NASH OR RAMSEY INFLATION IN THE U.S.?

THOMAS SARGENT, NOAH WILLIAMS, AND TAO ZHA

ABSTRACT. In this paper we develop and estimate a structural model of learning with an optimizing government that is able to explain much of the rise and fall of inflation in the US. In our self-referential learning model, the government's prior beliefs are a key element which affect the dynamics. By estimating the parameters of the government's beliefs along with the structural parameters of the model, we are able to explain the evolution of US monetary policy through the evolution of policymakers' beliefs.

I. INTRODUCTION

Has post-war monetary policy changed from the 1970s through the 80s into the 90s? If so, are such changes responsible for the rise and fall of post-war inflation in the US? There are three views on answers to these questions. Clarida, Gali, and Gertler (2000) and Primiceri (2003b) present empirical evidence that monetary policy has been drastically different in the 80s and 90s from the 70s. A second view is that there may have been changes in monetary policy, but these changes are difficult to detect statistically (as in Cogley and Sargent (2003) and Primiceri (2003a)). The third view is that there was no significant systematic shift in monetary policy. For example, Sims and Zha (2004) have found that models with coefficient changes fit the data much worse than models with time-varying shock variances only. The resolution of this debate has clear implications for whether the improved economic performance in the 1980s and 90s can be sustained.

To understand the current debate, Sargent and Williams (2003) develop a model of Federal Reserve policymakers who learn over time. Building on Sargent (1999) and Cho, Williams, and Sargent (2002), they show that policymakers who act based on a misspecified model which is adapted over time will occasionally alternate between a high inflation (Nash) policy and a low inflation (Ramsey) policy. The particular contribution of Sargent and Williams (2003) is to analyze how the government's prior beliefs affect the mean convergence to a high inflation equilibrium as well as the escape dynamics from Nash to Ramsey outcome. Moreover, different settings of the prior beliefs can lead to dramatically different outcomes. In earlier work, Sims (1988) suggested that good inflation outcomes

Date: July 12, 2004.

Key words and phrases. Learning, beliefs, self-confirming equilibrium, escape, nonconvergence.

We thank Will Roberds for helpful discussions and Jordi Mondria for excellent research assistance. The views expressed herein do not necessarily reflect those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

may be sustainable in such a learning model. The difference between the convergence results discussed in Cho, Williams, and Sargent (2002) and the non-convergence outcome found by Sims (1988) can be explained by different prior beliefs of the government. It is therefore essential that these beliefs, along with the model's other parameters, be estimated.

There has been some previous work along these lines, which we generalize. In particular, Chung (1990) and Sargent (1999) estimated different versions of the model, but their reported fit to the data was poor. An overriding task of our paper is to estimate a natural rate model that fits to the data very well and thus is capable of explaining the rise and fall of both inflation and unemployment in the US. One crucial place where we differ from previous work is that we estimate key components of the government's prior beliefs. This flexibility is crucial and allows us to fit the data much better than the previous papers. By limiting their attention to particular prior specifications, these papers had limited the amount of variation in the data that could be explained by policymaker learning. Our flexible specification shows that almost all of the evolution of inflation may be explained by changes in policymakers' beliefs.

Unlike Chung and Sargent, moreover, we also provide measures of how sharp the estimation is. Taking into account the parameter uncertainty is important. As Sargent and Williams (2003) show, whether monetary policy is off or on the Nash equilibrium and how it evolves over time are sensitive to the model's parameters (especially the government's belief in the covariance matrix for the drifting coefficients). We show that the key structural parameters are estimated sharply, but that some small variations in parameters may lead to dramatically different limiting behavior.

All these exercises shed light on the debate over monetary policy changes and on important questions such as whether the Fed stopped trying to exploit the Phillips curve long ago (e.g., by the early 1970s), as some have suggested. We find that the historical policy performance may be best thought of as a process of continual learning. In our model, policymakers allowed inflation to drift up in the 1970s, but were continually forecasting a return to lower inflation rates. In fact, in the long run our model suggest low and stable inflation, although there may be a long and volatile transition to those long run outcomes.

The rest of the paper is organized as follows. In Section II we lay out the model and discuss the theoretical characterizations of it. In Section III we discuss the estimation of the model and interpret our results. Section IV then describes and implements our sampling scheme for small sample inference. We then turn to some implications of the model. Section V examines the performance of the model in forecasting over intermediate horizons, while Section VI analyzes the long-run behavior of the model. Finally, Section VII concludes. Three appendices describe the data and provide technical detail on the setting of our prior distribution for estimation and the sampling scheme for inference.

II. The Model

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

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

$$\pi_t = x_{t-1} + \sigma_2 w_{2t}, \tag{2}$$

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

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

subject to (2) and

$$u_t = \hat{\alpha}'_{t|t-1}\Phi_t + \sigma w_t, \tag{3}$$

where π^* and u^{**} are the targeted levels of inflation and unemployment, both $\hat{\alpha}_{t|t-1}$ and Φ_t are $r \times 1$ vectors, and w_t is an i.i.d. standard normal random variable. In practice the vector Φ_t of regressors consists of lags of unemployment and inflation. By comparing (3) with (1) we see that the government fails to account for the role of expectations in determining the unemployment rate. Here $\hat{E}$ represents expectations with respect to the government's subjective model, and the subscript $t-1$ means that the government updates $\hat{\alpha}_{t|t-1}$ and optimally chooses x_{t-1} before observing π_t and u_t. Thus the government sets policy based on its misspecified Phillips curve (3), not the true Phillips curve (1). A self-confirming equilibrium emerges when the government optimizes based on its beliefs, and its beliefs are consistent with what it observes. In this model, the self-confirming equilibrium outcomes agree with the Nash equilibrium in which policymakers set inflation at a higher level than the socially optimal Ramsey level (see Sargent 1999).

While a self-confirming equilibrium is a population concept which restricts beliefs, at any point in history the government updates its beliefs as it learns. In particular, the government bases $\hat{\alpha}_{t|t-1}$, its mean estimate of the drifting parameter vector α_t, on the observations up to and including time $t-1$ from the following (misspecified) econometric model:

$$u_t = \alpha'_t \Phi_t + \sigma w_t, \tag{4}$$

$$\alpha_t = \alpha_{t-1} + \Lambda_t, \tag{5}$$

where Λ_t, uncorrelated with w_t, is an i.i.d. Gaussian random vector with mean 0 and covariance matrix V. Thus the government believes that the true economy drifts over time, so it continually adapts its parameter estimates. The innovation covariance matrix V governs the rate of drift, and is a key component of the model. The mean estimate of α_t for the econometric model (4)-(5) is

$$\hat{\alpha}_{t|t-1} \equiv \hat{E}(\alpha_t | \mathscr{I}_{t-1}),$$
$$\mathscr{I}_t \equiv \{u_1, \pi_1, \dots, u_t, \pi_t\}.$$

Let

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

As is well-known, the mean estimates can be updated via the Kalman filter. Given $\hat{\alpha}_{1|0}$ and $P_{1|0}$, the Kalman filter algorithm updates $\hat{\alpha}_{t|t-1}$ with the following formula:

$$\hat{\alpha}_{t+1|t} = \hat{\alpha}_{t|t-1} + \frac{P_{t|t-1}\Phi_t(u_t - \Phi_t'\hat{\alpha}_{t|t-1})}{\sigma^2 + \Phi_t' P_{t|t-1}\Phi_t}, \tag{6}$$

$$P_{t+1|t} = P_{t|t-1} - \frac{P_{t|t-1}\Phi_t\Phi_t' P_{t|t-1}}{\sigma^2 + \Phi_t' P_{t|t-1}\Phi_t} + V. \tag{7}$$

