The American Economic Review 96(4), 2006

Shocks and Government Beliefs: The Rise and Fall of American Inflation

By THOMAS SARGENT, NOAH WILLIAMS, AND TAO ZHA*

We use a Bayesian Markov Chain Monte Carlo algorithm to estimate the parameters of a "true" data-generating mechanism and those of a sequence of approximating models that a monetary authority uses to guide its decisions. Gaps between a true expectational Phillips curve and the monetary authority's approximating nonexpectational Phillips curve models unleash inflation that a monetary authority that knows the true model would avoid. A sequence of dynamic programming problems 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 United States to an intricate interaction between the monetary authority's beliefs and economic shocks. Shocks in the 1970s made the monetary authority perceive a tradeoff between inflation and unemployment which ignited big inflation. The monetary authority's beliefs about the Phillips curve changed in ways that account for former Federal Reserve Chairman Paul Volcker's conquest of U.S. inflation. (*JEL* E24, E31, E52, N12)

Today, many statesmen and macroeconomists believe that inflation can largely be determined by a government monetary authority. Why, then, did the Federal Reserve Board preside over high U.S. inflation during the late 1960s and the 1970s? This paper answers these questions by estimating a model that allows discrepancies between a true data generating mechanism and a monetary authority's approximating model. Our model features a process that makes a sequence 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 computes 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 and uses the same structural model. But its estimates of that model change.[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 ways from the true data generating model, a version of Robert E. Lucas Jr.'s (1973) aggregate supply function used by Kydland and Prescott (1977) and many others. The first deviation is

* Sargent: Department of Economics, New York University, 269 Mercer Street, New York, NY 10003 (e-mail: thomas.sargent@nyu.edu); Williams: Department of Economics, Princeton University, Princeton, NJ 08544-1021 (e-mail: noahw@princeton.edu); Zha: Research Department, Federal Reserve Bank of Atlanta, Atlanta, GA 30309-4470 (e-mail: tzha@mindspring.com). We thank seminar participants at the NBER Summer Institute, Robert Chirinko, Timothy Cogley, Eric Leeper, James Nason, Ricardo Reis, William Roberds, Frank Schorfheide, Lars Svensson, and especially Christopher Sims for helpful criticisms and Jordi Mondria for excellent research assistance. We also thank two anonymous referees for detailed and helpful comments and suggestions. Sargent and Williams thank the National Science Foundation for (separate) research support. The views expressed herein do not necessarily reflect those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

[1] Sargent (1999) called it a Phelps problem.

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

[3] As does Finn E. Kydland and Edward C. Prescott's (1977) model of time-consistent suboptimal inflation.

that the monetary authority omits the public's rational expectation of inflation from its Phillips curve. By itself, this omission need not prevent the outcomes of our model from coinciding with those predicted by Kydland and Prescott, nor need it imply that the government's model is wrong in a way that could be detected even from an infinite sample. The reason is that, depending on the history of outcomes, the constant term and lagged rates of inflation and unemployment can stand in perfectly for the expected rate of inflation that the government has omitted from its Phillips curve.[4] If the monetary authority were to believe that the coefficients of its Phillips curve were 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 U.S. 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.[5]

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's rule. The covariance matrix **V** and the initial condition for the regression parameters in the monetary authority's Phillips curve become hyperparameters of a model that shapes evolution of the monetary authority's beliefs.[6] 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, actually governs inflation-unemployment dynamics. We use a Bayesian Markov Chain Monte Carlo (MCMC) algorithm to estimate statistics that describe the posterior distribution of these parameters of our model. We obtain a much better explanation of the monetary authority's inflation choices than earlier efforts to estimate similar models had achieved.

We find that our model fits the data better than a benchmark time series model. We use several criteria to compare the empirical performance of our theoretical model with those of some atheoretical Bayesian vector autoregressions (BVARs). Our model has better forecasting performance than BVARs over one-month, two-year, and four-year horizons. Formal model selection criteria, such as the Schwarz criterion and Bayes factors, strongly favor our model. Equally important, our model outperforms the BVARs in predicting several key turning points in the inflation time series. Finally, while the fit of our model is competitive with statistical models, our results yield important insights that help to *understand* the U.S. inflation experience, something a purely statistical model cannot. One essential feature accounting for our model's success in fitting the data is how our estimation procedure exploits the cross-equation restrictions that the government's Phelps problem imposes on the sequence of government beliefs about the empirical Phillips curve. These restrictions are very informative for estimating the key government belief parameters in **V**.

With particular a priori settings of the parameter innovation covariance matrix **V**, Christopher A. Sims (1988), Heetaik Chung (1990), Sargent (1999), and In-Ko Cho et al. (2002) all studied versions of our model.[7] When Chung and Sargent estimated their a-priori-fixed-**V**

[4] See Kydland and Prescott (1977) for a heuristic argument and Sargent (1999) for a demonstration that the outcome in a self-confirming equilibrium is identical with Kydland and Prescott's time-consistent outcome.

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

[6] As is true in a rational expectations model, the monetary authority's beliefs are outcomes, not free parameters.

[7] Sargent and Williams (2005) is an extended 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.

versions of our model, they obtained discouraging results. They did not come close to explaining the rise and fall of U.S. inflation in terms of a process of the monetary authority's learning about its Phillips curve.[8]

This paper estimates settings for **V** that attain substantial improvements in the model's ability to rationalize the choices made by the U.S. monetary authority.[9] The MCMC algorithm finds values for **V** that allow the model to reverse engineer a sequence of government beliefs about the Phillips curve which, through the intermediation of the Phelps problem, capture both the acceleration of U.S. 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's rule, and its decisions, via successive solutions of its Phelps problem. The monetary authority's views about parameter drift and its application of Bayes's rule add a source of history dependence to its procession of decisions which is absent in either the SCE or the Markov perfect equilibrium of Kydland and Prescott's model. The resulting interactions of shocks and monetary authority's beliefs form the basis for our explanation of the rise and fall of U.S. inflation.

The rest of the paper is organized as follows. Section I relates our findings to other work. In Section II, we lay out the model and discuss theoretical characterizations of it. Section III develops an econometric methodology for estimation, and Section IV reports the estimated results. In Section V, we present further empirical results, stress the importance of cross-equation restrictions via the Phelps problem, examine the forecasting performance of the model, conduct some counterfactual exercises, and explore some important implications. Section VI discusses the estimated model's long-run properties. Section VII concludes. Four appendices describe the data and provide technical details about our prior distribution and the posterior sampling scheme.

I. Relation to Recent Literature

Timothy Cogley and Sargent's (2005a) explanation of U.S. post-WWII inflation also features the interaction of a government learning process and a sequence of Phelps problems.[10] Cogley and Sargent's government applies Bayes's rule recurrently to estimate three Phillips curve models, only one of which is a rational expectations version of a natural rate model. Cogley and Sargent focus on the role of model uncertainty in policy making and take no stand on the true data generating mechanism. Their story is about how an almost discredited model that has even a very small posterior probability will nevertheless be very influential if it leads to very bad outcomes under the policies that would have been recommended if its posterior probability were *exactly* zero.[11]

Giorgio Primiceri (2005b) also develops a learning model to explain the rise and fall of U.S. inflation. He estimates his model on U.S. 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 story is that the monetary authority's mismeasurement of the natural rate of unemployment caused policy to be looser than policymakers intended. Primiceri's

[8] Previous 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.

[9] To rationalize the U.S. inflation experience, the random coefficients specification of the values of **V** and $\mathbf{P}_{1|0}$ that we estimate attributes beliefs to policymakers that differ from ones that they would have attained by running least squares period by period.

[10] Christopher Sims has pointed out that neither the present paper, nor any of the papers in the literature we survey, provides what one would really like: a statistically respectable measure of the uncertainty that attaches to these alternative explanations of the history of U.S. inflation. Because the models that we survey assume the form of a likelihood function, in principle one could treat each of them as submodels and calculate a posterior probability distribution over them. We agree that such a project would be interesting.

[11] The monetary authority in our model updates only one model, an outmoded one at that, in light of the rational expectations revolution. From their readings of minutes of the Federal Open Market Committee (FOMC), Christina Romer and David Romer (2002) infer that the Fed's learning process was confined to a primitive Phillips curve specification like the one we impute to the monetary authority. Their story about the evolution of Fed beliefs assigns no influence to rational expectations ideas.

main focus is on a backward-looking Keynesian model that has no explicit role for private sector expectations that respond to the government's decision rule, unlike our true model, which has private sector expectations responding in the best way to government decisions.[12] In the SCE of Primiceri's model, the government learns that the coefficients on inflation in the Phillips curve sum to one. This restriction is correct in his specification, but need not hold under rational expectations.[13] Primiceri finds that the model's transient learning dynamics are able to reproduce low-frequency features of post-WWII U.S. inflation-unemployment dynamics.

Our paper differs substantially from Primiceri's because our true data generating mechanism (DGM) is a rational expectations natural rate model. Policymakers' misspecified model can eventually converge to attain a self-confirming equilibrium in which the inflation-unemployment dynamics generated by the true DGM agree with those expected by the government along the equilibrium outcome process. Unlike Primiceri, we view the rational expectations natural rate theory and the associated SCE as a useful starting point. We build on Sims (1988), Chung (1990), Sargent (1999), and Cho et al. (2002), as generalized by Sargent and Williams (2005). These studies a priori adopted parameter specifications that opened a substantial gap between a Ramsey inflation outcome (the one that would be chosen by a government that knew the correct DGM) and the Nash inflation outcome that emerges from the SCE. The latter three contributions discovered *mean dynamics* which on average push outcomes toward the Nash inflation level and *escape dynamics* which recurrently push them toward the Ramsey outcome.

The present paper estimates key parameters that control the mean dynamics and the escape dynamics. Our empirical estimates teach us to deemphasize the empirical relevance of *both* the mean dynamics and the escape dynamics and instead to focus on 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 Alan S. Blinder's (1998) skepticism about whether that gap is quantitatively important for the monetary authority's decision problem.

II. The Model

We extend the model of Sargent and Williams (2005). There is a Lucas natural-rate version of the Phillips curve and a true inflation process:

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

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

where u_t is the unemployment rate, u^{**} is the natural rate of unemployment, 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 in which systematic monetary policy has neither short-run nor long-run effects on unemployment.[14] Equation (1) embodies a stronger form of "policy irrelevance" than do many of today's New Keynesian Phillips curves. In this paper, we ignore the nonneutralities present in those models and aim to reverse engineer a set of government beliefs that can explain the low-frequency swings in U.S. data while insisting that the true DGM have the strong policy irrelevance of the Lucas supply function. Section IV shows that our reverse-engineering succeeds quantitatively in tracking the post-WWII inflation data.

Equation (2) states that the government determines inflation up to a random shock. The public has rational expectations, so that $E_{t-1}\pi_t = x_{t-1}$. The government dislikes inflation and unemployment. The policy decision x_{t-1} solves the tth component of the following sequence of "Phelps problems":

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

[13] See Sargent (1999) for a discussion.

[14] If $\mathrm{abs}(\theta_0) > \mathrm{abs}(\theta_1)$, (1) becomes a version of a natural-rate Phillips curve that allows a serially correlated disturbance (Sargent, 1999).

$$\min_{\{x_{t-1}\}_{t=0}^{\infty}} \hat{E} \sum_{j=0}^{\infty} \delta^j((\pi_{t+j} - \pi^*)^2 + \lambda(u_{t+j} - u^*)^2), \tag{3}$$

where $\hat{E}$ represents the expectation formed with respect to the model (2) and

$$u_{t+j} = \hat{\boldsymbol{\alpha}}'_{t+j|t+j-1}\boldsymbol{\Phi}_{t+j} + \sigma w_{t+j}, \tag{4}$$

where π^* and u^* are the targeted levels of inflation and unemployment, both $\hat{\boldsymbol{\alpha}}_{t|t-1}$ and $\boldsymbol{\Phi}_t$ are $r \times 1$ vectors, w_t is an i.i.d. standard normal random variable, and (4) is the monetary authority's model of inflation-unemployment dynamics. The vector $\boldsymbol{\Phi}_t$ of regressors consists of lags of unemployment and inflation. The government's policy at time t is the best response function $x_{t-1} = h(\hat{\boldsymbol{\alpha}}_{t|t-1})'\boldsymbol{\Phi}_t$. By comparing (4) with the true DGM (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 $\hat{\boldsymbol{\alpha}}_{t|t-1}$ 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 estimated Phillips curve (4), not the true Phillips curve (1). An SCE is a vector of government beliefs $\boldsymbol{\alpha}_{\text{SCE}}$ that is consistent with what it observes in the sense of satisfying the population least squares orthogonality condition

$$E[\boldsymbol{\Phi}_t(u_t - \boldsymbol{\Phi}_t'\boldsymbol{\alpha}_{\text{SCE}})] = 0, \tag{5}$$

where the mathematical expectation is evaluated with respect to the probability distribution of u_t, π_t, and x_{t-1} induced by (1), (2), and the decision rule implied by the Phelps problem associated with $\boldsymbol{\alpha}_{\text{SCE}}$.

Self-confirming equilibrium outcomes agree with the time-consistent Nash equilibrium outcomes in which policymakers set inflation higher than the socially optimal Ramsey level (see Sargent, 1999).[15] Nash inflation is

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

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 object that restricts beliefs to be time-invariant and that forms a benchmark—and as it can turn out, a limit point—for our model. Unlike an SCE, the government updates its beliefs at each date in our model. In particular, the government bases $\hat{\boldsymbol{\alpha}}_{t|t-1}$, its mean estimate of the drifting parameter vector $\boldsymbol{\alpha}_t$, on the observations up to and including time $t - 1$ from the following (misspecified) econometric model:

$$u_t = \boldsymbol{\alpha}_t'\boldsymbol{\Phi}_t + \sigma w_t, \tag{7}$$

$$\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t-1} + \boldsymbol{\Lambda}_t, \tag{8}$$

where $\boldsymbol{\Lambda}_t$, uncorrelated with w_t, is an i.i.d. Gaussian random vector with mean zero and covariance matrix $\mathbf{V}$. Thus, the government believes that the true economy drifts over time. That is why it continually adapts its parameter estimates with nonvanishing weight on new observations. The innovation covariance matrix $\mathbf{V}$ governs the perceived volatility of increments to the parameters, and is a key component of the model. The mean estimate of $\boldsymbol{\alpha}_t$ for the econometric model (7)–(8) is

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

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

Let

$$\mathbf{P}_{t|t-1} \equiv \text{Var}(\boldsymbol{\alpha}_t | \mathcal{I}_{t-1}).$$

Given the government's model, the mean estimates are optimally updated via the special case of Bayes's rule known as the Kalman filter. Given $\hat{\boldsymbol{\alpha}}_{1|0}$ and $\mathbf{P}_{1|0}$, the Kalman filter algorithm updates $\hat{\boldsymbol{\alpha}}_{t|t-1}$ with the following formula:[16]

[15] As explained by Sargent (1999, chap. 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.

