Evolution and Intelligent Design

Thomas J. Sargent‡

October 18, 2007

Abstract

This paper discusses two sources of ideas that influence monetary policy makers today. The first is a set of analytical results that impose the rational expectations equilibrium concept and do ‘intelligent design’ by solving Ramsey and mechanism design problems. The second is a long trial and error learning process that first taught us how to to anchor the price level with a gold standard, then how to replace the gold standard with a fiat currency system wanting nominal anchors. Models of out-of-equilibrium learning tell us that such an evolutionary process will converge to a self-confirming equilibrium (SCE). In an SCE, a government’s probability model is correct about events that occur under the prevailing government policy, but possibly wrong about the consequences of other policies. That leaves room for more mistakes and useful experiments than exist in a rational expectations equilibrium.

Keywords: Rational expectations equilibrium, mechanism design, model misspecification, learning, evolution, observational equivalence, self-confirming equilibrium. (JEL).

*New York University and Hoover Institution. Email: ts43@nyu.edu.
†This draft of my presidential address to the American Economic Association in January 2008 continues a long conversation I have had with Chris Sims (see Sims (1982)). After Bob Lucas read Sargent (1984), he wrote me that “With friends like you, Chris doesn’t need enemies.” Maybe it is more complicated than that. I thank Gadi Barlevy, Francisco Barillas, Marco Bassetto, Alberto Bisin, William Branch, In-Koo Cho, Timothy Cogley, Lars Peter Hansen, Kenneth Kasa, Narayana Kocherlakota, Larry Jones, Athanasios Orphanides, Carolyn Sargent, Hyun Shin, Christopher Sims, Noah Williams, François Velde, Carl Walsh, Peyton Young, and Tao Zha for helpful comments. I thank the National Science Foundation for research support.
1 Introduction

... the ideas of economists and political philosophers, both when they are right and when they are wrong, are more powerful than commonly understood. Indeed the world is ruled by little else. Practical men, who believe themselves to be quite exempt from any intellectual influences, are usually the slaves of some defunct economist. Keynes (1936, p. 383).

Today leading practical men and women at important institutions in my field, the Federal Reserve, the Bank of England, and various constituents of the European Central Bank, are distinguished academic economists. They have used prevailing academic ideas about macroeconomics to choose policy actions. Some designed new institutions. For example, in 1997 Mervyn King and some others created decision making protocols for Britain’s monetary policy committee virtually from scratch.

This essay is about two important sources of prevailing ideas in macroeconomics. The first is a collection of powerful theoretical results and empirical methods that apply the rational expectations equilibrium concept to design optimal macroeconomic policies intelligently. The second is an evolutionary process embedded in an historical economic record littered with discarded ideas and policies. but that can be hoping for too much.

The rational expectations equilibrium concept equates all subjective distributions with an objective distribution. It is useful to distinguish the step of equating subjective beliefs of all agents from the step of equating subjective beliefs also to the distribution that govern outcomes. By equating subjective distributions for all agents within a model, the rational expectations hypothesis makes agents’ beliefs disappear as extra components of a theory and sets the stage for a variety of powerful theoretical results and intelligent policy design exer-

A younger Keynes was less optimistic about the influence of economists’ ideas:

Financiers of this type [Lord Rothschild, Lord Avebury, Lord Swaythling] will not admit the feasibility of anything until it has been demonstrated to them by practical experience. It follows, therefore, that they will seldom give their support to what is new. Keynes (1913, pp. 24-25)
cises. Equating subjective distributions to an objective distribution facilitates econometrics.

The assumption that agents share common beliefs underpins influential doctrines about whether observed inflation-unemployment dynamics can be exploited by policy makers, the time inconsistency of benevolent government policy, the capacity of reputation to substitute for commitment, the incentives for a policy maker of one type to emulate another, and the wisdom of making information public. The common beliefs assumption is stressed especially in modern theories of optimal macroeconomic policy that focus on how a benevolent government optimally shapes expectations. This intelligent design approach to macroeconomic policy perfects an older econometric policy evaluation method that Lucas (1976) criticized because it imputed different beliefs to the government and the other agents.

Intelligent design is normative (`what should be') economics, but when it influences policy makers, it becomes positive (`what is') economics (as asserted in the epigraph from Keynes). Some researchers in the intelligent design tradition ignore the distinction between positive and normative economics from the start. Thus, a standard tool for understanding observed time series properties of government debt and taxes is to apply a normative analysis, e.g., Barro (1979), Lucas and Stokey (1983), and Aiyagari et al. (2002). It is also true that some policy advisors have enough faith that evolution produces good outcomes to recommend copying best practices (for example, see Keynes (1913)). If only good things survive the tests of time and practice, evolution produces intelligent design.

Theories of out-of-equilibrium learning tell us not always to expect that. A system of adaptive agents converges to a self-confirming equilibrium in which all agents have correct forecasting distributions for events that occur often enough along an equilibrium path, but possibly mistaken views about policies and outcome paths that will not be observed. This matters because intelligent macroeconomic policy design of rational expectations equilibria hinges on knowing and manipulating expectations about events that will not be observed. Self-confirming equilibria allow models to survive that imply mistaken policies even though they match historical data well. I devote section 4 to a framework for thinking about learning,
then use it in section \ref{sec:undirected} and appendix \ref{app:undirected} to describe some undirected evolutionary processes that show how ideas that were once prevalent, but have now been discarded, shaped policies and generated experiments that brought us to where we are.

2 Intelligent design with common beliefs

By solving Pareto problems in which a planner and all agents optimize in light of information and incentive constraints and a common probability model, what I call intelligent design is a coherent response to the Lucas (1976) indictment of pre-rational expectations macroeconomic policy design procedures. Lucas accused those procedures of incorporating private agents’ decision rules that were not best responses to government policy under an equilibrium probability measure. The cross-equation restrictions of common belief models fix that problem.

Throughout this paper, I use $f$ to denote a probability density and $x^t$ to denote a history $x_t, x_{t-1}, \ldots, x_0$. It is convenient to use the partition $x_t = [y_t, v_t]'$ where $v_t$ is a vector of decisions taken by a government and $y_t$ is a vector of all other variables. Let $f(y^\infty, v^\infty|\rho)$ be a joint density conditional on a parameter vector $\rho \in \overline{R}$. The joint density is a model of the economy that, among other things, describes best responses of private agents who choose some of the components $y_t$, where it is assumed that for all agents in the model those best responses maximize their expected utilities under the density $f$. Government chooses a sequence $h$ of functions

$$v_t = h_t(y_t|\rho), \quad t \geq 0, \quad (1)$$

to maximize a Pareto criterion that can be expressed as expected utility under density $f(x^\infty|\rho)$:

$$\int U(y^\infty, v^\infty|\rho)f(y^\infty, v^\infty|\rho)d(y^\infty, v^\infty). \quad (2)$$

Modern intelligent design in macroeconomics solves government programming problems \cite{2} with models $f$ that impute common beliefs and best responses to all of the agents who
inhabit the model. The common beliefs assumption used to construct the macroeconomic model makes parameters describing agents beliefs about endogenous variables disappear from the vector $\rho$.

The common beliefs assumption underlies a long list of interesting results in modern macro. The following have especially influenced thinking within central banks.

1. *Expected versus unexpected government actions.* Lucassen (1972b) drew a sharp distinction between the effects of foreseen and unforeseen monetary and fiscal policies when the government and the public share a probability model. That idea defines the terms in which central bankers now think about shocks and systematic policies.

2. *Optimal fiscal and monetary policy cast as Ramsey and mechanism design problems.* A literature summarized and extended by Clarida et al. (1999) and Woodford (2003) uses dynamic macroeconomic models with sticky prices to design monetary policy rules by solving Ramsey plans like (2) and finding attractive ways to represent and implement them. A new dynamic public finance literature aims to refine the Ramsey literature by focusing on a tradeoff between efficiency and incentives that emerges from the assumption each individual alone observes his own skills and effort, a feature that imposes constraints on the allocations that the planner can implement relative to ones he could achieve if he had more information.\footnote{See for example Golosov et al. (2003), Kocherlakota (2005), and Golosov et al. (2007).}

3. *Time consistency.* The availability of the rational expectations equilibrium concept enabled Kydland and Prescott (1977) and Calvo (1978) to explain how alternative timing protocols affect a benevolent government’s capacity to manipulate and then confirm prior expectations about its actions.\footnote{While technical treatments of the time consistency problem rely heavily on the rational expectations equilibrium concept, all that is needed to spot the problem is that private agents care about future government actions. In a discussion at the U.S. Constitutional Convention about whether the Federal government should be prohibited from fiduciary currency on August 16, 1787, Gouverneur Morris, Oliver Ellsworth, and James Madison recognized a time consistency problem, while Edmund Randolph and George Mason raised doubts about tying the hands of the government because no one can not foresee all contingencies. See Madison (1987, pp. 470-471).} The time consistency ‘problem’ is the
observation that equilibrium outcomes in a representative agent economy depend on the timing protocol for decision making that nature or the modeler imposes on a benevolent government. Better outcomes emerge if the government chooses a history-contingent plan once-and-for-all at time 0 than if it chooses sequentially. By choosing future actions at time 0, the government can take into account how expectations about its actions at times \( t > 0 \) influence private agents’ actions at all dates between 0 and \( t \). A government must ignore those beneficial expectations effects if it is forced to choose sequentially.

4. **Reputation can substitute for commitment.** A credible public policy is an equilibrium system of expectations that gives a government incentives to confirm prior expectations about its future actions, actions to which it cannot commit because it chooses sequentially. The key object is a history-dependent government strategy like (II). There are multiple equilibrium strategies, i.e., multiple systems of common expectations that a government would want to confirm, with good and bad equilibria being tied together via incentive constraints. A government strategy plays two roles, first, as a *decision rule* for the government and, second, as a *system of private sector expectations* about government actions that the government always wants to confirm.\(^4\)

---

\(^4\)See the credible public plans models of Stokey (1989, 1991) and Chari and Kehoe (1993b,a). By making an intrinsically ‘forward-looking’ variable, a promised discounted value for the representative household, also be a ‘backward-looking’ state variable that encodes history, Abreu et al. (1986, 1990) tie past and future together in a subtle way that exploits the common beliefs equilibrium concept. For some applications, see Chang (1998), Phelan and Stacchetti (2001), and Liungqvist and Sargent (2004, ch. 22).

