Benefits from U.S. Monetary Policy Experimentation in the Days of Samuelson and Solow and Lucas*

Timothy Cogley[†]
Riccardo Colacito[‡]
Thomas J. Sargent[§]

March 12, 2005

Abstract

A policy maker knows two models of inflation-unemployment dynamics. One implies an exploitable trade-off, the other does not. The policy maker's prior probability over the two models is part of his state vector. Bayes' law converts the prior probability into a posterior probability at each date and gives the policy maker an incentive to experiment. For two models calibrated to U.S. data through the early 1960s, we isolate the component of government policy that is due to experimentation by comparing the outcomes from two Bellman equations, the first of which 'experiments and learns', the second of which 'learns but doesn't experiment'. We interpret the second as an 'anticipated utility' model and study how well its outcomes approximate those from the 'experiment and learn' Bellman equation. The approximation is good. For our calibrations, the benefits from purposeful experimentation are small.

Key words: Learning, model uncertainty, Bayes' law, Phillips curve, opportunism.

*We thank Klaus Adams, Patrick Bajari, Lars Hansen, and Martin Schneider for useful discussions.

[†]University of California, Davis. Email: twcogley@ucdavis.edu
[‡]New York University. Email: rc652@nyu.edu
[§]New York University and Hoover Institution. Email: ts43@nyu.edu

1 Introduction

We estimate the value of deliberate experimentation to a policy maker when two models that fit the historical data approximately equally well have sharply different operating characteristics. Two competing models of inflation-unemployment dynamics differ with respect to whether they imply an exploitable Phillips curve.

To capture a debate that raged between advocates of the natural unemployment hypothesis and others who thought that there was an exploitable unemployment-inflation trade-off, we imagine that a monetary policy authority has the following two models of inflation-unemployment dynamics:[1]

- Model 1 (Samuelson-Solow):

$$U_t = .0023 + .7971U_{t-1} - .2761\pi_t + .0054\eta_{1,t}$$
$$\pi_t = v_{t-1} + .0055\eta_{3t}$$

- Model 2 (Lucas):

$$U_t = .0007 + .8468U_{t-1} - .2489(\pi_t - v_{t-1}) + .0055\eta_{2,t}$$
$$\pi_t = v_{t-1} + .0055\eta_{4t}$$

where U_t is the deviation of the unemployment rate from an exogenous measure of a natural rate U_t^*, π_t is the quarterly rate of inflation, v_{t-1} is the rate of inflation that at time $t-1$ the monetary authority and private agents had both expected to prevail at time t, and, for $i = 1, 2, 3, 4$, η_{it} is an i.i.d. Gaussian sequence with mean zero and variance 1. The monetary authority has a Kydland-Prescott (1977) loss function $E\sum_{t=0}^{\infty} \beta^t r_t$, where $r_t = -.5(U_t^2 + \lambda v_t^2)$.[2] The monetary authority sets v_t as a function of time t information.[3] The monetary authority attaches probability

[1]We use these specifications mainly as a device to get good fitting models while keeping the dimension of the state to the minimum required to represent 'natural rate' and 'non-natural rate' theories of unemployment. See appendix D for details.

[2]Alan Blinder (1998) has stressed that this objective function forces a conflict between the policy maker (who prefers an unemployment rate lower than the natural rate) and the public (which wants unemployment to equal the natural rate) that induces the time consistency problem for inflation described by Kydland and Prescott (1977).

[3]Under this timing protocol, there is no time-consistency problem in Kydland and Prescott's model. See Stokey (1989).

α_0 to model 1 and probability $1 - \alpha_0$ to model 2.[4,5] Although they fit the U.S. data from 1948:3-1963:I almost equally well, under our loss function these two models call for very different policies toward inflation. Model 1, whose main features many have attributed to Samuelson and Solow (1960), has an exploitable tradeoff between v_t and subsequent levels of unemployment. Having operating characteristics advocated by Lucas (1972, 1973) and Sargent (1973), model 2 has no exploitable Phillips curve: variations in the predictable part of inflation v_t affect inflation but not unemployment. If $\alpha_0 = 0$, then our decision maker should implement the trivial policy $v_t = 0$ for all t. However, if $\alpha_0 > 0$, the policy maker should set $v_t \neq 0$, partly to exploit a probable inflation-unemployment tradeoff, and partly to refine α.

We formulate and solve the pertinent Bellman equation then use the optimal decision rule to study the following questions:

1. Suppose that the Samuelson-Solow model generates the data and that before date T the government had assigned probability $\alpha = 1$ to the Samuelson-Solow model and had used the corresponding optimal policy for a long time, so that the economy is in a stochastic steady state at date $T - 1$. Having been persuaded by an advocate of the natural rate hypothesis, at date T the government suddenly lowers α to a number $\alpha \in (0, 1)$ even though, unbeknownst to the government, the Samuelson-Solow model is true. Under these assumptions, we use our model to quantify the adverse effects on government policy that follow from its attaching some weight to the Lucas model at dates $t \geq T$. Lucas's model is subversive in the sense that it leads to higher unemployment than would have prevailed had it never been invented. We ask how much higher is unemployment, and how long does it take for the government to forget the Lucas model?

2. Suppose that the Lucas model generates the data and that before date T the government had assigned probability $1 - \alpha = 1$ to the Lucas model and had used the corresponding optimal policy for a long time, so that the economy is

[4]We assume that model parameters are known because we want to reduce the dimension of the monetary authority's posterior distribution. If we were to treat the parameters as unknown, probability distributions for those parameters would be part of the monetary authority's prior, increasing the dimension of the state beyond what we can manage computationally. See Wieland (2000a,b) and Beck and Wieland (2002) for analysis of the Bellman equation for a decision maker who experiments to learn about parameter values. See El-Gamal and Rangarajan (1993) for an analysis of convergence in a class of models in which agents are learning. Kenneth Kasa (1999) adapts results that earlier researchers had obtained for a monopolist who could learn, but chooses not to learn, his demand curve. Kasa thereby creates a model in which the Fed chooses not to learn objects that could be learned with a different strategy.

[5]Hansen and Sargent's (2005a) T^1 operator, which we apply in a companion paper, allows us to analyze robustness to model perturbations that can be interpreted as coefficient uncertainty.

in a stochastic steady state at date $T-1$. At date T, having been persuaded by advocates of the Samuelson-Solow model, the government suddenly lowers $1-\alpha$ to a number in $(0,1)$ even though, unbeknownst to the government, the Lucas model is true. We use our model to quantify the effects on government policy that follow from its putting some weight on the Samuelson-Solow model. Samuelson and Solow's model is pernicious in leading to higher inflation but no lower unemployment than if it had never been thought of. We study how much more inflation is produced by this scenario and how long it takes for the data to discredit the Samuelson-Solow model.

3. We want to quantify the role of 'active' as opposed to 'passive' experimentation. We do this by comparing the decision rule and value function for the problem that includes α as a state variable and Bayes' law as a transition equation with the decision rule and value function that come from another problem that suppresses α as a state variable and Bayes' law as a transition equation. By comparing the associated decision rules, we identify a component of time t decisions that is attributable to intentional experimentation. Another perspective on these calculations is that they allow us to evaluate how well 'anticipated utility' decision rules approximate the decision rules that do experiment optimally.[6]

1.1 Organization

Section 2 formulates Bellman equations, one for a decision maker who intentionally experiments, another for an 'anticipated utility' decision maker who does not consciously experiment. These Bellman equations describe alternative states of mind for the policy maker. Section 3 describes alternative ways of modelling how the true data generating model relates to the policy maker's state of mind, and circumstances under which the government eventually learns the truth. Section 4 discusses our numerical approximations to the value functions and decision rules. Section 5 describes quantitative experiments designed to answer the three questions asked above, as well as a variety of statistics on 'waiting times' to learn the truth. Section 6 adds some concluding remarks. Four appendixes contain technical details about how we solved the Bellman equations and calibrated the two models. In a companion paper Cogley, Colacito, Hansen, and Sargent (2005), we study how the decision maker responds to concerns about two distinct sources of misspecification of his model, namely, misspecification of his prior over the two models, and misspecification of each of those

[6] Cogley and Sargent (2005) use an anticipated utility approximation to compute decision rules in the context of a model whose state is so large that calculations like those in this paper succumb to the curse of dimensionality. See Kreps (1998) for a justification of anticipated utility models.

approximating models.

2 Two formulations of the policy problem under model uncertainty

We map our example into a general setup, then state Bellman equations for the government under two alternative assumptions about the government's response to the opportunity to experiment.

2.1 The models

The policy maker has two models

$$s_{t+1} = A_i s_t + B_i v_t + C_i \epsilon_{i,t+1}, \tag{1}$$

$i = 1, 2$, where s_t is a state vector, v_t is a control vector, and $\epsilon_{i,t+1}$ is an i.i.d. Gaussian process with mean zero and contemporaneous covariance matrix I. Let $F(\cdot)$ denote the c.d.f. of this normalized multivariate Gaussian distribution. At time t, the policy maker has observed a history of outcomes $s^t = s_t, s_{t-1}, \ldots, s_0$ and assigns probability α_t to model 1 and probability $(1 - \alpha_t)$ to model 2. By applying Bayes' Law, the policy maker updates α_t:

$$\alpha_{t+1} = B(\alpha_t, s_{t+1}). \tag{2}$$

In equations (32) and (36) in appendix A, we provide a formula for $B(\alpha_t, s_{t+1})$. The policy maker wants a policy for setting v_t that maximizes

$$E_0 \sum_{t=0}^{\infty} \beta^t r(s_t, v_t), \quad \beta \in (0, 1), \tag{3}$$

where E_0 is a mathematical expectation with respect to the distribution over future outcomes induced by the models (1) and the policy maker's opinions about them.