[16] Many learning models such as Sargent (1999) have focused on a recursive least squares learning (RLS) rule that is closely related to the Kalman filter. Sargent and Williams

(9) $$\hat{\boldsymbol{\alpha}}_{t+1|t} = \hat{\boldsymbol{\alpha}}_{t|t-1} + \frac{\mathbf{P}_{t|t-1}\boldsymbol{\Phi}_t(u_t - \boldsymbol{\Phi}_t'\hat{\boldsymbol{\alpha}}_{t|t-1})}{\sigma^2 + \boldsymbol{\Phi}_t'\mathbf{P}_{t|t-1}\boldsymbol{\Phi}_t},$$

(10) $$\mathbf{P}_{t+1|t} = \mathbf{P}_{t|t-1} - \frac{\mathbf{P}_{t|t-1}\boldsymbol{\Phi}_t\boldsymbol{\Phi}_t'\mathbf{P}_{t|t-1}}{\sigma^2 + \boldsymbol{\Phi}_t'\mathbf{P}_{t|t-1}\boldsymbol{\Phi}_t} + \mathbf{V}.$$

An important issue is whether the learning process will converge to a self-confirming equilibrium in which the discrepancy between the government's model and the true DGM vanishes for outcomes that occur thereafter with positive probability. To summarize what we have known about this, we scale the innovation covariance matrix as $\mathbf{V} = \varepsilon^2\hat{\mathbf{V}}$, for $\varepsilon > 0$. Key analytical results from Sargent and Williams (2005) that highlight possible outcomes of the government's learning process are:

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

III. Empirical Methodology

The theoretical results indicate how very different outcomes can emerge from different government beliefs. The 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 $\mathbf{V}$, 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 $\hat{\boldsymbol{\alpha}}_{1|0}$ to avoid overparameterization. We have set the parameters δ, λ, and π^* to values taken from the existing literature. We shall discuss the fixed values of u^* and $\hat{\boldsymbol{\alpha}}_{1|0}$ in Section IV.

Group all other free structural parameters as

$$\boldsymbol{\phi} = \{v^*, \theta_0, \theta_1, \tau_1, \zeta_1, \zeta_2, u(\mathbf{C}_P), u(\mathbf{C}_V)\},$$

where $v^* = u^{**}(1 - \tau_1)$, $\mathbf{C}_P$ and $\mathbf{C}_V$ are upper triangular such that $\mathbf{P}_{1|0} = \mathbf{C}_P'\mathbf{C}_P$ and $\mathbf{V} = \mathbf{C}_V'\mathbf{C}_V$, and $\zeta_1 = 1/\sigma_1^2$ and $\zeta_2 = 1/\sigma_2^2$ represent the precisions of the corresponding innovations. The notation $u(\mathbf{C}_P)$ or $u(\mathbf{C}_V)$ means that only the upper triangular parts of $\mathbf{C}_P$ or $\mathbf{C}_V$ are among the free parameters. Notice that among the parameters in $\boldsymbol{\phi}$, $\{v^*, \theta_0, \theta_1, \tau_1, \zeta_1, \zeta_2\}$ describe the true data generating mechanism while $\{u(\mathbf{C}_P), u(\mathbf{C}_V)\}$ describe the government's beliefs.

The structural parameter $\zeta = 1/\sigma^2$ is not free. It is clear from (9), (10), and (14) that if we scale $\mathbf{V}$ and $\mathbf{P}_{1|0}$ 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. Sargent and Williams (2005) impose the restriction $\zeta = \zeta_1$, a normalization that implies that the policymakers correctly decompose the observed variation in the unemployment into variation in the regressors and variation due to exogenous shocks.[17] Note, however, that an SCE requires the orthogonality conditions (5), but not necessarily the equality restriction $\zeta = \zeta_1$. Indeed, the examples of Sims (1988) allow $\zeta \neq \zeta_1$. In what follows we set $\kappa = 0.01$, which makes the variability of our estimated $\mathbf{V}$ the same order of magnitude as the variability of the data. That implies that the standard deviation of the government's regression error σ is smaller by a factor of ten than the standard deviation exogenous unemployment shocks σ_1. As noted

(2005) show that RLS can be approximated by a Kalman filter with $\mathbf{V}$ proportional to $\sigma^2 E(\boldsymbol{\Phi}\boldsymbol{\Phi}')^{-1}$.

[17] This normalization has the advantage that it makes limiting results easier to derive.

above, this implies that large escapes from the SCE may frequently occur.

As we have noted, Sargent and Williams (2005) show that whether monetary policy stays close to a path associated with a self-confirming equilibrium (and when it does not), and how it evolves over time are both very sensitive to the model's parameters (especially the government's belief about the covariance matrix for the drifting coefficients). This sensitivity is what enables us sharply to estimate key structural parameters, including the elements of $\mathbf{V}$.

To take into account parameter uncertainty, we employ the Bayesian method and develop an MCMC algorithm that breaks $\boldsymbol{\phi}$ into three separate blocks: $\boldsymbol{\theta}$, $\{\zeta_1, \zeta_2\}$, and $\boldsymbol{\varphi}$, where

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

and $\boldsymbol{\varphi} = \{u(\mathbf{C}_P), u(\mathbf{C}_V)\}$. The prior pdf of $\boldsymbol{\phi}$ can be factored as:

$$p(\boldsymbol{\phi}) = p(\boldsymbol{\theta})p(\boldsymbol{\varphi})p(\zeta_1, \zeta_2).$$

The prior distributions of both $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$ take the Gaussian form:

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

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

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

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

From equations (1) and (2) one can see that the Jacobian transformation from $\sigma_1 w_{1t}$ and $\sigma_2 w_{2t}$ to u_t and π_t is equal to 1. It follows that the likelihood function is:

$$\mathcal{L}(I_T|\boldsymbol{\phi}) = \frac{\zeta_1^{T/2}\zeta_2^{T/2}}{(2\pi)^{T/2}} \exp\left\{-\frac{1}{2}\sum_{t=1}^{T}[\zeta_1 z_{1t}^2 + \zeta_2 z_{2t}^2]\right\}, \tag{14}$$

where z_{1t} and z_{2t} are the functions of $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$:

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

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

where the optimal decision rule depends on $\boldsymbol{\varphi}$.

The posterior pdf of $\boldsymbol{\phi}$ is proportional to the product of the likelihood (14) and the prior $p(\boldsymbol{\phi})$:

$$p(\boldsymbol{\phi}|I_T) \propto \mathcal{L}(I_T|\boldsymbol{\phi})p(\boldsymbol{\phi}). \tag{15}$$

The posterior distribution of $\boldsymbol{\phi}$ can be simulated by using a Gibbs sampler, i.e., by alternately sampling from the following conditional posterior distributions:

$$p(\boldsymbol{\theta}|I_T, \zeta_1, \zeta_2, \boldsymbol{\varphi}),$$

$$p(\zeta_1, \zeta_2|I_T, \boldsymbol{\theta}, \boldsymbol{\varphi}),$$

$$p(\boldsymbol{\varphi}|I_T, \boldsymbol{\theta}, \zeta_1, \zeta_2).$$

Appendix C tells how to sample from each of these conditional distributions.

IV. Reverse Engineering Estimation

In this section, we present our results. Using the monthly U.S. data described in Appendix A and the prior specified in Appendix B, we estimate $\boldsymbol{\phi}$ 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 $u^* = 1$. Kydland and Prescott (1977) set $\pi^* = 0$. Because in practice central banks seem to target positive inflation rates, we set $\pi^* = 2$. The value of u^* is set at a value low enough to allow Nash inflation to be higher than

TABLE 1—POSTERIOR ESTIMATES OF MODEL PARAMETERS

Maximum log value of likelihood (multiplied by prior): 564.92

Estimates of coefficients in true Phillips curve and inflation process with 68-percent and 90-percent 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 $\mathbf{P}_{1|0}$

0.1087	0.1432	0.0225	−0.2540	−0.0093	−0.1015
0.1432	0.1937	0.0296	−0.3398	−0.0119	−0.1359
0.0225	0.0296	0.0047	−0.0526	−0.0019	−0.0211
−0.2540	−0.3398	−0.0526	0.5990	0.0213	0.2396
−0.0093	−0.0119	−0.0019	0.0213	0.0008	0.0085
−0.1015	−0.1359	−0.0211	0.2396	0.0085	0.0958

Estimate of $\mathbf{V}$

0.0823	−0.0778	0.0092	0.0498	−0.0081	−0.4141
−0.0778	0.0814	0.0003	−0.0509	0.0194	0.6859
0.0092	0.0003	0.0299	0.0012	0.0370	0.7207
0.0498	−0.0509	0.0012	0.0320	−0.0105	−0.3996
−0.0081	0.0194	0.0370	−0.0105	0.0514	1.0064
−0.4141	0.6859	0.7207	−0.3996	1.0064	25.8831

Ramsey inflation.[18] Setting the unemployment target closer to the natural rate has no effect on our main results.

We set the initial belief $\hat{\boldsymbol{\alpha}}_{1|0}$ at the regression estimate obtained from the presample data from January 1948 to December 1959.[19] We tried to fix $\mathbf{P}_{1|0}$ at the value that scales up and down the presample regression estimate $\hat{\sigma}^2(\boldsymbol{\Phi}'\boldsymbol{\Phi})^{-1}$, but the fit was bad. Similarly, fixing $\mathbf{V}$ at the value estimated from a presample-estimated covariance matrix with different scales does not improve the poor fit. Departing from Sargent (1999), therefore, we estimate the government's prior beliefs $\mathbf{P}_{1|0}$ and $\mathbf{V}$ within the sample. Our MCMC or maximum likelihood algorithm reverse engineers the empirical Phillips curve at each date that, in conjunction with the Phelps problem, rationalizes that date's inflation rate. Estimating $\mathbf{P}_{1|0}$ and $\mathbf{V}$ gives us the flexibility to succeed in this reverse engineering. Moreover, this flexibility is arguably reasonable. We take the view that the presample data are informative about the government's subjective point estimates (which we fix), but that they substantially understate the government's subjective uncertainty about coefficient innovation volatility $\mathbf{V}$ (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 $\mathbf{V}$.

We report the posterior estimate of $\boldsymbol{\phi}$ (evaluated at the peak of the posterior pdf) in Table 1, along with the 68-percent and 90-percent probability intervals around the estimate.[20] In

[18] Blinder (1998) emphasizes that the source of time inconsistency in Kydland and Prescott's (1977) Phillips curve example is their specification that $u^{**} \neq u^*$ in the monetary authority's preferences (3). His experience as vice chairman of the Federal Reserve led Blinder to question whether the FOMC perceived there to be much of a gap between u^{**} and u^*.

[19] In an earlier draft, we followed Chung (1990) and estimated this belief from 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.

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

our estimation and inference, the regressor vector in the government's Phillips regression (4) is:

$$\mathbf{\Phi}_t = [\pi_t \quad \pi_{t-1} \quad u_{t-1} \quad \pi_{t-2} \quad u_{t-2} \quad 1]'.$$

Accordingly, the corresponding regression coefficient vector in (4) is $[\alpha_1 \ldots \alpha_6]'$ where the subscript $t|t-1$ is suppressed here.

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, we estimate the natural rate of unemployment u^{**} in equation (1) to be 6.1 with wide probability intervals, a finding that is consistent with the confidence interval in the statistical model of Douglas Staiger et al. (1997). We find only weak responses of unemployment to inflation surprises (θ_0 and θ_1) and they are statistically insignificant according to 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^*. Therefore, outcomes close to those associated with the limit point of the mean dynamics are close to the Ramsey outcome. Unemployment is by itself a persistent series and the persistence is tightly estimated.

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 estimate (evaluated at the peak of the posterior pdf) is outside the 90-percent interval.

The estimated $\mathbf{P}_{1|0}$ 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 fourth element) implies that the government is relatively uncertain about this coefficient, which affects the uncertainty about other coefficients, even though their marginal variances are relatively small.

The estimated $\mathbf{V}$ shows strong correlations among the innovations to the coefficients on current and lagged inflation variables. As discussed above, the scale of $\mathbf{V}$ is pinned down only *relative* to the government's regression error variance σ^2. Our $\mathbf{V}$ is large relative to σ^2 (which, recall, is $0.01\sigma_1^2$), implying that the government is willing to adjust its beliefs quickly in response to recent data. The constant term has the largest variance, which can be interpreted as reflecting its uncertainty about the natural rate of unemployment. The uncertainty in the constant affects the coefficients on the lagged unemployment variables because of their high correlations, but it has a small influence on the coefficients on inflation in the government's model. Because $\mathbf{V}$ is not small (or equivalently σ^2 *is*), 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 a negligible difference between the SCE and π^*. We show in Section VIB that 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 $\mathbf{V}$ permits frequent escapes to low inflation rates.

The inflation path produced by the government's inflation policy is plotted against the actual path in Figure 1, and one-step forecasts of unemployment are plotted against the actual path in Figure 2.[21] It is evident from these figures that the model explains the low-frequency movements of inflation well; so well, in fact, that it is difficult to discern the difference between the series.[22] 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.

These forecast errors are comparable to those from BVAR models with the standard prior settings proposed by Sims and Zha (1998). The root mean square error (RMSE) and the mean absolute error (MAE) are 0.225 and 0.179 for our model, 0.397 and 0.297 for the BVAR with one lag (BVAR(1)), and 0.357 and 0.272 for the BVAR with 13 lags (BVAR(13)). As another comparison model, we consider the case where the government's model is a Keynesian Phillips

[21] Inflation policy (x_{t-1}) is sharply estimated. Although we do not plot the error bands around the estimated x_{t-1} to avoid visually clustering Figure 1, they are quite tight and track the rise and fall of actual inflation well.