\(^5\)The theory is silent about who chooses an equilibrium system of beliefs, the government (after all, it is the government’s decision rule) or the public (but then again, they are the private sector’s expectations). This and the multiplicity of equilibria make it difficult to use this theory to formulate advice to policy makers about actions that can help it to earn a good reputation. Instead, the theory is about how a government comes into a period confronting a set of private sector expectations about its actions that it will want to confirm. Blinder (1998, pp. 60-62) struggles with this issue when he describes the pressures he perceived as Fed Vice Chairman not to disappoint the market. While Blinder’s discussion can be phrased almost entirely within the rational expectations paradigm, the account by Bernanke (2007) of the problems the Fed experiences in anchoring private sector expectations cannot. Bernanke argues in terms away from a rational expectations equilibrium.

\(^6\)The theory of credible public policy seems to explain why some policy makers who surely knew about better decision rules chose instead to administer ones supporting bad outcomes. Chari et al. (1998) and Albanesi et al. (2002) interpret the big inflation of the 1970s and its stabilization in the 1980s in terms of the actions of benevolent and knowledgeable policy makers who became trapped within but, thanks to a sunspot,
These theoretical rational expectations results have influenced the way monetary policy is now discussed within central banks. Because central banks want to implement solutions of Ramsey problems like (2) in contexts like (1) in which the distinction between the effects of foreseen and unforeseen policy actions is important, a time consistency problem like (3) arises, prompting them to focus on ways like (4) of sustaining a good reputation.\footnote{See \textit{Blinder} (1998) and \textit{Bernanke et al.} (2001).}

### 2.1 Justifications for equating objective and subjective distributions

These and many other theoretical results hinge on the part of the rational expectations equilibrium concept that equates subjective distributions for all agents inside a model. To gain empirical content, the rational expectations assumption also takes the logically distinct step of equating subjective distributions to the data generating distribution. I shall use asset pricing theory to illustrate two justifications for taking that step, one based on an argument that agents with beliefs closer to the truth will eliminate others, another on empirical convenience.

\textit{Hansen and Singleton} (1983) and many others have generated restrictions on the covariation of consumption and time $t+1$ returns $R_{j,t+1}(x_{t+1})$ for asset $j$ by starting with consumer $i$’s Euler equation

$$1 = \beta \int_{x_{t+1}} u_i'(c_{i,t+1}(x^{t+1})) \frac{u_i'(c_{i,t}(x^t))}{u_i'(c_{i,t}(x^t))} R_{j,t+1}(x_{t+1})f_i(x_{t+1}^t|x^t)dx_{t+1}$$  \hspace{1cm} (3)

where $f_i(x_{t+1}|x^t)$ is consumer $i$’s subjective one-step-ahead transition density for a state vector $x_{t+1}$ that determines both returns and time $t+1$ consumption $c_{i,t+1}$ and $u_i'(c_{i,t+1}(x^{t+1}))$ is consumer $i$’s marginal utility of consumption.

In a finite-horizon setting, \textit{Harrison and Kreps} (1979) showed that, when there are eventually managed to escape expectations traps within subgame perfect or Markov perfect equilibria.
plete markets, the **stochastic discount factor**

\[
m_{t+1} = \beta \frac{u_i'(c_{i,t+1}(x^{t+1}))}{u_i'(c_{i,t}(x^t))} \frac{f_i(x_{t+1}|x^t)}{f(x_{t+1}|x^t)}
\]  

(4)

is unique. Here \(f(x_{t+1}|x^t)\) is a common physical conditional measure that does not depend on \(i\). Because offsetting differences in marginal utility functions and probabilities leave the left side of (4) fixed, the uniqueness of the stochastic discount factor makes room for different densities \(f_i\). Suppose that \(f\) is the measure that actually governs outcomes. Then Blume and Easley (2006) showed that in complete markets economies with Pareto optimal allocations and an *infinite* horizon, the measures \(f_i(x^\infty)\) merge for agents who survive in the limit, and they merge to the density that is closest to the truth \(f(\infty)\). Merging means that the measures agree about probabilities of tail events. If \(f_i(x^\infty) = f(x^\infty)\) for some agent, then for a complete markets economy with a Pareto optimal allocation, this survival result implies the rational expectations assumption provided that we also assume that at time 0 agents have access to an infinite history of observations.

Grossman and Shiller (1981), Hansen and Singleton (1983), Hansen and Richard (1987) wanted an econometric framework to apply when markets are incomplete, in which case it is not enough to appeal Blume and Easley's market survival justification for assuming beliefs that are common or eventually common. Hansen and Singleton (1983) and Hansen and Richard (1987) did not let that stop them. They simply imposed rational expectations directly and made enough stationarity assumptions to validate a Law of Large Numbers that gives GMM or maximum likelihood estimation good asymptotic properties. Under the rational expectations assumption, (3) imposes testable restrictions on the empirical joint distribution of returns and either individual or aggregate consumption.

---

8 Closest as measured by Kullback and Leibler's relative entropy.

9 In the context of a complete markets economy with a Lucas tree, Sandroni (2000) argued that a disagreement about tail events would present some consumers with arbitrage opportunities that cannot exist in equilibrium.
2.2 An empirical reason to allow belief heterogeneity

Many have followed Hansen and Singleton (1983) and Hansen and Richard (1987) by imposing rational expectations and specifying theories of the stochastic discount factor defined in terms of aggregate or individual consumption, for example, by letting \( u(c) \) be a constant relative risk aversion utility function \( u(c) = \frac{c^{1-\gamma}}{1-\gamma} \) and defining the stochastic discount factor as the intertemporal marginal rate of substitution

\[
m_{t+1} = \frac{\beta u'(c_{t+1})}{u'(c_t)}. \tag{5}\]

The aggregate consumption data have treated (3) and (5) badly under the rational expectations assumption that \( f = f_i \). One reaction has been to stick doggedly to rational expectations but to add backward-looking (see Campbell and Cochrane (1999)) or forward-looking (see Epstein and Zin (1989)) contributions to time \( t \) felicity. Another reaction has been to let disparate beliefs contribute to the stochastic discount factor. Hansen and Jagannathan (1991) opened the door to such an approach when they treated the stochastic discount factor \( m_{t+1} \) as an unknown nonnegative random variable and deduced what observed returns \( R_{j,t+1} \) and the restriction

\[
1 = \int_{x_{t+1}} m_{i,t+1}(x_{t+1})R_{j,t+1}(x_{t+1})f(x_{t+1}|x^t)dx_{t+1} \tag{6}
\]

imply about the first and second moments of admissible stochastic discount factors (with incomplete markets, there exist multiple stochastic discount factors). Their idea was that before specifying a particular theory about the utility function and beliefs that and that link \( m_i \) to real variables like consumption, it could be useful to characterize the mean and standard deviation that an empirically successful \( m_i \) must have. This approach leaves open the possibility that a successful theory of a stochastic discount factor will assign a nontrivial role to a probability ratio \( \frac{f(x_{t+1}|x^t)}{f(x_{t+1}|x^t)} \) even for a representative agent economy. The likelihood ratio creates a wedge relative to the Euler equation that has usually been fit in the ratio-
nal expectations macroeconomic tradition originating in Hansen and Singleton (1983) and Mehra and Prescott (1985). Likelihood ratio wedge approaches have been investigated by Bossaerts (2002, 2004) and Hansen (2007), and Hansen and Sargent (2006), among others. The art in work like that recommended by Hansen (2007) is to relax rational expectations enough to understand the data better while also retaining the econometric discipline that rational expectations models acquire by eliminating free parameters that characterize agents’ beliefs.

In sections 4, 5, and 6, I will describe stories and models that feature a divergence between an objective distribution and the subjective distribution of a government that solves an intelligent design problem. To set the stage, it is helpful to use some ideas from rational expectations econometrics.

3 Rational expectations econometrics

This section reviews econometric methods that allow an outsider to learn about a rational expectations equilibrium. Doing that allows me to introduce some objects and possibilities that will be in play when we move on in later sections about models that contain agents who are also learning about the equilibrium.

A rational expectations equilibrium is a joint probability distribution $f(x^t | \theta_o)$ over histories $x^t$ indexed by free parameters $\theta_o \in \Theta$ that describe preferences, technologies, endowments, and information. For reasons that will become clear in subsequent sections, I have called the parameter vector $\theta$ rather than $\rho$ as in section 2. Rational expectations econometrics tells an econometrician who is outside the model how to learn $\theta$. The econometrician knows only a parametric form for the model and therefore initially knows less about the equilibrium joint probability distribution than nature and the agents inside the model. The econometrician’s tools for learning the parameter vector $\theta$ are (1) a likelihood function, (2) a

$^{10}$Hansen (2007) bears only one new free parameter, a scalar $\theta$ that governs how much a representative agent’s beliefs are exponentially twisted vis-a-vis the data generating mechanism.
time series or panel of observations drawn from the equilibrium distribution, and (3) a Law of Large Numbers, a Central Limit Theorem, and some large deviations theorems that can be used to characterize the convergence, rate of convergence, and tail behavior of estimators. With enough data and a correct likelihood function, the econometrician can learn \( \theta_0 \).

Another name for a rational expectations equilibrium evaluated at a particular history is a likelihood function

\[
L(\theta|x^t) = f(x^t|\theta) = f(x_t|x^{t-1}; \theta)f(x_{t-1}|x^{t-2}; \theta) \cdots f(x_1|x_0; \theta)f(x_0|\theta).
\]  

(7)

The factorization on the right side displays the restrictions that a rational expectations model imposes on a possibly nonlinear vector autoregression \( f(x_t|x^{t-1}; \theta) \).