2.2 Intentional experimentation

The policy maker's belief α_t is a component of the time t state vector (s_t, α_t). In choosing v_t, it is in the policy maker's interest to recognize the revisions of his beliefs that he foresees will occur through equation (2). Let $V(s_t, \alpha_t)$ be the optimal value in state (s_t, α_t). The Bellman equation is

$$
\begin{aligned}
V(s_t, \alpha_t) = \max_{v_t} \{ & r(s_t, v_t) \\
& + \beta \alpha_t \int V(A_1 s_t + B_1 v_t + C_1 \epsilon_{1,t+1}, B(\alpha_t, A_1 s_t + B_1 v_t + C_1 \epsilon_{1,t+1})) dF(\epsilon_{1,t+1}) \\
& + \beta (1-\alpha_t) \int V(A_2 s_t + B_2 v_t + C_2 \epsilon_{2,t+1}, B(\alpha_t, A_2 s_t + B_2 v_t + C_2 \epsilon_{2,t+1})) dF(\epsilon_{2,t+1}) \}
\end{aligned} \tag{4}
$$

The optimal decision rule can be represented recursively as

$$
v_t = v(s_t, \alpha_t) \tag{5}
$$

$$
\alpha_{t+1} = \alpha(s_t, \alpha_t). \tag{6}
$$

Repeated substitution of (6) into (5) yields the policy maker's strategy in the form of a sequence of functions

$$
v_t = \sigma_t(s^t, \alpha_0), \tag{7}
$$

where $s^t = (s_t, s_{t-1}, \dots, s_0)$. The presence of $B(\alpha_t, A_i s_t + B_i v_t + C_i \epsilon_{t+1})$, $i = 1, 2$, on the right side of (4) imparts a motive to experiment. To choose v_t is to design experiments.

2.3 Bellman equation in detail

Appendix A derives the function $B(s_t, \alpha_t)$ and thereby obtains a particular version of (4) that we approximate numerically. Let $\Omega_i = C_i C_i'$, $R_t = \frac{\alpha_t}{1-\alpha_t}$, and define

$$
\begin{aligned}
g(\epsilon_{1,t+1}; s_t, \alpha_t) = \log R_t - \frac{1}{2} \log |\Omega_1| + \frac{1}{2} \log |\Omega_2| & - \frac{1}{2} (C_1 \epsilon_{1,t+1})' \, \Omega_1^{-1} (C_1 \epsilon_{1,t+1}) \\
+ \frac{1}{2} & [(A_1 - A_2) s_t + (B_1 - B_2) v_t + C_1 \epsilon_{1,t+1}]' \\
\times \Omega_2^{-1} & [(A_1 - A_2) s_t + (B_1 - B_2) v_t + C_1 \epsilon_{1,t+1}]
\end{aligned} \tag{8}
$$

and

$$
\begin{aligned}
h(\epsilon_{2,t+1}; s_t, \alpha_t) = \log R_t - \frac{1}{2} \log |\Omega_1| + \frac{1}{2} \log |\Omega_2| & + \frac{1}{2} (C_2 \epsilon_{2,t+1})' \, \Omega_2^{-1} (C_2 \epsilon_{2,t+1}) \\
- \frac{1}{2} & [(A_2 - A_1) s_t + (B_2 - B_1) v_t + C_2 \epsilon_{2,t+1}]' \\
\times \Omega_1^{-1} & [(A_2 - A_1) s_t + (B_2 - B_1) v_t + C_2 \epsilon_{2,t+1}].
\end{aligned} \tag{9}
$$

Using (32) in Appendix A, we obtain a law of motion for α_{t+1} under the two models. Then Bellman equation (4) becomes

$$
\begin{aligned}
V(s_t, \alpha_t) = \max_{v_t} \Bigg\{ & r(s_t, v_t) + \beta\alpha_t \int V\left(A_1 s_t + B_1 v_t + C_1 \epsilon_{1,t+1}, \frac{e^{g(\epsilon_{1,t+1})}}{1 + e^{g(\epsilon_{1,t+1})}}\right) dF(\epsilon_{1,t+1}) \\
& + \beta(1-\alpha_t) \int V\left(A_2 s_t + B_2 v_t + C_2 \epsilon_{2,t+1}, \frac{e^{h(\epsilon_{2,t+1})}}{1 + e^{h(\epsilon_{2,t+1})}}\right) dF(\epsilon_{2,t+1}) \Bigg\}. \qquad (10)
\end{aligned}
$$

Appendix B describes how we approximate the solution of (10).

2.4 Attitudes toward experimentation

Despite knowing (4), prominent macroeconomists have advised against exploiting the opportunity (or succumbing to the temptation) to experiment identified by the right side of Bellman equation (4). Blinder (1998, p. 11) asserts that

> "while there are some fairly sophisticated techniques for dealing with parameter uncertainty in optimal control models with learning, those methods have not attracted the attention of either macroeconomists or policymakers. There is a good reason for this inattention, I think: You don't conduct policy experiments on a real economy solely to sharpen your econometric estimates."

Lucas (1981, p. 288) agrees, remarking that

> "Social experiments on the grand scale may be instructive and admirable, but they are best admired at a distance. The idea, if the marginal social product of economics is positive, must be to gain some confidence that the component parts of the program are in some sense reliable prior to running it at the expense of our neighbors."

Perhaps Blinder and Lucas suspect that the decision maker has too few models on the table (e.g., that neither of models in Bellman equation (4) is correct), or that it would be difficult to specify a prior over such models, and that therefore the decision problem is misspecified. Cogley, Colacito, Hansen, and Sargent (2005) describe how to address such concerns by applying the methods of Hansen and Sargent (2005a).

Another reason for not deliberately experimenting is that it is very difficult to approximate the solution of the Bellman equation that corresponds to (4) when there are more dimensions of uncertainty, e.g., unknown coefficients and more models. To sidestep that problem, researchers like Cogley and Sargent (2005) have appealed to Kreps's (1998) 'anticipated utility' model to justify an adaptive approach that abstracts from deliberate experimentation.

A third possible reason for being skeptical about experiments is related to the previous two. We can interpret the fact that Bellman equation (4) is difficult to solve as saying that it is difficult to design optimal experiments. The value function that obeys (4) is maximized over all possible experiments. Many suboptimal experiments would attain lower values than those delivered by the 'don't experiment' rule that solves an alternative Bellman equation that we now describe.

2.5 Unintentional experimentation

Comparing (4) with another Bellman equation lets us quantify how much the policy maker sacrifices by abstaining from the opportunity to experiment. We formulate an optimum problem that ignores the opportunity to experiment by replacing the law of motion (2) for α_t dictated by Bayes' law with the "don't experiment on purpose" specification

$$\alpha_{t+s} = \alpha_t \quad \forall s \geq 0. \tag{11}$$

When he makes a decision at time t, the policy maker pretends that he cannot or will not learn about the model from future data. One interpretation of this assumption is that the policy maker believes that nature will draw next period's s_{t+1} from an α-weighted mixture of models 1 and 2. Another interpretation is that the policy maker plans not to revise his views. Under either interpretation, a policy maker with this fixed-α view chooses a policy

$$v_t = w(s_t; \alpha_t). \tag{12}$$

that attains maximizes the right side of the following Bellman equation:[7]

$$\begin{aligned} \widetilde{W}(s_t; \alpha_t) = \max_{v_t} \{ & r(s_t, v_t) \\ & + \beta \alpha_t \int \widetilde{W}(A_1 s_t + B_1 v_t + C_1 \epsilon_{t+1}; \alpha_t) dF(\epsilon_{1,t+1}) \\ & + \beta (1 - \alpha_t) \int \widetilde{W}(A_2 s_t + B_2 v_t + C_2 \epsilon_{t+1}; \alpha_t) dF(\epsilon_{2,t+1}) \}. \end{aligned} \tag{13}$$

[7] Appendix C describes our algorithm for solving (13).

Notice how , by using (11) instead of (2) as the law of motion for α, (13) suppresses the motive to experiment deliberately. But in the spirit of the adaptive control literature, suppose that the policy maker does indeed revise α_t by applying Bayes' Law even though he uses a policy (12) derived by solving the abstain-from-learning Bellman equation (13). Then his actual decisions can be represented recursively as (2) and (12).

These decisions would emerge from a '*don't* experiment but *do* learn' prescription.[8] Equations (2), (12) can be solved by repeated substitution to yield the policy maker's strategy in the form of a sequence of functions

$$v_t = \tilde{\sigma}_t(s^t, \alpha_0). \tag{14}$$

Thus, although in formulating his policy, the policy maker ignores the motion in α_t impelled by Bayes' law (2), Bayes' law nevertheless affects the value that he can expect to attain under the policy (12), which evidently satisfies the Bellman equation

$$\begin{aligned} W(s_t;\alpha_t) &= r(s_t, w(s_t,\alpha_t)) \\ &+ \beta\alpha_t \int W(A_1 s_t + B_1 w(s_t,\alpha_t) + C_1\epsilon_{1,t+1}, B(\alpha_t, A_1 s_t + B_1 w(s_t,\alpha_t) + C_1\epsilon_{1,t+1}))dF(\epsilon_{1,t+1}) \\ &+ \beta(1-\alpha_t) \int W(A_2 s_t + B_2 w(s_t,\alpha_t) + C_2\epsilon_{2,t+1}, B(\alpha_t, A_2 s_t + B_2 w(s_t,\alpha_t) + C_2\epsilon_{2,t+1}))dF(\epsilon_{2,t+1}). \end{aligned} \tag{15}$$

Because (12) is a feasible policy for the decision maker of subsection 2.2 who is willing to experiment, it follows that

$$V(s,\alpha) \geq W(s;\alpha) \tag{16}$$

for all values of (s,α). The gap

$$V(s,\alpha) - W(s;\alpha) \tag{17}$$

measures the value of experimentation and the difference

$$v(s,\alpha) - w(s;\alpha) \tag{18}$$

measures the component of the policy choice that can be attributed purely to the policy maker's motive to experiment.

2.5.1 Anticipated utility as an approximation

In addition to representing a stylized '*don't* experiment but *do* learn' view, rules like (2)-(12) have been recommended as an alternative or approximation to (5)-(6) to

[8]We interpret Blinder (1998, chapter 1) as advocating this point of view.

be used in situations in which the curse of dimensionality prevents the policy maker or the analyst from solving the pertinent counterpart to Bellman equation (4). The appeal of this approximation is greatest when the dimension of the prior distribution is large.[9] We have assumed that A_i, B_i, C_i in (1) are known matrices. Had we assumed instead that the policy maker has a nontrivial prior probability distribution over those parameters, those distributions would enter the value function on the left side of (4). The Bellman equation for this value function would be difficult to solve because of the dimension of the state vector.

3 The truth

So far, everything we have said is about what the monetary authority believes and how it chooses v_t. We have yet to say how the economy actually works. Our description so far has been about ideas that are 'just in the head' of the monetary authority and how it responds to those ideas.

