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

Timothy Cogley†
Riccardo Colacito‡
Thomas J. Sargent§
February 15, 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 into a posterior at each date and gives the policy maker an incentive to experiment. For a model 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. We also compute decision rules that are robust to departures with respect to two classes of misspecifications of the decision maker’s model.

1 Introduction

We quantify the importance of deliberate experimentation when two models that fit the historical data equally well have sharply different operating characteristics

---

*We thank Klaus Adams, Patrick Bajari, 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
that are vital to a policy decision. As our laboratory, we analyze a case in which two competing models of inflation-unemployment dynamics differ with respect to whether they imply an exploitable Phillips curve.

In the late 1960’s, a debate raged between advocates of the natural unemployment hypothesis and those who thought that there is an exploitable unemployment-inflation trade-off. To capture this dispute, we imagine that a monetary policy authority has the following two models of inflation-unemployment dynamics:

- **Model 1 (Samuelson-Solow):**
  \[
  U_t = .0023 + .7971U_{t-1} - .2761\pi_t + .0054\eta_{1,t} \\
  \pi_t = v_{t-1} + .0055\eta_{2t}
  \]

- **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 \(U_t^*\), \(\pi_t\) 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 \(i = 1, 2, 3, 4\), \(\eta_{it}\) is an i.i.d. Gaussian sequence with mean zero and variance 1. The monetary authority has a Kydland-Prescott (1977) type of loss function \(E \sum_{t=0}^{\infty} \beta^t r_t\), where \(r_t = -.5(U_t^2 + \lambda v_t^2)\). The monetary authority sets \(v_t\) as a function of time \(t\) information. The monetary authority knows the parameters of each model for sure and attaches probability \(\alpha_0\) to model 1 and probability \(1 - \alpha_0\) to model 2.

$^{1}$We use these specifications mainly as a device to get good fitting models while keeping the dimension of the state of our model 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 lower than the natural rate) and the public (which would choose to set unemployment to the natural rate) that is essential to induce 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).

$^{4}$We assume that model parameters are known because we want to reduce to a minimum 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)
Although they fit the U.S. data from 1948:3-1963:I almost equally well, these two models call for very different policies toward inflation under our loss function. 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: systematic variations in 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 is willing to set $v_t \neq 0$ partly in order to exploit a probable inflation-unemployment tradeoff and partly in order to refine $\alpha$. After calibrating the two models to U.S. data before 1963, this paper imputes the same objective to the monetary authority that Kydland and Prescott (1977) used, then solves the Bellman equation.

We use the optimal decision rule to study the following questions:

1. Suppose that the Samuelson-Solow model actually governs 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. Having been persuaded by an advocate of the natural rate hypothesis, at date $T$ the government suddenly lowers $\alpha$ to a number $\alpha \in (0, 1)$ even though, unbeknownst to the government, the Samuelson-Solow model actually prevails. 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. Lucas’s model is subversive in leading 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 actually governs the data and that before date $T$ the government had assigned probability $1 - \alpha = 1$ to the Lucas model and used the corresponding optimal policy for a long time, so that the economy is 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 actually prevails. 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

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 through some different strategy.
inflation and no lower inflation 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 includes $\alpha$ as a state variable and Bayes’ law as a transition equation with another Bellman equation that suppresses $\alpha$ as a state variable and ignores 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.

1.1 Organization

Section 2 formulates Bellman equations, one for a decision maker who consciously 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. 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 studies 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 approximating models. Section 7 adds some concluding remarks. Four appendixes contain technical details about how we solved the Bellman equations and calibrated the two models.

2 Two formulations of the policy problem under model uncertainty

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

2.1 The models

The policy maker has two models

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

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

In equations (42) and (46) 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 \n
\begin{align*}
E_0 \sum_{t=0}^{\infty} \beta^t r(s_t, v_t), \quad \beta \in (0, 1), \tag{3}
\end{align*}

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 \( \alpha_t \) is a component of the time \( t \) state vector \((s_t, \alpha_t)\). 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 \((s_t, \alpha_t)\). The Bellman equation is

\begin{align*}
V(s_t, \alpha_t) &= \max_{v_t} r(s_t, v_t) \tag{4} \\
+ &\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{align*}

The optimal decision rule can be represented recursively as

\begin{align*}
v_t &= v(s_t, \alpha_t) \tag{5} \\
\alpha_{t+1} &= \alpha(s_t, \alpha_t). \tag{6}
\end{align*}

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

\begin{align*}
v_t &= \sigma_t(s^t, \alpha_0), \tag{7}
\end{align*}

where \( s^t = (s_t, s_{t-1}, \ldots, s_0) \). The presence of \( B(\alpha_t, A_i s_t + B_i v_t + C_i \epsilon_{1,t+1}), i = 1, 2, \) 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 $\Omega_t = C_t C'_t$, $R_t = \frac{\alpha_t}{1 - \alpha_t}$, and define

$$g(\epsilon_{1,t+1}; s_t, \alpha_t) = \log R_t - \frac{1}{2} \log |\Omega_1| + \frac{1}{2} \log |\Omega_2| - \frac{1}{2} (C_1 \epsilon_{1,t+1})' \Omega^{-1}_1 (C_1 \epsilon_{1,t+1})$$

$$+ \frac{1}{2} [(A_1 - A_2)s_t + (B_1 - B_2)v_t + C_1 \epsilon_{1,t+1}]'$$

$$\times \Omega^{-1}_2 [(A_1 - A_2)s_t + (B_1 - B_2)v_t + C_1 \epsilon_{1,t+1}]$$

and

$$h(\epsilon_{2,t+1}; s_t, \alpha_t) = \log R_t - \frac{1}{2} \log |\Omega_1| + \frac{1}{2} \log |\Omega_2| + \frac{1}{2} (C_2 \epsilon_{2,t+1})' \Omega^{-1}_1 (C_2 \epsilon_{2,t+1})$$

$$- \frac{1}{2} [(A_2 - A_1)s_t + (B_2 - B_1)v_t + C_2 \epsilon_{2,t+1}]'$$

$$\times \Omega^{-1}_1 [(A_2 - A_1)s_t + (B_2 - B_1)v_t + C_2 \epsilon_{2,t+1}].$$

