Dynamic Modelling and Control of National Economies 1989. Selected Papers from the 6th IFAC Symposium, Edinburgh, UK, 27-29 June 1989, N. Christodoulakis (ed.), Oxford: Pergamon Press, 1990, 293-98.

NON–CONVERGENCE TO RATIONAL EXPECTATIONS AND OPTIMAL MONETARY POLICY IN MODELS WITH LEARNING

Bruce Mizrach, Department of Economics, Boston College, Chestnut Hill, MA USA

Abstract: A principal argument in the rational expectations literature is the optimality of predictable policy. This paper illustrates that this claim does not hold in a world of parametric uncertainty for two reasons: (1) completely noiseless policy may lead to non–convergence to the true model parameters; (2) highly predictable policy is not very informative about the structure of the model. A series of examples illustrate the ramifications for macroeconomic policy.

Keywords: Learning systems; Bayes methods; stochastic control; dynamic programming;

INTRODUCTION

This paper examines the implications of learning for macroeconomic policy. In an environment in which the parameters of a structural model are unknown to economic agents, optimal policy differs strongly from the conventional wisdom. The leading argument in the rational expectations literature is policy ineffectiveness; that is, policy doesn't matter as long as its predictable. This translates into simple dictates for the monetary authority like k% growth rules.

Learning reverses the conventional wisdom in two ways. First, policy may be so predictable that it prevents economic agents from arriving at the true parameters of the structural model. Policy must continually vary for limiting beliefs to converge to rational expectations. Second, in dynamic models, optimal multi–period actions may differ greatly from the one–period maximizing policy. There may be incentives for the policy makers to use control variables in such a way as to ensure convergence of beliefs. Through a series of counter–examples in the spirit of Sargent and Wallace(1975), we present the dramatic implications of these conclusions for popular macroeconomic models.

We draw heavily on a wide–ranging literature. In studying the convergence of sequentially updated beliefs, we draw on the work of Taylor(1974), Jordan(1985) and Kiefer and Nyarko(1987). The problem of stochastic control has its origins in Blackwell's(1953) work on experimentation. He reformulated this question into the dynamic programming framework in a subsequent paper, Blackwell(1965). The finite parameter case we study here is closely related to the bandit problem. Rothschild(1974) first introduced this to the economic literature. McLennan(1984) presented an example of incomplete learning by a monopolist facing an unknown demand curve.

Prescott(1972) and Chow(1981) discuss applications of stochastic control to economic policy. Drawing on the recent work of Easley and Kiefer(1988) and McLennnan(1987) we extend this to a more general setting and provide an explicit characterization of the optimal policy.

Section II states sufficient conditions for the almost sure convergence of parameter estimates for a Bayes learner. Learning requires fluctuations in the independent variables, andwhen this is lacking, we have non–convergence to rational expectations.

We consider two examples in Section III in which policy is the critical factor in learning: the Cagan hyper–inflationary model of money demand and an expectations model of the term structure of interest rates. Both fixed rate growth rules and interest pegging are shown to be barriers to learning on the part of agents. The paper poses this paradox: while predictability in the policy environemnt is desirable in that it eliminates uncertainty over expectations of policy variables, in the absence of precise knowledge of the model's parameters, the agent cannot come to learn them over time in these examples.

We turn in Section IV to the issue of optimal policy in a stochastic control situation. We prove conditions that will give the monetary authority incentive to learn the model's paramters. As in the bandit problems, the policy authority trades off current reward for future gains. We offer two counter–examples in which the solution under certainty is very different from the case with parametric uncertainty. Poole's(1970) work on choice of monetary instruments in a stochastic IS/LM model is given a strikingly different interpretation. We show that the monetary authority may choose to use a money stock policy even when the interest pegging policy loss is initially higher. The second model is the Lucas–Sargent–Wallace model of aggregate supply. Here we show that if the monetary authority finds it optimal to learn the true parameters, the k% growth rule can never be optimal. We summarize the implications for rational expectations theory of these findings in a closing section.

NON–CONVERGENCE

The Environment

Let Ω be a complete and separable metric space, let F be its Borel field, and let P be the set of probability measures on F.[1] Consider a stochastic process $\{y_t\}^{\infty}_{t=1}$ defined on the

[1] A useful reference for this material is Billingsley(1979).

probability space (Ω,F,P) , whose distribution, conditional on a choice variable, $x_t \in X$, depends upon parameters, $\theta \in \Theta$, unknown to the economic agent. Let $p(y|x,\theta)$ be the density with respect to the measure P on Ω.