Under the monetary authority's prior distribution over sequences for unemployment and inflation that is implied by our specification, α_t is a martingale. See appendix A, section A.1 for a proof. Because $\alpha_t \in [0,1]$, the martingale convergence theorem implies that α_t converges almost surely under that measure. To say what happens to α_t under the measure that actually generates the economy, we have to say what that true measure is. If we assume that one of our two models, either model 1 (Samuelson and Solow's) or model 2 (Lucas's), or some fixed-α mixture of them, governs the data, then α_t given by (28) converges almost surely to the true α.[10]

Our concern in the next section is to study the *rates* at which α_t converges to the true α under alternative assumptions about which model is the true data generating process and alternative initial conditions for (α, U). We design alternative scenarios to shed light on the questions stated in section 1 and to determine which of our two models is more difficult to learn about.

4 Value functions and decision rules

We have reported calibrated versions of our two models in section 1. For government preference parameters, we begin by setting $\beta = .995$ and $\lambda = 0.1$. (In section 5.4,

[9] See footnote 4.

[10] Our model has a feature that El-Gamal and Rangarajan (1993) identify as important in promoting convergence, namely, the presence of an exogenous component of randomness that generates 'natural experiments' that can help discriminate between models even if the policy maker decides not to experiment in setting his policy.

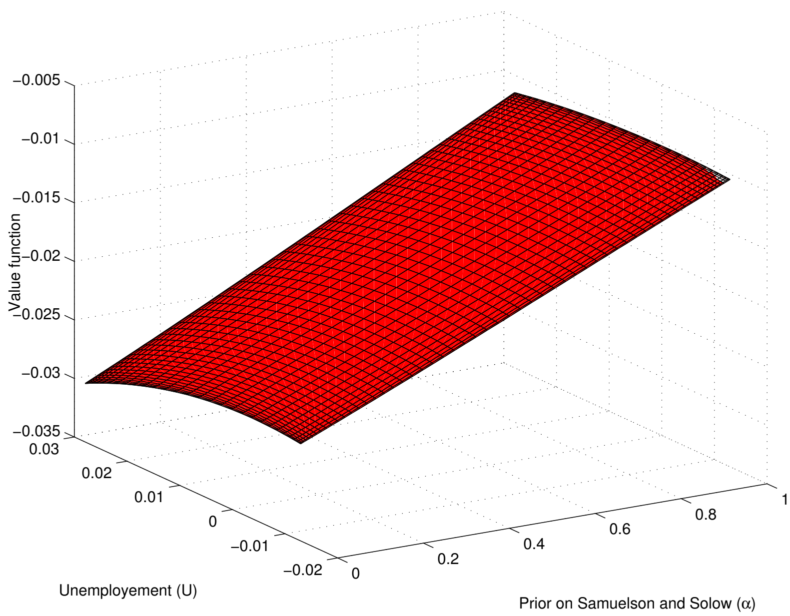


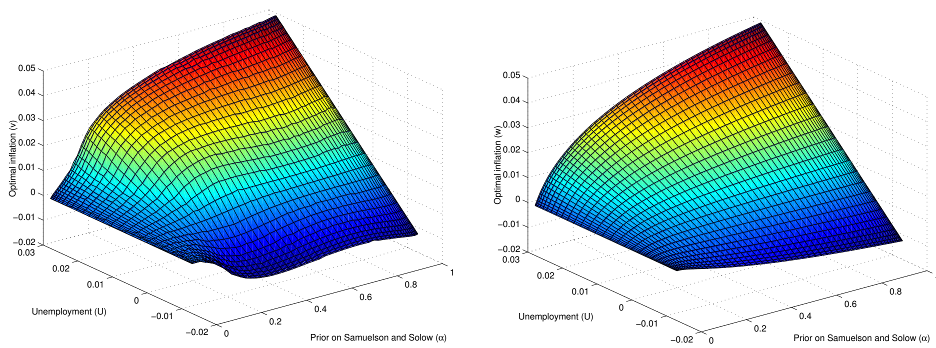
Figure 1: Two value functions: $W(U, \alpha) \leq V(U, \alpha)$.

we shall consider other values of λ.) A high value of β promotes experimentation because the costs of experimentation are paid up front while the benefits accrue in the future. A patient government is therefore more inclined to experiment. The parameter λ is the relative weight on inflation in the government's loss function. A low value means the monetary authority cares much more about unemployment, which may have been the case in the days of Samuelson and Solow.

Figures 1, 2, and 3 display value functions and decision rules associated with our two Bellman equations (4) and (13). Figure 1 shows both W and V, but they are so close that they cannot be distinguished. 3 shows their difference and confirms that $V(U, \alpha) > W(U, \alpha)$ except at the boundaries $\alpha = 1$ and $\alpha = 0$, where $V(U, \alpha) = W(U, \alpha)$.[11] This relationship of the 'experiment and learn' value function V to the 'don't experiment but learn' value function W is as expected: when $\alpha \in (0, 1)$, there is value to intentional experimentation. The policy functions in figure 2 and their difference in 3, panel b, show the different actions called for by the decision rules v and w associated with Bellman equations (4) and (13), respectively.

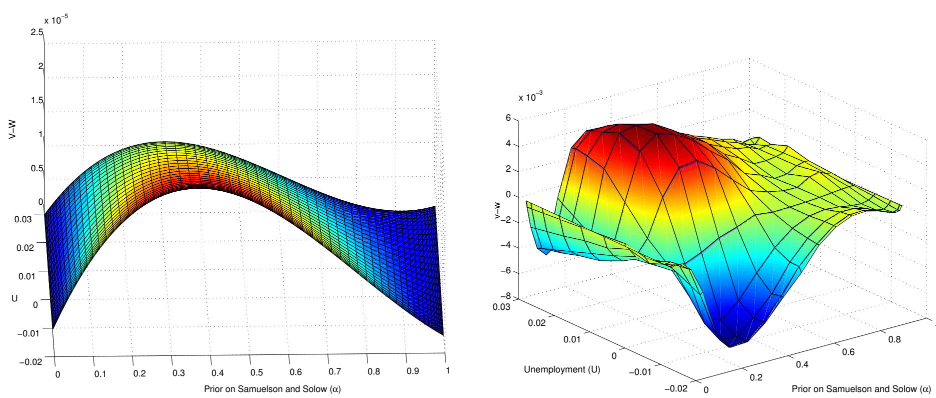
Overall, the differences between the value functions and the decision rules are

[11] The prior is dogmatic at the boundaries, where data never alter the policy maker's beliefs. Those who are sure they know the truth are uninterested in experimentation.



(a) $v(U, \alpha)$ (b) $w(U, \alpha)$

Figure 2: Policy functions with and without experimentation.



(a) $V(U, \alpha) - W(U, \alpha)$ (b) $v(U, \alpha) - w(v, \alpha)$

Figure 3: Differences in value functions and policies with and without deliberate experimentation.

both small. Therefore, in this example at least, the type of anticipated utility model used by Cogley and Sargent (2005), which is associated with Bellman equation (13), seems to provide a good approximation to the outcomes from the intentional experimentation model.[12] We study the quality of approximation more fully in section 5, where we analyze the questions posed in section 1.

To bring out their differences, figure 4 shows the decision rules $w(U, \alpha)$ and $v(U, \alpha)$ as functions of U for different values of α. As noted, the differences between v and w are always small, but the biggest differences occur for α's away from the boundaries of 0 and 1. The figures reveal that when α is well into the interior of $(0, 1)$, v's call for additional experimentation serves to make it nonlinear and to enhance the countercyclicality of inflation policy. That is, the v-inflation policy is higher than the w-inflation policy when U is high, and lower when U is low. This pattern reveals a kind of 'opportunism': the best time to experiment with Keynesian stimulus is when U is high.[13]

Another interesting feature of figure 4 is that for both the v and w decision rules, policy begins quickly to look more Keynesian even for $\alpha = 0.2$ (i.e., a small weight on the Samuelson-Solow model), while it continues to look quite Keynesian when there is a comparable small weight of $1 - \alpha = 0.2$ on the Lucas model. Thus, a little bit of doubt about the Lucas model makes the policy maker begin to behave like a Keynesian, while a Keynesian has to have bigger doubts about the Samuelson-Solow model to begin behaving as Lucas's model advises.[14] This follows from two features, namely, the policy ineffectiveness proposition inherent in the Lucas model and the assumption that $\lambda = 0.1$. The latter assumption means that the monetary authority cares mostly about real variables, and the policy-ineffectiveness proposition states that one policy rule is as good as another with respect to their effects on real variables. Accordingly, a little bit of doubt about the Lucas model is enough to make the policy maker acquiesce to Keynesian policy prescriptions. Section 5 revisits this issue for different values of λ.

4.1 Asymmetry of $V - W$

Also notice the asymmetry of $V(U, \alpha) - W(U, \alpha)$ shown in panel (a) of figure 3. This surface is more steeply sloped along the α-axis when α is close to zero than when

[12]See David Kreps (1998) for a broader defense of this modelling strategy in games and dynamic economic models.

[13]In contrast, Alan Blinder's opportunistic call for more deflation in recessions seems to have been motivated not by an appeal to optimal experimentation but as a way for the Fed to find political cover for reducing inflation. See Orphanides and Wilcox (2002) for an account of Blinder's argument.

[14]This feature of the decision rules conforms to the story about the conquest of American inflation told by Cogley and Sargent (2005).