Using (42) in Appendix A, we obtain a law of motion for $\alpha_{t+1}$ under the two models. Then Bellman equation (4) becomes

$$V(s_t, \alpha_t) = \max_{v_t} \left\{ r(s_t, v_t) + \beta \alpha_t \int V \left( A_1 s_t + B_1 v_t + C_1 \epsilon_{1,t+1}, \frac{e^{g(\epsilon_{1,t+1})}}{1 + e^{g(\epsilon_{1,t+1})}} \right) dF(\epsilon_{1,t+1}) \right. \right.$$  

$$+ \beta (1 - \alpha_t) \int V \left( A_2 s_t + B_2 v_t + C_2 \epsilon_{2,t+1}, \frac{e^{h(\epsilon_{2,t+1})}}{1 + e^{h(\epsilon_{2,t+1})}} \right) dF(\epsilon_{2,t+1}) \left. \right\}. \quad (10)$$

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. We shall address such concerns explicitly in section 6.

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 (2004) 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. Suboptimal experiments attain lower values. Many such suboptimal experiments would actually 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 $\alpha_t$ dictated by Bayes’ law with the “don’t experiment on purpose” specification

$$\alpha_t = \alpha \quad \forall t \geq 0.$$  

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 $\alpha$-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-view chooses a policy

$$v_t = w(s_t; \alpha).$$  \hfill (12)

that attains maximizes the right side of the following Bellman equation:\(^5\)

$$\tilde{W}(s_t; \alpha) = \max_{v_t} \{ r(s_t, v_t)$$  

$$+ \beta \alpha \int \tilde{W}(A_1 s_t + B_1 v_t + C_1 \epsilon_{t+1}; \alpha) dF(\epsilon_{1,t+1})$$

$$+ \beta (1 - \alpha) \int \tilde{W}(A_2 s_t + B_2 v_t + C_2 \epsilon_{t+1}; \alpha) dF(\epsilon_{2,t+1})\}.$$  

But in the spirit of the adaptive control literature, suppose that the policy maker does indeed revise \(\alpha_t\) 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

$$v_t = w(s_t; \alpha_t)$$  \hfill (14)

$$\alpha_{t+1} = B(s_t; \alpha_t).$$  \hfill (15)

These decisions would emerge from a ‘don’t experiment but do learn’ prescription.\(^6\) Equations (14), (15) can be solved by repeated substitution to yield the policy maker’s strategy in the form of a sequence of functions

$$v_t = \tilde{\sigma}_t(s^t; \alpha_0).$$  \hfill (16)

Thus, although in formulating his policy, the policy maker ignores the motion in \(\alpha_t\) impelled by Bayes’ law (2), Bayes’ law does affect the value that he can actually expect to attain under the policy (12), which evidently satisfies the Bellman equation

$$W(s_t; \alpha_t) = r(s_t, w(s_t, \alpha_t))$$

$$+ \beta \alpha_t \int W(A_1 s_t + B_1 w(s_t, \alpha_t) + C_1 \epsilon_{1,t+1}, B(\alpha_t, A_1 s_t + B_1 w(s_t, \alpha_t) + C_1 \epsilon_{1,t+1})) dF(\epsilon_{1,t+1})$$

$$+ \beta (1 - \alpha_t) \int W(A_2 s_t + B_2 w(s_t, \alpha_t) + C_2 \epsilon_{2,t+1}, B(\alpha_t, A_2 s_t + B_2 w(s_t, \alpha_t) + C_2 \epsilon_{2,t+1})) dF(\epsilon_{2,t+1})$$

---

\(^5\)Appendix C describes our algorithm for solving (13).

\(^6\)We interpret Blinder (1998, chapter 1) as advocating this point of view.
Because (12) is a feasible policy for the decision maker of subsection 2.2 who is willing to experiment, it follows that

\[ V(s_t; \alpha) \geq W(s_t; \alpha) \] (18)

for all values of \( s_t, \alpha \). The gap

\[ V(s_t, \alpha) - W(s_t; \alpha) \] (19)

measures the value of experimentation and the difference

\[ v(s_t, \alpha) - w(s_t; \alpha) \] (20)

measures the component of the time \( t \) 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 (14)-(15) have been recommended as an alternative or approximation to (5)-(6) to be used in situations in which the curse of dimensionality somehow prevents the policy maker or the analyst from solving Bellman equation (4) or the pertinent counterpart to it. The appeal of this approximation is greatest when the dimension of the prior distribution is large.\(^7\) 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 easy to write down but difficult to solve because of the dimension of the state vector.

3 The truth

So far our description has been about the views of the monetary authority that are summarized by equations (1) and \( \alpha_0 \in (0, 1) \). We have said everything about what the monetary authority believes and how it chooses \( v_t \), but nothing about how the economy actually works. Thus, our description so far is about ideas that are ‘just in the head’ of the monetary authority.