Assume the following: (A1) X, the action space, is a compact subset of R^k; (A2) $\Theta = \{\theta_0, \theta_1, ...\theta_m\}$ is a finite set; (A3) (i) The density function $p(.|)$ is jointly continuous; (ii) For all $(x,\theta) \in X \times \Theta$, $Y=\{y:p(y|x,\theta) > 0\}$; Prior beliefs are described by $\mu \in \Lambda(\Theta)$, where Λ is the set of probability measures on Θ. The agent is presumed to be a Bayes learner, meaning that he updates beliefs based upon observables according to $\Gamma: \Lambda(\Theta) \times X \times Y \rightarrow \Lambda(\Theta)$

$$\Gamma(\mu|\lambda,x,y) = \frac{\lambda(\theta)p(y|\theta,x)}{\Sigma_\Theta \lambda(\theta)p(y|\theta,x)} \quad (1)$$

Under our assumptions, McLennan(1987) proves continuity of the Bayes map Γ.

We will be working exclusively with the two parameter linear regression model

$$y_t = \alpha + \beta x_t + \varepsilon_t \quad (2)$$

which implies for (A1) k=1 and $\theta^*=(\alpha,\beta)$. We add the additional assumption: (A4) $\varepsilon_t \sim$ i.i.d. $(0,\sigma_\varepsilon^2 < \infty)$; This indicates that the disturbance term is a mean zero, independently distributed, random variable, with known, finite variance.

Conditions for Convergence of the Bayes Estimator

We state at the outset the main result. A formal proof may be found in Kiefer and Nyarko(1987). We offer an intuitive discussion below.

Proposition I: Consider the linear regression model (2). If $\{x_t\}^\infty_{t=1} \in R^\infty$ converges to some value $\bar{x}$, $\lim_{t\to\infty} x_t = \bar{x}$, then the limiting posterior distribution, μ_∞, will converge to a posterior process with support a subset of the set $\{(\alpha', \beta'); \alpha' + \beta'\bar{x} = \alpha + \beta\bar{x}]\}$, where α and β are the "true" parameter values. If x_t does not converge, then μ_t will converge to a point mass concentrated on the "true" parameter value.

Proof: Kiefer and Nyarko(1987).∎

The intuition[2] is best understood in terms of the of the sum of squared residuals. As noted in Zellner(1971), $\Sigma_{t=0}^\infty (x_t - E[x]))^2 = \infty$, is a sufficient condition for convergence of the posterior process to $\theta = (\alpha,\beta)$, the true parameter vector. In the least squares context, the condition is slightly stronger[3], but the intuition is that the if the sum of squared residuals goes to infinity, the Bayes estimator is strongly consistent.

For every realization of x_t, there will be an associated realization for y_t. The y's will be noisy though due to the disturbance term. For n realizations of a given x_t, say $\tilde{x}$, $\Sigma_{i=0}^n(\tilde{y}_i/n). = \alpha + \beta\tilde{x} + \Sigma_{i=0}^n(\varepsilon_i/n)$. From the strong law of large numbers on the disturbance term $\Sigma_{i=0}^n(\varepsilon_i/n) \rightarrow 0$ as $n\rightarrow\infty$, then if we define $\tilde{y}^* = \lim n\rightarrow\infty$ of $\Sigma_{i=0}^n(y_i/n) \rightarrow 0$, there exists some $\hat{\alpha}, \hat{\beta}$ satisfying $\tilde{y}^* = \hat{\alpha} + \hat{\beta}\tilde{x}$ by linearity. There are however, a continuum of values for $\hat{\alpha}$ and $\hat{\beta}$ that will satisfy this equation. This explains the first part of the proposition.

[2] We still rely heavily here on Kiefer and Nyarko(1987).

Consider another value for x, say x', such that $x' \neq \tilde{x}$. This determines a second equation with $\lim n\rightarrow\infty \Sigma_{i=0}^n(y'_i/n) = y'^*$. An alternating sequence of realizations for x between $\tilde{x}$ and x' is sufficient for the Zellner condition to be fulfilled. In our context, the two equations, one for $\tilde{y}^*$ and y'^*, are sufficient to define a unique $\hat{\alpha}^*$ and $\hat{\beta}^*$ satisfying the two equations. We can then be assured that $\hat{\alpha}^*$ and $\hat{\beta}^*$ are consistent for y.

In closing this section, we note that this equations and unknowns approach is utilized in a paper by Nyarko(1988) to extend some of these results to the multivariate case.

