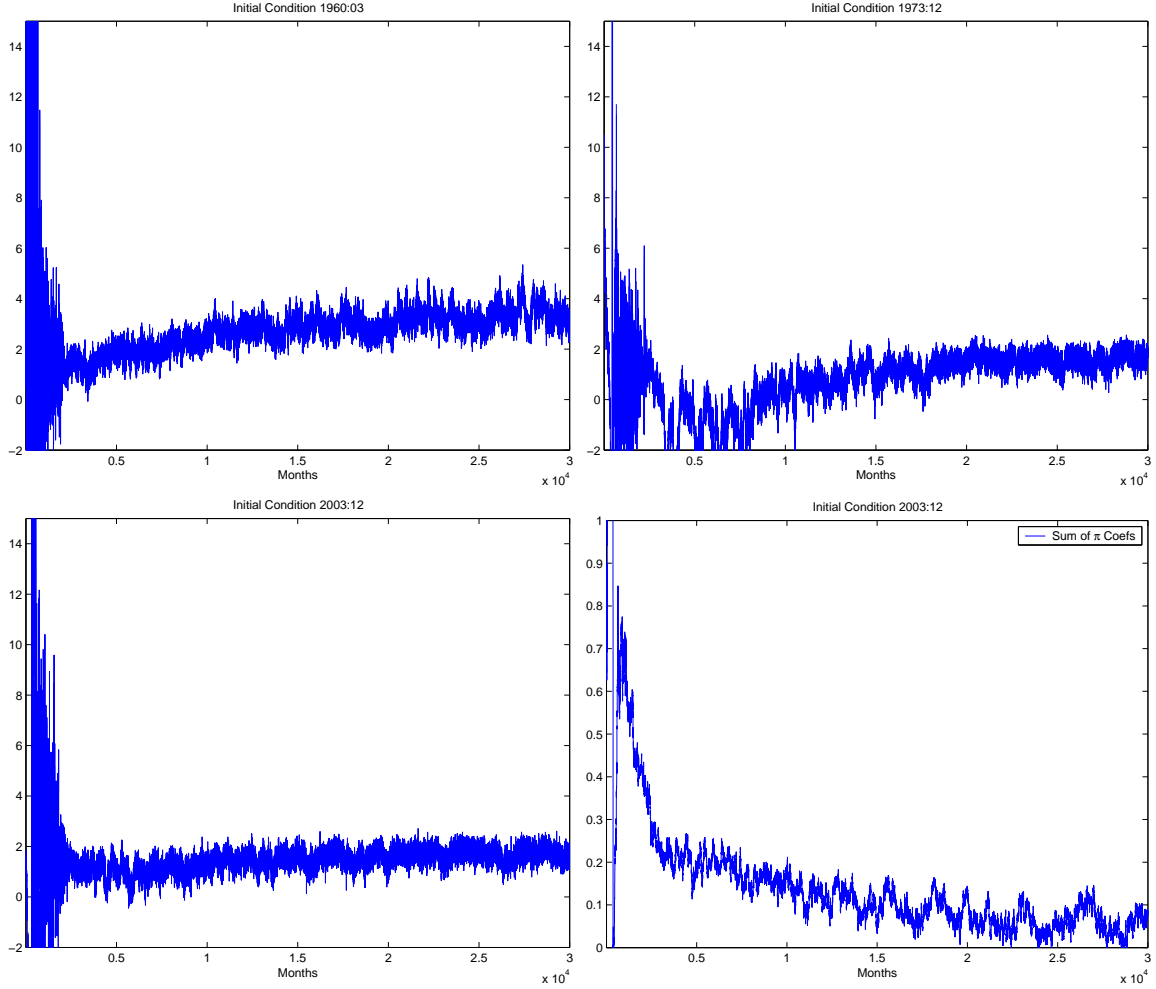


FIGURE 9. Government's inflation choice in long Monte Carlo simulations, using the different estimated initial conditions.

of the following ODEs:

$$\dot{\alpha} = PE [\Phi_t(u_t - \Phi_t' \alpha)] \quad (17)$$

$$\dot{P} = \sigma^{-2} \hat{V} - PE(\Phi_t \Phi_t') P, \quad (18)$$

where the expectations are calculated for fixed α . As we let the prior belief variance go to zero by shrinking ε , the government's beliefs track the trajectories of these differential equations. We call the ODEs (17)-(18) the *mean dynamics* because they govern the

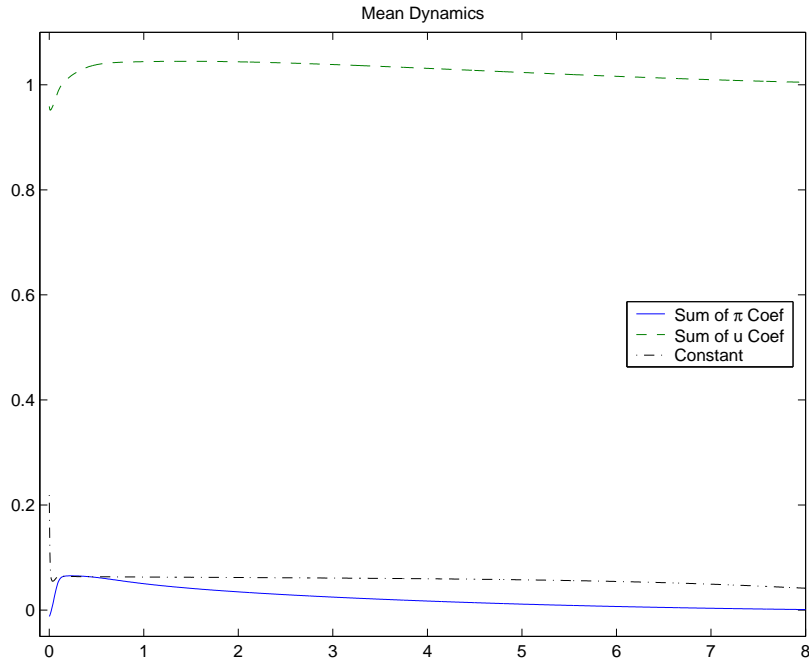


FIGURE 10. Mean dynamics for the baseline estimates, initialized at the start of the sample.

expected evolution of the government's beliefs. If the ODEs have a stable point $(\bar{\alpha}, \bar{P})$, then the government's beliefs will converge to it as $\varepsilon \rightarrow 0$ and $t \rightarrow \infty$. Note from (17) that the limiting beliefs satisfy the key least squares orthogonality condition (5) and hence comprise a self-confirming equilibrium. This orthogonality condition is the key identifying assumption in the government's subjective model, and in the limit it is satisfied by the true model.

In Figure 10 we plot trajectories of the mean dynamics for the government's beliefs starting from the initial conditions at the beginning of the sample. Here we see that the mean dynamics converge to a stable self-confirming equilibrium in the long run. The self-confirming equilibrium beliefs are:

$$\bar{\alpha} = [-0.0008 \quad -0.0000 \quad 0.9725 \quad 0.0000 \quad 0.0165 \quad 0.0688].$$

Thus, we see that in the SCE the government learns the true value of θ_0 , the effect of current inflation on unemployment. In the SCE, the government believes in a small tradeoff between inflation and unemployment, and so sets inflation slightly above the Ramsey level. In particular, the mean inflation rate in the SCE is 2.24% instead of the Ramsey level of 2%.