[22] These empirical results provide a formal justification for Ireland's (2005) key assumption that persistent changes occurred in the Federal Reserve's inflation target in the 1970s.

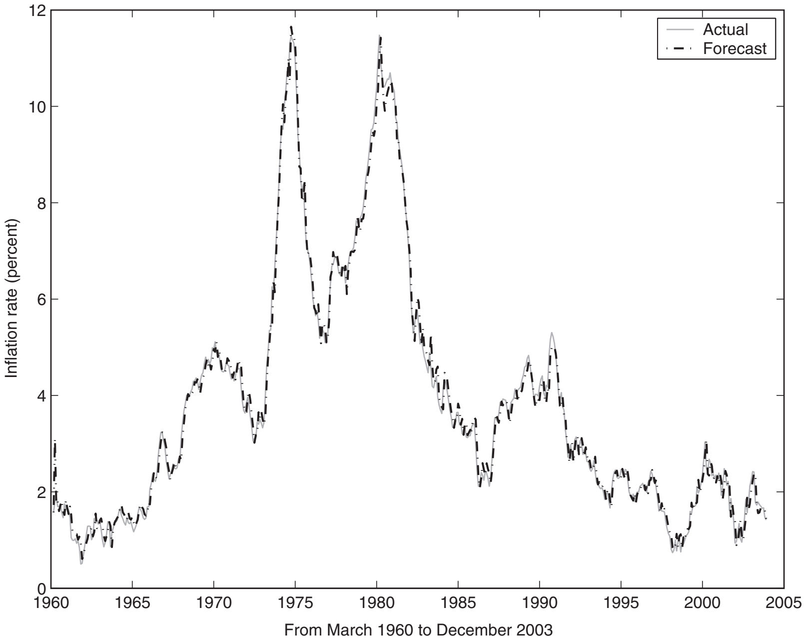

FIGURE 1. INFLATION: ACTUAL VERSUS ONE-STEP FORECAST (I.E., GOVERNMENT CONTROLLED INFLATION)

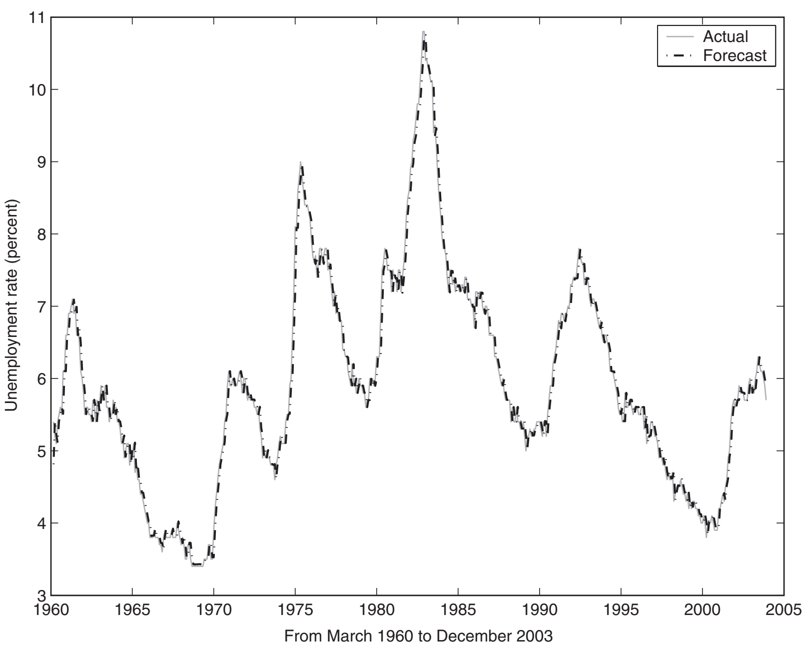

FIGURE 2. UNEMPLOYMENT RATE: ACTUAL VERSUS ONE-STEP FORECAST

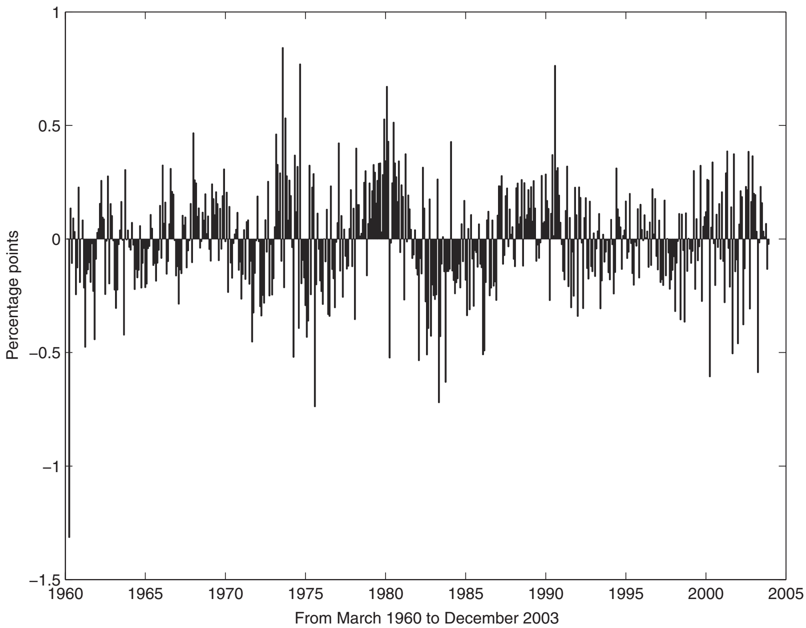

FIGURE 3. DIFFERENCES BETWEEN ACTUAL VALUES AND ONE-STEP FORECASTS OF INFLATION

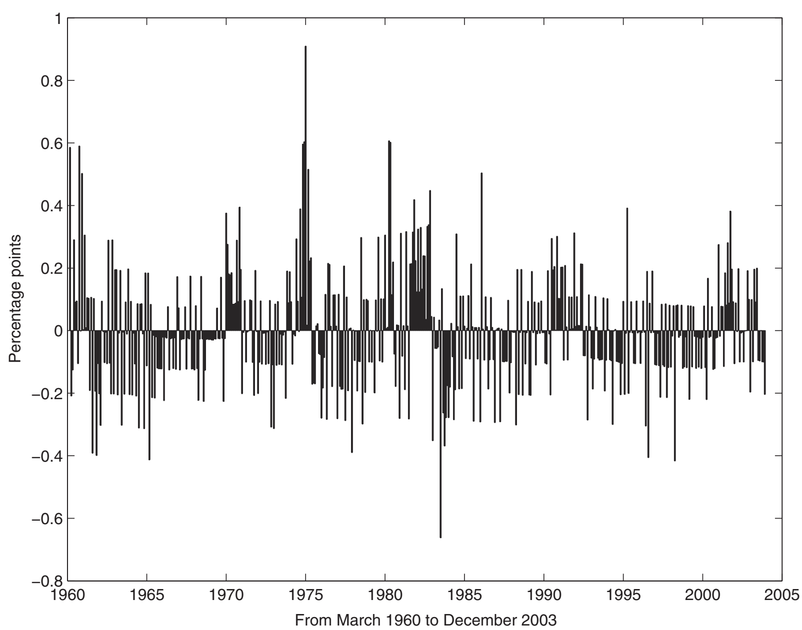

FIGURE 4. DIFFERENCES BETWEEN ACTUAL VALUES AND ONE-STEP FORECASTS OF UNEMPLOYMENT

curve that puts inflation on the left side of the regression, as discussed in Robert G. King and Mark W. Watson (1994) and Sargent (1999). The resulting RMSE is 0.478 and MAE is 0.332, substantially larger than our model's. The poor fit of the Keynesian model is consistent with the findings of Cogley and Sargent (2005a).

We also use the Schwarz criterion (SC) to compare maximum log values of the likelihood multiplied by the prior among our learning model, the BVAR(1), and the BVAR(13). The SC, sometimes called the Bayesian information criterion, adjusts the log likelihood by the number of degrees of freedom times log of sample size divided by two. This criterion is very useful because it can be readily computed from the estimates reported in Table 1 and because, under standard regularity conditions, it will converge to zero (or infinity) if the posterior odds ratio converges to zero (or infinity) as the sample size increases. We follow Sims and Zha (2006) and use the likelihood multiplied by the prior instead of likelihood itself, because models with a large number of parameters are better characterized by the likelihood multiplied by a prior. The same asymptotic reasoning that justifies the SC based on the likelihood applies to the likelihood multiplied by a prior. The SC value is 564.92 for our model, 313.98 for the BVAR(1), and 309.37 for the BVAR(13). Our learning model appears to dominate the two atheoretical models.

To see whether these asymptotic results hold in finite samples, we compute the marginal data density (MDD) for our learning model, using the modified harmonic mean method described in John Geweke (1999) and Propositions 1 and 2 in Appendix C. The log MDD value is 424.75. In comparison, the log MDD value is 172.05 for the BVAR(1) and 244.65 for the BVAR(13).[23] As measured by Bayes factors (which would put essentially zero weight on the BVARs), our learning model again dominates the BVARs.[24]

Higher Bayes factors, however, do not necessarily imply that our learning model outperforms BVARs in predicting the rise and fall of inflation. In Sections VC and VE, therefore, we compare the performances of both our learning model and BVARs in forecasting *longer-term* inflation. There we show that for forecasting low-frequency movements in inflation, our model performs as well as or better than the BVARs. 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 result had not been achieved in previous work (Sims, 1988; Chung, 1990; Sargent, 1999).

V. Further Empirical Results

A. *Shocks and Beliefs*

In our model, the rise and fall of inflation is driven by the Phelps problem in conjunction with the government's belief in an exploitable tradeoff between inflation and unemployment, which leads to a high inflation rate in the early 1970s. But then occasional sequences of stochastic shocks lead the government, at least temporarily, to believe that it can cut inflation with no rise in unemployment, which leads to rapid disinflation in the early 1980s. During these episodes, the government learns a version of the natural rate theory in which the sum of the coefficients on inflation is nearly zero in its model, reflecting a vertical long-run Phillips curve.

The evolution of the government's updated beliefs is displayed in Figure 5. The sum of the coefficients on inflation becomes very negative in the early 1970s and stays quite negative until the late 1970s. Although the sum of the inflation coefficients is still negative in the 1980s, it is small enough to induce policymakers to decide

[23] These results differ from the SC results. It is well known, however, that the Schwarz criterion tends to favor VAR models with shorter lags.

[24] The MDD values may be sensitive to the priors. But the differences between the MDD values of our learning model and the BVAR models are large enough for us to conclude that our model appears to dominate. The computed MDD for our model is based on one million posterior draws, which consume six days on a Pentium-IV PC desktop. Using the method of Whitney K. Newey and Kenneth K. West (1987), we obtain the numerical standard error for the log value of the estimated MDD, which is about 0.99. It is known, however, that the standard error computed this way may underestimate the uncertainty around the estimated marginal likelihood. From independent sequences of posterior draws, we find that the estimated value of log MDD can vary on the order of five.

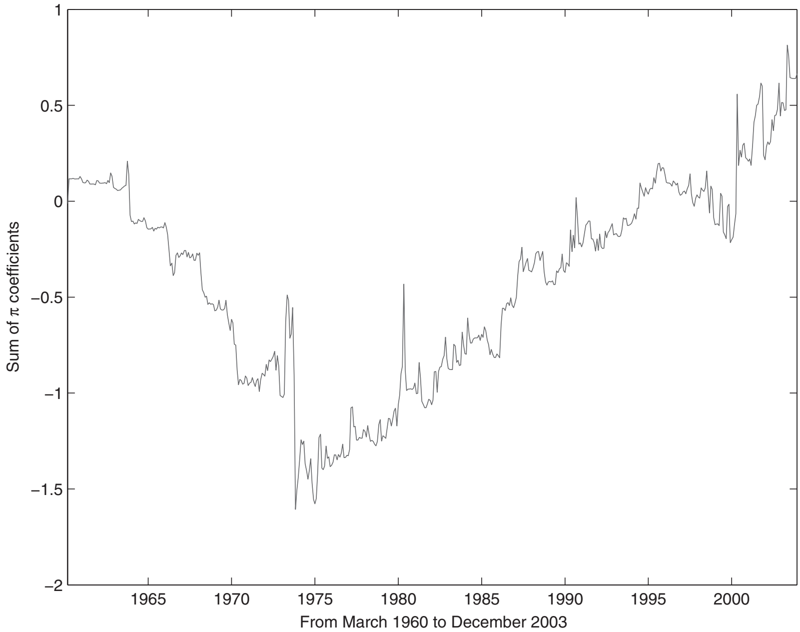

FIGURE 5. EVOLUTION OF THE GOVERNMENT'S BELIEFS

to cut inflation without worrying much about costs in unemployment.

Figure 6 displays the subjective covariations in the drift innovations of some key functions of 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).[25] 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 combination leads to a high perceived tradeoff between unemployment and inflation in December 1973.

In contrast, at the point associated with the SCE (indicated by the symbol "∘" in the second row of Figure 6), the estimated constant coefficient is small and the sum of the inflation coefficients is near zero, providing 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 that the constant coefficient and the sum of the inflation coefficients fall far within the northwest 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 postwar 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 1970s.

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. If $1 - \alpha_3 - \alpha_5$ and α_6 frequently have opposite signs because they are negatively correlated, the government would tend to predict unemployment below the natural rate because

[25] See Sargent (1999, chap. 5) for how the sum of coefficients on π affects the advice rendered by the Phelps problem.