EXAMPLES OF NON–CONVERGENCE

The Cagan Hyper–Inflation Model

Cagan(1956) postulates the following equation for money demand in a hyperinflationary economy:

$$\frac{M^d_t}{P} = K e^{-\eta\pi^*_t} \quad (3)$$

where M^d is nominal money demand, P is the price level, π^* is the expected rate of inflation, and K and η are parameters. Define:

$$\log(\frac{M^d_t}{P}) = m^d_t - p_t = \log(K e^{-\eta\pi^*_t}) = k - \eta\pi^*_t \quad (4)$$

$$\pi^*_t = E_t[\log(P_{t+1}) - \log(P_t)] = E_t[p_{t+1}] - p_t \quad (5)$$

Cagan assumes that the price level clears the money market every period, $p_t = m_t - k + \eta\pi^*_t$. Solving recursively, as in Mussa(1975)

$$E_t[p_{t+1}] = \frac{1}{1+\eta}\sum_{j=0}^{\infty}(E_t(m_{t+1+j}) - k)(\frac{\eta}{1+\eta})^j \quad (6)$$

Let us assume that the monetary authority follows a constant rate growth rule $m_t = \mu m_{t-1}$, then

$$\sum_{j=0}^{\infty} E_t(m_{t+1+j}) = \sum_{j=0}^{\infty}\mu^j m_{t+j} \quad (7)$$

We can then re–write (6) as

$$\frac{1}{1+\eta}\sum_{j=0}^{\infty} m_{t+1}(\frac{\mu\eta}{1+\eta})^j - \frac{1}{1+\eta}\sum_{j=0}^{\infty} k(\frac{\eta}{1+\eta})^j \quad (8)$$

Assuming that the processes in (8) are convergent, $(\mu\eta/1+\eta)$, $(\eta/1+\eta) < 1$, we have

$$E_t[p_{t+1}] = \frac{m_{t+1}}{1+\eta}\left(\frac{1}{1-\frac{\mu\eta}{1+\eta}}\right) - \frac{k}{1+\eta}\left(\frac{1}{1-\frac{\eta}{1+\eta}}\right) \quad (9)$$

If we take μ, η, and k to be constants, then $E_t[p_{t+1+n}] - p_{t+n}$ for n = 1,2,... will also be a constant.

We can now use Proposition I to prove that a representative agent will not converge to consistent beliefs about the unknown parameters k and η, since $\pi^*_t = \bar{\pi} \ \forall$ t. An anticipated monetary policy is still called for to help agents

[3] A sufficient condition for consistency of the least squares estimator is $\lambda_s(\mathbf{X'X}) \rightarrow \infty$, where λ_s is the smallest characteristic root of $(\mathbf{X'X})$. For proof, see Amemiya(1985).

probability space (Ω, F, P) , whose distribution, conditional on a choice variable, $x_t \in X$, depends upon parameters, $\theta \in \Theta$, unknown to the economic agent. Let $p(y|x,\theta)$ be the density with respect to the measure P on Ω.

Assume the following: (A1) X, the action space, is a compact subset of R^k; (A2) $\Theta = \{\theta_0, \theta_1, \ldots \theta_m\}$ is a finite set; (A3) (i) The density function $p(.|)$ is jointly continuous; (ii) For all $(x,\theta) \in X \times \Theta$, $Y = \{y: p(y|x,\theta) > 0\}$; Prior beliefs are described by $\mu \in \Lambda(\Theta)$, where Λ is the set of probability measures on Θ. The agent is presumed to be a Bayes learner, meaning that he updates beliefs based upon observables according to $\Gamma: \Lambda(\Theta) \times X \times Y \rightarrow \Lambda(\Theta)$

$$\Gamma(\mu| \lambda,x,y) = \frac{\lambda(\theta)p(y|\theta,x)}{\Sigma_\Theta \lambda(\theta)p(y|\theta,x)} \tag{1}$$

Under our assumptions, McLennan(1987) proves continuity of the Bayes map Γ.

We will be working exclusively with the two parameter linear regression model

$$y_t = \alpha + \beta x_t + \varepsilon_t \tag{2}$$

which implies for (A1) k=1 and $\theta^* = (\alpha,\beta)$. We add the additional assumption: (A4) $\varepsilon_t \sim$ i.i.d. $(0, \sigma_\varepsilon^2 < \infty)$; This indicates that the disturbance term is a mean zero, independently distributed, random variable, with known, finite variance.

Conditions for Convergence of the Bayes Estimator

We state at the outset the main result. A formal proof may be found in Kiefer and Nyarko(1987). We offer an intuitive discussion below.

Proposition I: Consider the linear regression model (2). If $\{x_t\}^\infty_{t=1} \in R^\infty$ converges to some value $\bar{x}$, $\lim_{t\to\infty} x_t = \bar{x}$, then the limiting posterior distribution, μ_∞, will converge to a posterior process with support a subset of the set $\{(\alpha', \beta'); \alpha' + \beta'\bar{x} = \alpha + \beta\bar{x}]\}$, where α and β are the "true" parameter values. If x_t does not converge, then μ_t will converge to a point mass concentrated on the "true" parameter value.