However, the mean dynamics and the self-confirming equilibrium govern the dynamics of our model only for small ε . In practice, ε must be quite small, on the order of 10^{-4} , for the asymptotic approximations to be accurate. Thus, for our baseline estimated V the mean dynamics do not fully characterize the evolution of beliefs. Loosely speaking, for any V we get convergence to a limit distribution. As $\varepsilon \rightarrow 0$ this limit distribution converges to a self-confirming equilibrium.

IV.5. Keynesian Direction of Fit. King and Watson (1994) discuss the importance of the direction in which a statistical Phillips curve is fit to the data. They use the adjectives Keynesian and Classical to denote Phillips curves with unemployment on the right and left sides, respectively.²¹ Unlike the version we consider above, the typical Samuelson-Solow Keynesian type of Phillips curve puts inflation on the left hand side of the regression. In this case, the government's Phelps problem becomes

$$\min_{x_t} \hat{E} \sum_{t=1}^{\infty} \delta^t ((\pi_t - \pi^*)^2 + \lambda_1 (u_t - u^{**})^2)$$

subject to (2) and

$$\begin{aligned} \pi_t = & \hat{\beta}_{0,t|t-1} u_t + \hat{\beta}_{1,t|t-1} u_{t-1} + \hat{\beta}_{2,t|t-1} \pi_{t-1} \\ & + \hat{\beta}_{3,t|t-1} u_{t-2} + \hat{\beta}_{4,t|t-1} \pi_{t-2} + \hat{\beta}_{5,t|t-1} + \sigma_K w_{Kt}, \end{aligned} \quad (19)$$

where w_{Kt} is an i.i.d. standard normal random variable.

²¹Sargent (1999, Chapter 7) discusses how the direction of fit affects the SCE that emerges.

With the normalization $\sigma_K = \sigma_1$, the government's beliefs $\hat{\beta}_{i,t|t-1}$ are updated using the usual Kalman filter. The government's Phelps problem is posed after inverting (19) to get

$$\begin{aligned} u_t = & \hat{\alpha}_{0,t|t-1} \pi_t + \hat{\alpha}_{1,t|t-1} \pi_{t-1} + \hat{\alpha}_{2,t|t-1} u_{t-1} \\ & + \hat{\alpha}_{3,t|t-1} \pi_{t-2} + \hat{\alpha}_{4,t|t-1} u_{t-2} + \hat{\alpha}_{5,t|t-1} - \sigma_t w_{Kt}, \end{aligned} \quad (20)$$

where

$$\begin{aligned} \hat{\alpha}_{0,t|t-1} &= 1/\hat{\beta}_{0,t|t-1}, \\ \hat{\alpha}_{1,t|t-1} &= -\hat{\beta}_{2,t|t-1}/\hat{\beta}_{0,t|t-1}, \\ \hat{\alpha}_{2,t|t-1} &= -\hat{\beta}_{1,t|t-1}/\hat{\beta}_{0,t|t-1}, \\ \hat{\alpha}_{3,t|t-1} &= -\hat{\beta}_{4,t|t-1}/\hat{\beta}_{0,t|t-1}, \\ \hat{\alpha}_{4,t|t-1} &= -\hat{\beta}_{3,t|t-1}/\hat{\beta}_{0,t|t-1}, \\ \hat{\alpha}_{5,t|t-1} &= -\hat{\beta}_{5,t|t-1}/\hat{\beta}_{0,t|t-1}, \\ \sigma_t &= \sigma_K/\hat{\beta}_{0,t|t-1}. \end{aligned}$$

For comparison with our results above, we re-estimate the model using the Keynesian direction of fit. The estimated values for the structural parameters are:

$$u^* = 5.7467, \theta_0 = -0.0287, \theta_1 = -0.0136, \tau_1 = 0.9881, \zeta_1 = 29.4602, \zeta_2 = 4.3244,$$

and the log value of the posterior kernel at its peak is -271.6802, which is much smaller than that in the classical model. These estimates are somewhat different from those for the classical model reported in Table 1. The biggest difference shows up in the estimate of ζ_2 . The significantly smaller value of ζ_2 in the Keynesian model implies large inflation shocks that tend to make the system unstable.²²

Figure 11 shows that even with the help of large inflation shocks, the fit is not nearly as good as the classical model, consistent with the findings of Cogley and Sargent (2004), who computed a sequence of Bayesian posterior odds for three models of the Phillips curve.

²²The results do not change much when we normalize σ_K differently. For example, instead of normalizing on σ_K directly, we let $\sigma_t = \sigma_1$. This normalization implies that σ_K is time-varying and $\sigma_{Kt} = \hat{\beta}_{0,t|t-1} \sigma_1$.

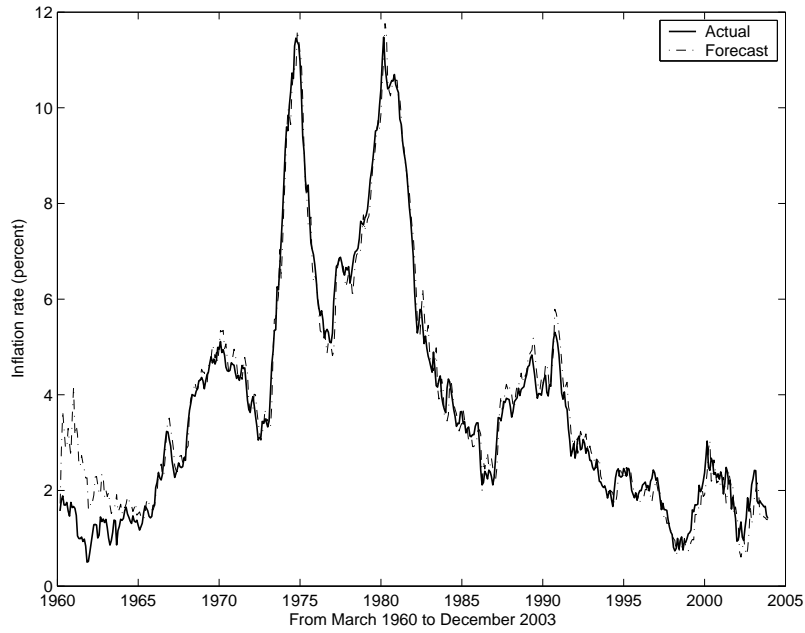


FIGURE 11. Keynesian model: actual inflation vs one-step forecast.