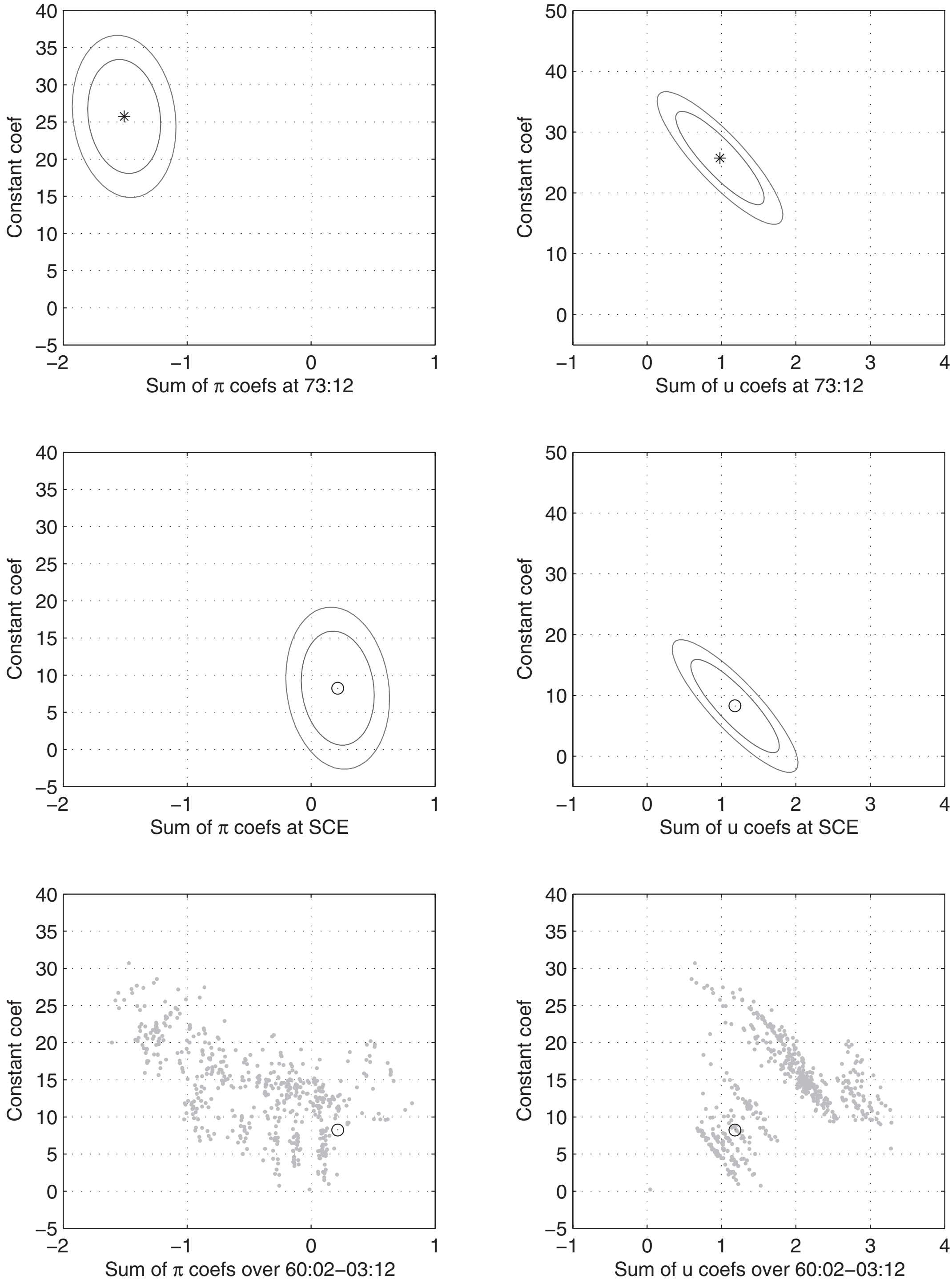

Figure 6. 68-Percent and 90-Percent Probability Ellipses about Key Parameters in the Government's Phillips Curve

Notes: 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 scatter 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 (∘) in the second and third rows depicts SCE values, which also equal limiting estimates from the mean dynamics.

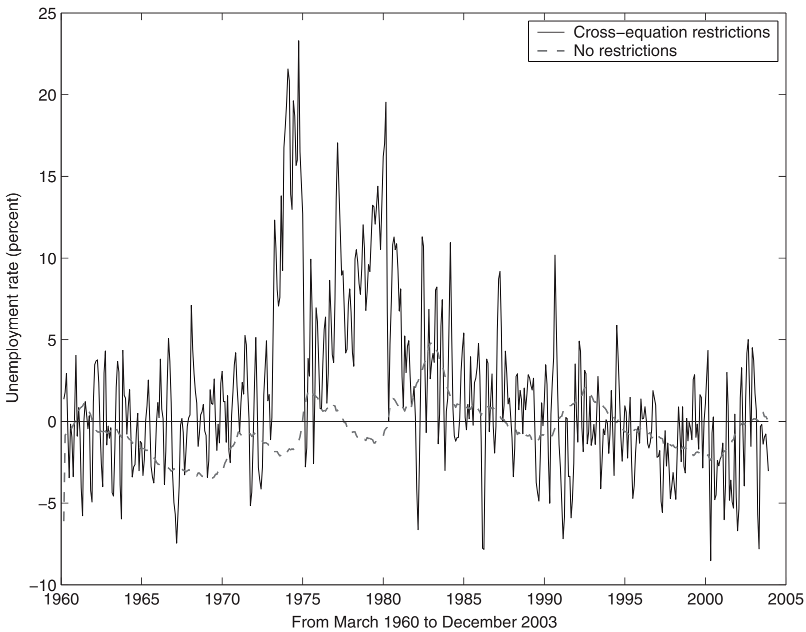

FIGURE 7. PERCEIVED EXCESS UNEMPLOYMENT UNDER THE RAMSEY POLICY OF 2-PERCENT INFLATION

Note: According to the government's beliefs, with and without imposing the cross-equation restrictions.

of a large value of α_6. This in turn would prompt the Phelps planner to disinflate a lot to stabilize his objective function, thereby causing volatile fluctuations of inflation. These volatile outcomes occur when these two parameters fall in the southeast and northwest quadrants. Fortunately for U.S. inflation outcomes, our historical estimates have been concentrated around the northeast quadrant, as shown in the bottom-right graph. It is only in out-of-sample simulations that we enter the more volatile regions. Exposure to those out-of-sample possibilities is a byproduct of the large **V** (relative to σ^2) which, in conjunction with the Phelps problem, Bayes's rule prompts us to use to reverse engineer the government's choice of inflation.

The belief parameters discussed above are key inputs to the government's perceived sacrifice ratio. To assess the government's perceived cost of having a stable low inflation policy, we construct an artificial time series of the unemployment rate that the government would have expected had it kept inflation constant at the Ramsey level of $\pi^* = 2$ percent throughout the sample.[26] Figure 7 plots the difference between this projected unemployment rate and our estimate of the natural rate. This provides a measure of the government's perceived sacrifice ratio, the expected unemployment in excess of the natural rate associated with the Ramsey inflation policy.[27] Here we see that, throughout the 1970s, the government's model implied that substantial increases in unemployment would result from a low inflation policy.[28] It wasn't until the early 1980s that this ratio fell nearly to zero for a sustained period of

[26] In particular, at each date we feed the actual past unemployment rates and 2-percent inflation into the government's Phillips curve and project the current unemployment rate.

[27] Note that our measure of the sacrifice ratio differs from a more conventional one that gives the cost of disinflating from a current inflation rate. Instead, ours is a full-sample measure that is independent of current inflation.

[28] A temporary drop in this sacrifice ratio around 1976 led to a temporary decline in inflation around that time. See Cogley and Sargent (2005a) for a story in which the government was deterred from stabilizing in the mid-1970s because it attached a small positive probability to a

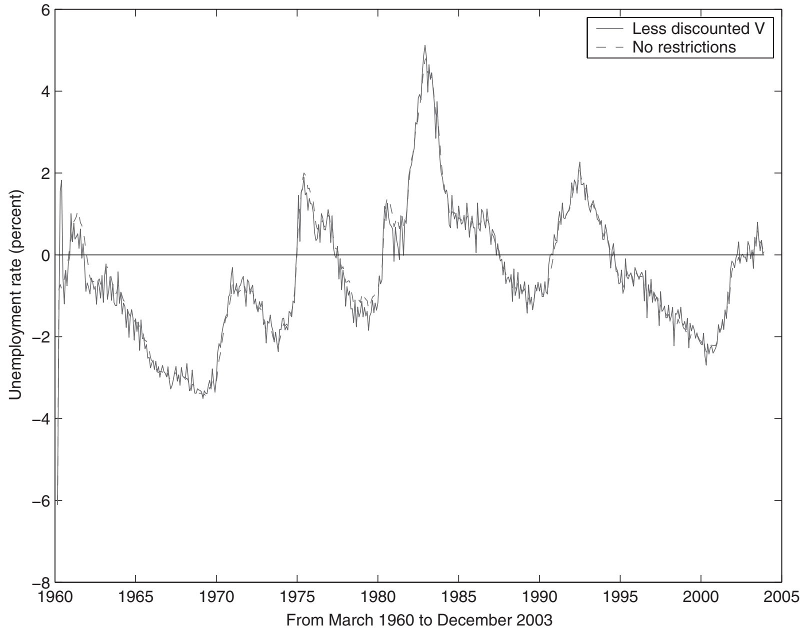

FIGURE 8. PERCEIVED EXCESS UNEMPLOYMENT UNDER THE RAMSEY POLICY OF 2-PERCENT INFLATION

Note: According to the government's beliefs with a scaled down **V**, and without imposing the cross-equation restrictions.

time, at which time the disinflation commenced. This point will be reiterated in Sections VE and VF where we present longer-term forecasts and counterfactual paths around that time.

Figure 7 also plots the corresponding sacrifice ratio when, as described in the next section, we do not impose the cross-equation restrictions. Here we see that the sacrifice ratio is much lower, even negative, for much of the sample, implying that such a model will not be able to reproduce the rise and fall of inflation that is observed in the data. We discuss this in more detail in the next section. The difference between our model and a model that does not impose the cross-equation restrictions is accounted for almost entirely by the large scale of our estimated **V**. If we scale down our **V** by a factor of 1×10^{-4}, the sacrifice ratios implied by the resulting beliefs are nearly identical to the ones obtained when we do not impose the cross-equation restrictions (see Figure 8). Our large estimated **V** makes the government's beliefs and its policy very sensitive to recent data, an essential key feature that allows our model to explain the evolution of U.S. inflation.

B. *Importance of Cross-Equation Restrictions*

As we've already noted, the flexibility that a large **V** (or small σ^2) gives our model is crucial for giving us the ability to reverse engineer government beliefs that, intermediated by the Phelps problem, account for the government's decisions about the predictable part of inflation x_{t-1}. In particular, our findings tell us to attribute the empirical failure of previous work with similar models by Chung and Sargent to the fact that they assumed a particular form for the key matrix **V** in (8) that governs the innovations to the parameters in the government's model.

To highlight the importance of **V**, we can estimate **V** (and $\mathbf{P}_{1|0}$) directly with (7) and (8), thereby abstaining from imposing the Phelps problem. These estimates can serve as a bench-

---

model that assigned high unemployment costs to a rapid disinflation.

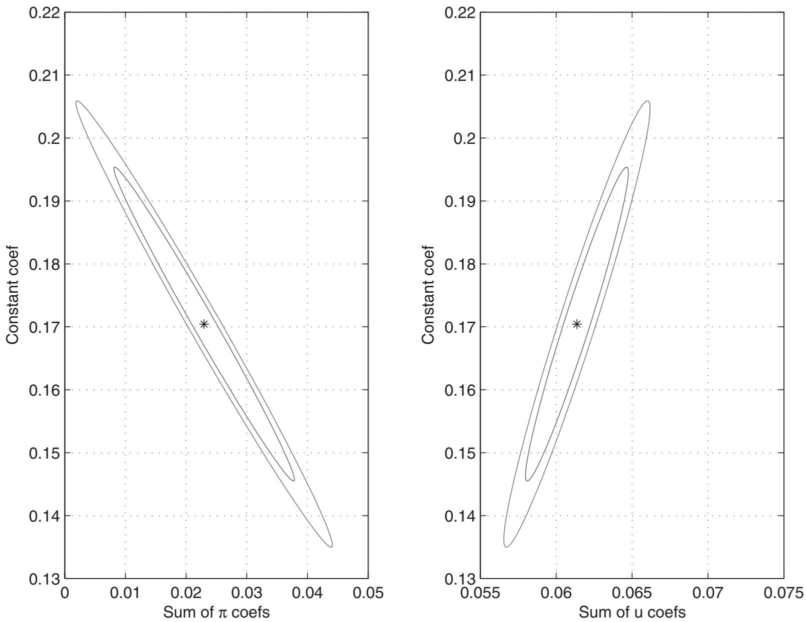

FIGURE 9. 68 PERCENT- AND 90 PERCENT-PROBABILITY ELLIPSES ABOUT KEY PARAMETERS IN THE GOVERNMENT'S PHILLIPS CURVE

Notes: Derived from the estimated **V** without imposing cross-equation restrictions. The asterisks mark the estimates of these belief parameters at 73:12 (when inflation was quite high).

mark for what impacts on **V** occur from our imposing the cross-equation restrictions via the Phelps problem discussed in Section VA.

Figure 9 displays the covariations in the key belief parameters when the restrictions from the Phelps problem are *not* imposed. Compared to Figure 6, where the restrictions from the Phelps problem are imposed, the ellipses in Figure 9 are very tight—so tight that if we were to combine Figure 9 and the first row of graphs in Figure 6, the tight Figure 9 ellipses would appear as short thin lines. Furthermore, the SCE values are far outside the Figure 9 ellipses.

We have already discussed theoretical reasons that make the **V** matrix so important and how different specifications of it affect the speed, direction, and stability of the learning dynamics. The **V** depicted in Figure 9 and those imposed by Chung and Sargent differ substantially from what we estimate when we impose the cross-equation restrictions induced by the Phelps problem. In particular, the **V**'s of Chung and Sargent are smaller in overall scale (again relative to σ^2) and have somewhat different correlations among parameters. These specifications constrain how learning could occur, and diminish the variation in the data that can be explained by evolving government beliefs.

Figure 10 shows what happens when we reestimate the model in the fashion of Chung and Sargent, imposing our estimate of **V** from Figure 9. The fit deteriorates substantially. The government's optimal policy completely misses the two peaks in inflation in the 1970s, which is what Sargent (1999) found.[29] This is consistent with the implied sacrifice ratio shown in Figure 7 above, which shows essentially no tradeoff between unemployment and inflation. Chung (1990) and Sargent (1999) found that with their

[29] If 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, the fit would be as poor as Figure 10. Similarly, if the originally estimated **V** in Section VA is scaled down by, say, 0.01 so that inflation dynamics are governed by the SCE, the implied inflation policy would completely miss the rise and fall of actual inflation.