Under the monetary authority’s prior distribution over sequences for unemployment and inflation that is implied by our specification, \( \alpha_t \) is a martingale. See appendix A, section A.1 for a proof. Because \( \alpha_t \in [0, 1] \), the martingale convergence

\(^7\)See footnote 4.
theorem implies that $\alpha_t$ converges almost surely under that measure. To say what happens to $\alpha_t$ 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-$\alpha$ mixture of them, governs the data, then $\alpha_t$ given by (38) converges almost surely to the true $\alpha$.

Our concern in the next section is to study the rates at which $\alpha_t$ converges to the true $\alpha$ under alternative assumptions about which model is the true data generating process and alternative initial conditions for $\alpha, U$. 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 set $\beta = .995$ and $\lambda = 0.1$. A high value of $\beta$ 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 $\lambda$ 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 $V(U, \alpha) = W(U, \alpha)$. 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 both small. Therefore, in this example at least, the type of anticipated utility model used by Cogley and Sargent (2004), which is associated with Bellman equation

---

8Our 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.

9The prior is dogmatic at the boundaries, where data never alter the central bank’s beliefs. Those who are sure they know the truth are uninterested in experimentation.
(13), seems to provide a good approximation to the outcomes from the intentional experimentation model.\textsuperscript{10} We study the quality of approximation more fully in the following subsections that 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 $\alpha$. As noted, the differences between $v$ and $w$ are always small, but the biggest differences occur for $\alpha$’s away from the boundaries of 0 and 1. The figures reveal that when $\alpha$ is well into the interior of $(0, 1)$, $v$’s call for additional experimentation serves to make it nonlinear and to enhance the countercyclical 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.\textsuperscript{11}

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.\textsuperscript{12} This follows from two features, 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 central bank acquiesce to Keynesian policy prescriptions.

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 $\alpha$-axis when $\alpha$ is close to zero than when $\alpha$ is near one. That means deliberate experimentation has more incremental value for a monetary authority which is leaning 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 $\alpha$ close to 1, the $w$-policy calls for inflation to vary energetically in response to unemployment. That

\textsuperscript{10}See David Kreps (1998) for a broader defense of this modelling strategy in games and dynamic economic models.

\textsuperscript{11}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.

\textsuperscript{12}This feature of the decision rules conforms to the story about the conquest of American inflation told by Cogley and Sargent (2004).
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 $\alpha$ close to one, the $w$-policy itself provides identifying information about the world. The central bank 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 $\alpha$ 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 central bank would have to rely solely on natural experiments to learn the truth. Thus, under the $w$-policies, learning is likely to take longer when $\alpha_0 \approx 0$ than when $\alpha_0 \approx 1.$\textsuperscript{13} 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 $\alpha \approx 1.$

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

\textsuperscript{13}Indeed, this emerges in the simulations reported below.
Figure 2: Policy functions with and without experimentation.

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

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

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

5 Experiments in forgetting pernicious ideas

We generate alternative scenarios by specifying an initial condition for $U$, government beliefs $\alpha$, and which of our two models actually generates the data. We use the policy functions in figure 2 to generate histories of outcomes, and we relate those outcomes to questions 1-3 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 the arrival of Lucas and his model prompt the policy maker erroneously to assign some probability to it. 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, Lucas’s model arrives and is assigned a positive probability. Starting from period 19, we model the behavior of three central banks. 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 central bank takes into account that this prior will be revised in subsequent periods (black continuous lines), while the third (red lines) does not.

Lucas’s idea is counterproductive in this scenario 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

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

Lucas’s idea is counterproductive in this scenario 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 differentials are visually 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.\footnote{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.} Thus, Lucas’s ideas are not too pernicious in this setting. The central bank 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, but the benefit is a sharper decrease in unemployment compared to the anticipated utility central bank. The second bank 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, under the influence of Samuelson and Solow, the monetary policy decision maker assigns a probability $\alpha = 0.75$ to their model. As with the previous subsection, 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 banks’ 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 $\alpha_t$ back to 0, appears to be slower than occurred our analysis in the previous subsection where the Samuelson-Solow model prevailed. Although learning is initially quite...
rapid, with $\alpha$ 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 bank. The adaptive bank chooses an inflation rate that is approximately 40 basis points higher and that gap persists.

![Figure 6: In the three panels: the dotted line represents the behavior of a central bank that attaches probability one to the Samuelson and Solow model, the bold continuous line is the experimenting central bank and the other line is the non-experimenting bank.](image)