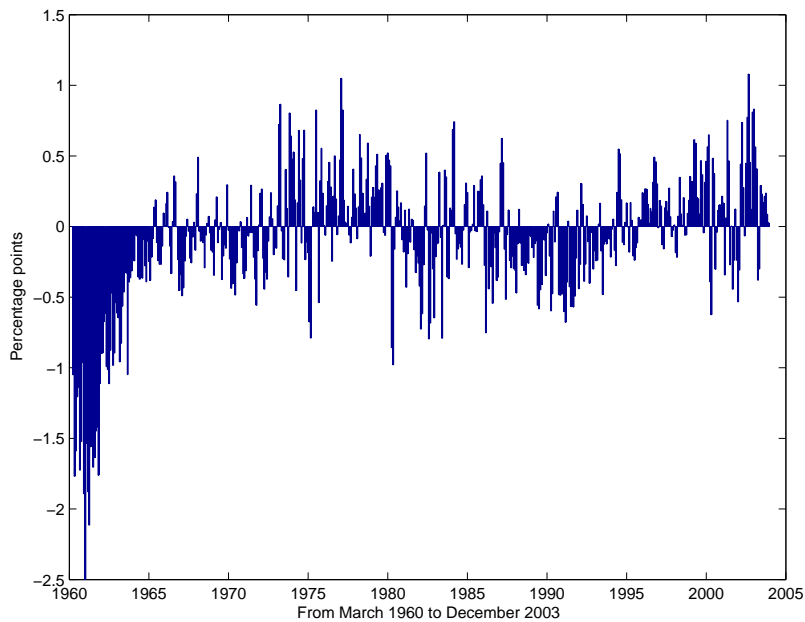


FIGURE 12. Keynesian model: differences between actual values and one-step forecasts of inflation.

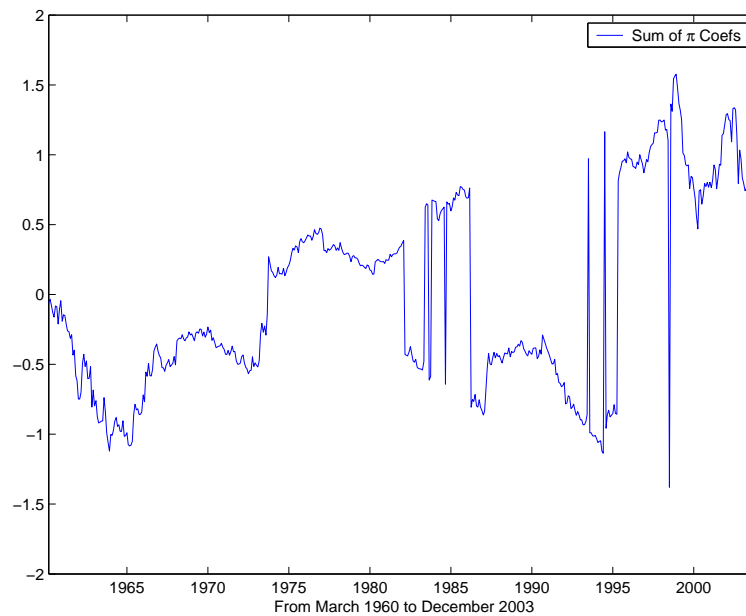


FIGURE 13. Keynesian model: evolution of the government's beliefs.

The forecast errors reported in Figure 12 reinforce this point. The persistent over-prediction of inflation in the early 60s is caused by the misperceived tradeoff between inflation and unemployment under the estimated Keynesian Phillips curve, as is shown in Figure 13. We see there that, throughout the 1960s, the government's Keynesian model implied an exploitable tradeoff. However this tradeoff dwindles in the early 70s and the sum of the coefficients on inflation reversed its sign sharply in the mid-to-late 70s. In particular, from 1975 to 1980 the government's model shows no sign of a tradeoff (Figure 13), and hence the model predicts lower inflation than actually occurred over this period (Figure 12). It is difficult to explain why the government did not cut inflation in the late 70s within the Keynesian story. One possibility is that because of large shocks in the government's Keynesian Phillips curve, policymakers decided to ignore the evidence that there was no tradeoff in the later part of 1970s. Instead, they produced high inflation based on an inherited Keynesian belief that the tradeoff must be there.²³

²³This interpretation was provided by James Nason, based on oral descriptions by Fed policymakers.

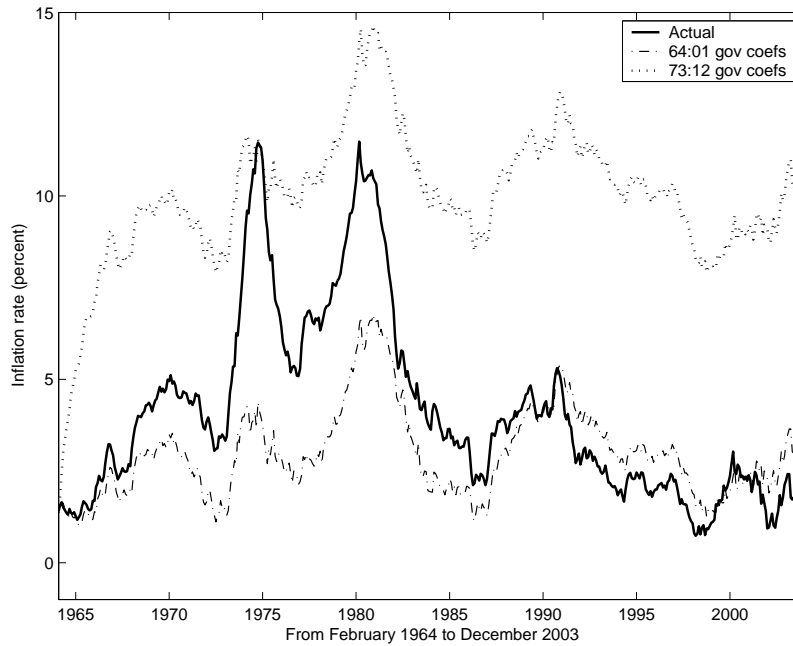


FIGURE 14. Inflation dynamics with fixed beliefs at the dates 64:01 and 73:12.

V. COUNTERFACTUAL RESULTS

As a way to quantify the role of econometric policy evaluation in the government's learning process, we use our estimated classical model to calculate what would have happened if the government's beliefs had differed from our estimates. First, we will look at belief drifts in the Phillips curve. Then, we will examine how the covariance matrix V affects these drifts.

V.1. Two Historical Episodes. All of the results in this section condition on our estimates of the historical shocks of unemployment and inflation inferred from our model estimates. We treat these shocks as random and exogenous in our counterfactual exercises.

The first episode begins in 1964:01. As seen from Figure 5, there is still little belief in the inflation-unemployment tradeoff in the early 1964, but by then end of 1973 the sum of

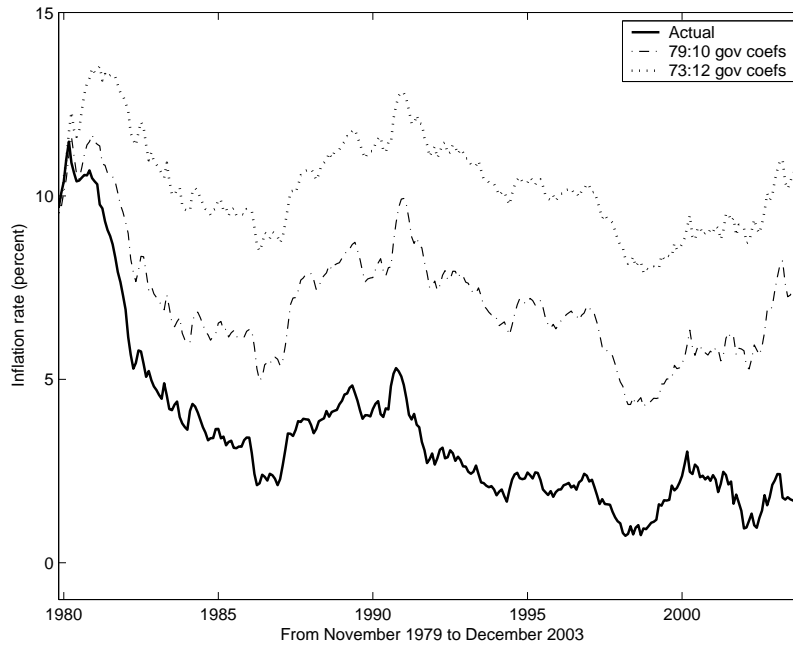


FIGURE 15. Inflation dynamics with fixed beliefs at the dates 79:10 and 73:12.