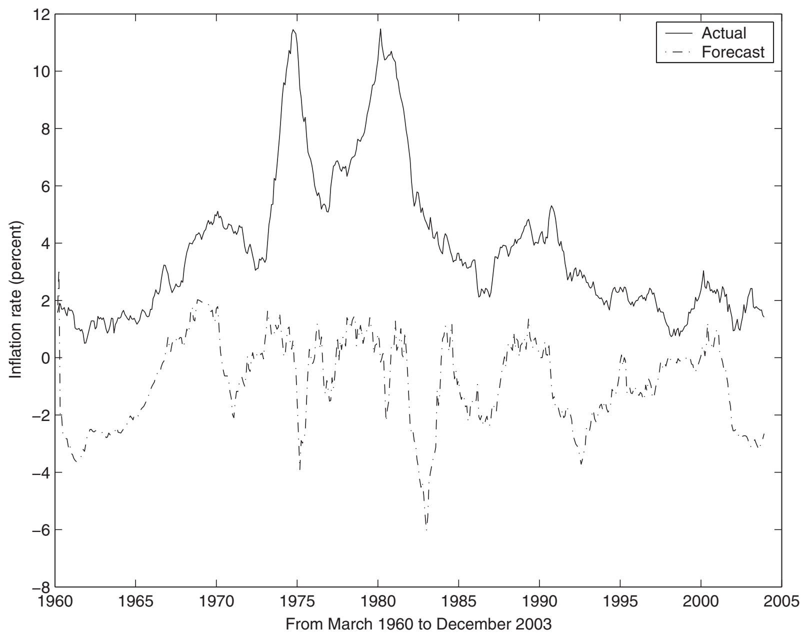

FIGURE 10. ACTUAL INFLATION AND ONE-STEP PREDICTION FROM THE BENCHMARK MODEL

Note: **V** is estimated without imposing the cross-equation restrictions.

choices of **V**, the government should have cut inflation much earlier than actually occurred.

Our results show how that outcome came from attributing to the government particular beliefs about how its model changes over time. By imputing to the government the necessary "openness to recent data" required by the cross-equation restrictions called for by the Phelps problem, the rise and fall of inflation can be much better explained by the evolution of government's beliefs in response to a particular sequence of shocks in the 1970s and 1980s. For someone who hopes or believes that the FOMC took the long view and did not highly discount data beyond the recent past, this large **V** (or again, small σ^2) could be viewed as disappointing or surprising. However, Brian Ironside and Robert J. Tetlow (2005) document large and consequential changes in the properties of the FRB/US model reported by the Fed staff from July 1996 to November 2003, including among them significant changes in the inflation-employment sacrifice ratio. Our large estimated value of **V** is consistent with the findings of Ironside and Tetlow (2005). For what it is worth, our large estimated **V** is also consistent with our own reading of the drifting views that we detect in our readings of FOMC transcripts.[30]

C. *Longer-Horizon Inflation Forecasts*

Longer-term forecasts of inflation play an important part in policy discussions at FOMC meetings. At each FOMC meeting, Federal Reserve economists prepare a report called the Greenbook which forecasts various economic variables over the two-year horizon. How well would our learning model do in producing two-year-ahead forecasts of inflation throughout the sample, as compared to the BVAR(13) model?

Figure 11 depicts the two-year-ahead median forecasts of inflation from our model and also

[30] One can infer from reading historical records of the FOMC that decision makers spent enormous amounts of time evaluating current economic conditions and that policy deliberations were dominated by interpretations of very recent changes in economic data. Even in the Greenspan era, policymakers' beliefs seemed to be heavily influenced by new developments (see various chapters in Henry W. Chappell et al., 2005). Our reverse-engineering estimate of **V** quantifies the FOMC's preoccupation with recent data in the context of a *formal* model.

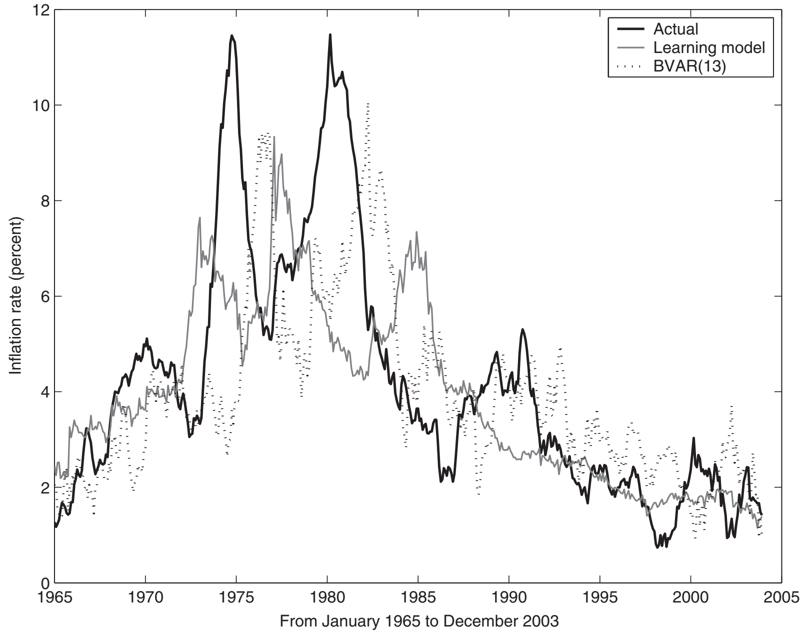

FIGURE 11. TWO-YEAR-AHEAD INFLATION FORECASTS: LEARNING MODEL VERSUS BVAR(13)

ones from a BVAR(13).[31] The forecast and actual values are aligned in such a way that if the forecast were on target, the values would coincide. As shown in the figure, the learning model produces forecasts that differ substantially from the BVAR.[32] Our learning model predicts the first two rises of inflation between about one and two years too early and a third small rise of inflation about five years too early. And it predicts a permanent fall of inflation after 1985 with less forecasting volatility than the BVAR(13). By contrast, the two-year-ahead forecasts of inflation from the BVAR(13) seem to lag the rise and fall of actual inflation.

Figure 12 traces these predicted accelerations of inflation to features of the government's beliefs that lead the Phelps planner to expect to "step on the gas" each of these three times. The left graph in the first row shows that the sum of inflation coefficients moves from left to right over time, getting more negative and prompting the government to step on the gas. In the right graph, one can see that the sum of unemployment coefficients and the constant coefficient are in the northeast quadrant, indicating that the inflation forecast is stable for this period, as discussed in Section VA. The second and third rows show similar patterns, with differences in how negative the sum of inflation coefficients gets over time. The left graph in the fourth row, however, reveals a completely different story. The sum of inflation coefficients moves toward zero over time and then passes into positive territory. Thus, the government faces, at most, weak inflation-unemployment trade-offs. These results explain why, after a third run-up of actual inflation between 1986 to 1990, the government would not want to step on the gas. Interestingly, the BVAR continues to predict a run-up even after 1990.

[31] At each time t, we first draw a sequence of structural shocks w_{1t+k} and w_{2t+k} defined in (1) and (2) for $k = 1, \ldots, 24$. Conditioning on the estimated values of the structural parameters, the estimated beliefs at t, and the data I_t defined in Section II, we then employ (1) and (2) to generate the forecasts u_{t+k} and π_{t+k} by recursively solving the inflation policy via the Phelps problem. We repeat this simulation 1,000 times and calculate the median of all simulated values of π_{t+24}. This computation takes about 40 hours on a Pentium-IV PC desktop.

[32] The RMSE and MAE are 1.9292 and 1.3233 for the learning model, 2.3939 and 1.6809 for the BVAR(1), and 2.0861 and 1.4617 for the BVAR(13).

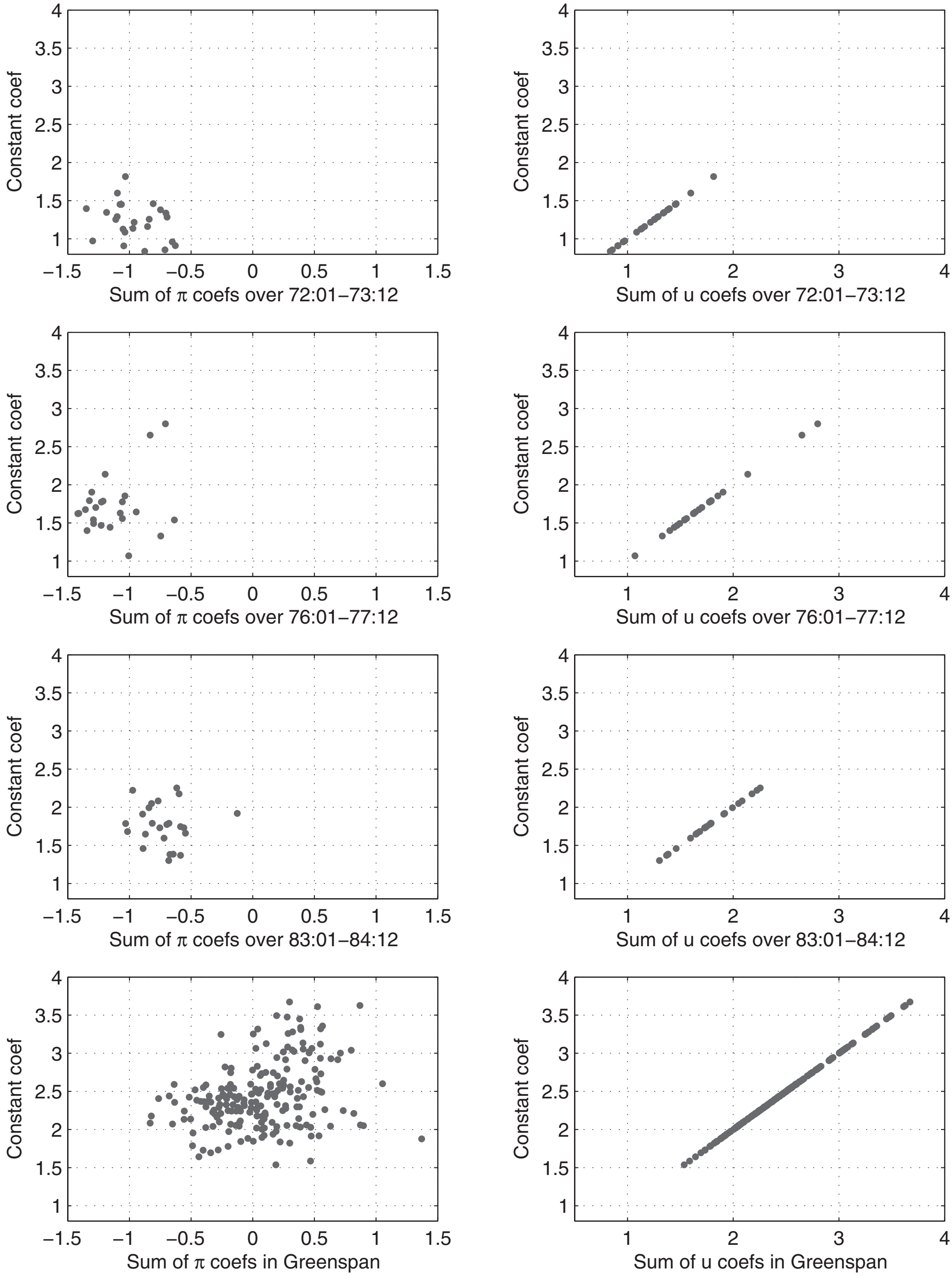

FIGURE 12. ESTIMATES OF KEY STEP-ON-GAS PARAMETERS OVER THE THREE PREDICTED RUN-UP PERIODS AND OVER THE GREENSPAN ERA

Notes: The first row shows the evolution of these belief parameters for 72:01–73:12 (the first predicted run-up period); the second row for 76:01–77:12 (the second predicted run-up period); the third row for 83:01–84:12 (the third predicted run-up period); and the fourth row for 87:07–03:12 (the Greenspan era).

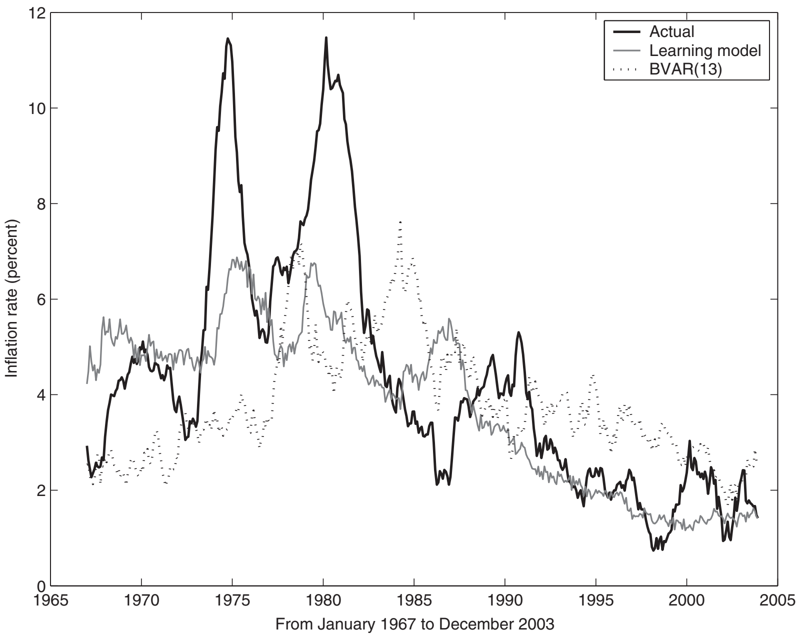

FIGURE 13. FOUR-YEAR-AHEAD INFLATION FORECASTS: LEARNING MODEL VERSUS BVAR(13)

D. *Good Low-Frequency Outcomes*

Figure 13 displays the four-year-ahead predictions from our model and the BVAR(13). Neither model predicts the magnitude of the rises of inflation that occurred. But our model captures the timings of the first two rises almost perfectly, while the predictions of the BVAR(13) again lag behind. The RMSE and MAE are 1.761 and 1.241 for our model, 2.838 and 2.195 for the BVAR(1), and 2.433 and 1.820 for the BVAR(13). Our model's four-year forecast errors are smaller than its two-year forecast errors, while the forecast errors from the BVARs are larger for the four-year horizon than for the two-year horizon.