Many studies of learning consider a learning rule known as recursive least squares (RLS) which is closely related to the Kalman filter. A key question under any learning rule is whether the learning process will converge to a self-confirming equilibrium (SCE). To address this issue, scale the innovation covariance matrix as $V = \varepsilon^2 \hat{V}$, for $\varepsilon > 0$. The key analytical results from Sargent and Williams (2003) that underscore the role of the government's learning from misspecified models are:

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

III. Estimation

The theoretical results illustrate that very different outcomes may emerge for different government beliefs. One main task of this paper is to fit the model to the data, estimating and quantifying the uncertainty of these prior belief parameters, σ^2 and V, jointly with the model's other structural parameters. In estimation, we fix the values of β, λ, π^*, and u^{**}.[1] Group all other free structural parameters as

$$\phi = \{v^*, \theta_0, \theta_1, \tau_1, \zeta_1, \zeta_2, z_{1|0}, C_P, C_V\},$$

where $v^* = u^*(1-\tau_1)$, $z_{1|0} = \hat{\alpha}_{1|0}$, C_P and C_V are upper triangular such that $P_{1|0} = C_P' C_P$ and $V = C_V' C_V$, and $\zeta_1 = 1/\sigma_1^2$ and $\zeta_2 = 1/\sigma_2^2$ represent the precisions of the corresponding innovations. The structural parameter $\zeta = 1/\sigma^2$ is not a free parameter. It is clear from (6), (7), and (8) that if we scale V and $P_{1|0}$ by κ and ζ by $1/\kappa$, the likelihood value remains the same. There would exist a continuum of maximum likelihood estimates (MLEs) if ζ were not restricted (i.e., the model is unidentified). Some normalization is necessary. Following Sargent and Williams, we impose the restriction $\zeta = \zeta_1$. This normalization implies that the variation that policymakers observe in the unemployment rate is correctly decomposed into variation in the regressors and variation due to exogenous shocks. It has an advantage because the limiting results are easier to derive.[2]

The likelihood function is:P

$$\mathscr{L}\left(\mathscr{I}_T|\phi\right) = \frac{\zeta_1^{T/2}\zeta_2^{T/2}}{(2\pi)^{T/2}} \exp\left\{-\frac{1}{2}\sum_{t=1}^{T}\left[\zeta_1 z_{1t}^2 + \zeta_2 z_{2t}^2\right]\right\}, \tag{8}$$

where:

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

The posterior pdf of ϕ is proportional to the multiplication of the likelihood (8) and the prior $p(\phi)$:

$$p(\phi|\mathscr{I}_T) \propto \mathscr{L}\left(\mathscr{I}_T|\phi\right)\, p(\phi). \tag{9}$$

Using the monthly U.S. data described in Appendix A and the prior specified in Appendix B, we estimate ϕ by maximizing the posterior density function (9). We obtained similar results using maximum likelihood, but as we discuss below the prior is crucial for small sample inference. The regressor vector in the government's misspecified Phillips regression is:

$$\Phi_t = \begin{bmatrix}\pi_t & \pi_{t-1} & u_{t-1} & \pi_{t-2} & u_{t-2} & 1\end{bmatrix}'.$$

---

[1] We set $\beta = 0.9936, \lambda = 1, \pi^{=}2$ and $u^{**} = 1$.

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

TABLE 1. Posterior estimates of model parameters

Log value of posterior kernel at its peak: -9.97759					
Estimates of coefficients in true Phillips curve and inflation process with 68% and 90% probability intervals in parentheses					
u^* : 6.3741 $(5.7041, 7.3310)$ $(5.0710, 8.6417)$					
θ_0 : -0.0702 $(-0.0981, -0.0356)$ $(-0.1183, -0.0142)$					
θ_1 : -0.0577 $(-0.0812, -0.0198)$ $(-0.1004, 0.0012)$					
τ_1 : 0.9873 $(0.9837, 0.9939)$ $(0.9804, 0.9972)$					
ζ_1 : 31.0258					
ζ_2 : 17.1587					
Estimate of $z_{1\|0}$ or $\hat{\alpha}_{1\|0}$:					
150.5168	−132.9030	−1.3263	62.9907	4.2185	−161.1007
Estimate of $P_{1\|0}$:					
205.5789	13.4584	−18.3878	−27.6365	44.1506	18.7960
13.4584	212.5153	35.4853	−1.3380	−21.1811	−15.3423
−18.3878	35.4853	353.9456	−7.7774	15.8083	−5.1209
−27.6365	−1.3380	−7.7774	340.2518	26.2907	7.8629
44.1506	−21.1811	15.8083	26.2907	322.7555	−6.9726
18.7960	−15.3423	−5.1209	7.8629	−6.9726	398.0110
Estimate of V:					
0.5094	0.4125	0.5838	−0.4672	−0.0947	0.6227
0.4125	146.7887	25.4800	71.9230	66.1744	13.7632
0.5838	25.4800	8.2430	11.3364	10.2849	2.4890
−0.4672	71.9230	11.3364	38.4990	32.8691	8.1652
−0.0947	66.1744	10.2849	32.8691	30.3237	5.7669
0.6227	13.7632	2.4890	8.1652	5.7669	89.6903

The posterior estimate of ϕ is reported in Table 1. The natural rate of unemployment is estimated to be 6.4. Responses of unemployment to inflation surprises are weak. Unemployment is by itself a persistent series. One important finding is that the coefficients on the government's Phillips regression at any time t show no trade-off between unemployment and inflation because the sum of all coefficients on current and lagged inflation variables is positive (e.g., see the estimate of $\hat{\alpha}_{1|0}$ in Table 1). This finding, consistent with the Lucas-Sargent model, implies that there is no incentive for the government to set inflationary policy in order to bring down unemployment even in the short run.

The one-step forecasts of inflation are plotted against the actual path in Figure 1, and one-step forecasts of unemployment are plotted against the actual path in Figure 2. It is evident

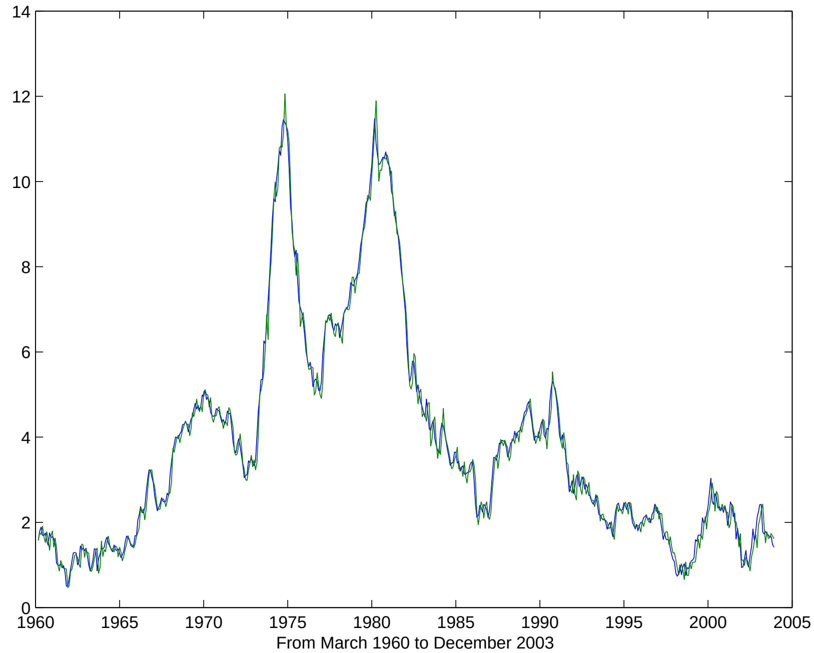

FIGURE 1. Inflation: actual vs one-step forecast (i.e, government controlled inflation)