the inflation coefficients is most negative.²⁴ Such continual adaptation of beliefs towards a bigger inflation-unemployment tradeoff gives the government an incentive, through the Phelps problem, to run a high inflation policy. This can be seen indirectly in Figure 7 as the perceived costs of low inflation rise dramatically in the early 1970s. To obtain a more direct comparison, suppose the government's beliefs had been frozen at the 64:01 initial condition. As shown in Figure 14, the inflation path would have been smoother and avoided much of the two large run-ups of actual inflation in the 70s. To take an opposite example, we replace the government's 64:01 beliefs with the 73:12 beliefs and fix them throughout the history. In this case, Figure 14 shows that inflation would have been much higher than was actually experienced throughout the sample and would have continued to stay around 10%.

²⁴See Sargent (1999, chapter 5) for how the sum of coefficients on π affects the advice rendered by the Phelps problem.

Figure 15 displays the second episode beginning with 1979:10, when Volcker's disinflation policy took place. As we've seen, if the government had held fixed to its 1973:12 perceived tradeoff, inflation would have stayed much higher. On the other hand, if the government's belief at 1979:10 had been fixed throughout the rest of the history, inflation would have come down to 5% by 1986 due to the sequence of historical shocks, but there would have been a tendency to return to a higher inflation level. These outcomes show the importance of the adaptation of the government's beliefs to achieve low inflation. With the same sequence of historical shocks, actual inflation came down and remained low as shown by the inflation path in Figure 15. Although the government's beliefs at the end of 1979 favored a disinflation, Figure 5 shows how the government's views continued to evolve to favor a low inflation policy. The experience of disinflation and continued low inflation led the government away from believing in an exploitable Phillips curve tradeoff.

These exercises suggest that while the rise of inflation in the 70s was caused by the government's misperceptions about the Phillips-curve relationship, the fall of inflation in the 1980s can be rationalized by an econometric policy evaluation procedure that embodies adapted beliefs.²⁵

V.2. Importance of V . As we've already noted, crucial to our empirical success is the flexibility of our model in fitting the government's beliefs. In particular, previous work of Sims, Chung, Sargent, and Cho, Williams, and Sargent (but not Sargent and Williams) had assumed a particular form for the key matrix V in (8) that governs the innovations to the parameters in the government's model. We have already discussed theoretical reasons why the V matrix is so important and how different specifications of it greatly affect the speed, direction, and stability of the learning dynamics. Previous work fixed V in ways that

²⁵Changes in beliefs do not necessarily imply changes in the linearized policy rule in which x_t is regressed on its own lagged values and current and lagged unemployment variables. In our case, because x_t tracks the actual inflation path so well, our results are consistent with reduced-form empirical findings that changes in the policy rule or the inflation process are difficult to detect statistically (Cogley and Sargent 2003, Primiceri 2003a, and Sims and Zha 2004).

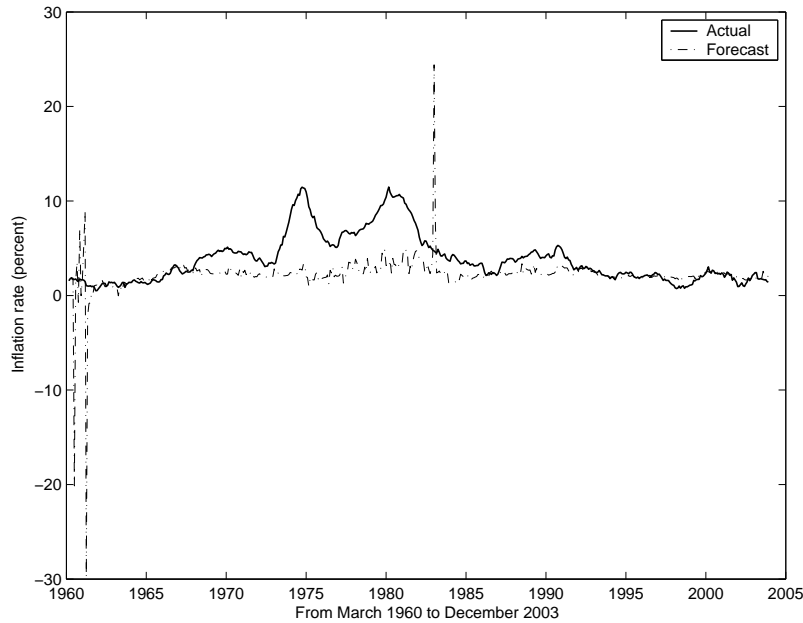


FIGURE 16. Actual and forecast inflation (government policy) with V derived from $E(\Phi\Phi')^{-1}$.

constrained how learning could occur, thereby diminishing the variation in the data that could be explained by evolving government beliefs.

One particular example of the importance of the V matrix is shown in Figure 16. There we fix the values of all the other structural parameters (including the estimated initial condition $P_{1|0}$) but only alter V . Most learning models such as Sargent (1999) have focused on a recursive least squares learning rule that is closely related to the Kalman filter. Sargent and Williams (2003) show that RLS can be approximated by a Kalman filter with V proportional to $\sigma^2 E(\Phi\Phi')^{-1}$. In the figure, therefore, we use the sample estimate of the second moment matrix and we choose the proportionality factor so that the new V matrix has the same norm as our estimate. This choice of V evidently leads to a substantial deterioration in fit. The government's optimal policy suggests very low inflation (with a few outliers) for almost the entire sample, which is what Sargent (1999) found. In particular, it completely misses the two peaks in inflation in the 1970s. This illustrates a point made by Chung (1990) and Sargent (1999), both of whom noted that their models implied that the

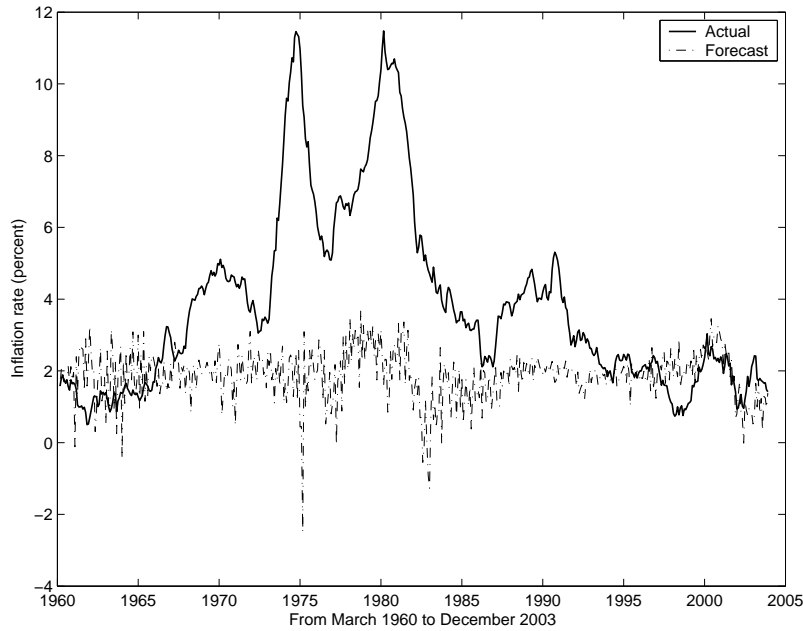


FIGURE 17. Actual and forecast inflation (government policy) with V being scaled down by 0.01.

government should have cut inflation much earlier than actually occurred in the data. However, that outcome relied on attributing to the government very particular beliefs about how its model changes over time. By allowing the data to inform our choice of these beliefs, we are able much better to explain the rise and fall of inflation.