Proof: Kiefer and Nyarko(1987).∎

The intuition[2] is best understood in terms of the of the sum of squared residuals. As noted in Zellner(1971), $\Sigma_{t=0}^\infty (x_t - E[x]))^2 = \infty$, is a sufficient condition for convergence of the posterior process to $\theta = (\alpha,\beta)$, the true parameter vector. In the least squares context, the condition is slightly stronger[3], but the intuition is that the if the sum of squared residuals goes to infinity, the Bayes estimator is strongly consistent.

For every realization of x_t, there will be an associated realization for y_t. The y's will be noisy though due to the disturbance term. For n realizations of a given x_t, say $\tilde{x}$, $\Sigma_{i=0}^n(\tilde{y}_i/n). = \alpha + \beta\tilde{x} + \Sigma_{i=0}^n(\varepsilon_i/n)$. From the strong law of large numbers on the disturbance term $\Sigma_{i=0}^n(\varepsilon_i/n) \rightarrow 0$ as $n \rightarrow \infty$, then if we define $\tilde{y}^* = \lim n \rightarrow \infty$ of $\Sigma_{i=0}^n(y_i/n) \rightarrow 0$, there exists some $\hat{\alpha}$, $\hat{\beta}$ satisfying $\tilde{y}^* = \hat{\alpha} + \hat{\beta}\tilde{x}$ by linearity. There are however, a continuum of values for $\hat{\alpha}$ and $\hat{\beta}$ that will satisfy this equation. This explains the first part of the proposition.

[2] We still rely heavily here on Kiefer and Nyarko(1987).

Consider another value for x, say x', such that $x' \neq \tilde{x}$. This determines a second equation with $\lim n \rightarrow \infty \ \Sigma_{i=0}^n(y'_i/n) = y'^*$. An alternating sequence of realizations for x between $\tilde{x}$ and x' is sufficient for the Zellner condition to be fulfilled. In our context, the two equations, one for $\tilde{y}^*$ and y'^*, are sufficient to define a unique $\hat{\alpha}^*$ and $\hat{\beta}^*$ satisfying the two equations. We can then be assured that $\hat{\alpha}^*$ and $\hat{\beta}^*$ are consistent for y.

In closing this section, we note that this equations and unknowns approach is utilized in a paper by Nyarko(1988) to extend some of these results to the multivariate case.

EXAMPLES OF NON-CONVERGENCE

The Cagan Hyper-Inflation Model

Cagan(1956) postulates the following equation for money demand in a hyperinflationary economy:

$$\frac{M^d_t}{P} = K\, e^{-\eta\pi^*_t} \tag{3}$$

where M^d is nominal money demand, P is the price level, π^* is the expected rate of inflation, and K and η are parameters. Define:

$$\log(\frac{M^d_t}{P}) = m^d_t - p_t = \log(K\, e^{-\eta\pi^*_t}) = k - \eta\pi^*_t \tag{4}$$

$$\pi^*_t = E_t[\log(P_{t+1}) - \log(P_t)] = E_t[p_{t+1}] - p_t \tag{5}$$

Cagan assumes that the price level clears the money market every period, $p_t = m_t - k + \eta\pi^*_t$. Solving recursively, as in Mussa(1975)

$$E_t[p_{t+1}] = \frac{1}{1+\eta}\sum_{j=0}^{\infty}(E_t(m_{t+1+j}) - k)(\frac{\eta}{1+\eta})^j \tag{6}$$

Let us assume that the monetary authority follows a constant rate growth rule $m_t = \mu m_{t-1}$, then

$$\sum_{j=0}^{\infty} E_t(m_{t+1+j}) = \sum_{j=0}^{\infty}\mu^j m_{t+j} \tag{7}$$

We can then re-write (6) as

$$\frac{1}{1+\eta}\sum_{j=0}^{\infty} m_{t+1}(\frac{\mu\eta}{1+\eta})^j - \frac{1}{1+\eta}\sum_{j=0}^{\infty} k(\frac{\eta}{1+\eta})^j \tag{8}$$

Assuming that the processes in (8) are convergent, $(\mu\eta/1+\eta)$, $(\eta/1+\eta) < 1$, we have

$$E_t[p_{t+1}] = \frac{m_{t+1}}{1+\eta}\left(\frac{1}{1-\frac{\mu\eta}{1+\eta}}\right) - \frac{k}{1+\eta}\left(\frac{1}{1-\frac{\eta}{1+\eta}}\right) \tag{9}$$

If we take μ, η, and k to be constants, then $E_t[p_{t+1+n}] - p_{t+n}$ for n = 1,2,... will also be a constant.