The most ambitious branch of rational expectations econometrics recommends maximizing a likelihood function or combining it with a Bayesian prior \( p(\theta) \) to construct a posterior \( p(\theta|x^t) \). In choosing \( \theta \) to maximize a likelihood function, a rational expectations econometrician in effect searches for a system of expectations that prompts the forward-looking artificial agents inside the model to make decisions that best fit the data. Taking logs in (7) gives

\[
\log L(\theta|x^t) = \ell(x^t|x^{t-1}; \theta) + \ell(x_{t-1}|x^{t-2}; \theta) + \cdots \ell(x_1|x_0; \theta) + \ell(x_0|\theta)
\]  

(8)

where \( \ell(x_t|x^{t-1}; \theta) = \log f(x_t|x^{t-1}; \theta) \). Define the score function as \( s_t(\theta) = \frac{\partial \ell(x_t|x^{t-1}; \theta)}{\partial \theta} \). In population, the first-order conditions for maximum likelihood estimation are the conditional moment conditions

\[
E[s_t|x^{t-1}] = 0,
\]  

(9)

which imply that the score is a martingale difference sequence, the starting point for a theory

\[\text{[11]}\text{For early applications of this empirical approach, see Sargent (1977), Sargent (1979), Hansen and Sargent (1980), Taylor (1980), and Dagli and Taylor (1984).}\]

\[\text{[12]}\text{As the econometrician searches over probability measures indexed by } \theta, \text{ he imputes to the agents inside the system of expectations implied by the } \theta \text{ under consideration.}\]
of statistical inference. By replacing the mathematical expectation $E$ in equation (9) with a sample average $T^{-1} \sum_{t=1}^{T}$, the econometrician finds a $\theta$ that allows him to approximate the equilibrium density very well as $T \to +\infty$.

The absence of free parameters that characterize decision makers’ beliefs underlies the cross-equation restrictions that identify parameters in rational expectations models.\footnote{See Imrohoroglu (1993) for a model that is an exception to the letter but not the spirit of the statement in the text. Cross-equation restrictions allow Imrohoroglu to use maximum likelihood estimation to pin down parameters including one that indexes a continuum of sunspot equilibria. Imrohoroglu usefully distinguishes econometric identification, which prevails, from uniqueness of equilibrium, which does not.} By re-orienting econometric attention away from parameterized decision rules to deeper parameters characterizing preferences, technology, endowments, and information, rational expectations econometrics provides an important part of a compelling response to the Lucas (1976) critique.

3.1 Using a misspecified model to estimate a better one

Lucas (1976) warned researchers not to use good-fitting non-structural models for policy analysis. But it can be wise to use the first-order conditions for estimating the parameters of a good fitting nonstructural model to make good inferences about parameters of a structural economic model.

Indirect inference assumes that a researcher wants to estimate a parameter vector $\rho$ of a structural rational expectations model for which (1) analytical difficulties prevent directly evaluating a likelihood function $f(x^t|\rho)$, and (2) computational methods allow simulating time series from $f(x^t|\rho)$ at given vector $\rho$. See Gourieroux et al. (1993), Smith (1993), and Gallant and Tauchen (1996). Indirect inference carries along two models, the model of interest with the untractable likelihood function, and an auxiliary model with a tractable likelihood function that fits the historical data well. The parameters of the economist’s model $\rho$ are interpretable in terms of preferences, technologies, and information sets, while the parameters $\theta$ of the auxiliary model $f(x^t|\theta)$ are data fitting devices. The idea of Gallant and Tauchen (1996) is first to estimate the auxiliary model by maximum likeli-
hood, then to use the score functions for the auxiliary model and the first-order conditions in equation (9) to define a criterion for a GMM estimator that can be used in conjunction with simulations of the economic model to estimate the parameters $\rho$. Thus, let the auxiliary model have a log likelihood function given by equation (8) and, for the data sample in hand, compute the maximum likelihood estimate $\hat{\theta}$. Then for a given artificial data set $\{x_t(\rho)\}$ from the economic model, evaluate the score function for the auxiliary model $s_t(x_t(\rho)|x^{t-1}(\rho), \hat{\theta})$ for each $t$. For different $\rho$’s, simulate paths $x_\tau(\rho)$ for $\tau = 1, \ldots, N$ from the economic model. Gallant and Tauchen estimate $\rho$ by setting the average score

$$\frac{1}{N} \sum_{\tau=1}^{N} s_\tau(x_\tau(\rho)|x^{\tau-1}(\rho), \hat{\theta})$$

(10)

as close to zero as possible when measured by a quadratic form of the type used in GMM in the sense that if the auxiliary model fits well, this method gives good estimates of the parameters $\rho$ of the economic model. If the economic model is true, the indirect estimator is as efficient as maximum likelihood when the economic and auxiliary models are observationally equivalent.

### 3.2 A troublesome possibility

This ideal case raises the following question: what happens if macroeconomic policy makers incorrectly uses what from nature’s point of view is actually an auxiliary model? Historical data can give the government no indication that it should abandon its model. Nevertheless, the government can make major policy design mistakes because its misunderstands the consequences of policies that it has not chosen. The possibility that the government uses what, unbeknownst to it, is just an auxiliary model, not a structural one, sets the stage for self-confirming equilibria.

---

14 This description fits their Case 2.
15 See [Lucas (1976)](#), [Sargent (1999, ch. 7)](#), and [Fudenberg and Levine (2007)](#).
4 Learning about an equilibrium

The learning literature constructs what Bray and Kreps (1987) call theories of learning about (as opposed to within) a rational expectations (or Nash) equilibrium. By saying about and not within, Bray and Kreps emphasize that the challenge is to analyze how a system of agents can come to learn an endogenous objective distribution by using adaptive algorithms that do not simply apply Bayes’ law to a correct probability model. We can’t appeal to the same econometrics that lets a rational expectations econometrician learn about an equilibrium because an econometrician is outside the model and his learning is a side-show that does not affect the data generating mechanism. It is different when the people learning about an equilibrium are inside the model. Their learning affects decisions and alters the distribution of endogenous variables over time, making other adaptive learners aim at moving targets. This feature of learning about an equilibrium makes it different from the rational expectations econometrician’s problem.

This section summarizes findings from a literature that studies systems of agents who use forward looking decision algorithms based on those temporary models that they update using recursive least squares algorithms (see Marcet and Sargent (1989a), Evans and Honkapohja (1999, 2001), Woodford (1990), and Fudenberg and Levine (1998)).

4.1 Self-confirming equilibrium

It is useful to begin by defining some population objects helpful for understanding the limiting behavior of an adaptive model.

\footnote{A difficult challenge in the machine learning literature is to construct an adaptive algorithm that learns dynamic programming. For a recent significant advance based on the application of the adjoint of a resolvent operator and a law of large numbers, see Meyn (2007, ch. 11).}

\footnote{Bray and Kreps’s ‘about’ versus ‘within’ tension also pertains to Bayesian theories of convergence to Nash equilibria. Marimon (1997) said that a Bayesian knows the truth from the beginning. Young (2004) pointed out that the absolute continuity assumption underlying the beautiful convergence result of Kalai and Lehrer (1993, 1994) requires that players have substantial prior knowledge of their opponents’ strategies. Young is skeptical that Kalai and Lehrer have answered the question “... can one identify priors [over opponents strategies] whose support is wide enough to capture the strategies that one’s (rational) opponents are actually using, without assuming away the uncertainty inherent in the situation?” Young (2004, p. 95)}

\footnote{Appendix A describes a related literature on learning in games.}
A true data generating process and an approximating model, respectively, are

\[ f(y^\infty, v^\infty|\rho) \land f(y^\infty, v^\infty|\theta). \]  \tag{11} 

A decision maker has preferences ordered by

\[ \int U(y^\infty, v^\infty) f(y^\infty, v^\infty|\theta) d(y^\infty, v^\infty) \]  \tag{12} 

and chooses a history-dependent plan

\[ v_t = h_t(y^t|\theta), \quad t \geq 0 \]  \tag{13} 

that maximizes (12). This gives rise to the sequence of decisions \( v(h|\theta)^\infty \). The difference between this problem and the canonical intelligent design problem in section 2 is the presence of the approximating model \( f(y^\infty, v^\infty|\theta) \) rather than the true model in (13). I call maximizing (12) a “Phelps problem” in honor of a particular version of a government control problem of this type that was solved by Phelps (1967) and that will play an important role in the empirical work to be discussed in subsection 6.

**Definition 4.1.** A self-confirming equilibrium (SCE) is a parameter vector \( \theta_o \) for the approximating model that satisfies the data-matching conditions

\[ f(y^\infty, v(h(\theta_o)^\infty|\theta_o) = f(y^\infty, v(h(\theta_o)^\infty|\rho). \]  \tag{14} 

An SCE builds in (1) optimization of (12) given beliefs indexed by \( \theta_o \), and (2) a \( \theta = \theta_o \) that satisfies the data matching conditions (14). Data matching prevails for events that occur under the equilibrium policy \( v(h|\theta_o)^\infty \), but it is possible that

\[ f(y^\infty, v^\infty|\theta_o) \neq f(y^\infty, v^\infty|\rho) \]  \tag{15} 

14
for $v^\infty \neq v(h|\theta)^\infty$. In an SCE, the approximating model is observationally equivalent with the true model for events that occur under the policy implied by equilibrium decisions, but not necessarily under other policies.