As discussed in Section IV.4.2, the self-confirming equilibrium (SCE) governs the inflation dynamics of our model only for a scaled-down value of V . Figure 17 shows another example of the importance of the V matrix where the originally estimated V is scaled down by 0.01. Again, this inflation policy completely misses the rise and fall of actual inflation. In this case, the model remains close to the self-confirming equilibrium throughout the sample, and we've seen that inflation is low in the SCE.

What are the long-run implications of the estimated V ? The large value of V suggests that one expect escapes from SCE to be frequent even if the inflation rate at the Nash equilibrium is much higher. To illustrate this point, we change θ_0 from its estimated value of -0.0008 to -1.0 while keeping all the other estimates fixed. This implies that the Nash

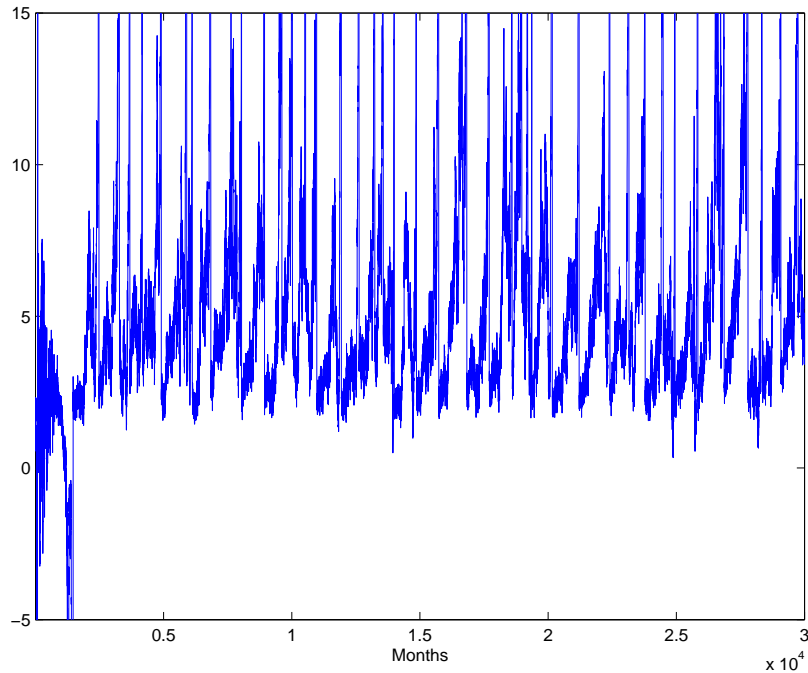


FIGURE 18. Long-run inflation dynamics with $\theta_0 = -1.0$, using the end-of-sample estimate as the initial condition.

inflation rate is around 10%, while the socially optimal Ramsey level remains at 2%. As can be seen from Figure 18, inflation tends to be high, but the large time-variation of the drifting beliefs implied by V allows the dynamics to escape to low inflation repeatedly, and there is no tendency for inflation to stay very long at the high level. Thus, our V matrix is consistent with repeated escapes in the long run, but they are difficult to detect under our estimates because we estimated a low sacrifice ratio.

VI. RELATION TO OTHER LITERATURE

Sargent (1999) fashions two competing stories about the acceleration of US inflation during the 1970s and Volcker's conquest of inflation during the early 1980s. The first story, called the 'triumph of the natural rate theory', explains the run up of inflation as reflecting the monetary authority's yielding to the temptation to exploit a non-expectational Phillips

curve, while Volcker's conquest of inflation comes from the monetary authority first learning a correct rational expectations version of the natural unemployment rate hypothesis, then somehow managing to commit itself to the Ramsey policy. The competing story, called the 'vindication of econometric policy evaluation', shares the 'triumph' story's account of the run up of inflation but differs about Volcker's stabilization. According to the vindication story, the Fed never learned a sharp rational expectations version of the natural rate hypothesis. Instead, sequentially refitting the *misspecified* non-expectational Phillips curve led the Fed to discover inflation-unemployment dynamics that prompted it, through a Phelps problem, to stabilize inflation by the early 1980s. According to the vindication argument, allowing the data to speak continuously, even through estimates of a misspecified Phillips curve model, did a good enough job to prompt the Fed to stabilize.²⁶

Sargent (1999), Cho, Williams, and Sargent (2002), and Sargent and Williams (2003) added structure and further features to the vindication story by analyzing versions of the model in this paper. They a priori adopted parameter specifications that opened a substantial gap between the Ramsey and Nash inflation outcomes, and their work discovered both the mean dynamics that on average push outcomes toward the Nash inflation level and the escape dynamics that recurrently push it toward the Ramsey outcome.

The present paper produces estimates of the key parameters that ignite the mean dynamics and escape dynamics. As we have stressed, our empirical estimates deemphasize the roles of *both* the mean dynamics and the escape dynamics and instead emphasize the short-term impacts of shocks on government beliefs. In addition, our estimate of a small gap between the Nash and Ramsey inflation levels supports Blinder's (1998) skepticism about whether that gap is central quantitatively to the monetary authority's decision problem.

Cogley and Sargent (2004) provides an alternative econometric explanation of US post WWII inflation that also features the interaction of a government learning process and a sequence of Phelps problems. Cogley and Sargent's learning process has the government

²⁶The oral tradition in which the vindication story originates, which Sargent heard from Albert Ando, saw it as a puzzle that Volcker waited so long to stabilize.

apply Bayes' rule as it estimates three Phillips curve models, one of which a rational expectations version of a natural rate model.

Primiceri (2003b) also develops a learning model to explain the rise and fall of US inflation. He estimates his model on US data and finds that its fit is comparable to an atheoretical VAR as a description of the data. Like us, he emphasizes that inflation remained high in the 1970s due to the government's perception that disinflation was too costly. Unlike us, a key component of his model is the mismeasurement of the natural rate of unemployment, which caused policy to be looser than policymakers thought. Our paper differ substantially from his in its theoretical structure. Our true economy is a rational expectations natural rate model, and policymakers can eventually converge to a self-confirming equilibrium, a type of rational expectations equilibrium. Primiceri's main focus is on a backward-looking Keynesian model, with no role for expectations.²⁷ We view the rational expectations natural rate theory as a useful starting point in analyzing the behavior of inflation.

VII. CONCLUSION

Our estimates attribute the differing inflation outcomes over the post-war period to changes over time in the monetary authority's beliefs. According to our estimated model, that learning process explains the delayed reaction to rising inflation in that period.