We can now use Proposition I to prove that a representative agent will not converge to consistent beliefs about the unknown parameters k and η, since $\pi^*_t = \bar{\pi}\ \forall$ t. An anticipated monetary policy is still called for to help agents

[3] A sufficient condition for consistency of the least squares estimator is $\lambda_s(\mathbf{X'X}) \rightarrow \infty$, where λ_s is the smallest characteristic root of $(\mathbf{X'X})$. For proof, see Amemiya(1985).

learn k and η, but without a policy that varies the monetary growth rate between periods, Proposition I assures us of non–convergence to rational expectations. Furthermore, a change in the money growth rate, after agent's beliefs have settled down in the regime $\mu_t = \bar{\mu}$, would lead to non–rational forecasts.

A Simple Expectations Model of the Term Structure

In this example, pegging the interest rate will be the policy barrier, preventing agents from learning the slope of the term structure equation. An approximation of the relation between long and short term interest rates, R_t and r_t is given by[4]

$$R_t = \theta + (1-\gamma) \sum_{k=0}^{\infty} E_t r_{t+k} + \varepsilon_t \tag{10}$$

where θ is interpreted as a constant term premium, $0 < \gamma < 1$ is a constant discount factor, and ε is a white noise disturbance term.. Let us assume that agents form their expectations of the short rate using a distributed lag on previous short rates:

$$E_t[r_t] = r^e_t = \sum_{k=0}^{T} \phi_k r_{t-k} \qquad \sum_{k=0}^{T} \phi_k = 1 \tag{11}$$

where the ϕ's are unkown to the econometrician. Suppose the monetary authority begins to peg the interest rate at $r_t = \bar{r}$. According to (11), it will take T periods, but eventually the forecast of all future short rates will be $\bar{r}$ as well. We will now show two results: (i) the discount rate will not be estimated consistently in a single pegging regime, (ii) it can only be consistently estimated with k–period pegging switches, where k corresponds to the length of the lag in (11).

The estimation problem here is very tricky without some help from the policy makers. The usual assumption made is that agents have rational expectations. In which case the econometricianestimates

$$R_{t+1} - R_t = \frac{-(1-\gamma)\theta}{\gamma} + \frac{1-\gamma}{\gamma}(R_t - r_t) + \frac{v_{t+1}}{\gamma} \tag{12}$$

where $v_{t+1} = (1-\gamma)\Sigma_{k=1}{}^{\infty}(\gamma^k(E_{t+1}r_{t+k} - E_t r_{t+k}))$ represents the revision in the agent's expectations. Consistent estimation of (12) requires the assumption that v_{t+1} be uncorrelated with the right hand side regressors. (11) violates this assumption (and our A4 as well). Only with the long rate constant can we ignore the ϕ's; successive k–period regime switches will then enable us to estimate (10) with realized short rates.

SOME ISSUES IN STOCHASTIC CONTROL

We shift our focus now slightly away from non–existence counterexamples to examples of otpimal policy in stochastic environments. The critical intuiton is the tradeoff presented in discounted control situations where the model parameters are unknown. While large variation in the control variable is undesirable from the point of view of the usual loss function, it is desirable in gaining knowledge over the system's parameters. For low discount rates, the policy authority can trade some losses in the present for reduction of uncertainty in the future.

[4] See Shiller(1979) for this approximation which is a linearization of a more complicated non–linear expectations model.

The Environment

The monetary authority is assumed to have a loss function $r(x_t,y_t)$. The expected loss $w{:}\Lambda(\Theta)\to R$, given beliefs μ and action $x \in X$ is

$$w(x_t,y_t) = \Sigma_\Theta \int_Y r(x_t,y_t)\, p(y_t|x_t,\theta_i) dP \lambda(\theta_i) \tag{13}$$

where we assume that: (A5) $w(x,\mu)$ has a unique minimum in x for beliefs $\mu \in \Lambda(\Theta)$; In the examples that follow r(.,.) will be quadratic, and (A5) will be satisfied.

The authority's objective function is to maximize the expected present discounted value of the negation of losses

$$E[-\sum_{t=0}^{\infty} \delta^t w(x_t,\mu_t)] \tag{14}$$

To facilitate analysis of this problem, we want to tranform this into a dynamic programming formulation, with beliefs forming the state space and a sequence of money balances forming the action space.

In choosing an action today, the policy maker must consider not only the present reward but the effect this will have on the distribution of posterior beliefs. The optimal policy will trade losses today for more information. We proceed towards characterizing the set of optimal policies by writing down the value function. Define

$$V(\mu) \equiv \max_{x\in X} [-w(x,\mu) + \delta V(\Gamma(\mu,x,y))[\Sigma_\Theta \lambda(\theta_i)p(y|x,\theta_i)]dP] \tag{15}$$

This expression gives us maximum future welfare given optimal actions for beliefs μ. This is a standard problem for which we know, given our assumptions, (a) the value function is well–defined; and (b) it possesses a continuous, unique solution. See Blackwell(1965) for proof. It is not however the case that a unique optimal policy exists. Our discussion will instead focus on the issue of whether or not the policy authority will pursue actions that will lead to learning the parameters θ^*.