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
\( \alpha \) 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. A number of things can be learned from this table. First, as expected, the experimenting central bank learns the truth faster than an anticipated utility bank. Deliberate experimentation reduces the median weighting 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, for 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 bank must rely solely on natural experiments to learn. In contrast, when \( \alpha_0 = 0.99 \), the policy rule calls for countercyclical movements in inflation, so the central bank 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 \( \alpha \) 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 central bank attaches a higher weight to inflation in the loss function. Figure ?? shows how the value of experimentation changes with \( \lambda \). It appears that as \( \lambda \) rises from 0.1 to 1, the gain that can be obtained under the experimenting policy is significantly higher. It is inter-
<table>
<thead>
<tr>
<th>True Model</th>
<th>$\alpha_0$</th>
<th>$U_0$</th>
<th>Waiting Time</th>
<th>Experimentation</th>
<th>No Experimentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS</td>
<td>0.01</td>
<td>0</td>
<td>247</td>
<td>[149,486]</td>
<td>[156,500+]</td>
</tr>
<tr>
<td>SS</td>
<td>0.01</td>
<td>0.025</td>
<td>272</td>
<td>[154,498]</td>
<td>[171,500+]</td>
</tr>
<tr>
<td>Lucas</td>
<td>0.99</td>
<td>0</td>
<td>97</td>
<td>[37,242]</td>
<td>[40,244]</td>
</tr>
<tr>
<td>Lucas</td>
<td>0.99</td>
<td>0.025</td>
<td>85</td>
<td>[21,213]</td>
<td>[26,216]</td>
</tr>
<tr>
<td>SS</td>
<td>0.5</td>
<td>0</td>
<td>39</td>
<td>[20,80]</td>
<td>[23,92]</td>
</tr>
<tr>
<td>SS</td>
<td>0.5</td>
<td>0.025</td>
<td>21</td>
<td>[6,54]</td>
<td>[8,69]</td>
</tr>
<tr>
<td>Lucas</td>
<td>0.5</td>
<td>0</td>
<td>87</td>
<td>[32,160+]</td>
<td>[35,160+]</td>
</tr>
<tr>
<td>Lucas</td>
<td>0.5</td>
<td>0.025</td>
<td>65</td>
<td>[14,160+]</td>
<td>[18,160+]</td>
</tr>
<tr>
<td>SS</td>
<td>0.28</td>
<td>0</td>
<td>52</td>
<td>[27,160+]</td>
<td>[35,160+]</td>
</tr>
<tr>
<td>SS</td>
<td>0.28</td>
<td>0.025</td>
<td>35</td>
<td>[15,142]</td>
<td>[19,148]</td>
</tr>
<tr>
<td>Lucas</td>
<td>0.28</td>
<td>0</td>
<td>80</td>
<td>[26,160+]</td>
<td>[28,160+]</td>
</tr>
<tr>
<td>Lucas</td>
<td>0.28</td>
<td>0.025</td>
<td>71</td>
<td>[20,160+]</td>
<td>[22,160+]</td>
</tr>
</tbody>
</table>

Table 1: Waiting times for various data generating processes and initial $(\alpha_0, U_0)$ pairs.
esting to notice that the peak of the curve shifts toward the middle of the support of $\alpha$. When the policy maker is more concerned about unemployment than about inflation a conspicuous amount of unintentional experimentation will be created when the Samuelson and Solow model is likely to be the data generating process. However, as inflation is attached a higher weight in the reward function, the degree of unintentional experimentation is bound to drop, explaining the increase in the difference of the value functions associated with the two policies.

When $\lambda$ is as high as 16, the value of following 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 $\alpha^{15}$ and experimentation usually calls for a lower inflation rate. However, $\lambda$ is so high that the programmed inflation resulting from the Bayesian and anticipated utility approach are extremely low and almost undistinguishable. Hence the very small benefit from experimentation shown in the picture.

In Table 2 we report the relative behavior of unemployment and inflation mean and variance for the Bayesian and anticipated utility central banks for the cases of misplaced experimentation when either Samuelson and Solow or Lucas are correct. In each experiment, we generated 500 sample paths of length 40 (ten years) for the three values of $\lambda$ discussed in this section. Several things are worth noticing in this table. First notice that inflation is always more volatile when the central bank is following the Bayesian rule$^{16}$. When $\lambda = 16$ the differences are mild and this 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 central bank 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.

6 Robustness

Our model falls within the framework of models for which Hansen and Sargent (2005a, 2005b) formulate robust filtering and control problems.$^{17}$ Hansen and Sargent induce robust decision rules by replacing each of two expectation operators in

\[ ^{15} \text{Pictures are available upon request.} \]
\[ ^{16} \text{There is only one exception that we attribute to numerical error.} \]
\[ ^{17} \text{See Epstein and Schneider (2003) and the literature they discuss for formulations of decision problems in which the decision maker expresses ambiguity over dynamic models by having multiple prior distributions.} \]
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 \((\alpha_0, U_0)\) pair, the first row reports the average inflation (unemployment) without experimentation relative to the average inflation (unemployment) with experimentation for different values of \(\lambda\). 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 \(\lambda\). In both panels: 500 samples of ten years (40 quarters) were generated.
Figure 7: $V(U, \alpha) - W(U, \alpha)$ for different values of $\lambda$. 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).

Bellman equation (4) (i.e., those that average over the distributions of $\epsilon_j$ and those that average over the distribution of model probabilities $\alpha$) with appropriate risk-sensitivity operators. This section computes and interprets decision rules that are robust with respect to various combinations of the two types of misspecification.