4.2 Learning converges to an SCE

An SCE is a possible limit point of an adaptive system. Suppose that an adaptive learner begins with an initial estimate $\hat{\theta}_0$ at time 0 and uses a recursive least squares learning algorithm

$$
\hat{\theta}_{t+1} - \hat{\theta}_t = e_{\theta}(\hat{\theta}_t, \hat{R}_t, y^t, v^t, t)
$$

$$
\hat{R}_{t+1} - \hat{R}_t = e_{R}(\hat{\theta}_t, \hat{R}_t, y^t, v^t, t).
$$

As in the models of learning in games of Foster and Young (2003) and Young (2004, ch. 8), we assume that decision makers mistakenly regard their time $t$ model indexed by $\hat{\theta}_t$ as permanent and form the sequence of decisions

$$
\hat{v}(h)_t = h_t(y^t|\hat{\theta}_t)
$$

where $h_t(y^t|\theta)$ is the same function that solves the original Phelps problem under the model $f(y^\infty, v^\infty|\theta)$. Under this scheme for making decisions, the joint density of $(y^\infty, v^\infty, \hat{\theta}^\infty)$ is

$$
f(y^\infty, \hat{v}(h)^\infty, \hat{\theta}^\infty|\rho).
$$

\footnote{Cho and Kasa (2006) create a model structure closer to the vision of Foster and Young (2003). In particular, Cho and Kasa’s model has the following structure: (1) one or more decision makers take actions at time $t$ by solving a dynamic programming problem based on a possibly misspecified time $t$ model, (2) the actions of some of those decision makers influence the data-generating process; (3) the decision maker shows that he is aware of the possible misspecification of his model by trying to detect misspecifications with an econometric specification test, (4) if the specification test rejects the model, the decision maker selects an improved model, while (5) if the current model is not rejected, the decision maker formulates policy using the model under the assumption (used to formulate the dynamic programming problem) that he will retain this model forever. Cho and Kasa define useful mathematical senses in which the same stochastic approximation and large deviations results that pertain to a least-squares learning setup also describe the outcomes of their model-validation setup.}
The learning literature states restrictions on the estimator $e$ and the densities $f(\cdot|\theta)$ and $f(\cdot|\rho)$ that imply that

$$\hat{\theta}_t \to \theta_o,$$

where convergence can be either almost surely or in distribution, depending on details of the estimator $e$ in (16).

### 4.3 REE or SCE?

Sometimes researchers have specified an approximating model to equal a true one, meaning that there exists a value $\theta_o$ for which $f(y^\infty, v^\infty|\rho) = f(y^\infty, v^\infty|\theta_o)$ for all plans $v^\infty$, not just equilibrium ones. This specification prevails in adaptive models in which least squares learning schemes converge to rational expectations equilibria, like those used by Woodford (1990) and Marcet and Sargent (1989b). When $f(y^\infty, v^\infty|\rho) \neq f(y^\infty, v^\infty|\theta_o)$ for some choices of $v$, the most that can be hoped for is convergence to an SCE.

### 4.4 SCE-REE gaps and policy design

Why is a gap between a rational expectations equilibrium and a self-confirming equilibrium important for a macroeconomist? Macroeconomists build models with many small agents and a small number (often one) of large decision makers called governments. Small private agents can take aggregate laws of motion as given within a recursive competitive equilibrium.

---

20 For example, so-called ‘constant gain’ algorithms give rise to convergence in distribution, while estimators whose gains diminish at the proper rates converge almost surely. See Williams (2004). A few papers have studied rates of convergence. There are examples in which convergence occurs at a $\sqrt{T}$ rate, but also examples where convergence occurs markedly more slowly.

21 Sargent (1999, ch. 6) works with a weaker notion of an SCE that Branch and Evans (2005, 2006) call a misspecification equilibrium. Branch and Evans construct misspecification equilibria in which agents $i$ and $j$ have different models parameterized, say, by $\theta_i$ and $\theta_j$, and in which $f(x^t|\theta_i) \neq f(x^t|\theta_j) \neq f(x^t|\rho)$, where again $\rho$ parameterizes the data generating mechanism. A misspecification equilibrium imposes moment conditions on agents’ approximating models that imply parameters $\theta_i$ that give equal minimum mean square error forecast errors $E_{\theta_i}(x_{t+1} - E_{\theta_i}(x_{t+1}|x^t))(x_{t+1} - E_{\theta_i}(x_{t+1}|x^t))^2$ for all surviving models. Branch and Evans use this setup to model equilibria in which beliefs and forecasts are heterogeneous across agents, though they have equal mean squared errors. They provide conditions under which recursive least squares learning algorithms converge to a subset of the possible misspecification equilibria. The models of Brock and Hommes (1997) and Brock and de Fontnouvelle (2000) are early versions of misspecification equilibria.
It is sufficient for them that their views are correct along the equilibrium path. If a small agent has access to a long enough history of observations drawn from a self-confirming equilibrium, he can form unimprovable forecasts by simply taking appropriate (conditional) averages of past outcomes. It doesn’t matter to a small agent that his views may be incorrect views off the equilibrium path.

But it can matter very much when a government, a large agent, has incorrect views off the equilibrium path because in designing its policy, we suppose that a government solves a Ramsey problem in which it contemplates the consequences of off-equilibrium path experiments. Wrong views about off-equilibrium path events shape government policy and the equilibrium path. Self-confirming equilibria leave ample room for mistaken policies, unintended consequences, disagreements about macroeconomic theories and the value of macroeconomic experimentation.

4.5 Uses of adaptive learning models in macroeconomics

One important use of adaptive models in macroeconomics has been to select among equilibria in models having several rational expectations equilibria (see Evans and Honkapohja (2001) for many examples). Another has been to choose among alternative possible representations of policy rules from Ramsey problems, a subset of which are stable under adaptive learning (see Evans and Honkapohja (2003)). Another use has been to suggest empirical specifications for improving the fit of models of asset pricing and improve the fit of rational expectations models by modelling the evolution of a gap between the objective density and investors' densities (e.g., Adam et al. (2006)). In the remainder of this paper, I confine myself to illustrating yet another application, namely, situations in which a government solves an intelligent design problem by using a misspecified model. As my laboratory, I focus on more or less formal descriptions of the process that has taught us the ideas in the heads of the central bankers mentioned in section 1. Section 5 is narrative account of a millenium of monetary policy made with misconceived models that experience revised, purposeful and
inadvertent experiments, unintended consequences, and eventual discoveries of ideas that we take now take for granted. I see self-confirming equilibria at work everywhere in this story, for example, in the way a commodity money standard worked to hide the quantity theory from empiricists by suppressing the variation in price levels and money supplies needed to identify it. That story takes us to the threshold of the 20th century experiment that sought to implement a well-managed fiat currency. Then section 6 fast-forwards to the 1960s and 1970s and summarizes the outcomes of efforts to use adaptive learning models to model how the U.S. monetary authorities struggled to understand inflation-unemployment dynamics as they sought to meet their dual mandate of giving the U.S. high output growth and low inflation.

5 Learning monetary policy over a millenium

5.1 From commodity to token to fiat money

Appendix B describes a 700 year process of theorizing and experimenting that transformed the prevailing commodity money system from one with many nominal anchors – mint-point, melt-point pairs for full bodied coins of all denominations – to a system that retained gold points for only one standard full bodied coin and used government-issued convertible token coins and notes for other denominations. After another 100 years, governments abolished the gold points for the standard coin too, leaving the nominal anchor to be the monetary authorities’ good intentions and their knowledge of the quantity theory of money. The appendix notes how a commodity money system concealed the quantity theory of money because the purpose of the gold and silver points was to make the price level a low variance, small trend exogenous variable and the money supply into a low variance, small trend endogenous variable. Eventually, some policy mistakes generated data that revealed the quantity theory to empiricists, and that set the stage for monetary experts like Keynes and Fisher to advocate

\[\text{Fetter (1978, p. 16) and Friedman (1991, pp. 150-151) discuss how concerns about small denomination coins shaped the gold standard.}\]
a well-managed fiat system.

5.2 Two threats to a well managed fiat money system

Friedman (1991, pp. 249-252) noted how our present fiat money system is historically unprecedented and repeated the warning of Fisher (1926, p.131) that “Irredeemable paper money has almost invariably proved a curse to the country employing it.” Two obstacles obstruct the path to a well managed fiat currency: (1) political pressures to use fiat money to finance the government expenditures, and (2) temptations to exploit a Phillips curve (Friedman (1991, p. 207)). Empirical learning models have been used to interpret monetary authorities’ struggles to understand and avoid these obstacles. Marcet and Nicolini (2003) and Sargent et al. (2006a) have constructed adaptive models that focus on (1) by featuring private agents learning. The models in those papers both select among rational equilibria and modify their outcomes enough to fit data from big inflations in Latin America. In the remainder of this paper, I choose to focus on statistical models that feature monetary authorities’ struggles with Friedman’s obstacle (2).

6 Learning inflation-unemployment dynamics

The next three subsections describe three stories about how the U.S. monetary authorities learned to understand and influence inflation-unemployment dynamics after World War II. These stories accept that a monetary authority can control inflation if it wants. Then why did the U.S. monetary authority allow inflation to rise in the late 1960s and 1970s, and why did it choose to bring inflation down in the 1980s and 1990s? If we assume that the monetary authority’s purposes did not change, and that it always disliked inflation and unemployment, then it is natural to focus on changes over time in the monetary authority’s understanding of inflation-unemployment dynamics. I’ll describe stories associated with three empirical models that highlight either temporary or permanent discrepancies between a government’s
model and a true data generating mechanism, and a government that each period solves a misspecified intelligent design problem and revises its parameter estimates to align them with new data.  

It is natural to impute popular contemporary models to the government. The ‘revisionist history’ of the U.S. Phillips curve by King and Watson (1994) provides a good source for these. King and Watson studied how econometric directions of fit (i.e., should you regress inflation on unemployment or unemployment on inflation?) affects government decisions. To make contact with studies from the 1970s, King and Watson call inflation on unemployment the Keynesian direction and unemployment on inflation the classical direction.

6.1 The (temporary) conquest of U.S. inflation

This story is about generating sufficient variation in the data to allow a government’s misspecified model to detect that there is no exploitable trade-off between inflation and unemployment. The story is told in terms of a model in which the only way a government’s model can discover that there is no exploitable tradeoff is for it falsely to infer that there is no trade-off whatsoever. That dooms any stabilization of inflation to be at best temporary.