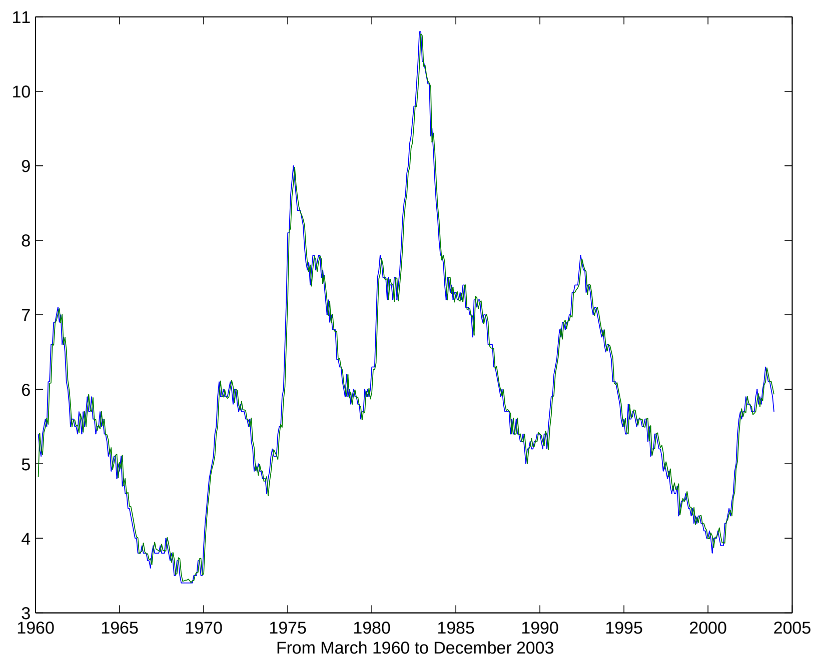

FIGURE 2. Unemployment rate: actual vs one-step forecast

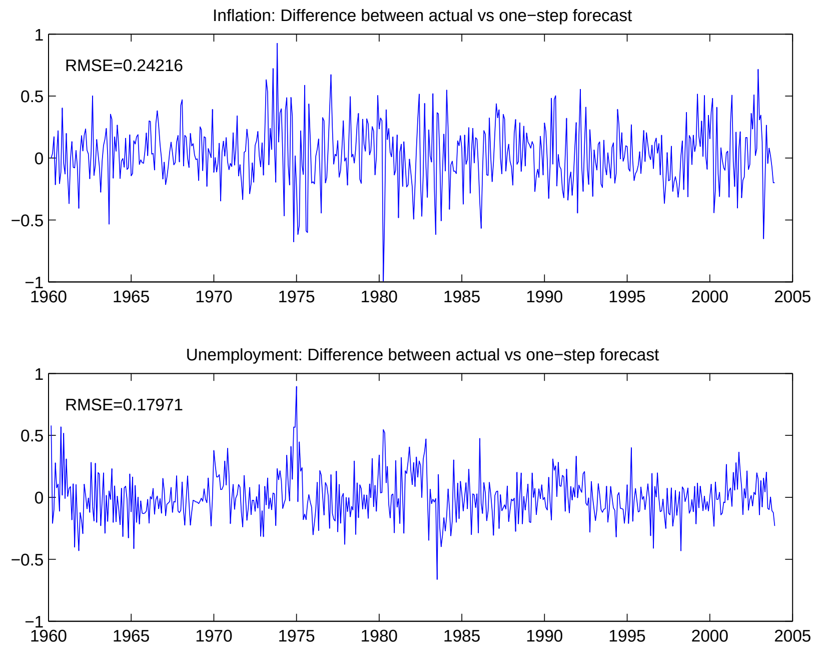

FIGURE 3. Difference between actual values and one-step forecasts of inflation and unemployment

from these figures that the model fits the data remarkably well. So well, in fact, that it is difficult to discern the difference between the series. Figure 3 plots the one-step forecast error, showing that for most of the sample the forecasts are within one half a percentage point of the realized value. We also list there the one-step root mean squared forecast error. The RMSE of our model is comparable to (if very slightly inferior to) a random walk.[3] The government's inflation policy explains, almost entirely, the rise and fall of post-war American inflation (Figure 1). This kind of result has not been achieved in previous work (e.g., Sims 1988, Chung 1990, and Sargent 1999).

As we've already noted, the rise and fall of inflation in our model is driven by a mechanism which is different from what Sargent (1999) stresses. In Sargent's work (and the follow up work of Cho, Williams and Sargent, 2002) the economy converges to self-confirming equilibrium in which the government believes in a an exploitable tradeoff between inflation and unemployment, which leads to a high equilibrium inflation rate. But then occasional sequences of stochastic shocks lead the government to temporarily discover that it can cut inflation and there will be no rise in unemployment, which leads to rapid disinflations. In particular, during these episodes government learns a version of the natural rate theory in which the sum of the coefficients on inflation in its model is zero,

[3]The one-step RMSE for a random walk (i.e. constant) forecast is 0.2343 for inflation and 0.1793 for unemployment.

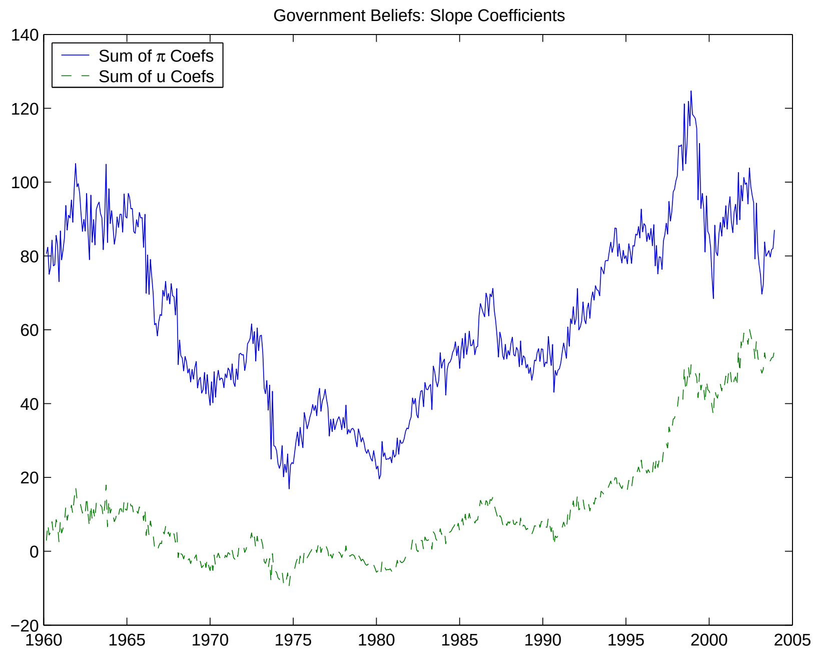

FIGURE 4. Evolution of the government's beliefs

reflecting a vertical long-run Phillips curve. By contrast, Figure 4 shows the evolution of the government's estimates over time in our model. We see that the sum of the coefficients on inflation is high at all times. Moreover we see that they decline slightly during the 1970's high inflation episodes. Later in the paper we show that the inflation experience in our model is best explained as a very slow drift in the government's model, rather than the more rapid convergence and escape which has been previously analyzed.