E. *Two Peaks and an Enduring Decline*

To reinforce the results in the last section, we now analyze in further detail how the model forecasts the two peaks of inflation in the 1970s and the sharp decline in the early 1980s. We look at both the point forecasts and the associated distributions at various forecast horizons, conditioning on the estimated values of the structural parameters. 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 repeat 5,000 simulations of 50 periods.[33] We then plot the actual experienced inflation and the median forecast along with 68-percent and 90-percent probability bands. In each plot, the initial condition is shown as date zero, from which we look forward 50 periods.

The left column of graphs in Figure 14 reports the forecasts from our model. The top panel on the left starts in January 1973 when inflation was at a very low level (3.3 percent). This is also when we say that the government most overestimated the trade-off between inflation and unemployment (see Figure 5). According to the model, the government exploited the trade-off and pushed up inflation to lower unemployment. The model predicts a steadily rising inflation path as high as 10 percent toward the end of the four-year horizon (the upper-90-percent band), and gives little probability to a lower inflation rate in the medium run.

Due to a sequence of shocks, the inflation path reached its peak earlier than the model

[33] Adding uncertainty in the parameters would widen our forecast bands only a little.

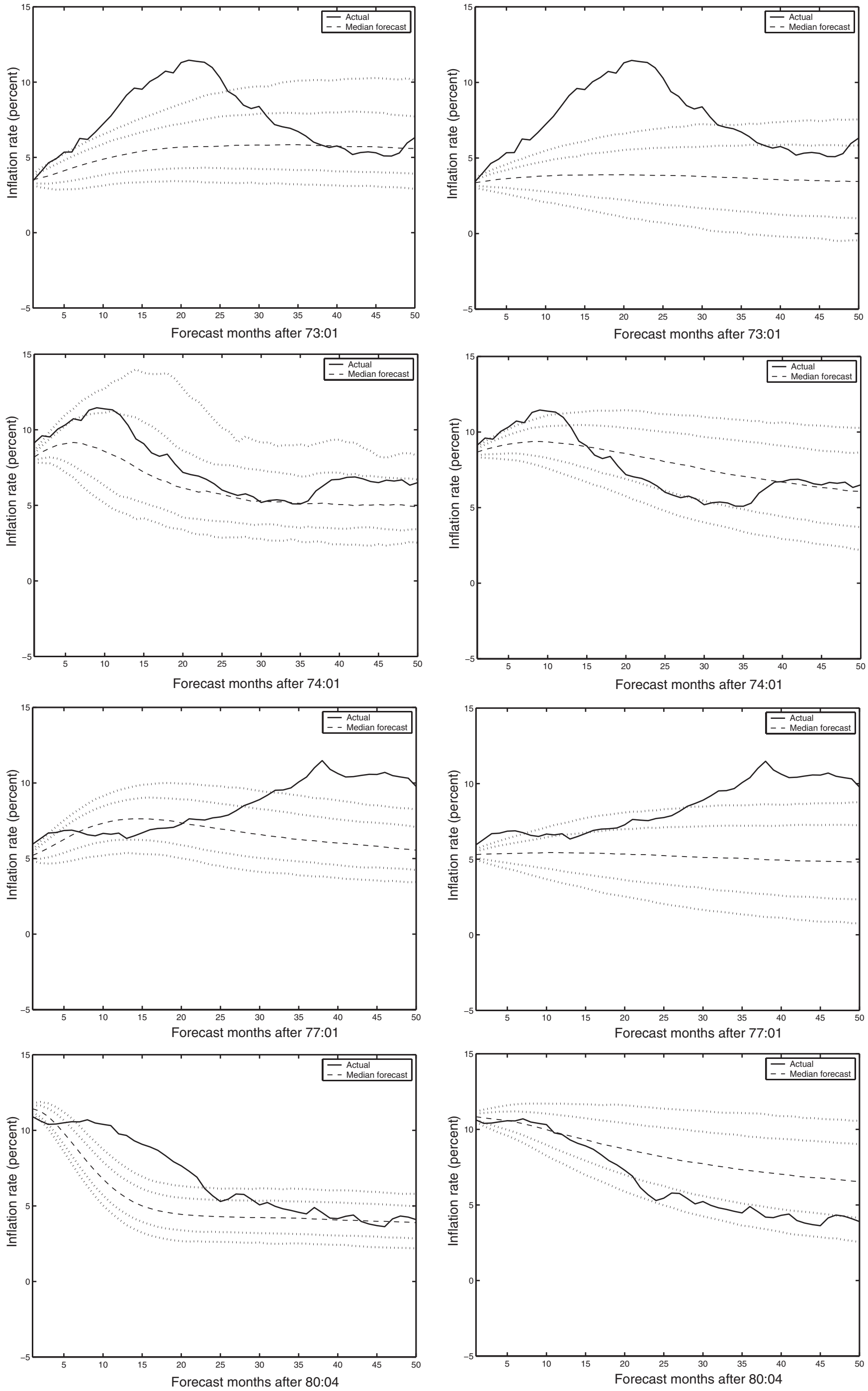

FIGURE 14. DYNAMIC FORECASTS OF INFLATION WITH 68-PERCENT AND 90-PERCENT ERROR BANDS

Notes: From our learning model (left column of graphs) and from BVAR(13) (right column of graphs), using as initial estimated conditions at 73:01, 74:01, 77:01, and 80:04. The two inner dotted lines around the dashed line represent the pointwise 0.68 probability bands; the two outer dotted lines correspond to the 0.90 probability bands.

predicts. But this is a treacherous period in which to predict, and as we show later in this section, our model's prediction of rising inflation compares favorably to predictions coming from alternative statistical models.

A year later, in January 1974, which is shown in the second panel from the top in the left column of Figure 14, inflation had continued upward, now reaching 8.4 percent. 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, shown in the third panel from the top, 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 a later date, the model assigns an overwhelming probability to higher inflation and the upper 90 percent reaches as high as 10 percent.

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 predicted that inflation was very likely to go higher than it actually did, or they failed to predict the fall of inflation. The bottom panel on the left column of Figure 14 displays the forecast from our learning model. While actual inflation declines at a somewhat slower speed than the model predicts in 1980 and 1981, the forecast of a fast decline in inflation is remarkable. The model's prediction is especially good farther along in the forecasting period. Unlike many forecasting models, our model gives almost no probability to rising inflation in the medium horizon, because the trade-off between inflation and unemployment by then is not high enough for the government to pursue double-digit inflation.

We now compare the model's forecasts with those from the BVARs. Above, we compared the fit of our model against BVARs with 1 and 13 lags. Although by some measures of fit the BVAR(1) performs better than the BVAR(13), with only one lag there are relatively few dynamics in the predictions. Thus we focus here on the BVAR(13). The right column of graphs in Figure 14 shows the forecasts of inflation at the various dates from the BVAR(13), as in the counterparts from our model in the left column of the figure. The 68-percent and 90-percent 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 we applied to our learning model. The forecasts at 73:01 from the BVAR(13) clearly fail to predict any rise of inflation with a significant probability. Moreover, the error bands are relatively wide, giving half a 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 give probability half to a decline of inflation, while the forecasts from our learning model in the left column place a vast majority of probability on rising inflation. For the forecast at 84:04, the BVAR(13) predicts a decline of inflation. But our learning model predicts a much sharper decline of inflation with narrow bands, while the BVAR assigns considerable probability to higher inflation than the actual path. Overall, our learning model performs as well or better than the BVARs in explaining the rise and fall of inflation at these crucial dates.

F. *Counterfactual Exercises*

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. All the results in this section condition on estimates of the historical shocks of unemployment and inflation that we infer 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 early 1964, but by the end of 1973 the sum of the inflation coefficients is mostly negative. Such continual adaptation of beliefs toward 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

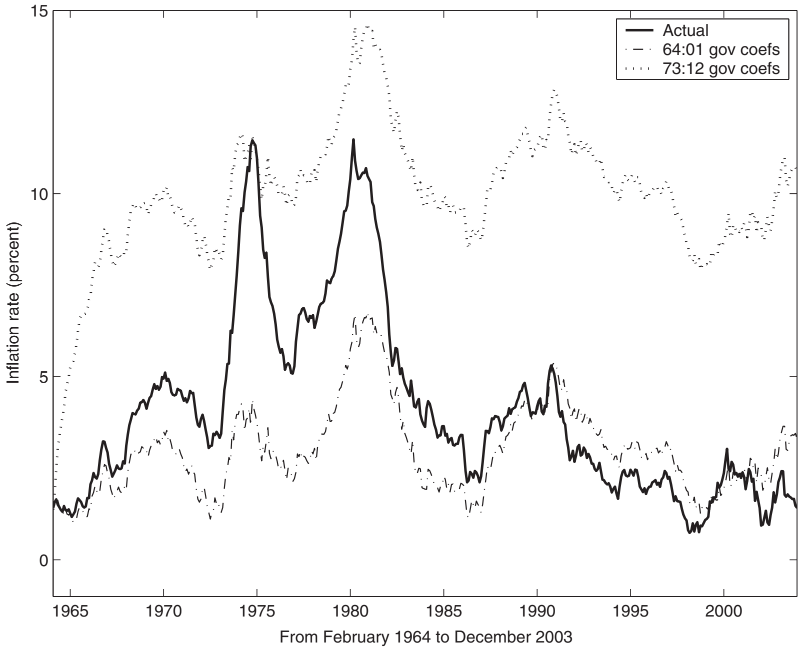

FIGURE 15. INFLATION DYNAMICS WITH FIXED BELIEFS AT THE DATES 64:01 AND 73:12

government's beliefs had been frozen at the 64:01 initial condition. As shown in Figure 15, the inflation path would have been smoother and avoided much of the two large run-ups of actual inflation in the 1970s. 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 15 shows that inflation would have been much higher than was actually experienced throughout the sample and would have continued to stay around 10 percent.

Figure 16 displays a 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. Figure 16 shows that if the government's belief at 1979:10 had been fixed throughout the rest of the history, inflation would have come down to 5 percent 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 important role that we assign to adapting government beliefs in the process of achieving lower inflation. With the same sequence of historical shocks, actual inflation came down and remained low, as shown by the inflation path in Figure 16. 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 trade-off.

These exercises suggest that while the rise of inflation in the 1970s was caused by the government's misperceiving the Phillips-curve relationship, the fall of inflation in the 1980s can be explained by an econometric policy evaluation procedure that embodies adaptive beliefs. The changes in beliefs over time that we estimate do not necessarily imply changes over time in a linear policy rule that we could estimate by regressing x_t on its own lagged values and on the lagged values of inflation and unemployment. Because our estimate of 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, 2005b; Primiceri, 2005a; Sims and Zha, 2006).

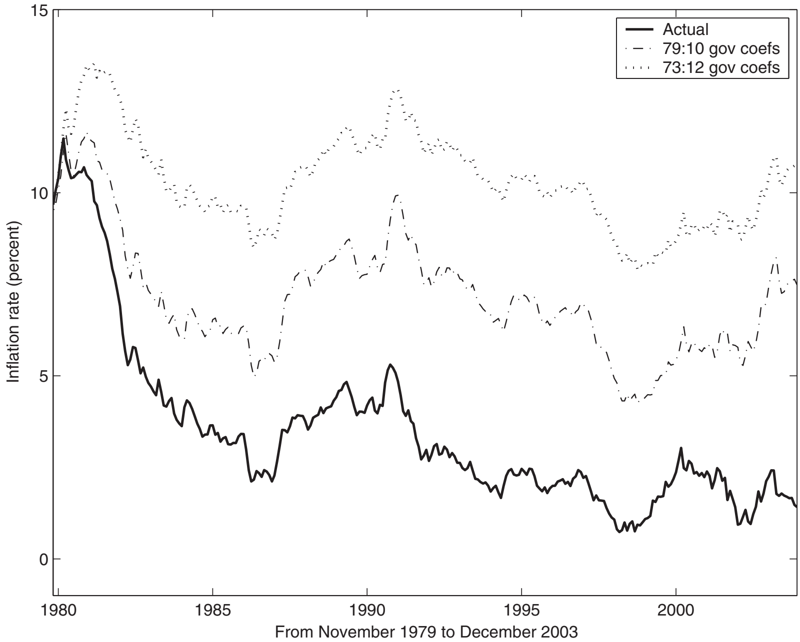

FIGURE 16. INFLATION DYNAMICS WITH FIXED BELIEFS AT THE DATES 79:10 AND 73:12

VI. Model Properties

We have shown that our model is capable of tracking the post-WWII inflation data. It is also important to examine the long-run properties of the model to see if the government's adapting beliefs will eventually support near-Ramsey outcomes. We first use the small variation limits of Sargent and Williams (2005) to obtain analytical asymptotic results. Then we discuss the convergence of our baseline model to a limit distribution.

A. *Small Variation Limits*

While it is difficult to obtain explicit convergence results for arbitrary $\mathbf{V}$, for small $\mathbf{V}$'s the beliefs drift at a slower rate, allowing us to approximate their evolution with a differential equation. In particular, as in Sargent and Williams (2005), we let $\mathbf{V} = \varepsilon^2\hat{\mathbf{V}}$ and study limits as $\varepsilon \to 0$. However, $\mathbf{P}_{t|t} \to \mathbf{0}$ as $\varepsilon \to 0$, so we define a scaled matrix $\hat{\mathbf{P}}_{t|t} = \mathbf{P}_{t|t}/\varepsilon$ that does not vanish. Sargent and Williams show that as $\varepsilon \to 0$, the sequence $\{\boldsymbol{\alpha}_{t|t}, \hat{\mathbf{P}}_{t|t}\}$ generated by (9)–(10) converges weakly to the solution of the following ODEs:

$$\dot{\boldsymbol{\alpha}} = \mathbf{P}E[\boldsymbol{\Phi}_t(u_t - \boldsymbol{\Phi}_t'\boldsymbol{\alpha})], \tag{16}$$

$$\dot{\mathbf{P}} = \sigma^{-2}\hat{\mathbf{V}} - \mathbf{P}E(\boldsymbol{\Phi}_t\boldsymbol{\Phi}_t')\mathbf{P}, \tag{17}$$

where the expectations are calculated for fixed $\boldsymbol{\alpha}$. 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 (16)–(17) the *mean dynamics* because they govern the expected evolution of the government's beliefs. If the ODEs have a stable point, then the government's beliefs will converge to it as $\varepsilon \to 0$ and $t \to \infty$. Note from (16) 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 when the data are generated by the true DGM.