Complete Learning

The recursivity of the value function will enable us to work with the much simpler two period problem. Consider the invariant action $x_t = \bar{x}\ \forall\ t$. The expected discounted sequence of rewards then is easy to calculate as the agent does not expect his current beliefs μ_0 to change:

$$E[-\sum_{t=1}^{\infty} \delta^t w(\bar{x},\mu_t) \mid \bar{x},\mu_0] = E[-\sum_{t=1}^{\infty} \delta^t w(\bar{x},\mu_0)] = [-\delta/(1-\delta)]w(\bar{x},\mu_0) \tag{16}$$

We can now readily see the link between the two–period and multi–period problems:

Lemma 1: Let $\bar{x}$ be the unique maximum to $-w(x,\mu_0)$. If $\bar{x}$ does not solve

$$\max_{x\in X}\left(-w(x,\mu_0) + \rho V^1(\Gamma(\mu_0,x,y))[\Sigma_\Theta \lambda(\theta_i)p(y|x,\theta_i)]dP\right) \tag{17}$$

where $V^1(\Gamma(.)) = -w(a^1(\mu_0),\mu_0)$ and $\rho = \delta/(1-\delta)$, then it does not solve (15).

Proof: Our proof follows Easley and Kiefer(1988). Note that from (16) for an invariant action, the multi–period value function (15) reduces to learn k and η, but without a policy that varies the monetary growth rate between periods, Proposition I assures us of non–convergence to rational expectations. Furthermore, a change in the money growth rate, after agent's beliefs have settled down in the regime $\mu_t = \bar{\mu}$, would lead to non–rational forecasts.

A Simple Expectations Model of the Term Structure

In this example, pegging the interest rate will be the policy barrier, preventing agents from learning the slope of the term structure equation. An approximation of the relation between long and short term interest rates, R_t and r_t is given by[4]

$$R_t = \theta + (1-\gamma)\sum_{k=0}^{\infty} E_t r_{t+k} + \varepsilon_t \qquad (10)$$

where θ is interpreted as a constant term premium, $0 < \gamma < 1$ is a constant discount factor, and ε is a white noise disturbance term.. Let us assume that agents form their expectations of the short rate using a distributed lag on previous short rates:

$$E_t[r_t] = r^e_t = \sum_{k=0}^{T}\phi_k r_{t-k} \qquad \sum_{k=0}^{T}\phi_k = 1 \qquad (11)$$

where the ϕ's are unkown to the econometrician. Suppose the monetary authority begins to peg the interest rate at $r_t = \bar{r}$. According to (11), it will take T periods, but eventually the forecast of all future short rates will be $\bar{r}$ as well. We will now show two results: (i) the discount rate will not be estimated consistently in a single pegging regime, (ii) it can only be consistently estimated with k–period pegging switches, where k corresponds to the length of the lag in (11).

The estimation problem here is very tricky without some help from the policy makers. The usual assumption made is that agents have rational expectations. In which case the econometricianestimates

$$R_{t+1} - R_t = \frac{-(1-\gamma)\theta}{\gamma} + \frac{1-\gamma}{\gamma}(R_t - r_t) + \frac{v_{t+1}}{\gamma} \qquad (12)$$

where $v_{t+1} = (1-\gamma)\Sigma_{k=1}{}^{\infty}(\gamma^k(E_{t+1}r_{t+k} - E_t r_{t+k}))$ represents the revision in the agent's expectations. Consistent estimation of (12) requires the assumption that v_{t+1} be uncorrelated with the right hand side regressors. (11) violates this assumption (and our A4 as well). Only with the long rate constant can we ignore the ϕ's; successive k–period regime switches will then enable us to estimate (10) with realized short rates.

SOME ISSUES IN STOCHASTIC CONTROL

We shift our focus now slightly away from non–existence counterexamples to examples of otpimal policy in stochastic environments. The critical intuiton is the tradeoff presented in discounted control situations where the model parameters are unknown. While large variation in the control variable is undesirable from the point of view of the usual loss function, it is desirable in gaining knowledge over the system's parameters. For low discount rates, the policy authority can trade some losses in the present for reduction of uncertainty in the future.

---

[4] See Shiller(1979) for this approximation which is a linearization of a more complicated non–linear expectations model.