This story uses specifications \( f(y^\infty, v^\infty|\rho) \neq f(y^\infty, v^\infty|\theta) \) to capture how a monetary authority misrepresents how its decisions impact on private agents’ expectations about inflation and, therefore, on the joint distribution of unemployment and inflation. Outcomes converge to a self-confirming equilibrium that makes inflation be higher than it would be if the government understood how its decisions affect private sector expectations. A run of random shocks can lead to a temporary steepening of the Phillips curve that activates a process in which the government chooses inflation rates that approach the Ramsey outcome.
under the true model. A conquest of inflation associated with such an escape from an SCE is temporary because mean dynamics will eventually push inflation back towards the SCE. If this is a good parable for the Volcker-Greenspan stabilization, we should be vigilant.

I illustrate the forces at work with the following simplified version of the type of model that Sims (1988), Cho et al. (2002), and Sargent and Williams (2005) have studied and that Chung (1990), Sargent (1999), and Sargent et al. (2006b) have fit to U.S. data. The true model is

\[
U = \rho_0 - \rho_1 \sigma_2 w_2 + \sigma_1 w_1, \\
\pi = v + \sigma_2 w_2
\]

(20) \hspace{1cm} (21)

where \(U\) is the unemployment rate, \(\pi\) is the rate of inflation, \(v\) is the systematic part of the inflation rate chosen by the monetary authority, \(w\) is a \(2 \times 1\) Gaussian random vector with mean zero and identity covariance, and \(\rho_0 > 0, \rho_1 > 0\), where \(\rho_0\) is the natural rate of unemployment and \(\rho_1\) is the slope of the Phillips curve. Through equation (20), which is the aggregate supply curve proposed by Lucas (1973), the model captures a rational expectations version of the natural rate hypothesis that asserts that the systematic component of inflation \(v\) does not affect the distribution of the unemployment rate conditional on \(v\). For an objective function \(-E(U^2 + \pi^2)\), the government’s best policy under the true model is \(v = 0\).

The government’s approximating model denies the natural rate hypothesis by asserting that \(v\) affects the probability distribution of \(U\) according to:

\[
U = \theta_0 + \theta_1 (v + \sigma_2 w_2) + \sigma_1 \tilde{w}_1, \\
\pi = v + \sigma_2 \tilde{w}_2
\]

(22) \hspace{1cm} (23)

where the random vector \(\tilde{w}\) has the same distribution as \(w\). Under the approximating model
The government’s best policy is

\[ v = h(\theta) = \frac{-\theta_1 \theta_0}{1 + \theta_1^2}. \]  

There exists a self-confirming equilibrium in which

\[ (\theta_0)_o = \rho_0 - \rho_1 h(\theta_o) \]  
\[ (\theta_1)_o = -\rho_1. \]  

The self-confirming equilibrium equals the time-consistent equilibrium of Kydland and Prescott (1977). An adaptive government’s estimates \( \hat{\theta}_t \) converge to the self-confirming equilibrium vector \( \theta_o \), and the systematic part of inflation converges to \( v = h(\theta_o) \).

The data-matching restriction (25) pinpoints how the government mistakenly ignores the effect of its policy choice \( v \), which equals the public’s expected rate of inflation, on the level of the Phillips curve. If \( v \) were generated randomly with enough variance, then even though it fits the wrong model, the government would estimate a Phillips curve slope \( \theta_1 \) of approximately zero and would according to (24) set \( v \) approximately to its optimal value of 0 under the true model. But within an SCE, \( v \) doesn’t vary enough for the government to estimate a \( \theta_1 \) close enough to zero for that to happen. Furthermore, the outcome that \( \hat{\theta}_t \rightarrow \theta_o \) means that the variation of \( v_t \) that occurs along transient paths is insufficient to allow the government’s model to approximate the data in a way that tells it to implement what would be the optimal policy under the true model.

However, that is not the end of the story because the adaptive model’s endogenous stochastic dynamics occasionally make \( v \) vary enough for the government to discover a (too

---

\( ^{25} \) The same suboptimal outcome occurs, but for a different reason from Kydland and Prescott (1977)’s. Our model in effect allows the government to choose the public’s expectation of inflation, so here the source of the sub optimality of the government’s choice originates in its misunderstanding of the economic structure, not the timing protocol under which it operates. The timing protocol in the model is such that if the government had the correct model, it would attain Ramsey outcomes, in the language of Stokey (1989). An inferior timing protocol is the problem isolated by Kydland and Prescott (1977).
strong) version of the natural rate hypothesis, too strong because it mistakenly asserts that there is no tradeoff whatsoever between $\pi$ and $U$. The adaptive system is destined to experience recurrent episodes in which ‘a most likely unlikely’ sequence of $w$ lowers the unconditional correlation between $U$ and $\pi$, which in turn prompts the government’s estimates $\hat{\theta}_t$ to induce the government to push $v_t$ downward from its self-confirming value. This process generates data that weakens the unconditional correlation between inflation and unemployment and drives $v$ even lower. The ultimate destination of this ‘escape’ from a self-confirming equilibrium is that the government estimates that $\theta_1$ is 0, prompting it to set $v_t$ at the optimal value 0. These escapes are more likely when the government’s estimator (16) discounts past data more heavily, for example, by using a so-called constant gain algorithm. An escape is temporary because the mean dynamics that drive the system toward the SCE vector $\theta_o$ are bound to reassert themselves and push inflation back toward the suboptimal SCE value of $h(\theta_o)$.

6.1.1 Details

The first implementations of this type of model imputed constant gain algorithms to the government. Simulations of Sims (1988) generated sample paths that seemed promising for explaining a Volcker-like stabilization prompted by the government’s being able to learn a good enough version of the natural rate hypothesis. However, formal econometric attempts to implement the model by Chung (1990) and Sargent (1999) failed to fit the U.S. data well, mainly because the government’s adaptive algorithm catches on to the adverse shifts in Phillips curve so quickly in the early 1970s that it tells the Phelps problem to tell the government to stabilize inflation much earlier than actually occurred. Sargent et al. (2006b) replaced the constant gain algorithm used in the earlier models with the Bayesian updating procedure implied by a drifting coefficients model with a covariance matrix $V$ for the innovations in the drifts to the coefficients. When they estimated $V$ along with the parameters of nature’s model by maximum likelihood, they found that could reverse engineer a drift-
ing set of government beliefs that when put into the Phelps problem each period produces a sequence of first period Phelps policy recommendations that do a good job of matching the actual inflation data. The estimated $V$ makes the intercept in the Fed’s quite volatile and thus makes contact with the account of Arthur Burns’s Fed, which according to Hetzel (1998), attributed much of the inflation of the 1970s to special factors that are akin to adding dummy variables to regression that capture intercept drift. It should be noted that the maximum likelihood estimate of $V$ is large and conveys the image of a government that expects coefficients to drift so much that it is very open to discounting past data heavily. The model’s conjuring up a Fed that over fits its models to recent data is food for thought for Fed watchers. The synthesized government beliefs succeed in rationalizing inflation ex post as a response to these government beliefs, and the beliefs themselves do a good job of forecasting inflation, thus capturing what seems to have been a remarkably good record of inflation forecasting by the Fed (see Bernanke (2007)).

6.2 A Keynesian account

The previous story is about how the troublesome possibility raised in subsection 3.2 plays out. The model of Primiceri (2006) returns to a world in which that is off the table because $f(y^\infty,v^\infty|\rho) = f(y^\infty,v^\infty|\theta_o)$ for all $v^\infty$ and an SCE equals an REE. All of the action in Primiceri’s model comes from calibrating an initial $\hat{\theta}_0 \neq \theta_o$ that leads to a stochastic path that converges to an SCE presided over by Greenspan and that mimics the post WWII U.S. data.

Primiceri’s model has a time invariant true data generating model featuring (i) an expectations augmented Phillips curve, (ii) a Cagan (1956)-Friedman (1956) adaptive expectations scheme that describes how the public forms the expectations of inflation that appear

\[26\text{But relative to available alternatives, the imputed beliefs do a poor job of forecasting unemployment, a deficiency of the model that hints that the reverse-engineering exercise may be imputing unrealistic views about joint inflation-unemployment dynamics to the Phelps problem in order to rationalize observed inflation outcomes.}\]
in (i)\(^{27}\) and (iii) an aggregate demand equation that describes how the time \(t\) value of an uninterpreted government policy instrument \(v_t\) affects current and future gaps between the unemployment rate \(u_t\) and a natural rate of unemployment \(v_t^N\).\(^{28}\) The model neatly allows the government’s misperception of the natural rate to influence policy, as advocated by Orphanides (2002, 2003). It also allows other potentially important government misperceptions to influence policy. Primiceri shows that the lower is the sum of the weights on lagged inflation in the expectational Phillips curve, and therefore the less persistent is inflation under a passive government policy, the less counterinflationary is the policy that emerges from the Phelps problem. A lower estimated persistence of inflation indicates to the government that mean reverting inflation will evaporate soon enough on its own. Coefficients that measure the strength of the feedback from unemployment to inflation also influence how actively counterinflationary is the policy called for by the time \(t\) Phelps problem.

Much of the action comes from how Primiceri calibrates initial government beliefs by using data between 1948 and 1960.\(^{29}\) These calibrated beliefs feature a level of persistence of inflation in the Phillips curve that is much lower than what prevails in the estimated model’s self-confirming equilibrium. In addition to these initial conditions, Primiceri sets two constant gain parameters, a separate one for the natural rate, another for all other coefficients in the government’s beliefs. These calibrated objects, the data, and the parameters of the structural relations pin down the government’s beliefs. There are no additional free parameters describing the government’s beliefs. Primiceri uses maximum likelihood to estimate parameters appearing in the government’s objective function and the time-invariant

---

\(^{27}\) Primiceri assumes that a fraction of agents form expectations this way, while the remainder have rational expectations. Primiceri’s specification imposes that the sum of weights on lagged inflation equals unity. Lucas (1972a) and Sargent (1971) argued that, except in a special case, the sum of the weights on lagged inflation being one is not a valid characterization of the natural rate hypothesis. See King and Watson (1994) and Sargent (1999).