Crucial to our empirical success is the flexibility of our model in fitting the government's beliefs. In particular, the previous work has basically assumed a particular form for the key matrix V which governs the innovations to the parameters in the government's model. We have already discussed the theoretical reasons why the V matrix is so important – different specifications of it greatly affect the speed, direction, and stability of the learning dynamics. Thus by fixing V the previous work has taken a strong stand on the role learning plays, and has minimized the amount of variation in the data that may be explained by evolving government beliefs.

One particular example of the importance of the V matrix is shown in Figure 5. There we fix the values of all the structural parameters, including the estimated initial conditions $z_{1|0}$ and $P_{1|0}$, as above and only alter V. As we noted previously, most learning models such as Sargent (1999) have focused on a recursive least squares learning rule which is closely related to the Kalman filter. Sargent and Williams (2003) show that RLS can be approximated by a Kalman filter with V proportional to $\sigma^2 E(\Phi\Phi')^{-1}$. In the figure, we

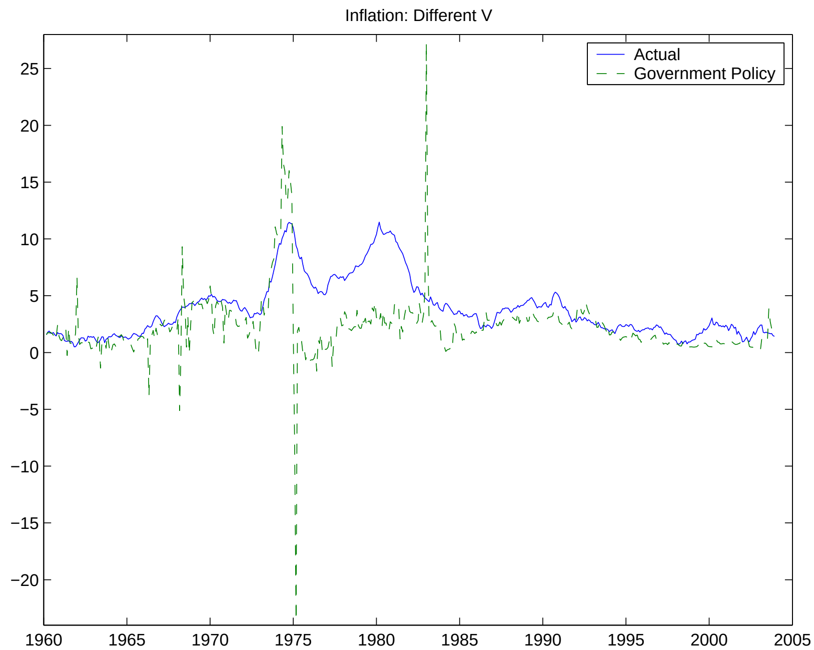

FIGURE 5. Actual and forecast inflation (government policy) with a different setting of V.

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. We see clearly that this choice of V leads to a substantial deterioration in fit. The inflation policy loosely tracks the rise in inflation in the early 1970s, but then suggests very low inflation (with a few outliers) for the rest of the sample. In particular, it completely misses the second peak in inflation in 1979-80. This illustrates a point made by Chung (1990) and Sargent (1999), as their models implied that the government should have cut inflation much earlier than actually occurred in the data. However this relies on attributing the government very particular beliefs about how its model changes over time. By allowing the data to inform our choice of these beliefs, we are able to much better explain the rise and fall of inflation.

IV. INFERENCE

The finite-sample inference of ϕ requires simulating the posterior distribution. We use a Bayesian Markov chain Monte Carlo (MCMC) algorithm to obtain the empirical posterior distribution of ϕ by sampling alternately from the following conditional posterior distributions (a Gibbs sampler):

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

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

The blocks θ and φ are defined in Appendix B. Let $\bar{\Sigma}_\theta$ be the prior covariance for θ, and $\bar{\alpha}$ and $\bar{\beta}$ be the hyperparameters in the prior Gamma distribution of ζ_1 and ζ_2 (Appendix B). From (9) and Appendix B one can show the following two propositions.

Proposition 1.

$$p(\theta \mid \mathscr{I}_T, \zeta_1, \zeta_2, \varphi) = \text{Normal}\,(\tilde{\theta}, \tilde{\Sigma}_\theta), \tag{10}$$

where

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

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

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

Proposition 2.

$$p(\zeta_1, \zeta_2 \mid \mathscr{I}_T, \theta, \varphi) = \text{Gamma}\,(\tilde{\alpha}_{\zeta_1}, \tilde{\beta}_{\zeta_1})\,\text{Gamma}\,(\tilde{\alpha}_{\zeta_2}, \tilde{\beta}_{\zeta_2}), \tag{11}$$

where the form of the pdf Gamma is described in Appendix B1, and

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

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

The government's optimization problem renders the conditional posterior pdf $p(\varphi \mid \mathscr{I}_T, \theta, \zeta_1, \zeta_2)$ to be of nonstandard form. To simulate from this distribution, therefore, we use a Metropolis algorithm as follows:

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

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

where ξ is randomly drawn from the normal distribution with mean zero and covariance $\mathfrak{c}\tilde{\Sigma}_\varphi$ specified in (C1). The scale factor $\mathfrak{c}$ will be adjusted to keep the acceptance ratio optimal (around $25\% - 40\%$).

(b) Compute

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

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

(d) If $\nu <= q$, accept φ^{prop} as the value of the current draw; otherwise, keep φ^{last} as the value of the current draw.

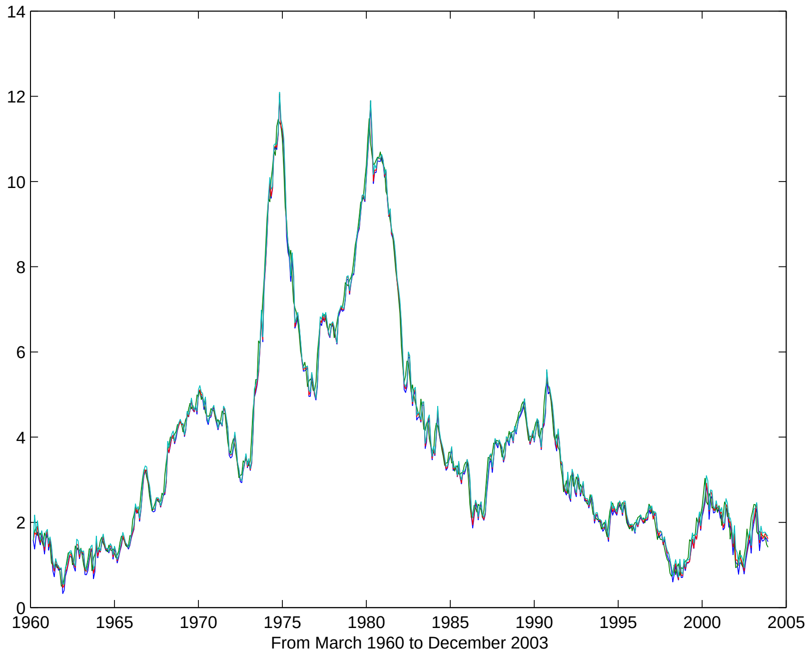

FIGURE 6. Inflation: actual vs one-step forecast with 90% error bands

A large number of MCMC samples alternately from these conditional posterior distributions will deliver, accurately, the empirical distribution of ϕ as though these samples were drawn directly from their own posterior distribution.[4]

The 68% and 90% probability intervals for the estimates of coefficients of the natural-rate Phillips curve are reported in Table 1.[5] It can be seen from this table that θ_0, θ_1, and τ_1 are all tightly estimated. The 90% probability interval for u^* is much wider, consistent with the confidence interval in the statistical model of Staiger, Stock, and Watson (1997). Inflation set by the government's policy (x_{t-1}), however, is sharply estimated, as shown by the error bands displayed in Figure 6. Again, the bands are so tight that they are difficult to distinguish. Figure 7 plots the data along with the forecast error bands from 1995 to the end of the sample. We see clearly that the bands are quite tight, and although the data frequently falls outside the bands, the predictions are relatively close to the actual data.