The Environment

The monetary authority is assumed to have a loss function $r(x_t,y_t)$. The expected loss $w:\Lambda(\Theta)\rightarrow R$, given beliefs μ and action $x \in X$ is

$$w(x_t,y_t) = \Sigma_\Theta \int_Y r(x_t,y_t)\ p(y_t|x_t,\theta_i)dP\lambda(\theta_i) \qquad (13)$$

where we assume that: (A5) $w(x,\mu)$ has a unique minimum in x for beliefs $\mu \in \Lambda(\Theta)$; In the examples that follow $r(.,.)$ will be quadratic, and (A5) will be satisfied.

The authority's objective function is to maximize the expected present discounted value of the negation of losses

$$E[-\sum_{t=0}^{\infty}\delta^t w(x_t,\mu_t)] \qquad (14)$$

To facilitate analysis of this problem, we want to tranform this into a dynamic programming formulation, with beliefs forming the state space and a sequence of money balances forming the action space.

In choosing an action today, the policy maker must consider not only the present reward but the effect this will have on the distribution of posterior beliefs. The optimal policy will trade losses today for more information. We proceed towards characterizing the set of optimal policies by writing down the value function. Define

$$V(\mu) \equiv \max_{x\in X}\ [-w(x,\mu) + \delta V(\Gamma(\mu,x,y))[\Sigma_\Theta\lambda(\theta_i)p(y|x,\theta_i)]dP] \qquad (15)$$

This expression gives us maximum future welfare given optimal actions for beliefs μ. This is a standard problem for which we know, given our assumptions, (a) the value function is well–defined; and (b) it possesses a continuous, unique solution. See Blackwell(1965) for proof. It is not however the case that a unique optimal policy exists. Our discussion will instead focus on the issue of whether or not the policy authority will pursue actions that will lead to learning the parameters θ^*.

Complete Learning

The recursivity of the value function will enable us to work with the much simpler two period problem. Consider the invariant action $x_t = \bar{x}\ \forall\ t$. The expected discounted sequence of rewards then is easy to calculate as the agent does not expect his current beliefs μ_0 to change:

$$E[-\sum_{t=1}^{\infty}\delta^t w(\bar{x},\mu_t)\ |\ \bar{x},\mu_0] = E[-\sum_{t=1}^{\infty}\delta^t w(\bar{x},\mu_0)] = [-\delta/(1-\delta)]w(\bar{x},\mu_0) \qquad (16)$$

We can now readily see the link between the two–period and multi–period problems:

Lemma 1: Let $\bar{x}$ be the unique maximum to $-w(x,\mu_0)$. If $\bar{x}$ does not solve

$$\max_{x\in X}\left(-w(x,\mu_0) + \rho V^1(\Gamma(\mu_0,x,y))[\Sigma_\Theta\lambda(\theta_i)p(y|x,\theta_i)]dP]\ \right) \qquad (17)$$

where $V^1(\Gamma(.)) = -w(a^1(\mu_0),\mu_0)$ and $\rho = \delta/(1-\delta)$, then it does not solve (15).

Proof: Our proof follows Easley and Kiefer(1988). Note that from (16) for an invariant action, the multi–period value function (15) reduces to

$$V(\mu_0) = -[w(\bar{x},\mu_0) + (\delta/(1-\delta))w(\bar{x},\mu_0)]$$
$$= -(1/(1-\delta))w(\bar{x},\mu_0) \quad (18)$$

Let $a^1(\mu_0) = \bar{x}$, and we see that for $\rho=\delta/(1-\delta)$, they are the same problem.∎

Next, we want to show the convexity of the value function. This will be instrumental in giving the policy authority the incentive to pursue actions that will yield uncertainty over future beliefs. This means a highly variable policy. We begin with

Lemma 2: The one–period value function is convex on $\Lambda(\Theta)$.

Proof: Our proof follows Marschak and Miyasawa(1968) and Easley and Kiefer(1988). Choose any two elements from the set of beliefs $\mu_a,\mu_b \in \Lambda(\Theta)$, and let μ_c be a convex combination: $\mu_c = \gamma\mu_a + (1-\gamma)\mu_b$. Let $\bar{x}(\mu)$ be the maximum of $-w(\bar{x},\mu)$ for beliefs μ.

$$\gamma V^1(\mu_a) + (1-\gamma)V^1(\mu_b) = \gamma V^1(\bar{x}(\mu_a),\mu_a) + (1-\gamma)V^1(\bar{x}(\mu_b),\mu_b)$$
$$\geq \gamma V^1(\bar{x}(\mu_c),\mu_c) + (1-\gamma)V^1(\bar{x}(\mu_c),\mu_c)$$
$$= V^1(\mu_c) \quad (19)$$

with the inequality based on revealed preference for a feasible alternative, and the equality resulting from the linearity of rewards in beliefs.∎

We depart from the literature on experimentation at this point. Prescott(1972) and Grossman, Kihlstrom and Mirman(1977) show that larger actions $x_1 > x_2$ are more informative[5]. This is sufficient to show that the policy authority will choose an action that exceeds the optimal one–period action. Our goal here, though, given the recent results of McLennan(1987) and Easley and Kiefer(1988), is to characterize conditions under which the policy authority will pursue actions that lead to beliefs converging to θ^*.