Our empirical results suggest an interpretation that differs from the work we build on. Sargent (1999) and Cho, Williams, and Sargent (2002) suggested that US experience could be explained by convergence to a high Nash inflation level coupled with occasional escapes to a lower Ramsey level. As discussed by Sargent and Williams (2003), these outcomes also occur in our model when we set parameters of the true Phillips curve to allow a larger gap between the Nash and Ramsey levels of inflation, and when we also impose what, relative to our estimates, is a scaled-down innovation volatility matrix V in the government's belief-drift dynamics (8). However, with our estimates, it appears that oscillations between the

²⁷Primiceri also considers a New Keynesian rational expectations model, but it fits substantially worse than his backward-looking specification.

Nash and Ramsey levels of inflation, driven alternately by the mean dynamics and then the escape dynamics of, say, the Sargent-Williams (2003) model, were not the main forces that accounted for the inflation process that the monetary authorities in the US chose to preside over during the post WWII years. For one thing, our estimates of the Nash level of inflation turned out to be near the Ramsey level.²⁸ Instead, the rise in inflation was driven by the interaction of shocks and government beliefs, and the fall in inflation was due to changes in those beliefs. If the US monetary authority remembers the lessons that prompted Volcker to disinflate in the early 1980s, then maybe Volcker's conquest of US inflation will endure.

APPENDIX A. DATA

The two monthly series employed in this paper are:

- Civilian unemployment rate, 16 years and older, seasonally adjusted (source: BLS);
- PCE chain price index (2000=100), seasonally adjusted (source: BEA).

Inflation is measured as an annual rate (12-month ended) of change of the PCE price index. The estimation sample (including lags) is from January 1960 to December 2003.

APPENDIX B. PROOF OF PROPOSITIONS 1 AND 2

From equations (1) and (2) it can be seen that the Jacobian transformation from w_{1t} and w_{2t} to u_t and π_t is equal to 1. It follows that the likelihood function is:

$$\mathcal{L}(\mathcal{I}_T|\phi) = \frac{\zeta_1^{T/2}\zeta_2^{T/2}}{(2\pi)^{T/2}} \exp \left\{ -\frac{1}{2} \sum_{t=1}^T [\zeta_1 z_{1t}^2 + \zeta_2 z_{2t}^2] \right\}. \quad (\text{B1})$$

²⁸Furthermore, if we arbitrarily set the parameters of the true expectational Phillips curve to create a big gap between the SCE-Nash and the Ramsey inflations, but retain estimated innovation covariance V , escapes from a SCE again occur frequently enough to vitiate any pattern of recurrent oscillations between the SCE and Ramsey levels of inflation.

As defined in the text, z_{1t} and z_{2t} are the functions of θ and φ :

$$\begin{aligned} z_{1t} &= u_t - u^* - \theta_0(\pi_t - x_{t-1}) - \theta_1(\pi_{t-1} - x_{t-2}) - \tau_1(u_{t-1} - u^*), \\ z_{2t} &= \pi_t - x_{t-1}, \end{aligned}$$

where the optimal decision rule depends on φ . The prior pdf of ϕ is

$$p(\phi) = p(\theta) p(\varphi) p(\zeta_1, \zeta_2),$$

where the pdfs on the right hand side of the above equation are defined in (11), (12), and (13).

The posterior pdf of ϕ is proportional to the product of the likelihood (B1) and the prior $p(\phi)$:

$$p(\phi | \mathcal{I}_T) \propto \mathcal{L}(\mathcal{I}_T | \phi) p(\phi). \quad (\text{B2})$$

Because x_{t-1} does not depend on θ , ζ_1 , and ζ_2 , it can be seen from (B2) and (B1) that the posterior distribution of θ conditional on all other parameters is Gaussian and that the posterior distribution of ζ_1 and ζ_2 is of Gamma. Algebra leads to the Gaussian form (14) and the Gamma form (15).

APPENDIX C. PRIOR SETTINGS

Our estimation results are quite similar to the maximum likelihood estimates. But the prior is essential for obtaining finite-sample inferences because the government belief parameter matrix V may not have a proper density function when there is no prior. The prior for θ is mostly based on economic theory. For example, the mass prior probability of θ_0 is in the negative region.

The prior mean for θ is set to

$$\begin{bmatrix} 0.12 \\ -0.20 \\ -0.16 \\ 0.98 \end{bmatrix},$$

which implies that the natural rate of unemployment is 6.0 with somewhat persistent unemployment. The prior mean of θ_1 is only slightly less than that of θ_0 in absolute value ($.16 < .20$), implying the low serial correlation of structural disturbances in Sargent's version of the Phillips curve (pp.70-71, Sargent 1999). The prior variance for θ is

$$\lambda_1 \begin{bmatrix} 0.06^2 \\ 0.10^2 \\ 0.08^2 \\ 0.01^2 \end{bmatrix},$$

where λ_1 controls the tightness of the prior variance. With $\lambda_1 = 1$, the prior standard deviation allows large variation but at the same time gives little probability to negative values of v^* , or positive values of θ_0 and θ_1 , or the value of τ_1 being greater than 1 (an explosive root).

For the prior of ζ_1 and ζ_2 , we set $\bar{\alpha} = 4$ and $\bar{\beta} = 12.5\lambda_2$. By setting $\lambda_2 = 1$, the prior mean for ζ_i becomes 50 and the prior variance becomes 25^2 , implying a quite loose prior for ζ_i .

The prior mean for C_P and C_V is 0. The prior variance is $5^2\lambda_3$ for the diagonals of C_P and C_V and 2.5^2 for the off-diagonal elements. The tightness control hyperparameter is set at 0.5.

In this paper, we have checked the robustness of our estimated results by varying the values of the tightness control parameters λ_1 , λ_2 , and λ_3 .

APPENDIX D. PROPOSAL DENSITY FOR THE METROPOLIS ALGORITHM

The key to the Metropolis algorithm for the posterior distribution φ is to obtain the covariance matrix for a normal proposal density. Since x_{t-1} is a function of φ , one can approximate it by a second-order Taylor expansion at the posterior estimate $\hat{\varphi}$. It can be

seen from (B2) that this approximation leads to the following covariance matrix for φ :

$$\begin{aligned} \tilde{\Sigma}_\varphi^{-1} = & (\zeta_1 \theta_0^2 + \zeta_2) \sum_{t=2}^T \frac{\partial x_{t-1}(\hat{\varphi})}{\partial \varphi} \frac{\partial x'_{t-1}(\hat{\varphi})}{\partial \varphi} + \zeta_1 \theta_1^2 \sum_{t=2}^T \frac{\partial x_{t-2}(\hat{\varphi})}{\partial \varphi} \frac{\partial x'_{t-2}(\hat{\varphi})}{\partial \varphi} \\ & + \zeta_1 \theta_0 \theta_1 \sum_{t=2}^T \left[\frac{\partial x_{t-1}(\hat{\varphi})}{\partial \varphi} \frac{\partial x'_{t-2}(\hat{\varphi})}{\partial \varphi} + \frac{\partial x_{t-2}(\hat{\varphi})}{\partial \varphi} \frac{\partial x'_{t-1}(\hat{\varphi})}{\partial \varphi} \right] + \bar{\Sigma}_\varphi^{-1}, \end{aligned} \quad (\text{D1})$$

where $\bar{\Sigma}_\varphi$ is the prior covariance matrix for φ .