\(^{28}\) Feature (ii) of Primiceri’s model embraces a Keynesian spirit of assuming that the authority influences output directly through the aggregate demand function, then inflation indirectly through the expectations-augmented Phillips curve. Contrast this with the classical specification adopted by Sims (1988), Chung (1990), Sargent (1999), Cho et al. (2002), and Sargent et al. (2006b).

\(^{29}\) Primiceri calibrates initial beliefs for the government about the value of the natural rate of unemployment and the coefficients in a reduced-form Phillips curve. The reduced form is derived by substituting the adaptive expectations scheme into the expectations augmented Phillips curve.
Figure 1: Evolution of policy-maker’s beliefs about: (a) the natural rate of unemployment; (b) the persistence of inflation in the Phillips curve; and (c) the slope of the Phillips curve. (Primiceri 2006, p. 882)
structural equations.

Primiceri accounts for the acceleration of inflation in the 1960s and 1970s, then the fall in the 1980s in terms of the government’s initial underestimates of the natural rate hypothesis as well as a temporal pattern of underestimates of the persistence of inflation and overestimates of the costs of disinflation coming from its estimated inflation-unemployment tradeoff. Figure [1] reproduces Primiceri’s figure II, which shows his estimates of the evolution of the Fed’s estimates of the natural rate of unemployment, the persistence inflation, and the slope of the Phillips curve. The Phelps problem attributes the acceleration of inflation to the monetary authority’s initial underestimates of both the natural rate and the persistence of inflation. After inflation had risen, the Phelps problem attributes the monetary authority’s reluctance to deflate to its overestimation of the costs of disinflation as captured by the slope of the Phillips curve. We will return to this point in subsection 6.3, where we link it to the conceptual issues about direction of fit raised by King and Watson (1994).

Under-estimates of the natural rate and over-estimates of the sacrifice ratio are connected. When the Fed under-estimates the natural rate and over-estimates the unemployment gap, it over-predicts the amount of disinflation. That causes it to revise its estimate of the slope of the Phillips curve towards zero. Thus, Orphanides’s story about the consequences of misestimating the natural rate of unemployment complements Primiceri’s story about sacrifice ratio pessimism.

6.3 An eclectic account

The models in the previous two sections take stands on what both the true and the government’s approximating models are. Cogley and Sargent (2005) performed an exercise that did not require them to specify a true data generating mechanism, it being enough for their purposes to consult the empirical distribution. But the government’s views about policy

Among many interesting features of Primiceri’s results are his estimate of \(k\), a parameter in the government objective function that allows Primiceri to evaluate the government’s temptation to deviate from the natural rate (he finds that the temptation is small) and the time series that he extracts for \(v_t\), which tracks a real interest rate very well after 1980.
choices not made play a key role. The government’s model $f(y^\infty, v^\infty|\theta)$ is a mixture of three submodels and $\hat{\theta}_t$ includes Bayesian posterior probabilities that the government uses to mix the three submodels.

A government entertains three models that Cogley and Sargent chose to capture specifications that had at one time or another received prominent support in the literature about U.S. unemployment-inflation dynamics described by King and Watson (1994). The models are (1) a Samuelson-Solow Phillips curve with King and Watson’s Keynesian direction of fit, a model that implies a long-run exploitable trade-off between inflation and unemployment, (2) a Solow-Tobin model with a Keynesian direction of fit that features a short-run but no long-run trade-off (albeit according to what Lucas (1972a) and Sargent (1971) claimed was a dodgy notion of long-run) between inflation and unemployment; and (3) a Lucas specification with a classical direction of fit that implies no exploitable trade-off between inflation and unemployment. If probability one is put on the Lucas model, the Phelps problem gives the trivial solution that the government should set the systematic part of inflation equal to zero. If probability one is put on either of the other models, inflation is a linear function of the state variables appearing in those exploitable dynamic Phillips curves. The government puts positive probability on all three models, so the Phelps problem brokers a compromise among the recommendations of the three models. But what kind of compromise? It depends on submodel probabilities times value functions.

The government starts with a prior with non-zero weights on all three models in 1960, estimates each sub model using Bayesian methods, and updates its prior over the three sub models. In each period, the government solves a Phelps problem that penalizes inflation and unemployment and that uses its time $t$ submodel probabilities to average over its time $t$ estimates of its three submodels. Cogley and Sargent put prior probabilities in 1960 of .98 on the Samuelson-Solow model and .01 each on the Solow-Tobin and the Lucas model. We put those low prior probabilities on the Lucas and Solow-Tobin models because only the Samuelson-Solow model existed in 1960. Putting U.S. inflation-unemployment data into
this machine, Cogley and Sargent computed time series of (1) the posterior model weights $\alpha_{t,t}$, and (2) the systematic part of the inflation rate set by the government in the Phelps problem.

Figures 2 and 3 taken from Cogley and Sargent (2005) frame the following puzzles. By the early 1970s, the data had moved the government’s prior to put probability approaching 1 on the Lucas model (see figure 2). Since the Lucas model recommends zero inflation, why nevertheless was actual inflation so high and variable in the 1970s? And why was the systematic part of inflation that emerges from the Phelps problem (see figure 3) even higher and more variable? Why does the Phelps planner seem to disregard the recommendations of the Lucas model and crank out high target inflation throughout the 1970s?

The answer is to be found in what the Samuelson-Solow and Solow-Tobin models say would happen if the Lucas zero-target-inflation policy were to be adopted, as summarized in figure 4. The Phelps problem weights the submodel posterior probabilities against losses associated with various off-taken-path recommendations. In the early 1970s, the coefficients in those submodels, with their Keynesian direction of fit, moved in ways that pointed to very
Figure 3: CPI inflation and recommendation from Phelps problem.

Figure 4: Loss from SS model (*) and ST model (o). When value is 1, it denotes infinite loss under Lucas zero inflation policy. When value is 0, it denotes finite loss.
high sacrifice ratios. Despite their low posterior probabilities, those models implied very high expected discounted losses if the Lucas policy recommendation were to be implemented immediately. In contrast, the high-probability Lucas model implied less adverse consequences if the recommendations of the Samuelson-Solow or Solow-Tobin models were allowed to prevail. So the Cogley and Sargent story is that the Lucas models policy recommendation did not prevail in the 1970s because there was a low probability that it would be disastrous. In order for a low-inflation recommendation to emerge from the Phelps problem, it was necessary that the estimated coefficients in the Samuelson-Solow and Solow-Tobin models adjust in ways that would render less adverse the consequences of a low-inflation policy. The data indicate that happened by the mid 1980s.

The direction-of-fit issue discussed by King and Watson (1994) is important for understanding how some of Primiceri’s results relate to Cogley and Sargent’s. Both models emphasize how monetary policy changed as the authorities updated their estimates, and Primiceri also attributes the inflation of the 1970s to the high perceived sacrifice ratio that Keynesian Phillips curve models presented to policy makers. But Primiceri assumes that the Fed relied exclusively on a version of the Solow-Tobin model and is silent about why the Fed disregarded the recommendations of the Lucas model. The central element of his story – the high perceived cost of disinflation or sacrifice ratio – is not a robust prediction across the three submodels used by Cogley and Sargent because it depends critically on the direction of fit, as documented by Cogley and Sargent (2005, p. 546-547). The reason that the sacrifice ratios differ so much across submodels comes from how the submodels interpret the diminished, near-zero contemporaneous covariance between inflation and unemployment.

31 The data also indicate that Bayes’ law sponsors comebacks for the Samuelson-Solow and Solow-Tobin models in the 1980s and 1990s. One reaction that a true believer in the Lucas model might have is that Bayes’ law is just too forgiving in still putting positive probability on those other models after the early 1970s data had come in, and that the inflation problem of the 1970s would have been solved by driving a stake through those other models. But no one has the authority to drive stakes, and models with operating characteristics much like those two survive today. The dispute between the fallacious (according to Friedman and Schwartz (1963, p. 191)) real bills doctrine and the quantity theory of money is mottled with repeated episodes having one of these doctrines being disposed of in favor of the other, then the other making a comeback. The real bills doctrine rides high in times like these when the Fed pegs a short term interest rate.
that had emerged by the mid 1970s. In a Keynesian Phillips curve, this diminished covariance flattens the short-term tradeoff, making the authorities believe that a long spell of high unemployment would be needed to bring inflation down, prompting Keynesian modelers to be less inclined to disinflate. But for a classical Phillips curve, the shift toward a zero covariance steepens the short-term tradeoff, making the authorities believe that inflation could be reduced at less cost in terms of higher unemployment. Thus, a classically-oriented policy maker would have been more inclined to disinflate.

7 Concluding remarks

Sawhill (1995) began her paper by citing Keynes (1936) in order to disagree with him and lament that politics not good economics drives decision making in Washington. For better or worse, in monetary economics today, the relationship between Washington policy makers and economists is so much easier than it was for Sawhill that I think it to acceptable to summarize the case for adaptive model by quoting one “madman in authority”:

The traditional rational-expectations model of inflation and inflation expectations has been a useful workhorse for thinking about issues of credibility and institutional design, but, to my mind, it is less helpful for thinking about economies in which (1) the structure of the economy is constantly evolving in ways that are imperfectly understood by both the public and policymakers and (2) the policymakers’ objective function is not fully known by private agents. In particular, together with the assumption that the central bank’s objective function is fixed and known to the public, the traditional rational-expectations approach implies that the public has firm knowledge of the long-run equilibrium inflation rate; consequently, their long-run inflation expectations do not vary over time in response to new information.