Let $z = \{0, 1\}$ be a binary hidden state variable that indexes which model prevails, with the interpretation being that model $z + 1$ is true (remember that we called Samuelson-Solow’s model 1 and Lucas’s model 2). Then let $s^*$’s denote next period values and express (4) as

$$v(s, \alpha) = \max_v \left\{ r(s, v) + E_z \left[ E_{s^*, \alpha^*} (\beta v(s^*, \alpha^*) | s, v, z) | s, v \right] \right\}$$

(21)
subject to

$$
\begin{align*}
    s^* &= \pi_s(s, v, z) \\
    \alpha^* &= \pi_\alpha(s, \alpha, v, z) \equiv B(\alpha, \pi_s(s, v, \alpha))
\end{align*}
$$

where $E_z$ denotes integration with respect to the distribution of the hidden state $z$, and $E_{s^*, \alpha^*}$ denotes integration with respect to $s^*, \alpha^*$ conditional on $(s, v, \alpha)$. Following Hansen and Sargent, to induce robust decision rules, in (21) we replace the expectation operator with respect to the distribution of $s^*, \alpha^*$ conditioned on $s, v, z$ with a risk-sensitivity operator $(T^1)(\theta_1)$ and the expectation operator with respect to $z$ conditioned on $s, v$ with a risk-sensitivity operator $(T^2)(\theta_2)$ to get a Bellman equation

$$
v(s, \alpha) = \max_v \left\{ r(s, v) + T^2 \left[ T^1(\beta v(s^*, \alpha^*)(s, v, z)) \right] \right\}. \tag{24}
$$

We shall now describe the two operators $T^1, T^2$ indexed by the two robustness parameters, $\theta_1 > 0, \theta_2 > 0$, respectively, and the types of misspecifications that the operators defend against. When both $\theta_i$ parameters are set to $+\infty$, the Bellman equations become identical with (4).

### 6.1 Misspecification of the prior over models: the $T^2$ operator

To indicate the types of misspecifications that the two operators protect against, we activate the operators one at a time. We begin by assuming that, conditional on knowing that it is true, the decision maker completely trusts each of his two models. However, he does not trust the prior probabilities that he has assigned to them. He models this distrust by considering a set of priors, i.e., a set of $\alpha$’s, whose size is inversely measured by $\theta_2$, and seeking a decision that works well over this set.\footnote{Hansen and Sargent (2005) describe how to measure the size of this set in terms of relative entropy and $\theta_2$.} To induce a decision rule that is robust to misspecification of $\alpha_t$, we replace (4) with the following Bellman equation that implicitly describes the $T^2$ operator:

$$
V(s_t, \alpha_t) = \max_{v_t} \left\{ r(s_t, v_t) \\
- \beta \theta_2 \log \left[ \alpha_t \exp \left( - \int V(A_1 s_t + B_1 v_t + C_1 \epsilon_{1,t+1}, B(\alpha_t, A_1 s_t + B_1 v_t + C_1 \epsilon_{1,t+1}))dF(\epsilon_{1,t+1}) \right) \right] \\
+ (1 - \alpha_t) \exp \left( - \int V(A_2 s_t + B_2 v_t + C_2 \epsilon_{2,t+1}, B(\alpha_t, A_2 s_t + B_2 v_t + C_2 \epsilon_{2,t+1}))dF(\epsilon_{2,t+1}) \right) \right\}. \tag{25}
$$
Application of the \((T^2)(\theta_2)\) operator implicitly defined on the right side of (25) implies a worst-case distortion to \(\alpha_t\) is

\[
d(s_t, \alpha_t) = \frac{n_1(s_t, \alpha_t)}{\alpha_t n_1(s_t, \alpha_t) + (1 - \alpha_t) n_2(s_t, \alpha_t)}
\]

where

\[
n_1(s_t, \alpha_t) = \exp\left(-\frac{\int V(A_1s_t + B_1v_t + C_1\epsilon_{1,t+1}, B(\alpha_t, A_1s_t + B_1v_t + C_1\epsilon_{1,t+1}))dF(\epsilon_{1,t+1})}{\theta_2}\right)
\]

\[
n_2(s_t, \alpha_t) = \exp\left(-\frac{\int V(A_2s_t + B_2v_t + C_2\epsilon_{2,t+1}, B(\alpha_t, A_2s_t + B_2v_t + C_2\epsilon_{2,t+1}))dF(\epsilon_{2,t+1})}{\theta_2}\right)
\]

where it is understood that \(v_t\) on the right side is evaluated at the robust decision rule for \(v_t\). The distorted probability over the hidden state is\(^{19}\)

\[
\tilde{\alpha}_t = d(s_t, \alpha_t) \alpha_t.
\]

For \(\lambda = .1\) and \(\theta_2 = .1\), the left panel of figure 8 plots the worst case distortion to \(\alpha\) and the right panel plots robust decision rule for inflation \(v_t\) as a function of \(U_t, \alpha_t\). We set \(\theta_2 = .1\) partly because this value gives noticeable effects of a concern about robustness on decision rules. We interpret the outcome that the worst-case \(\alpha\) is always less than \(\alpha\), recall from the value function displayed in (1) that for the policy maker, the Lucas model is the worse of the two models. Since \(\alpha\) is the probability attached to the Samuelson-Solow model, it is understandable that the worst case \(\alpha\) is always lower than \(\alpha\). Evidently, for intermediate values of \(\alpha\) but high values of \(U\), the robust decision maker sets a lower inflation target than does one who has no doubts about his prior probabilities. This is because the policy maker makes robust decisions by in effect increasing the prior weight that he attaches to the Lucas model, under which inflation is ineffective as a tool for affecting unemployment.