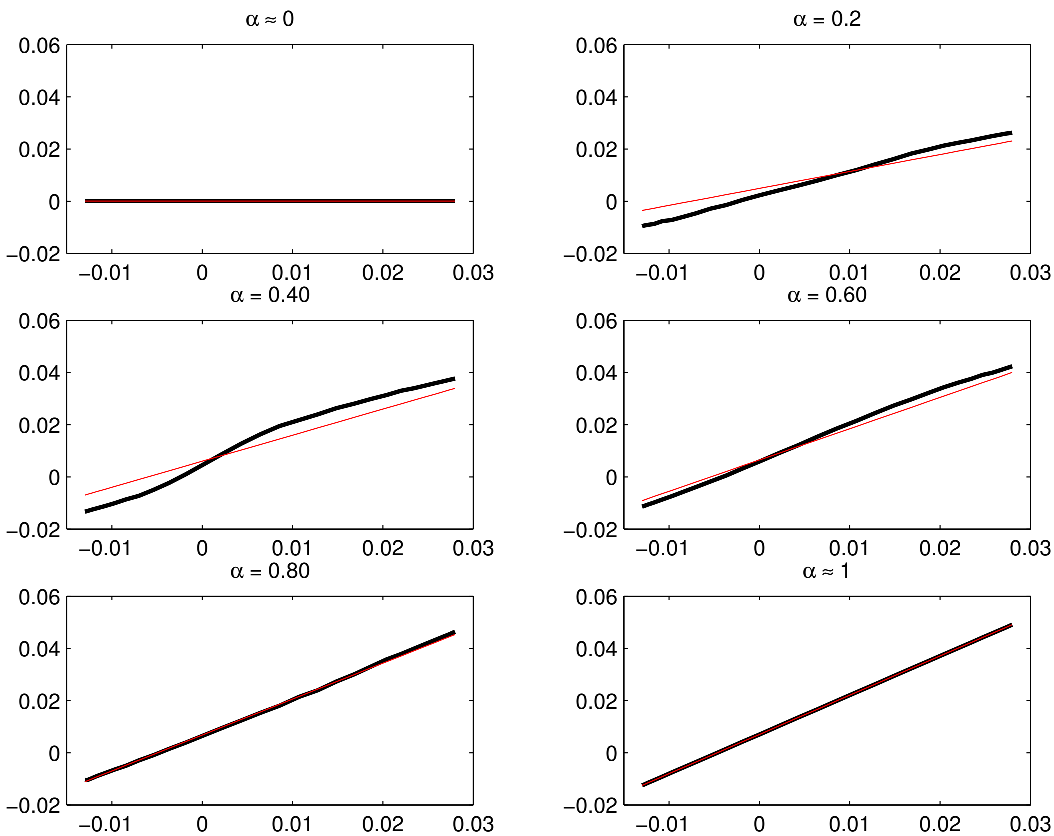


Figure 4: Slices of the optimal decision rules for inflation. The bold line is $v(U, \alpha)$ and the other line is $w(U, \alpha)$.

α is near one. That means deliberate experimentation has more incremental value for a monetary authority that leans toward the Lucas model than for one that leans toward the Samuelson-Solow model.

To understand why, think about the information content of the passive w-policy and how much intentional experimentation would add to it. For α close to 1, the w-policy calls for inflation to vary energetically in response to unemployment. That would eventually result in lower average and less variable unemployment if the Samuelson-Solow model were true, but would not alter the properties of unemployment if the Lucas model were true. So for α close to one, the w-policy itself provides identifying information about the world. The policy maker learns not only from natural experiments arising from shocks, but also from variation in inflation arising from its w-policy. That is not the case when α is close to zero, for then the w-policy always keeps expected inflation close to zero. Because the w-policy provides little identifying information in this case, the policy maker would have to rely solely on natural experiments to learn the truth. Thus, under the w-policies, learning is likely to take longer when $\alpha_0 \doteq 0$ than when $\alpha_0 \doteq 1$.[15] Other things equal, deliberate experimentation is less attractive when model uncertainty is likely to evaporate quickly on its own, and that explains why $V - W$ rises more slowly from $\alpha \doteq 1$.

These features of our policy rules and value functions will influence outcomes of the experiments that we report in the next section.

5 Forgetting pernicious ideas

We generate alternative scenarios by specifying initial conditions for U and government beliefs α and then making an assumption about which of our two models generates the data. We use the policy functions in figure 2 to generate histories of outcomes, and we relate those outcomes to questions 1 and 2 in the introduction.

5.1 Misplaced experimentation when Samuelson and Solow are correct

Assume that the data generating process is the Samuelson and Solow model. Figure 5 shows outcomes after policy maker erroneously assign positive probability to the Lucas model. For the first 19 periods, the policy maker had $\alpha = 1$ and therefore had optimally exploited the tradeoff between unemployment and inflation given by the Samuelson-Solow model. In period 19, the policy maker assigns a positive probability to Lucas's model. Starting from period 19, we model the behavior of three policy makers. As a benchmark, the first one (dotted line in the pictures) continues to assign probability one to the Samuelson-Solow model and therefore abstains from experimenting or learning. The second and the third ones attach a prior probability of 75% to the Lucas model being true. The second policy maker takes into account that this prior will be revised in subsequent periods (black continuous lines), while the third (red lines) does not.

In this scenario, Lucas's idea is destructive because it distorts the policy rule relative to the optimal $\alpha = 1$ policy. The Samuelson-Solow model offers a lever for maintaining low average unemployment and reducing its variability. The experimental policy initially calls for lower inflation relative to the optimal $\alpha = 1$ policy and a weaker countercyclical response to unemployment, and that results in higher average unemployment and greater cyclical variability. But notice how small the differences are. Unemployment is initially a bit higher, but the differences are visually

[15] Indeed, this emerges in the simulations reported below.

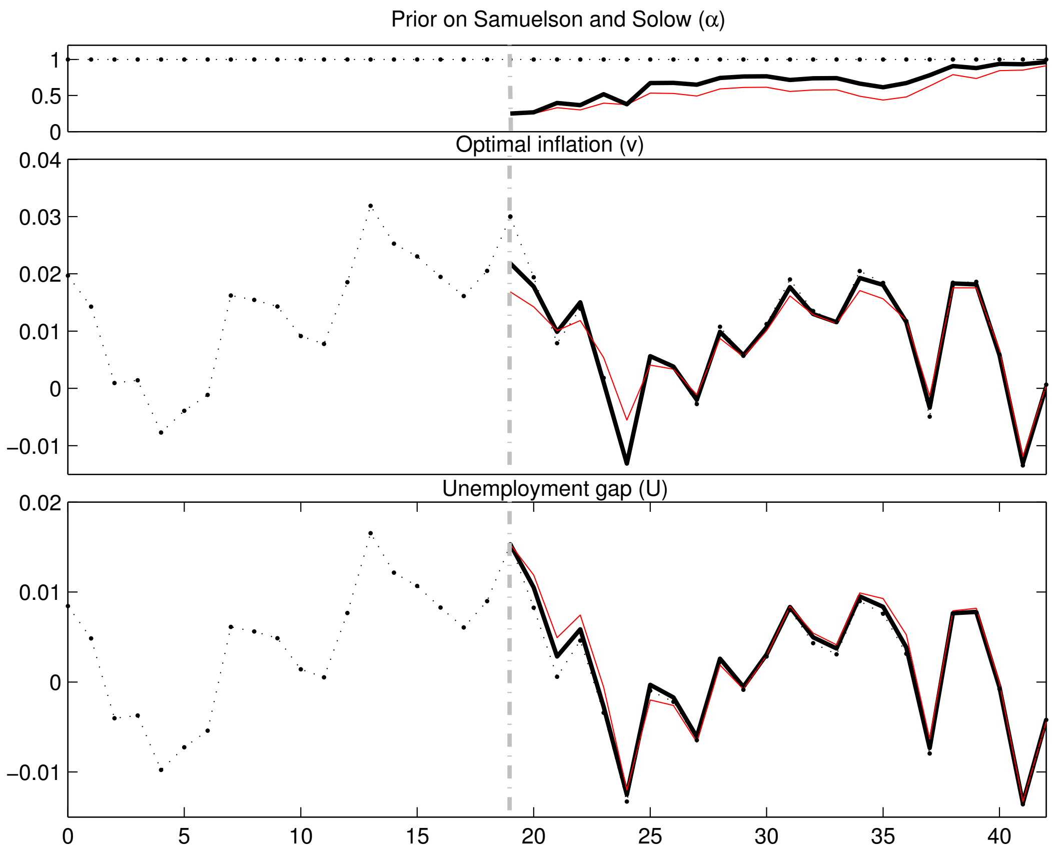


Figure 5: In the three panels: the dotted line represents the behavior of a policy maker that attaches probability one to the Samuelson and Solow model, the bold continuous line is the experimenting policy maker and the other line is the non-experimenting policy maker.

hard to detect. This follows from the Keynesian nature of the experimental policy. An optimal experiment does not involve a sudden, sharp drop to zero inflation. On the contrary, it still calls for countercyclical inflation variation, just tempered somewhat relative to the $\alpha = 1$ rule.[16] Thus, Lucas's ideas are not too pernicious in this setting. The policy maker does not forfeit the opportunity to accomplish its unemployment objectives when experimenting with the Lucas model; it just pursues them less energetically for a time.

The figure also compares outcomes of the experimental v-policy with those of the adaptive w-policy. The experimenting policy maker keeps inflation higher for a while

[16] I.e., in figure 4, the decision rule is less steeply sloped when $\alpha = 0.2$ than when $\alpha = 1$, but it still slopes upward.

in exchange for a sharper decrease in unemployment compared to the anticipated utility policy maker. The second policy maker evidently learns faster than the third. But once again, the outcomes are similar. At least in this instance, the adaptive policy well approximates the experimental policy, which suggests that the incremental benefit of deliberate experimentation is slight.

5.2 Misplaced experimentation when the Lucas model is true

Now assume that Lucas's is the true data generating mechanism. Figure 6 shows outcomes when after 19 periods of correct policy under Lucas's model, the monetary policy decision maker assigns a probability $\alpha = 0.75$ to Samuelson and Solow's model. We display paths for three types of decision makers. As a benchmark, the first continues to assign probability one to Lucas's model throughout and neither learns nor experiments. The second experiments and learns, while the third learns but does not intentionally experiment.

Unemployment behaves in the same way under the three policy makers' policies because Lucas's model is the true data generating process. Samuelson and Solow's advice is costly because it makes inflation higher and more variable without generating an offsetting benefit in terms of better unemployment outcomes. Furthermore, the process of forgetting the 'wrong model', as reflected in the convergence of α_t back to 0, appears to be slower than occurred in the previous subsection where the Samuelson-Solow model prevailed. Although learning is initially quite rapid, with α falling from 0.75 to around 0.15 in the first year, substantial model uncertainty remains for more than a decade, during which the experimental policy retains its countercyclical character.

Although the experimenting policy maker generates higher inflation than is optimal when $\alpha = 0$, it typically chooses a lower inflation rate than does the non-experimenting, adaptive policy maker. The adaptive policy maker chooses an inflation rate that is approximately 40 basis points higher and that gap persists.

5.3 How long it takes to learn

Table 1 presents summary statistics from several related experiments. The variable that we call waiting time represents the number of quarters that are needed for α to return to within a 0.01 neighborhood of what it should be under the data generating process. For each experiment, we report the true model, the initial prior, the initial unemployment rate, the median waiting time with and without experimentation and the 10%-90% confidence sets in square brackets. When a '+' appears next to a number it means that the waiting time exceeded the length of the simulated path.