APPENDIX E. STATISTICAL COMPARISON TO BVARs

We compute the marginal data density (MDD) for our learning model, using the modified harmonic mean method described in Geweke (1999). The log MDD value is about 27.50, which is substantially lower compared to monthly Bayesian vector autoregression (BVAR) models of unemployment and inflation. With the standard prior settings proposed by Sims and Zha (1998), the log MDD value is 172.05 for the BVAR with one lag and 244.65 for the BVAR with 13 lags. By the measure of Bayes factors, the BVARs (including the BVAR with only one lag) dominate our learning model. Note that this result may be sensitive to the priors used in estimating the models. The prior for our model was loose along several dimensions, which means that the MDD may be penalizing our model more heavily than the VAR.

Higher Bayes factors do not necessarily imply that BVARs outperform our learning model in explaining the rise and fall of inflation. Figure E1 displays the forecast from the BVAR(1). The 68% and 90% error bands are produced by simulating the VAR shocks while holding the parameter estimates fixed at those obtained using the 60:01-03:12 sample, the same procedure as we did for our learning model. As can be seen, this BVAR fails to predict any rise of inflation with a significant probability. And the upper 90% band is well below 10%.

Figure E2 shows the forecasts of inflation at the various dates from the BVAR(13), resembling Figure 8. The forecasts at 73:01 from the BVAR(13) are not so different from

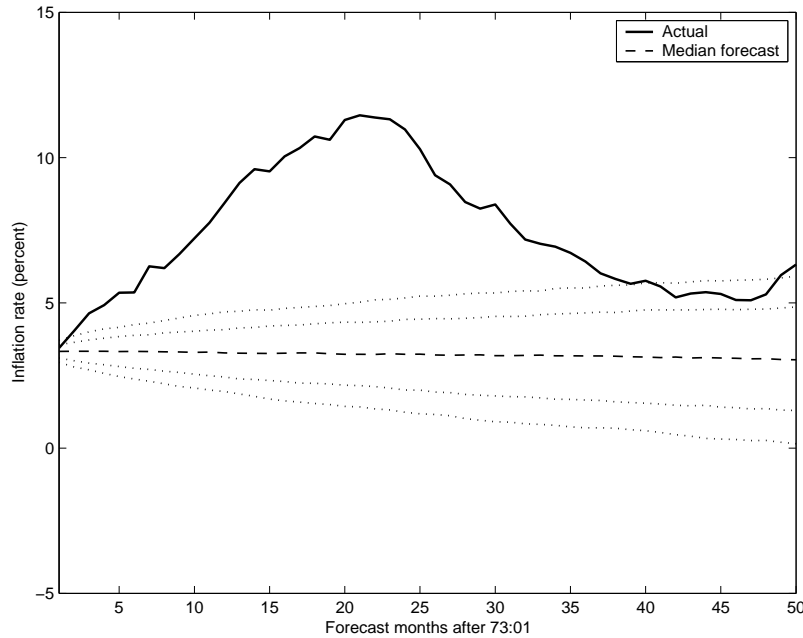


FIGURE E1. Dynamic forecasts of inflation from BVAR(1) at 73:01, with 68% and 90% probability bands.

those from the BVAR(1) except the error bands are much wider. It gives half probability to a decline of inflation.

For the forecasts at 74:01, the BVAR forecasts are comparable to those from our learning model.

The forecasts at 77:01 from the BVAR(13) again gives half probability to a decline of inflation while the forecasts from our learning model (Figure 8) put a vast majority of probability to rising inflation.

For the forecast at 84:04, the BVAR(13) predicts the decline of inflation like our learning model. But our learning model predicts a much sharper decline of inflation with narrow bands while the BVAR gives a considerable probability of higher inflation than the actual path with wider bands.

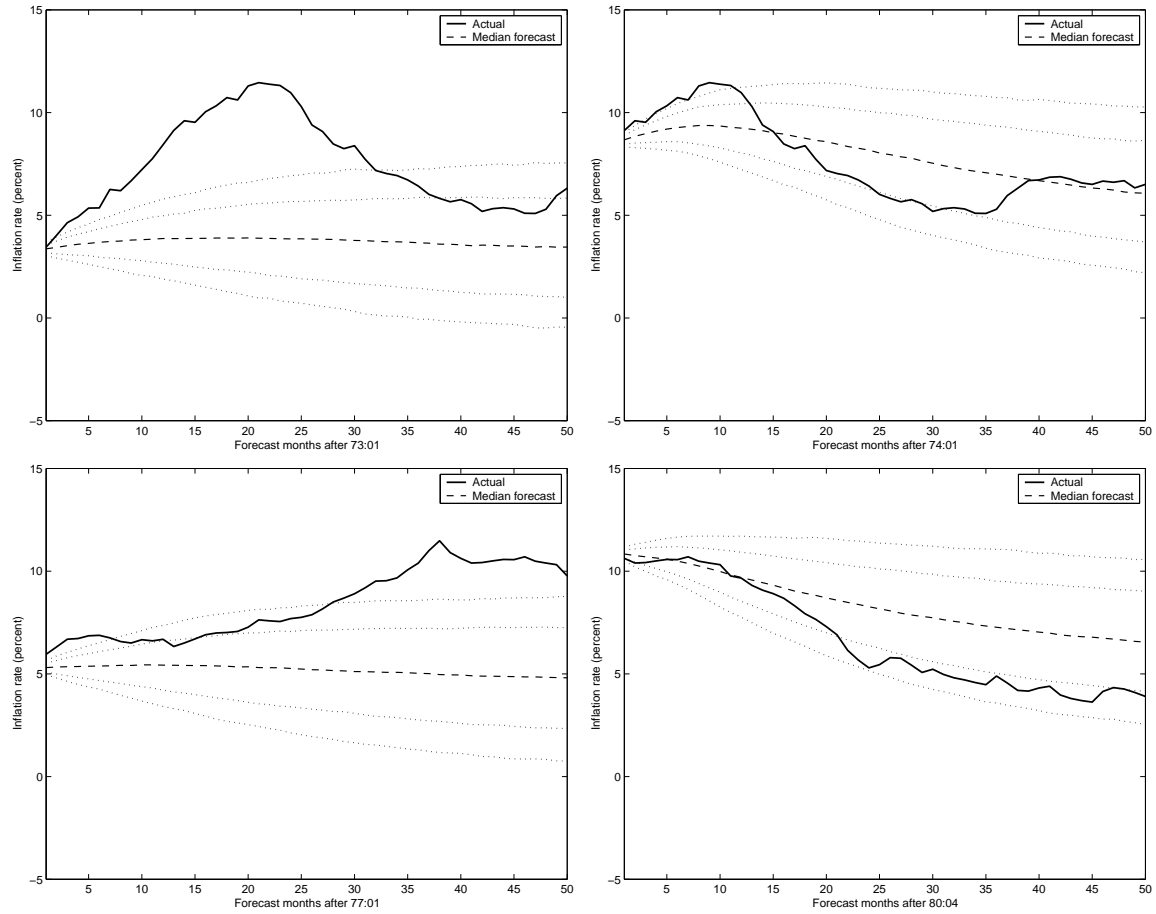


FIGURE E2. Dynamic forecasts of inflation from BVAR(13) at 73:01, 74:01, 77:01, and 80:04, with 68% and 90% error bands.