V. SHORT-RUN DYNAMICS

Thus far we've focused on the one-step properties of the model. Now we turn to some of the intermediate term implications, by looking at the dynamics of the model over three

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

[5]The small-sample posterior distribution is generated from a sequence of MCMC 30,000 draws.

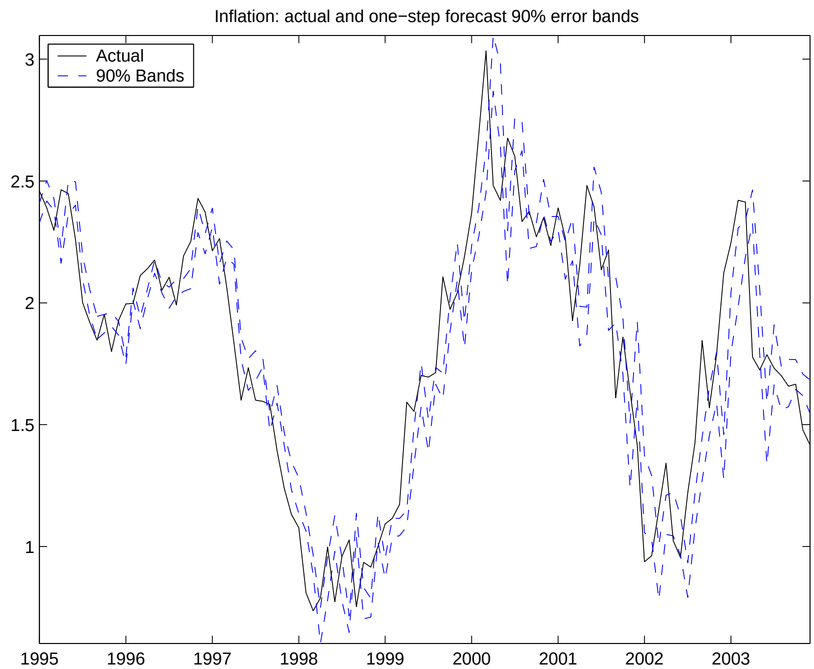

FIGURE 7. Inflation: actual and one-step forecast 90% error bands over the recent past.

year horizons. Then in the next section we analyze the long run and limiting behavior of the model. In particular, in this section we focus on the performance of the model in forecasting the two peaks of inflation in the 1970s. We use Monte Carlo simulations to assess the distribution of forecasts looking forward. For simplicity, we focus here on our baseline estimates and do not account for the parameter uncertainty. Acknowledging uncertainty in the parameters would certainly widen our forecast bands, but the tightness of the parameter estimates suggests that this effect may be rather small.

We now present a number of figures for the key episodes, each of which represents simulations going forward from different initial conditions. In each case, we take the estimated beliefs at the starting date, and draw 5000 simulations of 35 periods each. The figures then plot the means of the government controlled inflation (x_{t-1}) and inflation in the model, along with 68% and 90% probability bands on the government inflation, and the actual experienced inflation. In each plot, the initial condition is shown as date zero, from which we look backward three months and forward 35.

Figure 8 considers the first episode. The upper left panel starts in July 1973 when inflation had already risen from near 3% up to 5.4%. The model predicts that this inflation should stabilize and slowly decline, although of course inflation actually spiked up substantially over the forecast horizon before coming back down. However by six months later in January 1974, which is shown in the second panel, inflation had continued upward, now

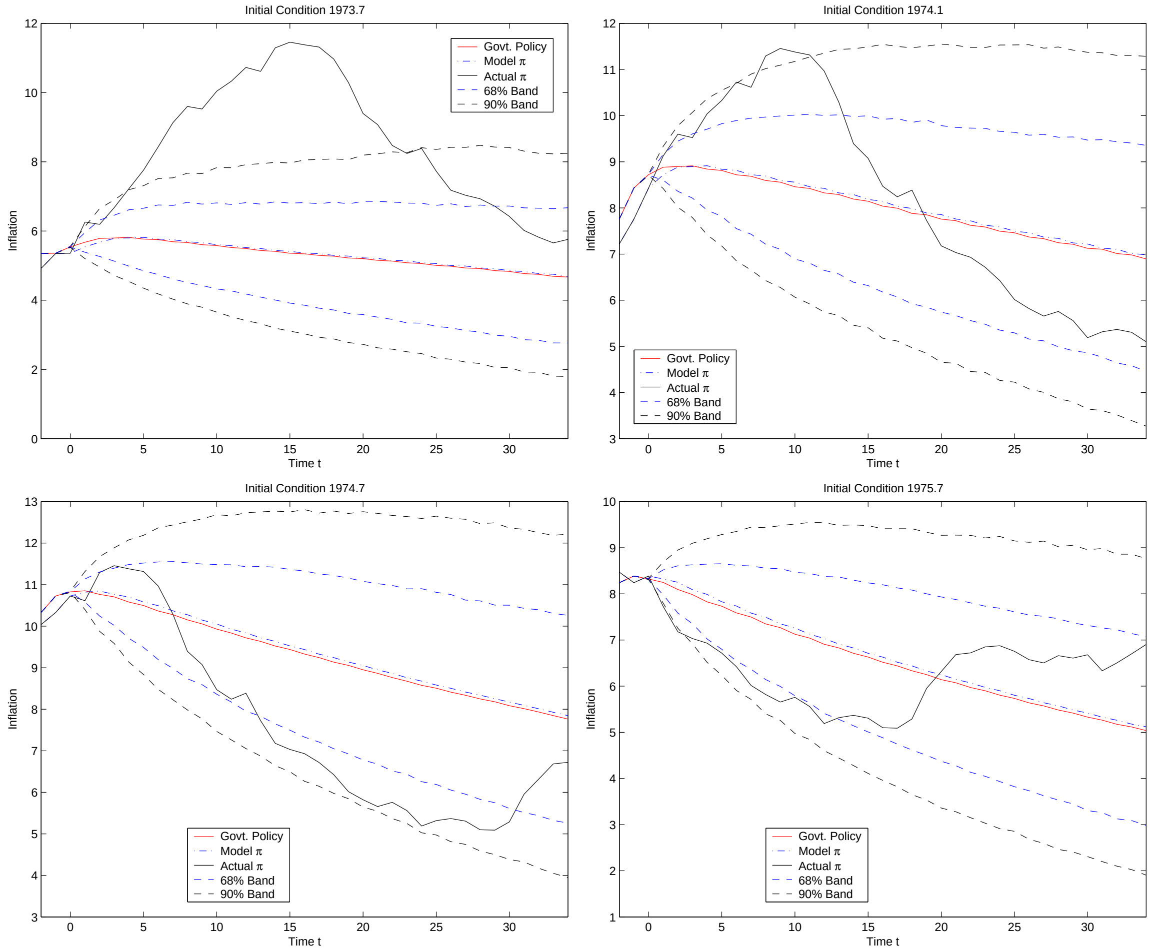

FIGURE 8. Monte Carlo simulation, using as initial conditions estimates over 1973-1975.

reaching 8.4%. Here we see that the model predicts a further increase in inflation prior to a return to lower levels. The mean path increases rather modestly, but the forecast bands (which are admittedly a bit wide) cover both the rise and fall in inflation for all but a few months. Inflation is near its peak six months later in July 1974, at which time the model predicts a slow decline. One year later in July 1975, inflation is well down from the peak, and the model matches the rate of decrease relatively well.