### 6.2 Fear of misspecification of each model: the \(T^1\) operator

We now activate a concern about misspecification for each of the two models themselves by using the parameter \(\theta_1 > 0\) that characterizes Hansen and Sargent’s’s \(T^1\) operator. To begin, we shall assume that the decision maker is confident about his prior and set \(\theta_2 = +\infty\). We activate the \(T^1\) operator by replacing the pertinent conditional expectation operator in (4) with the \(T^1\) operator, in particular,

\[
\int V(A_j s_t + B_j v_t + C_j \epsilon_{j,t+1}, B(\alpha_t, A_j s_t + B_j v_t + C_j \epsilon_{j,t+1}))dF(\epsilon_{j,t+1})
\]

\(^{19}\)Notice that Bayes’ law under the approximating model becomes part of the approximating model around which the decision maker puts statistical perturbations against which he seeks robust decisions. See Hansen and Sargent (2004a, 2004b).
in (4) with
\[
-\theta_1 \log \int \exp \left( -\frac{V(A_j s_t + B_j v_t + C_j \epsilon_{j,t+1}, B(\alpha_t, A_j s_t + B_j v_t + C_j \epsilon_{j,t+1}))}{\theta_1} \right) dF(\epsilon_{j,t+1}).
\]

The worst case distortion to the density of \( \epsilon_{j,t+1} \) is
\[
\phi(\epsilon_{j,t+1}, s_t, \alpha_t) = \frac{\exp \left( -\frac{V(A_j s_t + B_j v_t + C_j \epsilon_{j,t+1}, B(\alpha_t, A_j s_t + B_j v_t + C_j \epsilon_{j,t+1}))}{\theta_1} \right)}{\int \exp \left( -\frac{V(A_j s_t + B_j v_t + C_j \epsilon_{j,t+1}, B(\alpha_t, A_j s_t + B_j v_t + C_j \epsilon_{j,t+1}))}{\theta_1} \right) dF(\epsilon_{j,t+1})}
\]
where it is understood that \( v_t \) on the right side is evaluated at the robust decision rule. The distorted conditional density of \( \epsilon_{j,t+1} \) is then
\[
\bar{\phi}(\epsilon_{j,t+1}, s_t, \alpha_t) = \phi_n(\epsilon_{j,t+1})\phi(\epsilon_{j,t+1}, s_t, \alpha_t)
\]
where $\phi_n(\epsilon_{j,t+1})$ is the standard normal density.

Figure 9 displays the robust decision rule for $\theta_1 = 0.1$ and $\theta_2 = +\infty$,\(^{20}\) while figure (10) displays the conditional means and variances of the worst case conditional densities (28), where recall that model 1 indicates the Samuelson-Solow model, and figure 11 shows them for the Lucas model. Figure 10 shows that for intermediate values of $\alpha$, the worst case conditional mean and variance of $\eta_{1,t+1}$ are both increased when $U$ is high. The increase in the conditional mean when $U$ is high translates into higher persistence of unemployment in the Samuelson-Solow model. According to figure 9, the robust policy authority responds to this outcome by increasing target inflation relative to that set by a policy maker who fully trusts the specification of each model. Thus, a concern about the specification of each model, while retaining full confidence in his prior over the two classes of models, has an opposite effect from a concern about the prior alone, which we summarized in figure 8.

Figure 12 activates both types of concerns about the specification. As might be guessed from the complexion of the earlier results, turning on both source of concern about robustness yields a decision rule that is close to the one we obtained without any concerns about robustness. In effect, the worst case $\alpha$ shown in the left panel of figure 12 offsets the worst-case dynamics coming from the dependence of the worst-case conditional mean on $U_t, \alpha_t$. Thus, our optimal decision rule with experimentation calculated without explicit reference to robustness is robust to a mixture of concerns about the two types of misspecification captured by the $T^1$ and $T^2$ operators.

7 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.

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 $\alpha$ away from 1, than are the decisions of Lucas-style classical economist to small perturbations of $\alpha$ away from 0.

\(^{20}\)We set $\theta_1 = .1$ because this value delivers noticeable affects on the decision rule.
That lack of robustness of the classical recommendations to small doubts plays an essential feature in Cogley and Sargent’s accounting for 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 influencing the speed of convergence is the probability weight $\alpha$, for that affects the contribution of policy-induced variation to learning. When $\alpha$ is close to zero, policy keeps inflation close to zero, and the central bank must rely heavily on natural experiments for learning. When $\alpha$ is close to one, the bank still learns from natural experiments but also learns from policy-induced, countercyclical variation in inflation. That extra source of variation speeds learning.

To evaluate the lessons of our results, it is important to assess the reality that we impute to our policy maker to the doubts that one thinks should be in the minds of the policy maker. We have given the policy maker only two models, each
of which he knows for sure. Of course, the models have very different operating characteristics. While their differences are important, they are not subtle, so this makes easier 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, $\alpha$. 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 section 6 go part way in addressing some of those concerns.
Figure 11: Distorted shocks to Lucas model with $\theta_1 = .1$.

8 Appendixes

A Transition equation for $\alpha_t$

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 $\theta_i$ are known. By Bayes’s theorem, the posterior probability on model $i$ is

$$\alpha_{it} \equiv p(M_i|s_i^t, \theta_i) = \frac{p_i(s_i^t|\theta_i)p(M_i)}{\int p_i(s_i^t|\theta_i)p(M_i) dM_i}. \quad (29)$$

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
\[
\log w_{it} - \log w_{it-1} = \log p_i(s_i^t | \theta_i) - \log p_i(s_i^{t-1} | \theta_i).
\] (30)

Note that the prior model weight drops out of the recursion; $\alpha_{i0}$ initializes the sequence but the likelihood is all that matters for updates. Also notice that $\alpha$-updates depend only on the value of the likelihood at the given $\theta_i$. Usually the model probability updates would depend on a marginalized likelihood, but this drops out because we assume that $\theta_i$ 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
\[
\log p_i(s_i^t | \theta_i) = \sum_{s=1}^{t} \log p_i(s_{is} | s_i^{s-1}, \theta_i).
\] (31)
Subtracting the log-likelihood through \( t - 1 \) from that through \( t \), we get
\[
\log w_{it} = \log w_{it-1} + \log p_i(s_{it}|s_i^{t-1}, \theta_i).
\] (32)