APPENDIX F. ACCOUNTING FOR PARAMETER UNCERTAINTY

In this appendix, we report probability bands on key government policy variables. We first assess how sharply inflation set by the government's policy (x_{t-1}) is estimated. Figure F1 displays the error bands, which are generated with both parameter uncertainty and exogenous random inflation shocks. Clearly, the bands themselves track the rise and fall of actual inflation very well.

The key insight into the intricate interactions between shocks and government beliefs is the perceived long-run tradeoffs under the Ramsey policy. Figure F2 reports the estimated

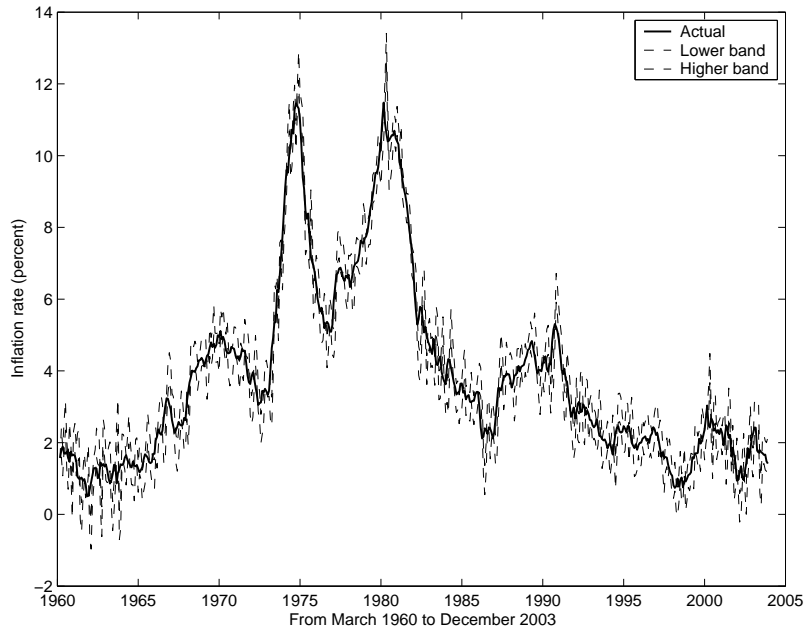


FIGURE F1. Inflation: actual vs one-step forecast with 90% error bands

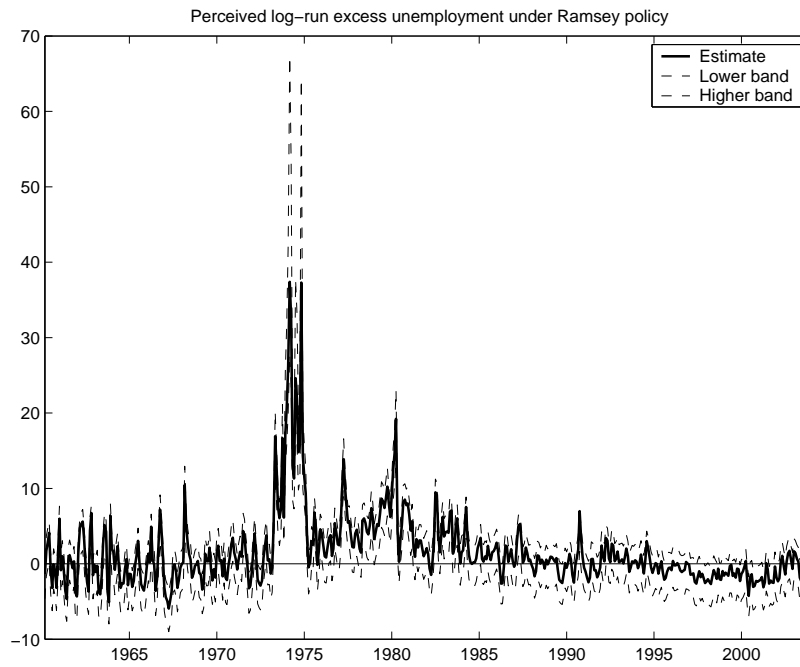


FIGURE F2. Perceived long-run excess unemployment under Ramsey policy of 2% with 90% error bands.

tradeoffs with 90% error bands. The large tradeoffs in the 70s are sharply estimated. Judging by the error bands, the small but sharp fluctuations of tradeoffs experienced in the 60s do not exist in the 90s; the government's perceived tradeoffs are smaller and stabler in the 90s.

REFERENCES

- BLINDER, A. S. (1998): *Central Banking in Theory and Practice*. MIT Press, Cambridge, Massachusetts.
- CHO, I.-K., N. WILLIAMS, AND T. J. SARGENT (2002): “Escaping Nash Inflation,” *Review of Economic Studies*, 69, 1–40.
- CHUNG, H. (1990): *Did Policy Makers Really Believe in the Phillips Curve? An Econometric Test*. University of Minnesota, Minneapolis, USA, ph.d. dissertation edn.
- COGLEY, T., AND T. J. SARGENT (2003): “Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.,” Manuscript, University of California (Davis) and New York University.
- (2004): “The Conquest of U.S. Inflation: Learning and Robustness to Model Uncertainty,” Manuscript, University of California (Davis) and New York University.
- GEWEKE, J. (1999): “Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication,” *Econometric Reviews*, 18(1), 1–73.
- HAMILTON, J. D., D. F. WAGGONER, AND T. ZHA (2003): “Normalization in Econometrics,” Manuscript, University of California (San Diego) and Federal Reserve Bank of Atlanta.
- IRELAND, P. (1999): “Does the Time-Consistency Problem Explain the Behavior of Inflation in the United States?,” *Journal of Monetary Economics*, 44(2), 279–292.
- KING, R. G., AND M. WATSON (1994): “The Post-War U.S. 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.
- PARKIN, M. (1993): “Inflation in North America,” in *Stabilizing the World Economy in the 1990s*, ed. by K. Shigehara. Bank of Japan, Tokyo.
- PRIMICERI, G. (2003a): “Time Varying Structural Vector Autoregressions and Monetary Policy,” Princeton University.

- (2003b): “Why Inflation Rose and Fell: Policymakers’ Beliefs and US Postwar Stabilization Policy,” Princeton University.
- SARGENT, T. J. (1999): *The Conquest of American Inflation*. Princeton University Press, Princeton, New Jersey.
- SARGENT, T. J., AND N. WILLIAMS (2003): “Impacts of Priors on Convergence and Escapes from Nash Inflation,” Manuscript, New York University and Princeton University.
- SIMS, C. A. (1988): “Projecting Policy Effects with Statistical Models,” *Revista de Analisis Economico*, 3, 3–20.
- SIMS, C. A., AND T. ZHA (1998): “Bayesian Methods for Dynamic Multivariate Models,” *International Economic Review*, 39(4), 949–968.
- (2004): “Were There Regime Switches in US Monetary Policy?,” Manuscript, Princeton University and Federal Reserve Bank of Atlanta.
- STAIGER, D., J. H. STOCK, AND M. W. WATSON (1997): “The NAIRU, Unemployment and Monetary Policy,” *Journal of Economic Perspectives*, Winter, 11(1), 33–49.

NEW YORK UNIVERSITY AND HOOVER INSTITUTION, PRINCETON UNIVERSITY AND NBER, FEDERAL RESERVE BANK OF ATLANTA