In summary, the model was initially caught off guard by the continuing rise in inflation. But once the inflation was entrenched, the model captured reasonably well the continued increase and then the decline in inflation. A similar story emerges from the spike in inflation in 1979-80, as shown in Figure 9. In each case the model forecasts reasonably well, although rise and fall in the data are sharper than the mean predictions of the model.

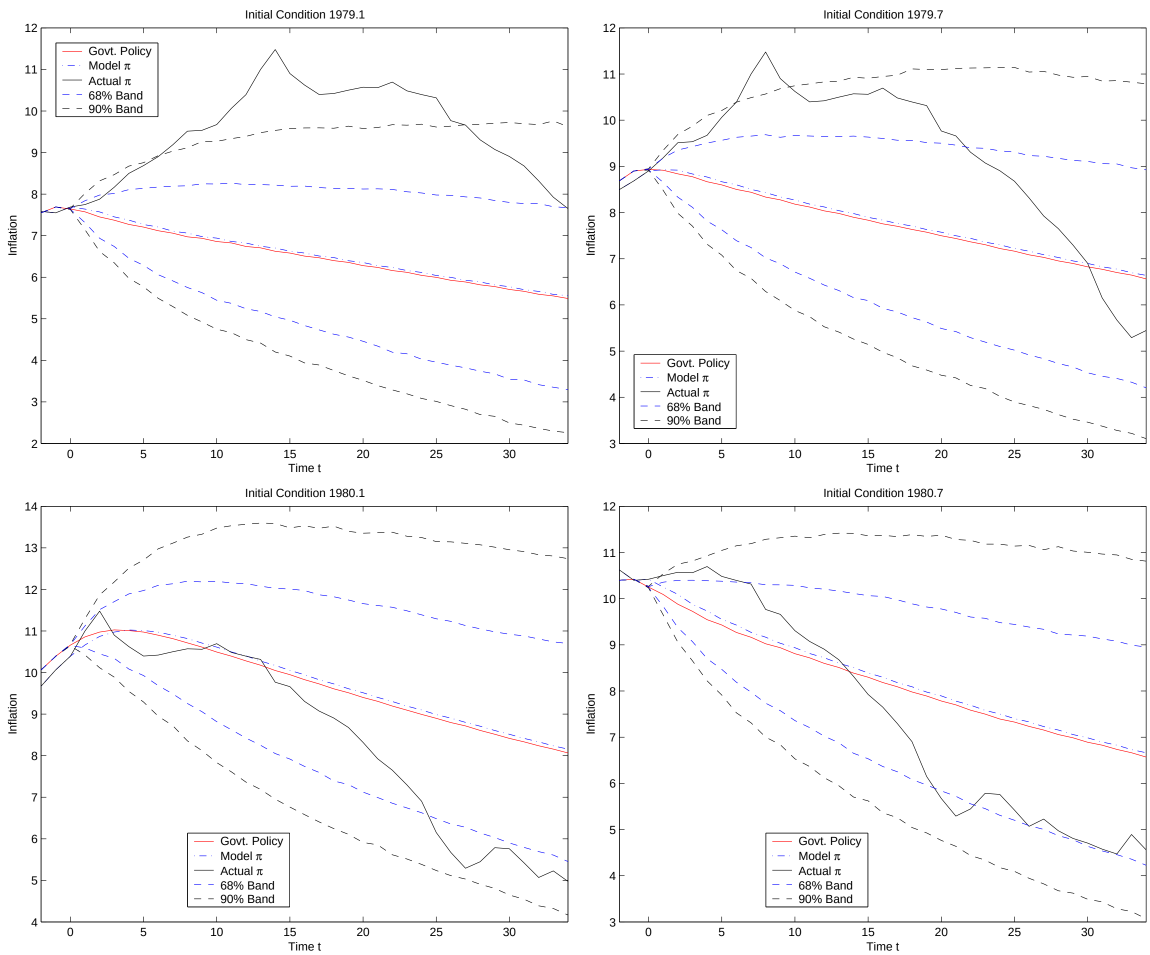

FIGURE 9. Monte Carlo simulation, using as initial conditions estimates over 1979-1980.

VI. LONG-RUN DYNAMICS

We now turn to some of the long run properties of the model. First we discuss the convergence of our baseline model to a limit distribution. Then we analyze the small variation limits as in Sargent and Williams (2003).

VI.1. **Long-Run Convergence.**

In the preceding section we've seen that the forecasts under our estimates typically imply falling inflation over three year horizons. This is a robust implication of the model – looking forward from any data point in the sample we expect inflation to trend downward, even if it starts at a relatively low level as in the late 1990s. In fact, starting from any of the data points in our sample we find that inflation trends downward to deflationary levels. Once the inflation rate becomes sufficiently negative, this

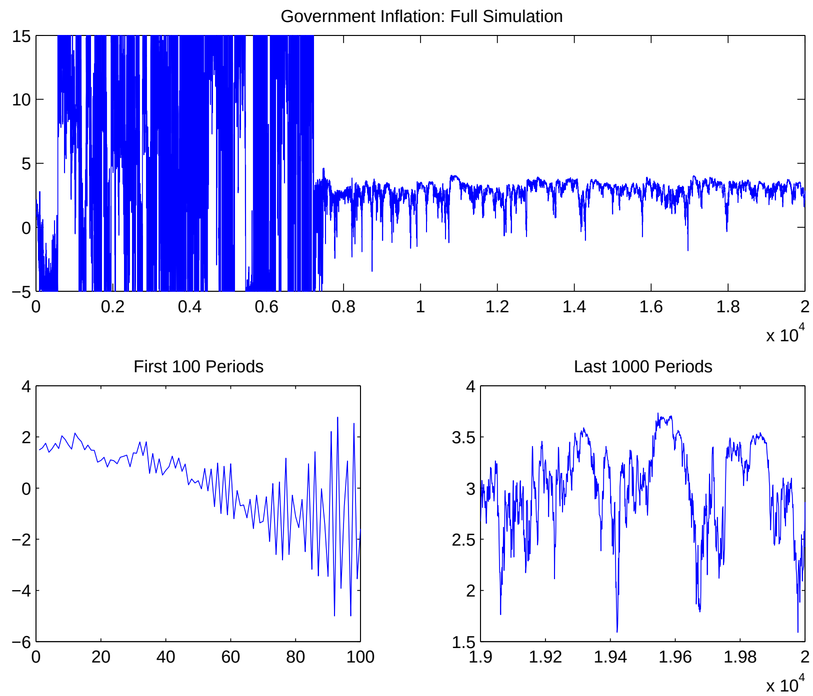

FIGURE 10. Time path of the government's inflation choice in one long Monte Carlo simulation, using the estimated initial conditions. Bounds of [-5,15] are imposed.

triggers some explosive oscillations in the government's inflation rate. However the government's beliefs are well behaved even in these oscillatory periods. To deal with the large oscillations, we imposed bounds which we took to be [-5,15] on the allowable rate of inflation that the government can choose. If the solution of the Phelps problem implies an expected inflation rate outside these bounds, we simply truncate at the bound. In this case, as shown in Figures 10 and 11 we have apparent convergence to a limit distribution in the (very) long run.