In Figure 17 we plot trajectories of the mean dynamics for some functions of the parameters describing the government's beliefs, starting from the initial conditions at the beginning of the sample. Evidently, the mean dynamics converge to a stable self-confirming

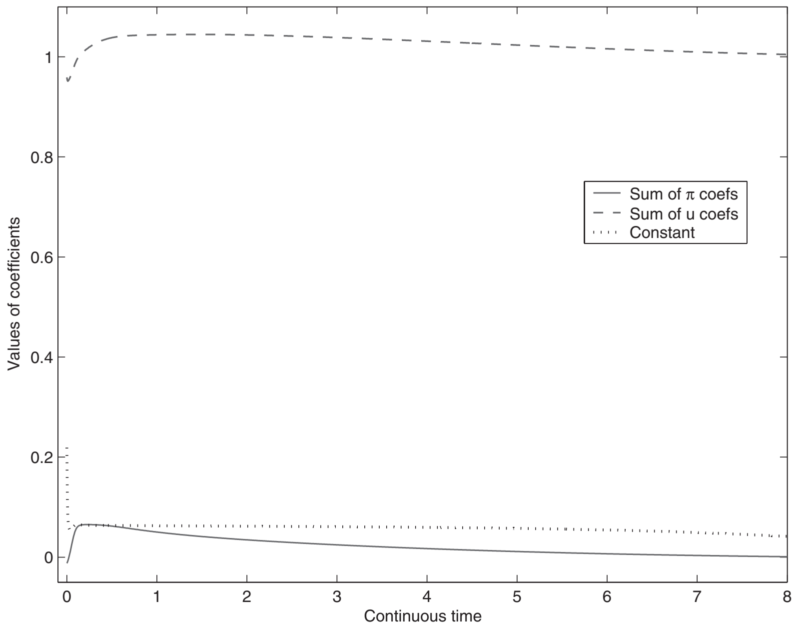

FIGURE 17. MEAN DYNAMICS FOR THE BASELINE ESTIMATES, INITIALIZED AT THE START OF THE SAMPLE

equilibrium. The self-confirming equilibrium beliefs are:

$$\boldsymbol{\alpha}_{\text{SCE}} = [-0.0008 \quad -0.0000 \quad 0.9725 \quad 0.0000 \quad 0.0165 \quad 0.0688].$$

In the SCE, the government knows the true value of θ_0, the effect of current inflation on unemployment. In the SCE, the government believes in a small trade-off between inflation and unemployment, and so sets inflation only slightly above the Ramsey level. In particular, the mean inflation rate in the SCE is 2.24 percent instead of the Ramsey level of 2 percent.

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

B. *Convergence to Near-Ramsey*

What are the long-run implications of the estimated $\mathbf{V}$? The large estimated value of $\mathbf{V}$ suggests that one would expect escapes from the SCE to be frequent even if the inflation rate at the Nash equilibrium were much higher. To illustrate this point, we change θ_0 from its estimated value of -0.0008 to -1.0, the value used by Sargent (1999), while keeping all other parameters fixed at the values we estimated. This implies that the Nash inflation rate is around 10 percent, while the socially optimal Ramsey level remains at 2 percent. As can be seen from Figure 18, inflation tends to be high, but the large time-variation of the drifting beliefs implied by our estimated $\mathbf{V}$ allows the dynamics to escape to low inflation repeatedly, and there is no tendency for inflation to stay for long at the high level. Thus, our $\mathbf{V}$ matrix is consistent with repeated escapes in the long run, but they are difficult to detect under our estimates because our estimate of θ_0 implies such a low sacrifice ratio.

To elaborate on this point, Figure 19 shows the inflation dynamics for simulations of 30,000 months starting at different estimated initial

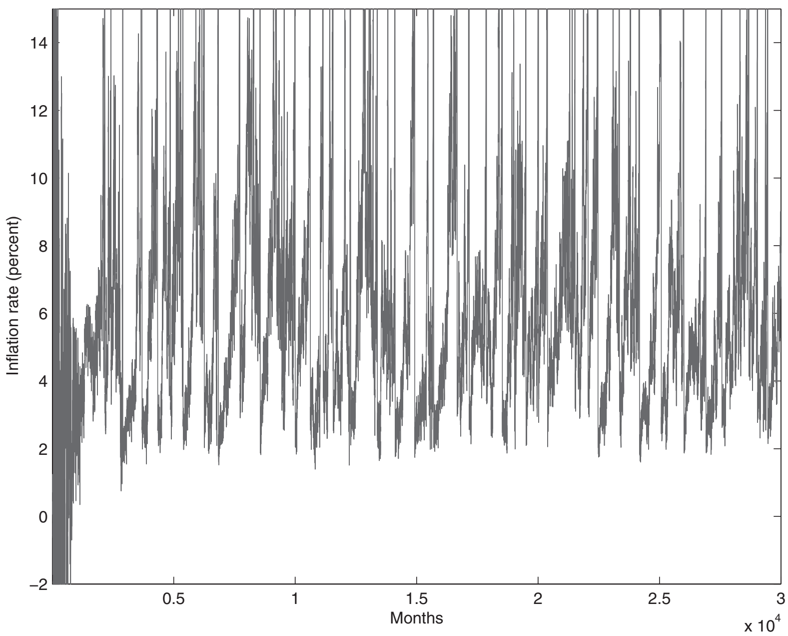

FIGURE 18. LONG-RUN INFLATION DYNAMICS WITH $\theta_0 = -1.0$ USING THE END-OF-SAMPLE ESTIMATE AS THE INITIAL CONDITION

conditions: 1960:03 (the beginning of the sample), 1973:12 (the date when both inflation and the perceived trade-off are quite high), 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 an American inflation process that was temporarily off the SCE equilibrium. As shown in the lower-left panel, we are likely to see some high inflation in the near future, but such high inflation will be caused purely by exogenous random shocks to inflation, so long as the government continues to see no tradeoff between inflation and unemployment (see the lower-right panel of Figure 19). The government's beliefs are volatile for a while, but eventually the sum of coefficients on inflation converges to near zero.[34] Consequently, the mean dynamics suggest that inflation converges to around 2 percent.

While the large inflation volatility in the out-of-sample forecasts is discomforting, the future outcomes would be much better if the government were to discount data less heavily. In particular, when we scale down our estimated **V** matrix by 1×10^{-4} and project out of sample, we obtain much more plausible predictions, as shown in Figure 20. Starting from the end of 2003, the model predicts a disinflation to roughly zero, followed by convergence to roughly 2 percent. These long-run properties foster a view of U.S. monetary history as a process of continual learning before inflation

[34] 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 northwest quadrant of the graph discussed in Section VA. If the sum of the inflation coefficients is negative, one can see 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.

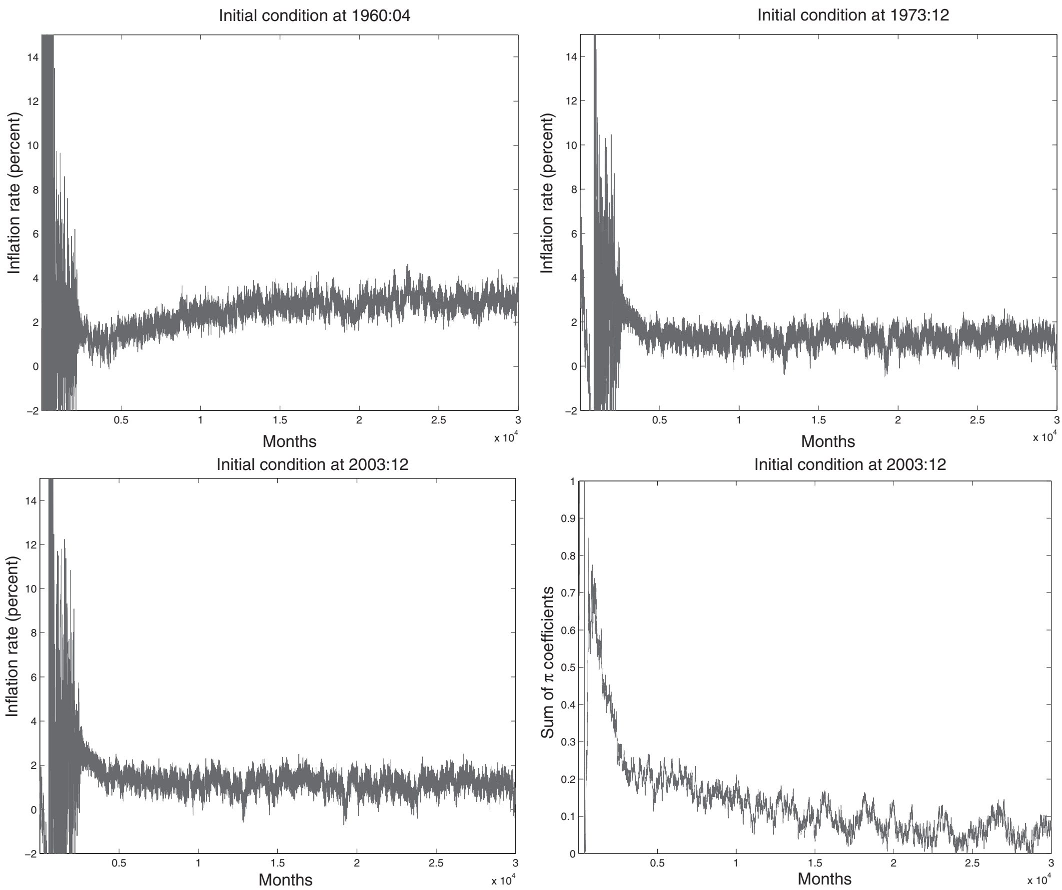

FIGURE 19. INFLATION DYNAMICS IN A LONG SEQUENCE OF MONTE CARLO SIMULATIONS, USING THE DIFFERENT ESTIMATED INITIAL CONDITIONS

becomes stable around the Ramsey outcomes. This transition process would be much improved if the government were now to take a longer horizon and put less weight on incoming data.

VII. Conclusion

Our estimates attribute the differing inflation outcomes over the postwar period to changes over time in the monetary authority's beliefs. Our empirical results suggest an interpretation that differs from the work we build on. Sargent (1999) and Cho et al. (2002) suggested that U.S. 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 (2005), these outcomes also occur in our model when we arbitrarily 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). With our estimates, however, it appears that oscillations between the Nash and Ramsey levels of inflation, driven alternately by the mean dynamics and then the escape dynamics of the Sargent-Williams (2005) model, were not the main forces that accounted for the inflation process that the monetary authorities in the United States chose to administer during the post-WWII years. Our estimates of the Nash level of inflation are so near the Ramsey level that it

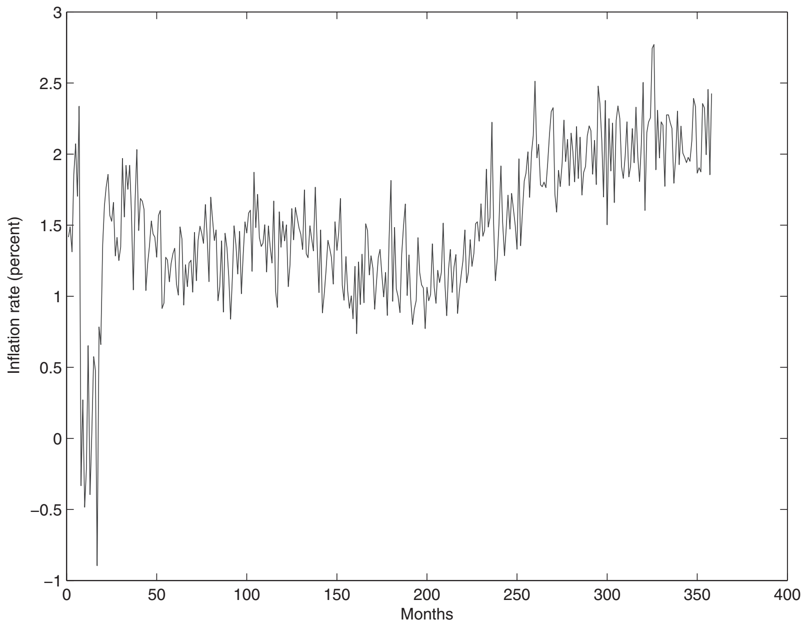

FIGURE 20. INFLATION DYNAMICS IN A SEQUENCE OF MONTE CARLO SIMULATIONS, USING A DOWNSCALED **V** MATRIX AND THE ESTIMATED $\mathbf{P}_{T|T-1}$ AT THE END OF SAMPLE

arrests most of the action coming from the mean dynamics and the escape dynamics.[35] Instead, the rise in inflation was driven by the interaction of government beliefs and the shocks impinging on the economy, and the fall in inflation was ultimately due to changes in those beliefs. If the U.S. monetary authorities remember the lessons that prompted Volcker to disinflate in the early 1980s, then maybe Volcker's conquest of U.S. inflation, sustained by Alan Greenspan, will endure.

APPENDIX A: DATA

The two monthly series employed in this paper are:

- Civilian unemployment rate, 16 years and older, seasonally adjusted (source: Bureau of Labor Statistics);
- Personal Consumption Expenditures (PCE) chain price index (2,000 = 100), seasonally adjusted (source: Bureau of Economic Analysis).

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: 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 $\boldsymbol{\theta}$ is mostly based on economic theory. For example, the mass prior probability of θ_0 is in the negative region.

The prior mean for $\boldsymbol{\theta}$ is set to

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

[35] Furthermore, 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.

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 ($0.16 < 0.20$), implying the low serial correlation of structural disturbances in Sargent's (1999) version of the Phillips curve (pp. 70–71). The prior variance for $\boldsymbol{\theta}$ 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 one (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 $\mathbf{C}_P$ and $\mathbf{C}_V$ is 0. The prior variance is $5^2\lambda_3$ for the diagonals of $\mathbf{C}_P$ and $\mathbf{C}_V$ 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 C: Conditional Posterior Distributions

Because x_{t-1} does not depend on $\boldsymbol{\theta}$, ζ_1, and ζ_2, it can be seen from (11)–(14) that the posterior distribution of $\boldsymbol{\theta}$ 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 following propositions.