True Model	α_0	U_0	Waiting Time: Experimentation	Waiting Time: No Experimentation
SS	0.01	0	247 [149,486]	267 [156,500+]
SS	0.01	0.025	272 [154,498]	278 [171,500+]
Lucas	0.99	0	97 [37,242]	107 [40,244]
Lucas	0.99	0.025	85 [21,213]	87 [26,216]
SS	0.5	0	39 [20,80]	45 [23,92]
SS	0.5	0.025	21 [6,54]	25 [8,69]
Lucas	0.5	0	87 [32,160+]	93 [35,160+]
Lucas	0.5	0.025	65 [14,160+]	68 [18,160+]
SS	0.28	0	52 [27,160+]	65 [35,160+]
SS	0.28	0.025	35 [15,142]	46 [19,148]
Lucas	0.28	0	80 [26,160+]	90 [28,160+]
Lucas	0.28	0.025	71 [20,160+]	73 [22,160+]

Table 1: Waiting times for various data generating processes and initial (α_0, U_0) pairs.

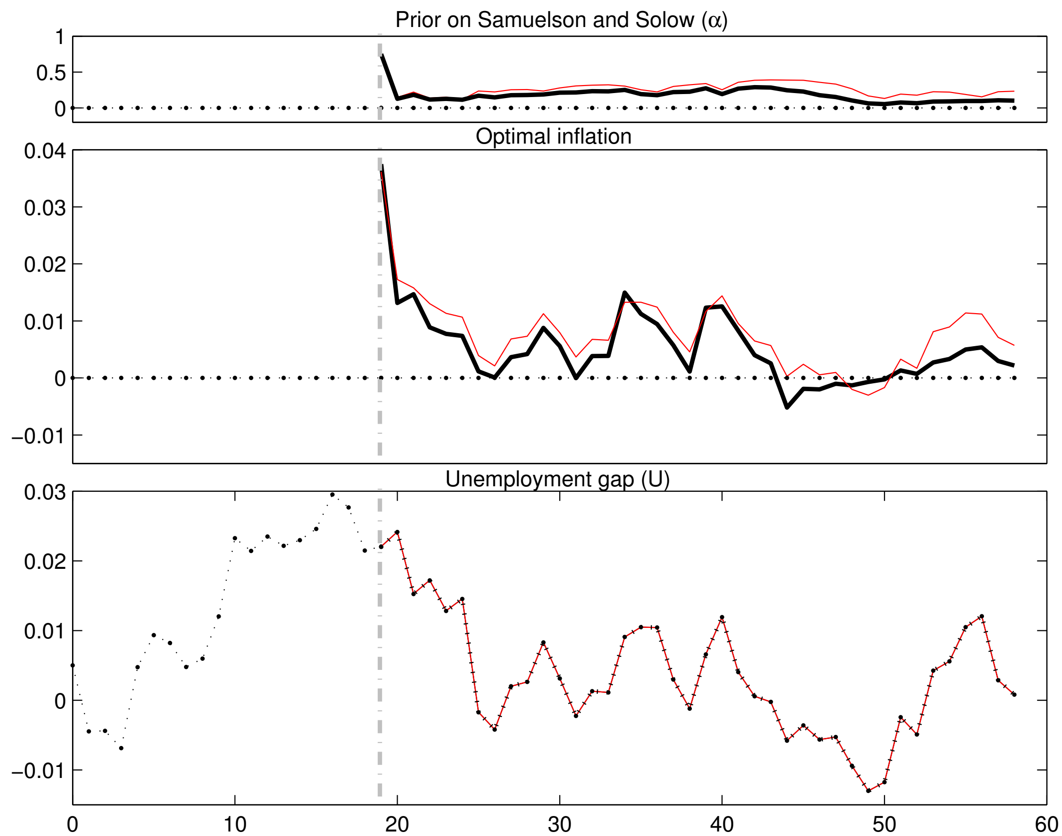


Figure 6: In the three panels: the dotted line represents the behavior of a policy maker that attaches probability one to the Samuelson and Solow model, the bold continuous line is the experimenting policy maker and the other line is the non-experimenting policy maker.

Several things can be learned from this table. First, as expected, the experimenting policy maker learns the truth faster than an anticipated utility policy maker. Deliberate experimentation reduces the median waiting time by approximately 2 years on average, or by 10 percent. For the most part, learning is also faster when unemployment is initially high. That reflects the cyclical opportunism of optimal policy: the best time to experiment with Keynesian stimulus is when unemployment is high. Paths that start with high unemployment therefore get a bigger dose of initial experimentation.

Second, if we start by attaching a probability of almost one to the wrong model, it is easier to learn when Lucas's model is true than when Samuelson and Solow's is true. In the latter case, with $\alpha_0 = 0.01$, the optimal policy involves little variation in inflation, and the policy maker must rely solely on natural experiments to learn.

In contrast, when $\alpha_0 = 0.99$, the policy rule calls for countercylical movements in inflation, so the policy maker can learn not only from shocks but also from policy-induced variation. That extra source of variation helps identify the true structure and speeds learning. Keynesian policy makers operating in a classical environment discover more quickly the error of their ways. Classical policy makers working in a Keynesian world update their beliefs more slowly.

Third, if we start with a 50-50 probability on the two models, the Samuelson and Solow model is easier to unveil. In this case, the optimal rule initially provides the same degree of policy-induced variation. But as beliefs are updated, that source of identifying information increases in the Samuelson-Solow economy (because α is increasing) and decreases in the Lucas economy. Thus, learning accelerates in the Samuelson-Solow economy and slows down in the Lucas economy.

Finally, for our baseline calibration the benefits of deliberate experimentation are maximized when $\alpha = 0.28$. We also simulated a number of paths starting at $\alpha_0 = 0.28$ to see whether that alters any of the results described above. The rows in the table corresponding to $\alpha_0 = 0.28$ show that the basic picture remains the same.

5.4 Alternative parameterizations

In this section we explore the possibility that the policy maker attaches a higher weight to inflation in the loss function. Figure 7 shows how the value of experimentation changes with λ.

As λ increases from 0.1 to 1, the gain that can be obtained under the experimenting policy rises significantly. It is also interesting to notice that the peak of the curve shifts toward the middle of the support of α. When $\lambda = 0.1$, the policy maker is more concerned about unemployment than about inflation, and a conspicuous amount of unintentional experimentation is created when the Samuelson and Solow model is likely to be the data generating process. However, as λ increases to 1, inflation gets a higher weight in the reward function, causing the degree of unintentional experimentation to drop. The policy authorities compensate for the reduced information flow from the w-policy by increasing the degree of intentional experimentation. Thus, the increase in the difference of the value functions associated with the two policies essentially reflects a substitution of deliberate for passive experimentation.

As λ continues to increase, the value of intentional experimentation eventually begins to fall. For example, when λ is as high as 16, the value of the policy rule described in equation (5) is lower than the two cases discussed earlier. For this parametrization, the value functions are negatively sloped in the direction of α,[17] and experimentation usually calls for a lower inflation rate. However, λ is so high

[17] Pictures are available upon request.

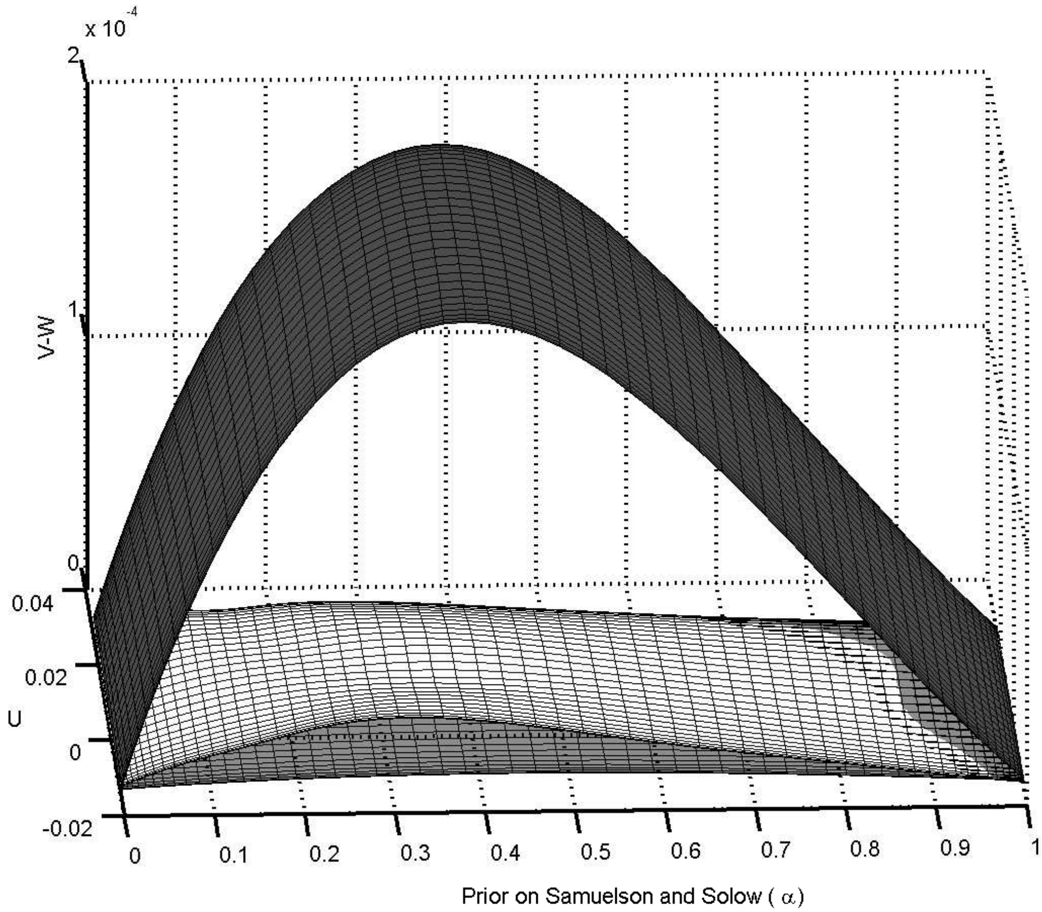


Figure 7: $V(U, \alpha) - W(U, \alpha)$ for different values of λ. The white curve is drawn for $\lambda = 0.1$ and the other two are for the cases of $\lambda = 1$ (top figure) and of $\lambda = 16$ (bottom figure).

that the programmed inflation resulting from the Bayesian and anticipated utility approaches are extremely low and almost undistinguishable. Hence the very small benefit from experimentation shown in the picture.