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 \( \alpha_t \) be the normalized probability weight for model 1,
\[
\alpha_t = \frac{w_{1t}}{w_{1t} + w_{2t}}.
\] (33)

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

The normalizing constant is a nuisance, so we eliminate it by taking the ratio,
\[
R_t \equiv \frac{\alpha_t}{1 - \alpha_t} = \frac{w_{1t}}{w_{2t}}.
\] (34)

The transition equation for log \( R_t \) follows from the transition equations for log \( w_{it} \),
\[
\log R_t = \log R_{t-1} + \log \frac{p_1(s_{1t}|s_1^{t-1}, \theta_1)}{p_2(s_{2t}|s_2^{t-1}, \theta_2)}.
\] (35)

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 \( \alpha_t \), we find
\[
\frac{\alpha_t}{1 - \alpha_t} = \frac{\alpha_{t-1} p_1(s_{1t}|s_1^{t-1}, \theta_1)}{1 - \alpha_{t-1} p_2(s_{2t}|s_2^{t-1}, \theta_2)},
\] (36)

or
\[
\alpha_t = \frac{\frac{\alpha_{t-1} p_1(s_{1t}|s_1^{t-1}, \theta_1)}{1 - \alpha_{t-1} p_2(s_{2t}|s_2^{t-1}, \theta_2)}}{1 + \frac{\alpha_{t-1} p_1(s_{1t}|s_1^{t-1}, \theta_1)}{1 - \alpha_{t-1} p_2(s_{2t}|s_2^{t-1}, \theta_2)}},
\] (37)

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

If the two models involve the same data, we can equate \( s_{1t} = s_{2t} \). In that case,
\[
\alpha_t = \frac{\alpha_{t-1} p_1(s_t|s_t^{t-1}, \theta_1)}{\alpha_{t-1} p_1(s_t|s_t^{t-1}, \theta_1) + (1 - \alpha_{t-1}) p_2(s_t|s_t^{t-1}, \theta_2)}.
\] (38)

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

The updating formula makes $\alpha_t$ 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 $\alpha_t$ with respect to the posterior at $\alpha_{t-1}$,

$$E_{t-1}B(\alpha_{t-1}, s_t) = \int B(\alpha_{t-1}, s_t) f_{t-1}(s_t|s^{t-1}) ds_t.$$  \hfill (39)

Because model parameters are assumed to be known, there is a single source of uncertainty about next period’s $\alpha_t$, 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,

$$f_{t-1}(s_t|s^{t-1}) = \alpha_{t-1} p_1(s_t|s^{t-1}, \theta_1) + (1 - \alpha_{t-1}) p_2(s_t|s^{t-1}, \theta_2).$$  \hfill (40)

Thus, the conditional expectation for $\alpha_t$ is

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

$$= \int \alpha_{t-1} p_1(s_t|s^{t-1}, \theta_1) ds_t,$$

$$= \alpha_{t-1} \int p_1(s_t|s^{t-1}, \theta_1) ds_t = \alpha_{t-1}. \hfill (41)$$

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

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

$\alpha_t$ can be obtained back through the following expression

$$\alpha_t = \frac{1}{1 + \exp \log R_t} \hfill (42)$$

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:

\[
C_2 \epsilon_{2,t+1} = (A_1 - A_2)s_t + (B_1 - B_2)v_t + C_1 \epsilon_{1,t+1} \tag{43}
\]

\[
C_1 \epsilon_{1,t+1} = -(A_1 - A_2)s_t - (B_1 - B_2)v_t + C_2 \epsilon_{2,t+1} \tag{44}
\]

Define \( \Omega_1 = C_1C_1^T, \Omega_2 = C_2C_2^T \). We use (43) and (44) to write the recursion for \( \log R_{t+1} \) under models 1 and 2. When model 1 is true, we have

\[
\log R_{t+1} = \log R_t - \frac{1}{2} \log |\Omega_1| + \frac{1}{2} \log |\Omega_2| - \frac{1}{2} (C_1 \epsilon_{1,t+1})' \Omega_1^{-1} (C_1 \epsilon_{1,t+1}) + \frac{1}{2} [(A_1 - A_2)s_t + (B_1 - B_2)v_t + C_1 \epsilon_{1,t+1}]' \\
\times \Omega_2^{-1} [(A_1 - A_2)s_t + (B_1 - B_2)v_t + C_1 \epsilon_{1,t+1}] \tag{45}
\]

When model 2 is true, we have

\[
\log R_{t+1} = \log R_t - \frac{1}{2} \log |\Omega_1| + \frac{1}{2} \log |\Omega_2| + \frac{1}{2} (C_2 \epsilon_{2,t+1})' \Omega_2^{-1} (C_2 \epsilon_{2,t+1}) - \frac{1}{2} [(A_2 - A_1)s_t + (B_2 - B_1)v_t + C_2 \epsilon_{2,t+1}]' \\
\times \Omega_1^{-1} [(A_2 - A_1)s_t + (B_2 - B_1)v_t + C_2 \epsilon_{2,t+1}] \tag{46}
\]