Although variations in the extent to which inflation expectations are anchored are
not easily handled in a traditional rational-expectations framework, they seem to fit quite naturally into the burgeoning literature on learning in macroeconomics. The premise of this literature is that people do not have full information about the economy or about the objectives of the central bank, but they instead must make statistical inferences about the unknown parameters governing the evolution of the economy. Bernanke (2007)

It is easy to agree with Sims (1980) that leaving the rational expectations equilibrium concept sends us into the “wilderness” because there is such a bewildering variety of ways to put discrepancies between objective and subjective distributions. Vis-a-vis some of the models that my friends in behavioral economics have taught me, the adaptive models described in this paper are very timid departures from rational expectations theory and from rational expectations econometrics. They can be viewed as statistical perturbations of rational expectations models. The timidity of our departure from rational expectations reflects our desire to retain as much of the discipline of rational expectations econometrics as possible. I have chosen to focus this paper on what can happens when a government solves an intelligent design problem while using a misspecified model. I view the very simple statistical models in section 6 as a parable for the situation that we are always in, that our probability models are misspecified. By stressing the possibility that learning has propelled us to a self-confirming equilibrium in which the government chooses an optimal policy based on a wrong model, the learning literature changes how we should think about promoting the novel policies that will allow misguided governments to break out of the lack-of-experimentation traps to which self-confirming equilibria confine them.

It has puzzled me that while Milton Friedman was willing to use expected utility in his work with Savage about why people gamble and buy insurance, he abstained from using it to form recommendations about macroeconomic policy. He also abstained from using rational

---

32 There are an infinite number of ways to be wrong, but only one way to be correct.
33 This is the starting point of calibration in macroeconomics, i.e., the refusal to use maximum likelihood because the model builder views it as an approximation.
expectations models, though early applications of them to the consumption function and the Phillips curve strengthened some of the policy recommendations that he had made. I do not know why he made that modelling choice, but Friedman’s works on monetary history are about the evolution of monetary authorities’ models and how wrong models sometimes led to bad policies and how arguments and evidence sometimes improved models. Friedman told us that evolution may not yet have taught everything about running a monetary system:

\[ \ldots \text{the world is now engaged in a great experiment to see whether it can fashion a different anchor, one that depends on government restraint rather than on the cost of acquiring a physical commodity} \ldots \text{The verdict is far from in on whether fiat money will involve a lower cost than commodity money} \ldots \text{Friedman (1991, p. 42).} \]

Nonetheless, it remains an open question whether the temptation to use fiat money as a source of revenue will lead to a situation that will ultimately force a return to a commodity standard \ldots The final answer will come only as history unfolds over the next decades. What that answer will be depends critically on our success in learning from historical episodes such as those that have been examined in this book. Such a learning process has been under way for centuries, ever since the first appearance of systematic analyses of money and monetary institutions. It has entered a new and urgent stage as the world ventures into hitherto unexplored terrain. \text{Friedman (1991, pp. 259-260).}
Appendixes

A Learning in games

In a game, a Nash equilibrium is the natural counterpart of a rational expectations equilibrium or a recursive competitive equilibrium. An extensive literature studies whether a system of adaptive players converges to a Nash equilibrium. A range of plausible adaptive algorithms have been proposed that are differentiated by how much foresight and theorizing they attribute to the players.\footnote{For a critical survey of this literature, see Young (2004).} At one extreme are adaptive models that have naive players who ignore strategic interactions and either play against histograms of their opponents past actions (this is called fictitious play) or alter their moves in directions that \textit{ex post} reduce their \textit{regret} at not having taken other actions in the past, given their opponents’ histories of actions. At the other extreme are models in which players construct statistical theories about their opponents’ behavior, use them for a while to make forward-looking decisions, occasionally subject their theories to hypothesis tests, discard rejected ones and choose new specifications.

This literature has sought plausible and robust algorithms that converge to a Nash equilibrium. Hart and Mas-Colell\footnote{Experimental economics has supplied data sets designed to check ideas from the literature on adaptive learning in games. It is remarkable that laboratory experiments using macroeconomics are rarer than those using microeconomics. See Duffy (2006) for an account of the existing experiments. I suspect that the main reason for fewer experiments in macro than in micro is that the choices confronting artificial agents within even one of the simpler recursive competitive equilibria used in macroeconomics are very complicated relative to the settings that experimentalists usually confront their subjects with.} tell us that this is a tall order:

It is notoriously difficult to formulate sensible adaptive dynamics that guarantee convergence to Nash equilibrium. In fact, short of variants of exhaustive search (deterministic or stochastic), there are no general results. Hart and Mas-Colell (2003, p. 1830) show that the source of the difficulty is that most adaptive schemes specify that adjustments in a player’s strategy do not depend on the payoff functions of other players, an uncoupling of the dynamics that in general dooms the system not to converge to a Nash equilibrium. Many examples of the adaptive schemes in the literature are uncoupled. Because many game theorists find uncoupled schemes desirable, parts of the literature have lowered the bar by looking for convergence to something weaker than Nash equilibria, namely, correlated equilibria or coarse correlated equilibria. Hart and Mas-Colell (2003, p. 1834) make the telling remark that “It is thus interesting that Nash equilibrium, a notion that does not predicate coordinated behavior, cannot be guaranteed to be reached in an uncoupled way, while correlated equilibrium, a notion based on coordination, can.”

Hart and Mas-Colell (2000, 2001, 2003) study adaptive schemes that are backward looking. For example, some of the most interesting ones have a player construct counterfactual historical payoffs that he would have received had he played other strategies, then compute
a measure of regret, then adjust his future play in directions that would have minimized his regret. These schemes impute little or no theorizing and foresight to the players.

For my present purposes, one of the most interesting contributions comes from part of the literature that attributes more sophistication to players, in particular, the work of Foster and Young (2003), which is also summarized in Young (2004, ch. 8). Their model has the following components: (1) each player has a large set of potential models that describe his opponents’ strategies; (2) players use a random device to select a particular model; (3) after that model is selected, there is an ‘act and collect data’ period during which a player (incorrectly) assumes that he will believe his current model forever; during this period, each player chooses his actions via a smoothed best response to what his model tells him about opponents’ actions (e.g., a quantal response function); (4) after a data collection period, a player compares the empirical pattern of his opponents’ play with that predicted by his model. He performs an hypothesis test that compares the theoretical and empirical distributions. If he rejects his current model, he randomly draws a new model from his set of models, then returns to step 2. If he accepts the model, he returns to step 3, waits a random number of periods, and then begins another data collection period.

With suitable assumptions about the lengths of testing periods and the tolerances of the hypothesis tests, Foster and Young (2003) show that behaviors eventually emerge that are often close to Nash equilibria. Their notion of hypothesis tests is sufficiently broad to include many plausible procedures. Their convergence result seems to be an uncoupled multi-agent learning scheme that actually approaches Nash equilibria, not something weaker like the coarse correlated equilibrium that the entirely backward-looking schemes mentioned above can approach. They avoid the conundrum of Hart and Mas-Colell partly by weakening the notion of convergence.

I like the Foster and Young (2003) parable about conventional wisdom followed by hypothesis testing.

B From commodity to fiat money

My theme is the data and the prior theorizing that shaped David Ricardo’s idea:

The introduction of the precious metals for the purposes of money may with truth be considered as one of the most important steps towards the improvement of commerce, and the arts of civilised life; but it is no less true that, with the advancement of knowledge and science, we discover that it would be another improvement to banish them again from the employment to which, during a less enlightened period, they had been so advantageously applied. Ricardo (1816, p. 65)

A long and disorderly process with “much backing and filling and confusion about purpose and power” led to Ricardo’s idea. Keynes and others made that idea the foundation of

36For a distinct but related approach, see Jehiel (1995, 1998). The Foster and Young (2003) model seems to me to capture some of the flavor of the anticipated utility framework advocated by Kreps (1998). The classifier models in Marimon et al. (1990) have a similar flavor.

37I borrowed the words in quotes from Friedman and Schwartz (1963, p.193), who used them to describe the evolution of the beliefs and policies of the Federal Reserve.
their proposals for a well managed fiat currency.

### B.1 Learning to supplement a commodity currency with tokens

Redish (1990, 2000) and Sargent and Velde (2002) described how it took 800 years to understand and cope with two imperfections that marred an ideal self-regulating commodity money system in which coins of all denominations were meant to exchange at values proportional to silver (or gold) content. In that ideal system, a government instructed a mint to offer to sell coins of different denominations for silver at prices proportional to their weights in silver. The mint did not buy coins for silver, but citizens were free to melt silver coins to recover silver. If minting and melting were costless, this self-regulating system would automatically adjust the denomination structure of coins to suit coin holder’s preferences by letting them melting coins of a denomination they wanted less of, then taking the silver to the mint to buy coins of the denomination they wanted to acquire. In the ideal system, a silver melt point equaled a silver mint point, denomination by denomination.

In practice, two imperfections hampered this system: (1) it was costly to produce coins; and (2) coins depreciated through wear and tear and sweating and clipping. The first imperfection gave rise to nonempty intervals between melt and mint points for gold or silver coins of each denomination – an upper point that indicated a melting point for that coin and a lower one that prompted minting. The proportionate spreads between minting and melting points differed because as a fraction of the value of the coin, it is cheaper to produce a large denomination coin than a small denomination coin. Unless the government were to subsidize the mint for producing low denomination coins, the spread between minting and melting points would be proportionately wider for low denomination coins. The second imperfection allowed underweight coins to circulate alongside full weight coins.

A nonempty interval between melting and minting points allowed coins to circulate by tale (i.e., by what is written on the coin rather than by weight) at an exchange value that exceeded their value by weight. Indeed, as Adam Smith pointed out, in the presence of costs of producing coins, the money supply mechanism provided incentives for people to purchase new coins from the mint only when their value in exchange exceeded their value by weight by enough to cover the mint’s brassage and seigniorage fees (Smith 1789, Book I, ch. 5). Nonempty intervals with proportionately wider widths for lower denomination coins and a consequent exchange rate indeterminacy allowed the intervals to shift over time and eventually to become so misaligned that they recurrently provided incentives to melt small denomination coins. That created the recurring shortages of small coins documented by Cipolla (1956, 1982).