In Table 2 we report the relative behaviors of the mean and variance of unemployment and inflation for the Bayesian and anticipated utility policy makers. In the top panel, the Samuelson-Solow model is correct, but the policy authorities put some prior weight on the Lucas model. The bottom panel reverses the status of the two models. In each experiment, we generated 500 sample paths of length 40 (ten years) for the three values of λ discussed in this section.

Several things are worth noticing in this table. First, when the Lucas model is true, mean inflation is significantly higher in the absence of deliberate experimen-

		Inflation			Unemployment		
α_0	U_0	$\lambda = 0.1$	$\lambda = 1$	$\lambda = 16$	$\lambda = 0.1$	$\lambda = 1$	$\lambda = 16$
0.25	0.025	1.0098	1.0695	1.0575	0.9759	0.9720	0.9984
		0.8791	0.6059	0.9612	1.0230	1.0345	1.0007
0.25	0	1.0372	1.0880	1.0710	0.8393	0.9538	0.9972
		0.7998	0.6046	0.9654	1.0258	1.0308	1.0001
0.50	0.025	1.0067	0.9889	1.0380	0.9725	1.0099	0.9980
		0.9372	0.7236	0.9496	1.0139	1.0325	1.0008
0.50	0	1.0145	1.0510	1.0465	0.8780	0.9522	0.9969
		0.8953	0.7170	0.9553	1.0159	1.0290	1.0002
0.75	0.025	1.0005	1.0039	1.0165	1.0005	0.9971	0.9988
		0.9737	0.9047	0.9762	1.0060	1.0130	1.0004
0.75	0	1.0080	1.0238	1.0173	0.8900	0.9678	0.9985
		0.9636	0.9100	0.9744	1.0061	1.0086	1.0001

		Inflation			Unemployment		
α_0	U_0	$\lambda = 0.1$	$\lambda = 1$	$\lambda = 16$	$\lambda = 0.1$	$\lambda = 1$	$\lambda = 16$
0.25	0.025	1.3601	1.2369	1.0794	1.0000	1.0000	1.0000
		0.8969	0.7040	0.9833	1.0000	1.0000	1.0000
0.25	0	1.3753	1.4365	1.1069	1.0000	1.0000	1.0000
		0.9256	0.7601	1.0017	1.0000	1.0000	1.0000
0.50	0.025	1.3037	1.1160	1.0610	1.0000	1.0000	1.0000
		0.9164	0.7018	0.9637	1.0000	1.0000	1.0000
0.50	0	1.3084	1.2951	1.0915	1.0000	1.0000	1.0000
		0.9088	0.7350	0.9858	1.0000	1.0000	1.0000
0.75	0.025	1.2307	1.0677	1.0325	1.0000	1.0000	1.0000
		0.9362	0.8116	0.9713	1.0000	1.0000	1.0000
0.75	0	1.2064	1.1236	1.0506	1.0000	1.0000	1.0000
		0.9211	0.8254	0.9703	1.0000	1.0000	1.0000

Table 2: Top panel: data are generated according to the Samuelson and Solow model. Bottom panel: data are generated according to the Lucas model. For each (α_0, U_0) pair, the first row reports the average inflation (unemployment) without experimentation relative to the average inflation (unemployment) with experimentation for different values of λ. The second row reports the standard deviation of inflation (unemployment) without experimentation relative to the standard deviation of inflation (unemployment) with experimentation for different values of λ. In both panels: 500 samples of ten years (40 quarters) were generated.

tation, in some cases by as much as 40 percent. Because the authorities put some weight on the Samuelson-Solow model, they adopt a positive target for average inflation, which shrinks to zero as α converges to zero. Intentional experimentation accelerates learning, so the Bayesian authorities move more rapidly to zero inflation than the anticipated-utility policy maker. Hence the w-policy results in higher average inflation during the 10-year transition. Also notice that mean unemployment is identical for the two policy rules. This follows from the policy-ineffectiveness proposition, which implies that deliberate experimentation does not matter for unemployment outcomes.

Second, inflation is always more volatile when the policy maker is following the Bayesian rule[18]. When $\lambda = 16$ the differences are mild, which is consistent with the small benefits of experimentations shown in Figure 7. The gap between the two volatilities is by far greatest when $\lambda = 1$. This can partly justify why the highest benefits from experimentation are coming from this parametrization. The Bayesian policy maker is using the programmed rate of inflation more dramatically in the first few periods in order to gather better information about the underlying model. This calls for a higher volatility of the inflation rate. The fact that for $\lambda = 1$ the variance gap is so big is suggestive of an intensive use of inflation for experimenting reasons.

When the Samuelson-Solow model is true, the additional variation in inflation under the v-policy pays off in terms of lower volatility of unemployment. But notice that the relative decline in unemployment volatility is substantially smaller than the increase in inflation volatility. When the Lucas model is true, the extra inflation volatility has no influence on unemployment volatility, again because of the policy-ineffectiveness proposition. Thus, in the Lucas model, the cost of deliberate experimentation is greater inflation volatility along the transition, and the benefit is more rapid convergence to zero inflation.

6 Concluding remarks

The value functions and decision rules in figures 1 and 2 reveal that in our example, an anticipated utility model does a good job of approximating outcomes of a Bayesian model in which the monetary policy maker exploits the opportunity to experiment. While the passive learner in the anticipated utility does not design policies in order to experiment, the outcomes of his policies induce enough variation in the data that he is able to discriminate between the two models almost as fast as the Bayesian agent. This outcome is related to features in the environment identified by El-Gamal and Rangarajan (1993), who show how the presence of sufficient 'natural experiments' promotes learning.

[18] There is only one exception that we attribute to numerical error.

Another interesting outcome is captured in the concavity of the decision rules in figure 2. This shape conveys that the decisions of a Samuelson-Solow style Keynesian are more robust to small doubts, i.e., perturbations of α away from 1, than are the decisions of Lucas-style classical economist to small perturbations of α away from 0. That lack of robustness of the classical recommendations to small doubts plays an important role in Cogley and Sargent's (2005) account of U.S. inflation policy during the 1970s.

Our calculations also reveal how long it takes to disabuse a doubtful monetary authority of the wrong model. A key factor that influences the speed of convergence is the probability weight α, for that affects the contribution of policy-induced variation to learning. When α is close to zero, policy keeps inflation close to zero, and the policy maker must rely heavily on natural experiments for learning. When α is close to one, the policy maker still learns from natural experiments but also learns from policy-induced, countercyclical variations in inflation. That extra source of variation speeds learning.

We have given the policy maker only two models, each of which he knows for sure. The models have very different operating characteristics. While their differences are important, they are not subtle, so this eases the task of generating or waiting for data to discriminate between them. In effect, we have assumed that the monetary authority's doubts are limited to ignorance of the 'correct' value of one hyperparameter, α. If in practice one thinks that the monetary authority's doubts are broader and vaguer, we have substantially understated the difficulty of the decision and learning problem that it faces. The robustness calculations in Cogley, Colacito, Hansen, and Sargent (2005) provide ways to address those concerns.

7 Appendixes

A Transition Equation For α_t

Let $\alpha_{i0} \equiv p(M_i)$ be the prior probability on model i, and let $p_i(s_i^t|\theta_i)$ represent its likelihood function. Here we abstract from parameter uncertainty by adopting the shortcut that the parameters θ_i are known. By Bayes's theorem, the posterior probability on model i is

$$\alpha_{it} \equiv p(M_i|s_i^t, \theta_i) = \frac{p_i(s_i^t|\theta_i)p(M_i)}{\displaystyle\int p_i(s_i^t|\theta_i)p(M_i)dM_i}. \tag{19}$$

The numerator is an unnormalized model weight, which we label w_{it}, and the denominator is a normalizing constant that ensures that model probabilities sum to

1. With a finite collection of models, the denominator is just the sum of the unnormalized model weights, $\sum_i w_{it}$.

We start with a simple recursion for the unnormalized weights w_{it}. After taking logs and first-differencing, we find

$$\log w_{it} - \log w_{it-1} = \log p_i(s_i^t|\theta_i) - \log p_i(s_i^{t-1}|\theta_i). \tag{20}$$

Note that the prior model weight drops out of the recursion; α_{i0} initializes the sequence but the likelihood is all that matters for updates. Also notice that α-updates depend only on the value of the likelihood at the given θ_i. Usually the model probability updates would depend on a marginalized likelihood, but this drops out because we assume that θ_i is known. We need only to evaluate the likelihood, not marginalize across unknown parameters.

To simplify further, use the prediction error decomposition of the likelihood to write

$$\log p_i(s_i^t|\theta_i) = \sum\nolimits_{s=1}^{t} \log p_i(s_{is}|s_i^{s-1}, \theta_i). \tag{21}$$

Subtracting the log-likelihood through $t-1$ from that through t, we get

$$\log w_{it} = \log w_{it-1} + \log p_i(s_{it}|s_i^{t-1}, \theta_i). \tag{22}$$

The date t update depends on the value of the conditional log-likelihood. An observation that is likely given the model raises the unnormalized model weight, and a puzzling observation (for that model) lowers it. Notice that $\log w_{it}$ is a martingale if the model residuals are serially uncorrelated.

Now let's specialize to a two-model model. Let α_t be the normalized probability weight for model 1,

$$\alpha_t = \frac{w_{1t}}{w_{1t} + w_{2t}}. \tag{23}$$

The probability weight on model 2 is $1 - \alpha_t$.