It is convenient to use \( \alpha_t \) rather than \( \log R_t \) as a state variable. So we want to transform (45) and (46) to get laws of motion for \( \alpha_t \) under the two models. For the purpose of doing this, define

\[
g(\epsilon_{1,t+1}; s_t, \alpha_t) = \log R_t - \frac{1}{2} \log |\Omega_1| + \frac{1}{2} \log |\Omega_2| - \frac{1}{2} (C_1 \epsilon_{1,t+1})' \Omega_1^{-1} (C_1 \epsilon_{1,t+1}) + \frac{1}{2} [(A_1 - A_2)s_t + (B_1 - B_2)v_t + C_1 \epsilon_{1,t+1}]' \\
\times \Omega_2^{-1} [(A_1 - A_2)s_t + (B_1 - B_2)v_t + C_1 \epsilon_{1,t+1}] \tag{47}
\]

and

\[
h(\epsilon_{2,t+1}; s_t, \alpha_t) = \log R_t - \frac{1}{2} \log |\Omega_1| + \frac{1}{2} \log |\Omega_2| + \frac{1}{2} (C_2 \epsilon_{2,t+1})' \Omega_2^{-1} (C_2 \epsilon_{2,t+1}) - \frac{1}{2} [(A_2 - A_1)s_t + (B_2 - B_1)v_t + C_2 \epsilon_{2,t+1}]' \\
\times \Omega_1^{-1} [(A_2 - A_1)s_t + (B_2 - B_1)v_t + C_2 \epsilon_{2,t+1}] \tag{48}
\]

33
Using (42), we get a law of motion for $\alpha_{t+1}$ under the two models. Then our Bellman equation can be expressed

$$V(s_t, \alpha_t) = \max_{v_t} \left\{ r(s_t, v_t) + \beta \alpha_t \int V \left( A_1 s_t + B_1 v_t + C_1 \epsilon_{1,t+1} + \frac{e^{\beta(\epsilon_{1,t+1})}}{1 + e^{\beta(\epsilon_{1,t+1})}} \right) d\epsilon_{1,t+1} \\
+ \beta(1 - \alpha_t) \int V \left( A_2 s_t + B_2 v_t + C_2 \epsilon_{2,t+1} + \frac{e^{\beta(\epsilon_{2,t+1})}}{1 + e^{\beta(\epsilon_{2,t+1})}} \right) d\epsilon_{2,t+1} \right\}.$$  

(49)

**B Approximating the Bellman equation**

Discretize the support of $\alpha$ and $s$ into $I_\alpha$ 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 $\alpha_i$ 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

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

(50)

$\forall i = 1, ..., I$ 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, $\Phi$ 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 (50) in the Bellman equation we get for each node $i \in \{1, ..., I\}$:

$$V_i = \max_{v_i} \left\{ r_i(v_i) + \beta \alpha_i \sum_{k=1}^K \sum_{j=1}^J w_k c_j \phi_j \left( s_{1,i,k}'(v_i), \frac{\exp[g_{k,i}(v_i)]}{1 + \exp[g_{k,i}(v_i)]} \right) \\
+ \beta(1 - \alpha_i) \sum_{k=1}^K \sum_{j=1}^J w_k c_j \phi_j \left( s_{2,i,k}'(v_i), \frac{\exp[h_{k,i}(v_i)]}{1 + \exp[h_{k,i}(v_i)]} \right) \right\}$$  

(51)

where

$$r_i(v_i) = r(s_i, v_i)$$

$$s_{1,i,k}'(v_i) = A_1 s_i + B_1 v_i + C_1 \epsilon_k$$

$$s_{2,i,k}'(v_i) = A_2 s_i + B_2 v_i + C_2 \epsilon_k$$
and $g_{k;i}(v_i)$ and $h_{k;i}(v_i)$ defined as in (47) and (48) respectively:

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

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 $(s, \alpha)$, compute the right hand side of equation (51) 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

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

or

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

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

$$L = -0.5E \sum_{t=0}^{\infty} \beta^t \left\{ \alpha s'_{1t}Rs_{1t} + (1 - \alpha)s'_{2t}Rs_{2t} + v'_tQv_t \right\}$$

or

$$L = -0.5E \sum_{t=0}^{\infty} \beta^t \left\{ s'_t \begin{bmatrix} \alpha R & 0 \\ 0 & (1 - \alpha)R \end{bmatrix} s_t + v'_tQv_t \right\}$$

Cogley and Sargent (2004) note that the problem of choosing a decision rule to maximize (55) with respect to (53) is an optimal linear regulator problem. The optimal decision rule is

$$v_t = -Fs_t = -F_1s_{1t} - F_2s_{2t}.$$
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 \( UR_t \),

\[
U^*_t = U^*_{t-1} + \mu(UR_t - U^*_{t-1}), \tag{57}
\]

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

\[
U_t \equiv UR_t - U^*_t = \frac{(1 - \mu)(1 - L)}{1 - (1 - \mu)L} UR_t. \tag{58}
\]

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](image)

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

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>$U_{t-1}$</th>
<th>$\pi_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.0023</td>
<td>0.7971</td>
<td>-0.2761</td>
</tr>
<tr>
<td>$\sigma_\beta$</td>
<td>0.0010</td>
<td>0.0699</td>
<td>0.1189</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>$U_{t-1}$</th>
<th>$\pi_t - v_{t-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.0007</td>
<td>0.8468</td>
<td>-0.2489</td>
</tr>
<tr>
<td>$\sigma_\beta$</td>
<td>0.0008</td>
<td>0.0674</td>
<td>0.1298</td>
</tr>
</tbody>
</table>

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


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.