PROPOSITION 1:

$$\text{(C1)} \quad p(\boldsymbol{\theta} | I_T, \zeta_1, \zeta_2, \boldsymbol{\varphi}) = \text{Normal}(\tilde{\theta}, \tilde{\boldsymbol{\Sigma}}_\theta),$$

where

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

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

$$\mathbf{y}_t = [1 \quad z_{2t} \quad z_{2t-1} \quad u_{t-1}]',$$

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

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

PROPOSITION 2:

$$\text{(C2)} \quad p(\zeta_1, \zeta_2 | I_T, \boldsymbol{\theta}, \boldsymbol{\varphi}) = \text{Gamma}(\tilde{\alpha}_{\zeta_1}, \tilde{\beta}_{\zeta_1})\text{Gamma}(\tilde{\alpha}_{\zeta_2}, \tilde{\beta}_{\zeta_2}),$$

where

$$\tilde{\alpha}_{\zeta_1} = \tilde{\alpha}_{\zeta_2} = \frac{T}{2} + \bar{\alpha},$$

$$\tilde{\beta}_{\zeta_i} = \frac{1}{0.5 \sum_{t=1}^{T} z_{it}^2 + \bar{\beta}^{-1}}, \qquad \forall\, i \in \{1, 2\}.$$

The government's optimization problem renders the conditional posterior pdf

$$p(\boldsymbol{\varphi} | I_T, \boldsymbol{\theta}, \zeta_1, \zeta_2)$$

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.*

(*i*) *Given the value $\boldsymbol{\varphi}^{\text{last}}$, compute the proposal draw*

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

where $\boldsymbol{\xi}$ is randomly drawn from the normal distribution with mean zero and co-

variance $\mathbf{c}\,\tilde{\boldsymbol{\Sigma}}_{\varphi}$ *specified in* (D1). *The scale factor* **c** will be adjusted to keep the acceptance ratio optimal (*around* 25 *percent*–40 *percent*).

(*ii*) *Compute*

$$q = \min\left\{\frac{p(\boldsymbol{\varphi}^{\text{prop}}|I_T, \boldsymbol{\theta}, \zeta_1, \zeta_2)}{p(\boldsymbol{\varphi}^{\text{last}}|I_T, \boldsymbol{\theta}, \zeta_1, \zeta_2)}, 1\right\}.$$

(*iii*) *Randomly draw* ν *from the uniform distribution* $U(0, 1)$.

(*iv*) *If* $\nu <= q$, *accept* $\boldsymbol{\varphi}^{\text{prop}}$ *as the value of the current draw; otherwise, keep* $\boldsymbol{\varphi}^{\text{last}}$ *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 $\boldsymbol{\phi}$ that emulates the posterior distribution.[36]

Appendix D: Proposal Density for the Metropolis Algorithm

The key to the Metropolis Algorithm for the posterior distribution $\boldsymbol{\varphi}$ is to obtain the covariance matrix for a normal proposal density. Since x_{t-1} is a function of $\boldsymbol{\varphi}$, one can approximate it by a second-order Taylor expansion at the posterior estimate $\hat{\boldsymbol{\varphi}}$. It can be seen from (15) that this approximation leads to the following covariance matrix for $\boldsymbol{\varphi}$:

(D1)

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

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

[36] For each draw of $\boldsymbol{\phi}$, ζ is normalized to be equal to ζ_1 before the government's inflation policy is solved. This normalization is consistent with the normalization principles discussed in James D. Hamilton et al. (2004).

REFERENCES

Blinder, Alan S. *Central banking in theory and practice*. Cambridge, MA: MIT Press, 1998.

Chappell, Henry W., Jr.; McGregor, Rob Roy and Vermilyea, Todd. *Committee decisions on monetary policy: Evidence from historical records of the Federal Open Market Committee*. Cambridge, MA: MIT Press, 2005.

Cho, In-Koo; Williams, Noah and Sargent, Thomas J. "Escaping Nash Inflation." *Review of Economic Studies*, 2002, *69*(1), pp. 1–40.

Chung, Heetaik. *Did policy makers really believe in the Phillips Curve? An econometric test.* Ph.D. diss., University of Minnesota, 1990.

Cogley, Timothy and Sargent, Thomas J. "The Conquest of U.S. Inflation: Learning and Robustness to Model Uncertainty." *Review of Economic Dynamics*, 2005a, *8*(2), pp. 528–63.

Cogley, Timothy and Sargent, Thomas J. "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S." *Review of Economic Dynamics*, 2005b, *8*(2), pp. 262–302.

Geweke, John. "Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication." *Econometric Reviews*, 1999, *18*(1), pp. 1–73.

Hamilton, James D.; Waggoner, Daniel F. and Zha, Tao. "Normalization in Econometrics." Federal Reserve Bank of Atlanta, Working Paper Series: No. 2004–13, 2004.

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

Ireland, Peter N. "Changes in the Federal Reserve's Inflation Target: Causes and Consequences." Boston College, Department of Economics, Working Papers in Economics: No. 607, 2005.

Ironside, Brian and Tetlow, Robert J. "Real-Time Model Uncertainty in the United States: The Fed from 1996–2003." Centre for Economic Policy Research, CEPR Discussion Papers: No. 5305, 2005.

King, Robert G. and Watson, Mark W. "The Post-War U.S. Phillips Curve: A Revisionist Econometric History." *Carnegie-Rochester Conference Series on Public Policy*, 1994, *41,* pp. 157–219.

Kydland, Finn E. and Prescott, Edward C. "Rules Rather Than Discretion: The Inconsistency of Optimal Plans." *Journal of Political Economy*, 1977, *85*(3), pp. 473–91.

Lucas, Robert E., Jr. "Some International Evidence on Output-Inflation Tradeoffs." *American Economic Review*, 1973, *63*(3), pp. 326–34.

Newey, Whitney K. and West, Kenneth K. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 1987, *55*(3), pp. 703–08.

Parkin, Michael. "Inflation in North America," in Kumiharo Shigehara, ed., *Price stabilization in the 1990s: Domestic and international policy requirements*. Tokyo: Bank of Japan, 1993, pp. 47–83.

Primiceri, Giorgio E. "Time Varying Structural Vector Autoregressions and Monetary Policy." *Review of Economic Studies*, 2005a, *72*(3), pp. 821–52.

Primiceri, Giorgio E. "Why Inflation Rose and Fell: Policymakers' Beliefs and US Postwar Stabilization Policy." National Bureau of Economic Research, Inc., NBER Working Papers: No. 11147, 2005b.

Romer, Christina D. and Romer, David H. "The Evolution of Economic Understanding and Postwar Stabilization Policy," in *Rethinking stabilization policy: Symposium proceedings*. Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, August 29–31, 2002. Kansas City: Federal Reserve Bank of Kansas City, 2002, pp. 11–78.

Sargent, Thomas J. *The conquest of American inflation*. Princeton: Princeton University Press, 1999.

Sargent, Thomas J. and Williams, Noah. "Impacts of Priors on Convergence and Escapes from Nash Inflation." *Review of Economic Dynamics*, 2005, *8*(2), pp. 360–91.

Sims, Christopher A. "Projecting Policy Effects with Statistical Models." *Revista de Analisis Economico*, 1988, *3,* pp. 3–20.

Sims, Christopher A. and Zha, Tao. "Bayesian Methods for Dynamic Multivariate Models." *International Economic Review*, 1998, *39*(4), pp. 949–68.

Sims, Christopher A. and Zha, Tao. "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review*, 2006, *96*(1), pp. 54–81.

Staiger, Douglas; Stock, James H. and Watson, Mark W. "The NAIRU, Unemployment and Monetary Policy." *Journal of Economic Perspectives*, 1997, *11*(1), pp. 33–49.

This article has been cited by:

1. ALINA BARNETT, MARTIN ELLISON. 2013. Learning by Disinflating. *Journal of Money, Credit and Banking* **45**:4, 731-746. [CrossRef]
2. Florin O. Bilbiie, Roland Straub. 2013. Asset Market Participation, Monetary Policy Rules, and the Great Inflation. *Review of Economics and Statistics* **95**:2, 377-392. [CrossRef]
3. Efrem Castelnuovo, Luciano Greco, Davide Raggi. 2012. POLICY RULES, REGIME SWITCHES, AND TREND INFLATION: AN EMPIRICAL INVESTIGATION FOR THE UNITED STATES. *Macroeconomic Dynamics* 1-23. [CrossRef]
4. GEORGE W. EVANS, SEPPO HONKAPOHJA, KAUSHIK MITRA. 2012. Does Ricardian Equivalence Hold When Expectations Are Not Rational?. *Journal of Money, Credit and Banking* **44**:7, 1259-1283. [CrossRef]
5. Guido Ascari, Lorenza Rossi. 2012. Trend Inflation and Firms Price-Setting: Rotemberg Versus Calvo*. *The Economic Journal* **122**:563, 1115-1141. [CrossRef]
6. Laurence Bloch. 2012. Product market regulation, trend inflation and inflation dynamics in the new Keynesian Phillips curve. *Economic Modelling* **29**:5, 2058-2070. [CrossRef]
7. Antonio E. Noriega, Carlos Capistrán, Manuel Ramos-Francia. 2012. On the dynamics of inflation persistence around the world. *Empirical Economics* . [CrossRef]
8. SETH PRUITT. 2012. Uncertainty Over Models and Data: The Rise and Fall of American Inflation. *Journal of Money, Credit and Banking* **44**:2-3, 341-365. [CrossRef]
9. David Gbaguidi. 2012. La courbe de Phillips : temps d'arbitrage et/ou arbitrage de temps. *L'Actualité économique* **88**:1, 87. [CrossRef]
10. Sergey Slobodyan, Raf Wouters. 2011. Learning in an estimated medium-scale DSGE model. *Journal of Economic Dynamics and Control* . [CrossRef]
11. Zheng Liu, Daniel F. Waggoner, Tao Zha. 2011. Sources of macroeconomic fluctuations: A regime-switching DSGE approach. *Quantitative Economics* **2**:2, 251-301. [CrossRef]
12. Marcus Hagedorn. 2011. Optimal disinflation in new Keynesian models. *Journal of Monetary Economics* . [CrossRef]
13. Kevin X. D. Huang, Guoqiang Tian. 2011. REPUTATION AND OPTIMAL CONTRACTS FOR CENTRAL BANKERS. *Macroeconomic Dynamics* 1-24. [CrossRef]
14. João Ricardo Faria, André Varella Mollick, Adolfo Sachsida, Le Wang. 2011. Do central banks affect Tobin's q?. *International Review of Economics & Finance* **22**:1, 1-10. [CrossRef]
15. Kuang-Liang Chang, Nan-Kuang Chen, Charles Ka Yui Leung. 2011. The dynamics of housing returns in Singapore: How important are the international transmission mechanisms?. *Regional Science and Urban Economics* . [CrossRef]
16. S. Boragan Aruoba,, Frank Schorfheide. 2011. Sticky Prices versus Monetary Frictions: An Estimation of Policy Trade-offs. *American Economic Journal: Macroeconomics* **3**:1, 60-90. [Abstract] [View PDF article] [PDF with links]
17. ROBERT J. GORDON. 2011. The History of the Phillips Curve: Consensus and Bifurcation. *Economica* **78**:309, 10-50. [CrossRef]
18. Gary Koop, Roberto Leon-Gonzalez, Rodney W. Strachan. 2010. Dynamic Probabilities of Restrictions in State Space Models: An Application to the Phillips Curve. *Journal of Business and Economic Statistics* **28**:3, 370-379. [CrossRef]

19. Kuang-Liang Chang, Nan-Kuang Chen, Charles Ka Yui Leung. 2010. Monetary Policy, Term Structure and Asset Return: Comparing REIT, Housing and Stock. *The Journal of Real Estate Finance and Economics* . [CrossRef]
20. George W. Evans, Seppo Honkapohja, Noah Williams. 2010. GENERALIZED STOCHASTIC GRADIENT LEARNING. *International Economic Review* **51**:1, 237-262. [CrossRef]
21. Cogley Timothy, Primiceri Giorgio E., Sargent Thomas J.. 2010. Inflation-Gap Persistence in the US. *American Economic Journal: Macroeconomics* **2**:1, 43-69. [Abstract] [View PDF article] [PDF with links]
22. Lars Peter Hansen, Thomas J. SargentWanting Robustness in Macroeconomics **3**, 1097-1157. [CrossRef]
23. Antonio E. Noriega, Manuel Ramos-Francia. 2009. The dynamics of persistence in US inflation. *Economics Letters* **105**:2, 168-172. [CrossRef]
24. George W. Evans, Seppo Honkapohja. 2009. Learning and Macroeconomics. *Annual Review of Economics* **1**:1, 421-449. [CrossRef]
25. Giacomo Carboni, Martin Ellison. 2009. The Great Inflation and the Greenbook#. *Journal of Monetary Economics* **56**:6, 831-841. [CrossRef]
26. Kevin X.D. Huang, Zheng Liu, Tao Zha. 2009. Learning, Adaptive Expectations and Technology Shocks. *The Economic Journal* **119**:536, 377-405. [CrossRef]
27. Timothy Cogley,, Argia M. Sbordone. 2008. Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve. *American Economic Review* **98**:5, 2101-2126. [Abstract] [View PDF article] [PDF with links]
28. Jesús Fernández-Villaverde. 2008. Horizons of Understanding: A Review of Ray Fair's Estimating How the Macroeconomy Works. *Journal of Economic Literature* **46**:3, 685-703. [Abstract] [View PDF article] [PDF with links]
29. Thomas J. Sargent. 2008. Evolution and Intelligent Design. *American Economic Review* **98**:1, 5-37. [Abstract] [View PDF article] [PDF with links]
30. F MILANI. 2007. Expectations, learning and macroeconomic persistence#. *Journal of Monetary Economics* **54**:7, 2065-2082. [CrossRef]
31. ROBERT J. TETLOW, BRIAN IRONSIDE. 2007. Real-Time Model Uncertainty in the United States: The Fed, 1996?2003. *Journal of Money, Credit and Banking* **39**:7, 1533-1561. [CrossRef]
32. R TETLOW. 2007. On the robustness of simple and optimal monetary policy rules#. *Journal of Monetary Economics* **54**:5, 1397-1405. [CrossRef]
33. Tao Zha. 2007. Comment on An and Schorfheide's Bayesian Analysis of DSGE Models. *Econometric Reviews* **26**:2, 205-210. [CrossRef]