Considering the inflation dynamics first, the upper panel of Figure 10 shows a Monte Carlo simulation of 20,000 periods starting at our estimated initial conditions. Here we see that there is quite a bit of volatility in the early part of the sample, with inflation hitting the upper and lower bounds regularly until about period 7000. After this time however, the inflation rate is relatively stable for the remainder of the sample. The lower left panel of the figure provides a detail of the first 100 periods of the simulation. There we see, that consistent with our discussion above that inflation trends downward from its initial value near 2% down to around -4%. The oscillations begin after this date, and would be essentially unbounded if we did not impose the bounds. Throughout the long period of volatile inflation, the government's beliefs continue to evolve, and the outcomes that finally

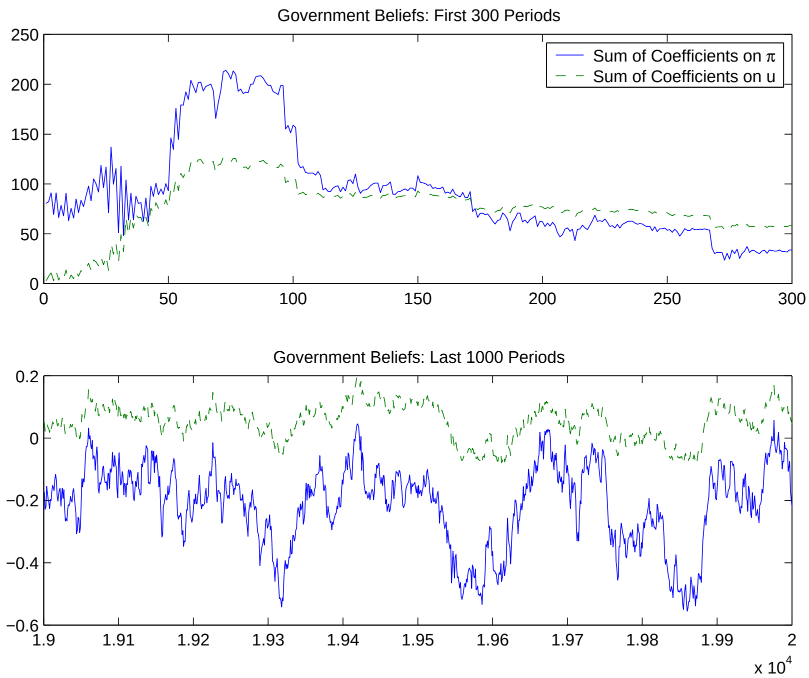

FIGURE 11. Time pats of the government's estimates from the long Monte Carlo simulation, using the estimated initial conditions.

emerge are relatively good. The lower right panel of the figure shows the last 1000 periods from the sample, and we see that during this period inflation is relatively stable, mostly in the range of 2-3%. Thus we see that in the long run the government's policy is actually quite good. However it takes quite some time, and quite a bit of volatility, before the long run is reached.

Figure 11, which plots the evolution of the government's estimates from this simulation, provides some insight on the causes of these dynamics. The top panel plots the sum of the coefficients on inflation and unemployment in the government's model in the first 300 of the 20,000 simulation periods, while the bottom panel focuses on the last 1000 periods. In the top panel we see that initially there is a large increase in the coefficients on inflation and unemployment, which is what eventually triggers the large oscillations. Once the lower bound on inflation is hit around period 100, beliefs start coming back down. In the bottom panel we see that beliefs are relatively stable over the last 1000 periods at relatively low levels. The sum of the coefficients on inflation is typically negative, implying that in the long run the government does come to believe, at least temporarily, in an exploitable Phillips curve tradeoff. However we see that this sum occasionally increases to zero, and these episodes correspond to the falls in inflation to near 2% that we see in the lower right panel of Figure 10. This is precisely the behavior that Sargent (1999) stressed, as we discussed above. Thus in the long run our model has features similar to Sargent's original

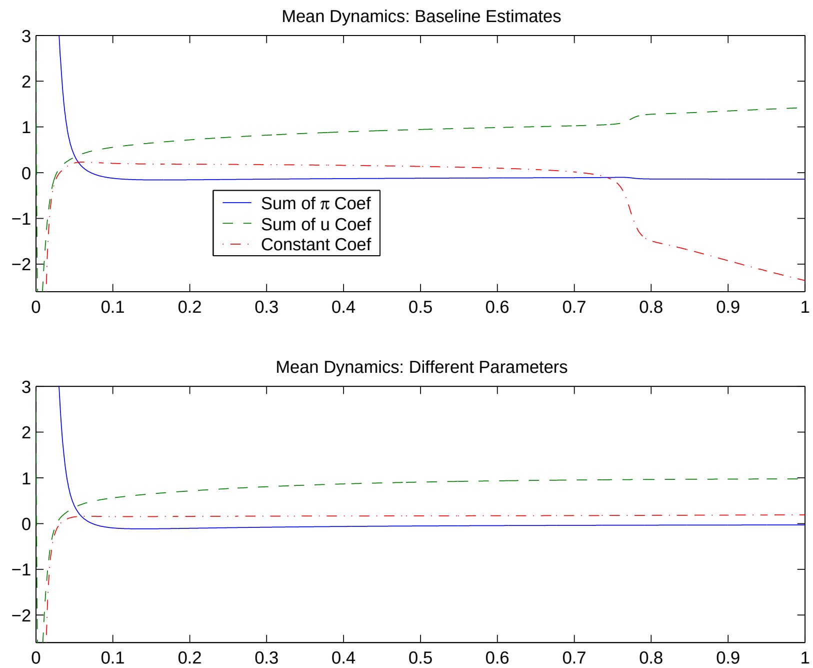

FIGURE 12. Mean dynamics for the baseline estimates and an alternative set of parameters.

model, but explaining the US experience requires that we start far from this limit. This suggests that we view the US monetary history as a process of continual learning.

VI.2. **Small Variation Limits.** In the previous section we saw some evidence in simulations that the economy converges to a limit distribution. In order to obtain more explicit analytic results, we consider small variation limits. While it is difficult to explicitly analyze the model for any arbitrary setting of V, for smaller V the beliefs drift at a slower rate and we can approximate their evolution via a differential equation. In particular, as in Sargent and Williams (2003) we let $V = \varepsilon^2 \hat{V}$ and $P_{t|t} = \varepsilon \hat{P}_{t|t}$. Then SW show that as $\varepsilon \rightarrow 0$ the sequence $\{\alpha_{t|t}, \hat{P}_{t|t}\}$ generated by (6)-(7) converges weakly to the solution of the following ODEs:

$$\dot{\alpha} = PE\left[\Phi_t(u_t - \Phi_t'\alpha)\right] \tag{12}$$

$$\dot{P} = \sigma^{-2}\hat{V} - PE(\Phi_t\Phi_t')P, \tag{13}$$

where the expectations are calculated for fixed α. As we let the prior belief variance go to zero by shrinking ε, the government's beliefs track the trajectories of these differential equations. We call the ODEs (12)-(13) the mean dynamics, as they govern the expected evolution of the government's beliefs. If the ODEs have a stable point $(\bar{\alpha}, \bar{P})$, then the government's beliefs will converge to it as $\varepsilon \rightarrow 0$ and $t \rightarrow \infty$. Note from (12) that the limiting beliefs satisfy the key least squares orthogonality condition:

$$E\left[\Phi_t(u_t - \Phi_t'\bar{\alpha})\right] = 0$$

and hence they comprise a self-confirming equilibrium. This orthogonality condition is the key identifying assumption in the government's subjective model, and in the limit it is satisfied by the true model.