The normalizing constant is a nuisance, so we eliminate it by taking the ratio,

$$R_t \equiv \frac{\alpha_t}{1-\alpha_t} = \frac{w_{1t}}{w_{2t}}. \tag{24}$$

The transition equation for $\log R_t$ follows from the transition equations for $\log w_{it}$,

$$\log R_t = \log R_{t-1} + \log \frac{p_1(s_{1t}|s_1^{t-1}, \theta_1)}{p_2(s_{2t}|s_2^{t-1}, \theta_2)}. \tag{25}$$

Thus, the updating rule for the log odds ratio depends only on the log-likelihood ratio for the two competing models. If we write this in terms of α_t, we find

$$\frac{\alpha_t}{1-\alpha_t} = \frac{\alpha_{t-1}}{1-\alpha_{t-1}} \frac{p_1(s_{1t}|s_1^{t-1}, \theta_1)}{p_2(s_{2t}|s_2^{t-1}, \theta_2)}, \tag{26}$$

or

$$\alpha_t = \frac{\frac{\alpha_{t-1}}{1-\alpha_{t-1}} \frac{p_1(s_{1t}|s_1^{t-1},\theta_1)}{p_2(s_{2t}|s_2^{t-1},\theta_2)}}{1 + \frac{\alpha_{t-1}}{1-\alpha_{t-1}} \frac{p_1(s_{1t}|s_1^{t-1},\theta_1)}{p_2(s_{2t}|s_2^{t-1},\theta_2)}}, \tag{27}$$

$$= \frac{\alpha_{t-1} p_1(s_{1t}|s_1^{t-1},\theta_1)}{\alpha_{t-1} p_1(s_{1t}|s_1^{t-1},\theta_1) + (1-\alpha_{t-1}) p_2(s_{2t}|s_2^{t-1},\theta_2)}.$$

If the two models involve the same data, we can equate $s_{1t} = s_{2t}$. In that case,

$$\alpha_t = \frac{\alpha_{t-1} p_1(s_t|s^{t-1},\theta_1)}{\alpha_{t-1} p_1(s_t|s^{t-1},\theta_1) + (1-\alpha_{t-1}) p_2(s_t|s^{t-1},\theta_2)}. \tag{28}$$

The right side of this equation spells out the function $B(\alpha_{t-1}, s_t)$.

A.1 Martingale Property of α_t

The updating formula makes α_t a martingale from the point of view of the Bayesian agent (this is an example of Doob's martingale result for Bayesian updating). To see why, take the expectation of α_t with respect to the posterior at α_{t-1},

$$E_{t-1} B(\alpha_{t-1}, s_t) = \int B(\alpha_{t-1}, s_t) f_{t-1}(s_t|s^{t-1}) ds_t. \tag{29}$$

Because model parameters are assumed to be known, there is a single source of uncertainty about next period's α_t, viz. what next period's s_t will be. Therefore the expectation is taken with respect to the agent's posterior predictive density for s_t, which we denote $f_{t-1}(s_t|s^{t-1})$. This density is a probability weighted average of the predictive densities for the two models,

$$f_{t-1}(s_t|s^{t-1}) = \alpha_{t-1} p_1(s_t|s^{t-1},\theta_1) + (1-\alpha_{t-1}) p_2(s_t|s^{t-1},\theta_2). \tag{30}$$

Thus, the conditional expectation for α_t is

$$\begin{aligned} E_{t-1}\alpha_t &= \int B(\alpha_{t-1}, s_t) \left[\alpha_{t-1} p_1(s_t|s^{t-1},\theta_1) + (1-\alpha_{t-1}) p_2(s_t|s^{t-1},\theta_2)\right] ds_t, \\ &= \int \alpha_{t-1} p_1(s_t|s^{t-1},\theta_1) ds_t, \\ &= \alpha_{t-1} \int p_1(s_t|s^{t-1},\theta_1) ds_t = \alpha_{t-1}. \end{aligned} \tag{31}$$

A.2 A Different State Space

To get a tractable Bellman equation, it is convenient to rewrite the problem so that the state transition equation is linear. Define:

$$\log R_t = \log \frac{\alpha_t}{1-\alpha_t}$$

then

$$\log R_{t+1} = \log R_t + \log \frac{f_1(s_{t+1}|s_t)}{f_2(s_{t+1}|s_t)}$$

α_t can be obtained back through the following expression

$$\alpha_t = \frac{1}{1+(\exp \log R_t)^{-1}} \tag{32}$$

In the Bellman equation, we take expectations of functions that involve the log likelihood ratio. These expectations involve the distribution of $\varepsilon_{2,t+1}$ under model 1 and viceversa. We can represent those distributions by exploiting the assumption that s_t is the same across models. This assumption means that the model innovations are related. After subtracting the transition equation for model 2 from that for model 1, we find:

$$C_2\epsilon_{2,t+1} = (A_1 - A_2)s_t + (B_1 - B_2)v_t + C_1\epsilon_{1,t+1} \tag{33}$$

$$C_1\epsilon_{1,t+1} = -(A_1 - A_2)s_t - (B_1 - B_2)v_t + C_2\epsilon_{2,t+1} \tag{34}$$

Define $\Omega_1 = C_1C_1', \Omega_2 = C_2C_2'$. We use (33) and (34) to write the recursion for $\log R_{t+1}$ under models 1 and 2. When model 1 is true, we have

$$\begin{aligned}\log R_{t+1} = {} & \log R_t - \frac{1}{2}\log|\Omega_1| + \frac{1}{2}\log|\Omega_2| - \frac{1}{2}\left(C_1\epsilon_{1,t+1}\right)'\Omega_1^{-1}\left(C_1\epsilon_{1,t+1}\right) \\ & +\frac{1}{2}\left[(A_1 - A_2)s_t + (B_1 - B_2)v_t + C_1\epsilon_{1,t+1}\right]' \\ & \times\Omega_2^{-1}[(A_1 - A_2)s_t + (B_1 - B_2)v_t + C_1\epsilon_{1,t+1}]\end{aligned} \tag{35}$$

When model 2 is true, we have

$$\begin{aligned}\log R_{t+1} = {} & \log R_t - \frac{1}{2}\log|\Omega_1| + \frac{1}{2}\log|\Omega_2| + \frac{1}{2}\left(C_2\epsilon_{2,t+1}\right)'\Omega_2^{-1}\left(C_2\epsilon_{2,t+1}\right) \\ & -\frac{1}{2}[(A_2 - A_1)s_t + (B_2 - B_1)v_t + C_2\epsilon_{2,t+1}]' \\ & \times\Omega_1^{-1}[(A_2 - A_1)s_t + (B_2 - B_1)v_t + C_2\epsilon_{2,t+1}]\end{aligned} \tag{36}$$

It is convenient to use α_t rather than $\log R_t$ as a state variable. So we want to transform (35) and (36) to get laws of motion for α_t under the two models. For the purpose of doing this, define

$$\begin{aligned} g(\epsilon_{1,t+1}; s_t, \alpha_t) &= \log R_t - \frac{1}{2}\log|\Omega_1| + \frac{1}{2}\log|\Omega_2| - \frac{1}{2}(C_1\epsilon_{1,t+1})'\,\Omega_1^{-1}(C_1\epsilon_{1,t+1}) \\ &\quad + \frac{1}{2}[(A_1 - A_2)s_t + (B_1 - B_2)v_t + C_1\epsilon_{1,t+1}]' \\ &\quad \times\Omega_2^{-1}[(A_1 - A_2)s_t + (B_1 - B_2)v_t + C_1\epsilon_{1,t+1}] \end{aligned} \tag{37}$$

and

$$\begin{aligned} h(\epsilon_{2,t+1}; s_t, \alpha_t) &= \log R_t - \frac{1}{2}\log|\Omega_1| + \frac{1}{2}\log|\Omega_2| + \frac{1}{2}(C_2\epsilon_{2,t+1})'\,\Omega_2^{-1}(C_2\epsilon_{2,t+1}) \\ &\quad - \frac{1}{2}[(A_2 - A_1)s_t + (B_2 - B_1)v_t + C_2\epsilon_{2,t+1}]' \\ &\quad \times\Omega_1^{-1}[(A_2 - A_1)s_t + (B_2 - B_1)v_t + C_2\epsilon_{2,t+1}] \end{aligned} \tag{38}$$

Using (32), we get a law of motion for α_{t+1} under the two models. Then our Bellman equation can be expressed

$$\begin{aligned} V(s_t, \alpha_t) &= \max_{v_t}\left\{ r(s_t, v_t) + \beta\alpha_t \int V\left(A_1 s_t + B_1 v_t + C_1\epsilon_{1,t+1}, \frac{e^{g(\epsilon_{1,t+1})}}{1 + e^{g(\epsilon_{1,t+1})}}\right) dF(\epsilon_{1,t+1}) \right. \\ &\quad \left. + \beta(1 - \alpha_t)\int V\left(A_2 s_t + B_2 v_t + C_2\epsilon_{2,t+1}, \frac{e^{h(\epsilon_{2,t+1})}}{1 + e^{h(\epsilon_{2,t+1})}}\right) dF(\epsilon_{2,t+1}) \right\}. \end{aligned} \tag{39}$$

B Approximating the Bellman Equation

Discretize the support of α and s into I_α and I_s points respectively, to get $I = I_\alpha \cdot I_s$ nodes $(\alpha, s)_i$, $\forall i = 1, ..., I$. In what follows, we will refer to α_i and s_i as the first and the second entry of $(\alpha, s)_i$ respectively. Specify J known linearly independent basis functions $\phi_j((\alpha, s)_i)$, $j \in \{1, ..., J\}$. In our solution, we employ a third order complete polynomial, implying that $J = 10$. The goal is to find basis coefficients c_j, $j = 1, ..., J$ that best approximate the value function

$$V_i = V((\alpha, s)_i) \approx \sum_{j=1}^{J} c_j \phi_j((\alpha, s)_i) = \sum_{j=1}^{J} c_j \phi_{j,i} \tag{40}$$

$\forall i = 1, ..., I$ or, in the equivalent matrix notation:

$$V \approx \Phi c$$

where V is the $I \times 1$ vector of approximated value functions at each node, Φ is the $I \times J$ collocation matrix and $c = [c_1, ..., c_J]'$ is the vector of approximation coefficients. We also discretize the support of the two shocks in K_1 and K_2 points and denote w_k the approximated probability mass associated to each of the resulting $K = K_1 \times K_2$ nodes. Using (40) in the Bellman equation we get for each node $i \in \{1, ..., I\}$:

$$\begin{aligned} V_i &= \max_{v_i} \Bigg\{ r_i(v_i) + \beta \alpha_i \sum_{k=1}^{K} \sum_{j=1}^{J} w_k c_j \phi_j \left(s'_{1,i,k}(v_i), \frac{\exp[g_{k,i}(v_i)]}{1+\exp[g_{k,i}(v_i)]} \right) \\ &\quad + \beta(1-\alpha_i) \sum_{k=1}^{K} \sum_{j=1}^{J} w_k c_j \phi_j \left(s'_{2,i,k}(v_i), \frac{\exp[h_{k,i}(v_i)]}{1+\exp[h_{k,i}(v_i)]} \right) \Bigg\} \end{aligned} \tag{41}$$

where

$$\begin{aligned} r_i(v_i) &= r(s_i, v_i) \\ s'_{1,i,k}(v_i) &= A_1 s_i + B_1 v_i + C_1 \varepsilon_k \\ x'_{2,i,k}(v_i) &= A_2 s_i + B_2 v_i + C_2 \varepsilon_k \end{aligned}$$

and $g_{k,i}(v_i)$ and $h_{k,i}(v_i)$ defined as in (37) and (38) respectively:

$$\begin{aligned} g_{k,i}(v_i) &= g(\varepsilon_k; s_i, \alpha_i, v_i) \\ h_{k,i}(v_i) &= h(\varepsilon_k; s_i, \alpha_i, v_i) \end{aligned}$$

We can now use the following algorithm to solve the Bellman equation recursively:

1. guess an initial vector of basis coefficients c^1
2. for each node $(s, \alpha)_i$ compute the right hand side of equation (41) using c^1 and call $v(c^1)$ the outcome
3. solve for $c^2 = (\Phi'\Phi)^{-1} \Phi' v(c^1)$
4. replace c^1 with c^2 and iterate until convergence.