Cipolla (1956) described a temporary practical remedy for these shortages. To cure a shortage of small denomination coins, the authorities debased them, thereby shifting the mint-melt intervals for small denomination coins in a direction that motivated citizens to purchase new coins from the mint. Monetary authorities throughout Europe used this method

---

38 Sargent and Velde (2002, p. 95) cited Bernando Davanzati, who in 1588 wrote that “metal should be worth as much in bullion as in coin, and be able to change from metal to money and money to metal without loss, like an amphibious animal.”

39 This multi-interval commodity money system in which coins circulate by tale is taken for granted by Smith (1789, book I, ch. 5).
for hundreds of years. There were repeated debasements in small denomination silver coins and secular declines in rates of exchange of small denomination for large denomination coins.

Many experiments, some inadvertent, others purposeful, were performed, and numerous theoretical tracts were written and disputed before what Cipolla (1956) called the ‘standard formula’ for issuing token small denomination coins was put into practice in the mid 19th century. It solved the problem of misaligned mint-melt intervals for coins of different denominations by, first, having only one large denomination full weight coin that the mint sold for a precious metal, and, second, having the government issue difficult-to-counterfeit small denomination token coins that it promised to convert on demand into the large denomination coin. This required a technology for manufacturing coins that were difficult to counterfeit.

As examples of inadvertent experiments, token monies were occasionally issued inside besieged cities and sometimes they worked. A document that prefigured later arguments of John Law, Adam Smith, and David Ricardo sparked a purposeful experiment. It advised King Ferdinand II of Spain that he could issue token copper coins that Spanish residents would voluntarily accept from the government in exchange for full bodied silver coins. It described how this could be done in a noninflationary way and how it would provide a fiscal boon to the Spanish treasury. Three successive Spanish Kings tried this experiment, which had all of the ingredients of the 19th century standard formula except convertibility. For 25 years, the experiment worked well, yielding the government substantial revenues without inflation. But eventually excessive issues of copper coins caused inflation, in the aftermath of which the Spanish monetary authorities pursued a fascinating sequence of experiments. They by restamped copper coins and manipulated the unit of account in order to adjust the price level and/or raise revenues for the Spanish government.

In a commodity money system, the quantity theory is mostly concealed because it can operate only in the limited interval between the mint and melt points for the precious metal. When the Spanish broke through those restrictions, they gave the British statistician Sir William Petty the data that he used to discover a quantity theory of money (see Hull (1899)). Other episodes created more data to substantiate the quantity theory of money, for example, the construction and collapse of John Law’s system (see Velde (2007)) and the overissuing of French assignats after the sales of the church lands that had initially backed them were suspended after war broke out in 1792 (see Sargent and Velde (1995)). But while those episodes lent vivid empirical support to a quantity theory, they also brought evidence that government monetary authorities could not be trusted to administer a pure fiat standard in ways that stabilized prices.

In 1660, the master of the British mint, Henry Slingsby, added an element missing from the Spanish experiment, namely, convertibility of token coins, and recommended what in

---

40 This process of shuttling through experiments, reformulations of theories, and further experiments reminds me of the hypothesis-testing learning models of Foster and Young (2003) and Cho and Kasa (2006), but I might be imagining things.


43 I suspect that is why later advocates for replacing the gold standard with ‘more scientific’ systems of managed currencies including Adam Smith and Ricardo to Keynes purposefully omitted references to some of the historical experiments that generated the data that were sources for the quantity theory of money. For example, Smith (1789) did not cite John Law’s theoretical writings as among the sources for his monetary recommendations.
the 19th century became the standard formula. But perhaps because the inflation accompanying the Spanish and some other similar experiments had given token coins a bad name, the British government ignored Slingsby’s recommendations. Many experts, including Locke (1691), continued to insist that token coins of any denomination were dangerous and that a good faith commodity money system required that coins of all denominations be full bodied. For a long time, that sentiment convinced national governments not to issues tokens, but that did not stop other entities from creating them. In seventeenth and eighteenth century Britain, hundreds of private firms and municipalities issued small denomination tokens that formed a substantial part of the country’s coinage. Between 1816 and 1836, the British government implemented the standard formula by nationalizing a token coin industry that had long existed.

B.2 Ricardo’s proposal

It required 156 years to take the short logical step from Slingsby’s 1660 standard formula for issuing convertible token subsidiary coins to David Ricardo’s 1816 recommendation. Ricardo proposed that a country’s domestic money supply should ideally consist of paper notes that the government promises to exchange at a pegged price for gold bullion bars, but that no gold coins should actually be minted. A variant of Ricardo’s scheme in which a government promises to redeem domestic notes for gold, but only for foreign residents, came to be practiced around 1900. This arrangement, by which “a cheap local currency [is] artificially maintained at par with the international standard of value,” (Keynes 1913, p. 25) was called the “gold exchange standard.” Keynes described how by 1913 this system had come to prevail in India through a sequence of haphazard administrative decisions that eventually produced a coherent system that no one had planned but that Keynes applauded. (Keynes 1913, p. 25) predicted that Ricardo’s scheme would be an essential part of “the ideal currency system of the future.”

The standard formula eliminates the gold or silver points for all coins except one standard coin and uses the mint and melt points for that coin to regulate the total quantity of money, while it uses its promise freely to convert tokens into that standard coin to produce the correct denomination composition. It was one more step from the standard formula or Ricardo’s proposal to the idea of Fisher (1920), Keynes, and others that well intentioned government officials should administer a fiat currency in ways that stabilize the price level. Doing that would allow them to to remove the mint and melt points for the one standard coin too. Discovering the quantity theory of money was an essential step in learning the conditions under which a fiat money system could be managed to provide greater price level

45Speaking of how a change in Indians’ preferences for holding gold could cause world-wide inflation in prices:

The time may not be far distant when Europe, having perfected her mechanism of exchange on the basis of a gold standard, will find it possible to regulate her standard of value on a more rational and stable basis. It is not likely that we shall leave permanently the most intimate adjustments of our economic organism at the mercy of a lucky prospector, a new chemical process, or a change of ideas [preferences for holding gold] in Asia. (Keynes 1913, p. 71)
stability than could be achieved with a gold standard.

As Keynes wanted, in the twentieth century governments throughout the world carried out the historically unprecedented experiment of managing currencies completely cut off from gold backing (see [Friedman (1991, p. 245)]). Figure 5 documents that, at least until very recently, the monetary authorities in four hard-currency countries failed to live up to Keynes’s high expectations for them and to deliver the kind of price stability that their predecessors had attained when they were restrained by that barbarous relic. Figures 6 and 7 show price indexes for Istanbul and Argentina, places with softer currencies (compare the vertical scales).

### C A monetary policy rules literature

The adaptive models described in section 6 explain the rise and fall of post WWII U.S. inflation in terms of monetary policy rules that drifted over time in response to drifts over time in the monetary authorities’ models of the economy. All three models embed very crude descriptions of the monetary policy rules and completely sidestep interesting questions about monetary policy transmission mechanisms. It is appropriate to say a few words about a related literature that uses time series data to infer the structure of post WWII U.S. monetary policy rules and how they have changed over time. The bottom line is that this literature has mixed evidence about whether monetary policy rules shifted enough to validate
Figure 6: Indices of prices in Istanbul.

Figure 7: Price index for Argentina.
stories along the lines of our three adaptive models.\footnote{This mixed news partly reflects the theoretical property of time series models that it is statistically difficult to detect drifts or shifts in the systematic part of a vector autoregression and much easier to detect changes in volatilities.}

Bernanke and Mihov (1998) developed an SVAR methodology for measuring innovations in monetary policy and their macroeconomic effects. They compared alternative ways of measuring monetary policy shocks and derived a new measure of policy innovations based on possibly time-varying estimates of the Fed’s operating procedures. They presented a measure of the overall stance of policy (see Bernanke and Mihov (1998, Fig. III, p. 899)) that is striking in how the distribution of tight and loose policies seems not to have changed much in the periods before and after 1980.

But Clarida et al. (2000) estimated a forward-looking monetary policy reaction function for the postwar United States economy before and after Volcker’s appointment as Fed Chairman in 1979 and found substantial differences in the estimated rules across periods. They found that interest rate policy in the Volcker-Greenspan period has been much more sensitive to changes in expected inflation than in the pre-Volcker period. They then extracted implications of the estimated rules for the equilibrium properties of inflation and output in a new Keynesian DSGE model and found that the Volcker-Greenspan rule is stabilizing, but that the earlier rule was not. Lubik and Schorfheide (2004) estimated a new Keynesian model in which the equilibrium is undetermined if monetary policy is passive and constructed posterior weights for the determinacy and indeterminacy region of the parameter space as well as estimates for the propagation of fundamental and sunspot shocks. They found that U.S. monetary policy post-1982 was consistent with determinacy but that the pre-Volcker policy was not, and also that before 1979 indeterminacy substantially altered the propagation of shocks.

In contrast, working in terms of less fully interpreted models, Sims and Zha (2006) estimated a multivariate regime-switching model for monetary policy and found that the best fit allows time variation in disturbance variances only. When they permitted the systematic VAR coefficients to change, the best fit was with change only in the monetary policy rule. They estimated three regimes that correspond to periods across which the folk-wisdom states that monetary policy differed. But they found that those differences among regimes were not large enough to account for the rise and decline of inflation of the 1970s and 1980s. Likewise, by estimating a time-varying VAR with stochastic volatility, Primiceri (2005) found that both the systematic and non-systematic components of monetary policy had changed. In particular, he found that the systematic responses of the interest rate to inflation and unemployment exhibited a trend toward a more aggressive behavior, while also having sizeable high frequency oscillations. But Primiceri concluded that those had small effects on the rest of the economy and that exogenous non-policy shocks were more important than interest rate policy in explaining the high inflation and unemployment episodes described above, thus coming down more on the ‘bad luck’ than the ‘bad policies’ side of the argument. I hope that conclusion is too pessimistic because we have learned to do better.
References