In Figure 12 we plot trajectories the mean dynamics for the government's beliefs starting from our estimated initial conditions. The top panel plots the paths under our baseline estimates, while in the bottom panel we reduce the magnitude of both the contemporary Phillips curve slope parameter θ_0 and the coefficient on lagged unemployment τ_1 to values within the 90% probability bands. In each case, to make the figures more interpretable set the plot limits so that the initial conditions fall outside of the plot. In the top panel, we see that under our baseline estimates there is not a stable self-confirming equilibrium. The parameter estimates appear to converge at first, but then diverge later on. We found similar results from many other initial conditions, being unable to find a self-confirming equilibrium. This suggests that under our estimates as we scale down the rate of variation in beliefs by letting ε go to zero, we would get long run divergence. By experimenting, we found that the existence of a stable SCE was very sensitive to the θ_0 and τ_1 parameters. This explains our choice of the different parameters for the bottom panel. For these parameters, the mean dynamics look very similar to our baseline estimates, but they stay in a SCE and do not diverge. Thus the existence of a stable self-confirming equilibrium is somewhat fragile in our model.

However the mean dynamics and the self-confirming equilibrium only govern the dynamics of our model for small ε. It turns out that in practice ε must be quite small, on the order of 10^{-4}, for the asymptotic approximations to be valid. Thus for our baseline estimated V it is not terribly important whether a stable SCE exists or not. As we've seen above, we do get apparent convergence in the long run in this case and found no evidence of the long run divergence that the mean dynamics suggest. Put somewhat differently, for any V we get convergence to a limit distribution. As $\varepsilon \to 0$ this limit distribution converges to a self-confirming equilibrium, if it exists, or else the limit distribution itself ceases to exist. The nonexistence of a SCE under our baseline estimates essentially puts a lower bound on the rate of parameter drift that the government can entertain and still lead to reasonable long run behavior.

VII. CONCLUSION

Overall, our model provides a remarkably good fit to the data and is able to capture much of the rise and fall in inflation in the US. The parameters of the model are estimated rather sharply, and the model provides good one-step forecasts, and reasonably accurate forecasts out to three years. While some of the long run behavior of the model is a bit more troubling, such as the need to truncate the government's inflation policy, overall the model is reasonably successful.

In this paper we have thus provided a new viewpoint to the debate over post-war policy changes. We developed a structural model of learning with an optimizing government that is able to explain much of the rise and fall of inflation in the US. Rather than being driven by exogenous or abrupt shifts in behavior or changes in policy rules, our model suggests that the differing inflation outcomes over the post-war period have resulted from changes in the government's beliefs over time. Moreover the learning process has taken place gradually over the entire sample, suggesting that the US monetary experience may be viewed as a process of continual learning.

Appendix A. Data

The two monthly series employed in this paper are:

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

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

Appendix B. Prior Settings

To carry out the MCMC algorithm efficiently, we break ϕ into three separate blocks: θ, ζ_1 and ζ_2, and φ where

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

and $\varphi = \{z_{1|0}, u(C_P), u(C_V)\}$. The notation $u(C_P)$ or $u(C_V)$ means that only the upper triangular part of C_P or C_V is used. The prior distributions of both θ and φ take a Gaussian form. The prior mean for θ is set to

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

which implies that the natural rate of unemployment is 6.0 with somewhat persistent unemployment. The prior mean of θ_1 is only slightly less than that of θ_0 in absolute value

(.16 < .20), implying the low serial correlation of structural disturbances in Sargent's version of the Phillips curve (pp.70-71, Sargent 1999). The prior variance for θ is

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

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

The prior probability density for the precisions ζ_1 and ζ_2 is typically of Gamma distribution form:

$$p(\zeta_1, \zeta_2) = \text{Gamma}(\bar{\alpha}, \bar{\beta}) = \prod_{i=1}^{2} \frac{1}{\Gamma(\bar{\alpha})\bar{\beta}^{\bar{\alpha}}} \zeta_i^{\bar{\alpha}-1} e^{-\frac{\zeta_i}{\beta}}, \tag{B1}$$

where $\bar{\alpha} = 4$ and $\bar{\beta} = \lambda_2 12.5$. When $\lambda_2 = 1$, the prior mean for ζ_i is 50 and the prior variance is 25^2, implying a quite loose prior for ζ_i.

The prior mean for $z_{1|0}$ is zero and the prior variance is $\lambda_3 100^2$. The prior mean is 20 for the diagonal of C_P and 0 for the off-diagonal elements. The prior variance is $\lambda_4 10^2$. The prior mean is 10 for the diagonal of C_V and 0 for the off-diagonal elements. The prior variance is $\lambda_5 5^2$. In this paper, the values of all the tightness control hyperparameters from λ_1 to λ_5 are set to 1.[6]

While the prior described here derives from economic theory (e.g., θ_0 should be negative) and does not change the characteristics of the likelihood implied by the model, it is essential for obtaining finite-sample inferences.

APPENDIX C. PROPOSAL DENSITY FOR THE METROPOLIS ALGORITHM

The key to the Metropolis algorithm for the posterior distribution φ is to obtain the covariance matrix for a normal proposal density. Since x_{t-1} is a function of φ, one can approximate it by a second-order Taylor expansion at the posterior estimate $\hat{\varphi}$. It can be seen from (9) that this approximation leads to the following covariance matrix for φ:

$$\begin{aligned} \tilde{\Sigma}_{\varphi}^{-1} =& (\zeta_1 \theta_0^2 + \zeta_2) \sum_{t=2}^{T} \frac{\partial x_{t-1}(\hat{\varphi})}{\partial \varphi} \frac{\partial x'_{t-1}(\hat{\varphi})}{\partial \varphi} + \zeta_1 \theta_1^2 \sum_{t=2}^{T} \frac{\partial x_{t-2}(\hat{\varphi})}{\partial \varphi} \frac{\partial x'_{t-2}(\hat{\varphi})}{\partial \varphi} \\ &+ \zeta_1 \theta_0 \theta_1 \sum_{t=2}^{T} \left[\frac{\partial x_{t-1}(\hat{\varphi})}{\partial \varphi} \frac{\partial x'_{t-2}(\hat{\varphi})}{\partial \varphi} + \frac{\partial x_{t-2}(\hat{\varphi})}{\partial \varphi} \frac{\partial x'_{t-1}(\hat{\varphi})}{\partial \varphi} \right] + \bar{\Sigma}_{\varphi}^{-1}, \end{aligned} \tag{C1}$$

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

[6] We also vary the values of these parameters and the results are quite similar.

REFERENCES

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

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

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

COGLEY, T., AND T. J. SARGENT (2003): "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.," Manuscript, University of California (Davis) and New York University.

HAMILTON, J. D., D. F. WAGGONER, AND T. ZHA (2003): "Normalization in Econometrics," Manuscript, University of California (San Diego) and Federal Reserve Bank of Atlanta.

PRIMICERI, G. (2003a): "Time Varying Structural Vector Autoregressions and Monetary Policy," Princeton University.

——— (2003b): "Why Inflation Rose and Fell: Policymakers' Beliefs and US Postwar Stabilization Policy," Princeton University.

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

SARGENT, T. J., AND N. WILLIAMS (2003): "Impacts of Priors on Convergence and Escapes from Nash Inflation," Manuscript, New York University and Princeton University.

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

SIMS, C. A., AND T. ZHA (2004): "Were There Regime Switches in US Monetary Policy?," Manuscript, Princeton University and Federal Reserve Bank of Atlanta.

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

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