C The 'Don't Experiment' Model

This appendix describes how to solve Bellman equation (13) by mapping the problem into what Cogley and Sargent (2004) called a 'Bayesian linear regulator'. Stack the two state space models from (1) as

$$\begin{bmatrix} s_{1,t+1} \\ s_{2,t+1} \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} s_{1,t} \\ s_{2,t} \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} v_t + \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix} \begin{bmatrix} \epsilon_{1,t+1} \\ \epsilon_{2,t+1} \end{bmatrix} \tag{42}$$

or

$$s_{t+1} = As_t + Bv_t + C\epsilon_{t+1} \tag{43}$$

Let $\alpha \in (0,1)$ be a fixed probability that the decision maker attaches to model 1. Express the time t loss as $r(s_t, v_t) = -.5(s_t' R s_t + v_t' Q v_t)$. The decision maker seeks to maximize

$$L = -.5E \sum_{t=0}^{\infty} \beta^t \Big\{ \alpha s_{1t}' R s_{1t} + (1-\alpha) s_{2t}' R s_{2t} + v_t' Q v_t \Big\} \tag{44}$$

or

$$L = -.5E \sum_{t=0}^{\infty} \beta^t \left\{ s_t' \begin{bmatrix} \alpha R & 0 \\ 0 & (1-\alpha)R \end{bmatrix} s_t + v_t' Q v_t \right\} \tag{45}$$

Cogley and Sargent (2004) note that the problem of choosing a decision rule to maximize (45) with respect to (43) is an optimal linear regulator problem. The optimal decision rule is

$$v_t = -Fs_t = -F_1 s_{1t} - F_2 s_{2t}. \tag{46}$$

D Description of the Empirical Specification

Here we briefly describe how the two policy models are estimated. Inflation is measured by the log difference of the chain-weighted GDP deflator, and unemployment is the civilian unemployment rate. Both series are seasonally adjusted and are sampled over the period 1948:1 to 1963:1. We stop the estimation there to represent the kind of model uncertainty that Federal Reserve officials would have faced in the years leading up to the Great Inflation.

Both Phillips curve specifications involve the gap between the unemployment rate and a time-varying natural rate of unemployment. In order to keep the size of the state space to a minimum, we approximate the natural rate of unemployment U_t^* by exponentially smoothing the actual unemployment rate UR_t,

$$U_t^* = U_{t-1}^* + \mu(UR_t - U_{t-1}^*), \tag{47}$$

with a constant gain parameter $\mu = 0.075$. That makes the unemployment gap a geometrically distributed lag of past changes in unemployment,

$$U_t \equiv UR_t - U_t^* = \frac{(1-\mu)(1-L)}{1-(1-\mu)L} UR_t. \tag{48}$$

This procedure approximates a one-sided high-pass filter that transforms unemployment into the unemployment gap. The decomposition is shown in the following figure.

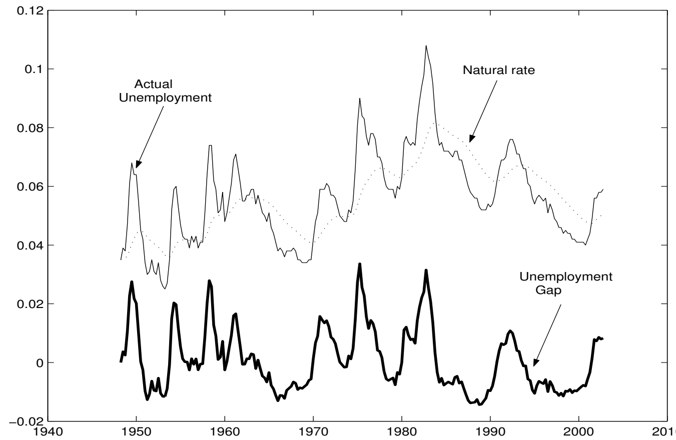


Figure 1: Decomposing Unemployment: The Natural Rate and the Gap

The blue line records actual unemployment, the red line depicts our proxy for the natural rate, and the green line is the unemployment gap, which is the variable that appears in the Phillips curves. This decomposition assigns most of the short-term variation in unemployment to the unemployment gap, and attributes long-term movements in the level to shifts in the natural rate. For the years over which we estimate the models, the natural rate increases only slightly, and most of the variation in UR_t is in the gap measure U_t.

For model 1, this is all we need for estimation. We simply project the current unemployment gap onto a constant, current inflation, and one lag of gap, and estimate parameters by OLS. For the period 1948:3-1963:1, the least-squares point estimates and standard errors are as follows.

Table 1: Estimates of Model 1, 1948:3-1963:1

	Intercept	U_{t-1}	π_t
$\hat{\beta}$	0.0023	0.7971	-0.2761
$\sigma_{\hat{\beta}}$	0.0010	0.0699	0.1189

In model 2, unemployment depends not on inflation but on unexpected inflation, $\pi_t - v_{t-1}$, so to estimate that model we also need a measure of expected inflation v_{t-1}. We construct that in the simplest way possible, by projecting current inflation on a constant along with one lag of inflation and unemployment. The fitted value from that regression is our measure of v_{t-1}, and the residual is our measure of unexpected inflation, $\pi_t - v_{t-1}$. Then we substitute that variable into the Phillips curve and estimate its parameters by least squares. The estimates and standard errors for model 2 are shown in the next table.

Table 2: Estimates of Model 2, 1948:3-1963:1

	Intercept	U_{t-1}	$\pi_t - v_{t-1}$
$\hat{\beta}$	0.0007	0.8468	-0.2489
$\sigma_{\hat{\beta}}$	0.0008	0.0674	0.1298

We use the point estimates in these tables to calibrate the two policy models. Our central bank takes the point estimates as if they were known with certainty and formulates policy by averaging across the models. Thus, it takes account of model uncertainty, but suppresses parameter uncertainty.

References

Beck, G. and V. Wieland, 2002, "Learning and Control in a Changing Environment," *Journal of Economic Dynamics and Control* 26, 1359-1377.

Blinder, A.S. 1998, *Central Banking in Theory and Practice*, Cambridge: MIT Press.

Cogley, T. and T.J. Sargent, 2005, "The Conquest of U.S. Inflation: Learning and Robustness to Model Uncertainty," *Review of Economic Dynamics*, April, in press.

Cogley, T., R. Colacito, L.P. Hansen, and T.J. Sargent, 2005, "Robustness and U.S. Monetary Policy Experimentation," Unpublished manuscript.

El-Gamal, Mahmoud A., and Rangarajan K. Sundaran, 1993, "Bayesian Economists . . . Bayesian Agents: An alternative approach to optimal learning," *Journal of Economic Dynamics and Control*, 17, 355-383.

Epstein, Larry and Martin Schneider, 2003, "Recursive Multiple Priors," *Journal of Economic-Theory*, Volume 113, Number 1, November, pp. 1-31.

Hansen, Lars Peter and Thomas J. Sargent, 2005a, "Recursive Robust Decisions with Hidden States," Unpublished manuscript, University of Chicago and New York University.

Hansen, Lars Peter and Thomas J. Sargent, 2005b, "Discounting, Commitment, and Recursive Formulations of Robust Estimation and Control," Unpublished manuscript, University of Chicago and New York University.

Kasa, Kenneth, 1999, "Will the Fed Ever Learn?" *Journal of Macroeconomics*. 21(2): 279-92.

Kreps, David, 1998, "Anticipated Utility and Dynamic Choice," 1997 Schwartz Lecture, in *Frontiers of Research in Economic Theory*, Edited by Donald P. Jacobs, Ehud Kalai, and Morton Kamien, Cambridge University Press, Cambridge, England.

Kydland, F.E. and E.C. Prescott, 1977, "Rules Rather Than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy* 85, 473-91.

Lucas, R.E., Jr., 1972, "Expectations and the Neutrality of Money," *Journal of Economic Theory* 4, 103-124.

Lucas, R.E., Jr., 1972, "Some International Evidence on Output-Inflation Trade-Offs," *American Economic Review* 63, 326-334.

Lucas, R.E., Jr., 1981, "Methods and Problems in Business Cycle Theory," in R.E. Lucas, Jr., (ed.) *Studies in Business-Cycle Theory*, Cambridge, Mass., MIT Press.

Orphanides, Athanasios, and David Wilcox, 2002, "The Opportunistic Approach to Disinflation", *International Finance*, 5(1), 47-71, Spring.

Samuelson, P.A., and R.M. Solow, 1960, "Analytical Aspects of Anti-Inflation Policy," *American Economic Review*, 50, 177–184.

Sargent, T.J., 1973, "Rational Expectations, the Real Rate of Interest, and the Natural Rate of Unemployment," *Brookings Papers on Economic Activity*, 429-472.

Stokey, N.L., 1989, "Reputation and Time Consistency," *American Economic Review Papers and Proceedings* 79, 134-139.

Wieland, V., 2000a, "Monetary Policy, Parameter Uncertainty, and Optimal Learning," *Journal of Monetary Economics* 46, 199-228.

Wieland, V., 2000b, "Learning by Doing and the Value of Optimal Experimentation," *Journal of Economic Dynamics and Control* 24, 501-534.