Proposition II: If $V^1(\mu)$ is convex, then there exists a $\delta^* < 1$, such that if $\delta \geq \delta^*$, $\mu_\infty = \theta^*$, almost surely.

Proof: We condense several propositions from Easley and Kiefer(1988). Let $S \subset \Theta$ be the set of beliefs invariant for any action $a^1(\mu)$. By the martingale convergence theorem, $\mu \rightarrow \mu_\infty$, and straightforward arguments give us that the support of μ_∞ will be S. Convexity of $V^1(\mu)$ along with (A2) give us that S is finite. We need then to consider δ's corresponding to each element in S.

Consider now some $\bar{\mu} \in S - \{\theta^*\}$. Let $a^1(\bar{\mu}) = \bar{x}$. From Lemma 2, we know that $\bar{x}$ does not solve $E[-w(\bar{x},\bar{\mu})]$. There then exist δ's: $0 \leq \delta^* \leq \delta < 1$, where δ^* is a critical discount factor for which learning is incomplete, such that $\bar{x}$ does not solve (17). By Lemma 1 then, it cannot solve the infinite horizon problem.∎

The optimal policy choice turns on the discount factor. It is similarly easy to show that for small δ (you discount the future highly), incomplete learning can be an optimal policy. This is just Rothschild's(1974) two–armed bandit example, where incomplete learning of the demand curve by an "experimenting" monopolist is optimal. Kiefer and Nyarko and McLennan(1984) also have examples of incomplete learning.

---

[5] The motivation for this word comes from Blackwell's(1953) theorem on the comparison of experiments. Blackwell defines an experiment x_1 to be more informative than x_2 if every loss vector $w(x_2,\mu)$ is also attainable with x_1. Clearly $x_1 > x_2$ is sufficient for x_1 to be more informative.

For large enough δ in which complete learning of θ^* is optimal, we have established that continually variable policy may be necessary for the policy authority to maximize welfare. We show below in two popular macroeconomic models, the implications of these results.

STOCHASTIC CONTROL EXAMPLES

We consider two examples. The first re–examines the conclusions for optimal policy in a stochastic IS/LM model, and the other seeks to do the same in a rational expectations model.

Optimal Policy in the IS/LM Model

Our first example is motivated by Poole's(1970) paper on the choice of instruments in an IS/LM model. We demonstrate that taking account of parametric uncertainty can reverse the standard conclusions of that paper. Following Poole, we pose a linear stochastic version of Hick's IS/LM model

$$y = a_0 + a_1 r + e_t \qquad a_1 < 0 \quad e_t \sim N(0,\sigma_e) \quad (20)$$
$$m = b_0 + b_1 Y + b_2 r + v_t$$
$$b_1 > 0,\ b_2 < 0 \quad v_t \sim N(0,\sigma_v) \quad (21)$$

where y is output, r is the interest rate, m is the money supply (we will not even purport to distinguish between real and nominal magnitudes), e_t and v_t are disturbance terms, not assumed independent, and a's and b's are parameters. Solving out for the reduced forms for income we get

$$y = a_0 + a_1 r + e_t \quad (22)$$

$$y = \alpha + \beta m_t + \varepsilon_t \quad (23)$$

where $\alpha = (a_0b_2 - a_1b_0)/(a_1b_1 + b_2)$, $\beta = a_1/(a_1b_1 + b_2)$, and $\varepsilon_t = (b_2e - a_1v)/(a_1b_1 + b_2)$.

The monetary authority's function here is to minimize a quadratic loss around some optimal y^*:

$$L(m,y) = (y^* - y)^2 \quad (24)$$

Given no parametric uncertainty, the expected losses with r and m as the instruments are

$$L_r = \sigma_e \quad (25)$$

$$L_m = \sigma_\varepsilon = (a^2{}_1\sigma_v - 2a_1b_2\sigma_{ev} + b^2{}_2)/(a_1b_1 + b_2)^2 \quad (26)$$

The standard conclusion in the static model is if $\sigma_e > \sigma_\varepsilon$ (which implies that $\sigma_v/\sigma_e < b_1$) then a money stock policy is to be preferred.

Consider now the case where a_0, a_1, α, β are unknowns. Assume that the agent holds a bi–variate normal prior,

$$\mu(\alpha,\beta|\bar{\alpha}_0, \bar{\beta}_0, \Sigma_0) = N\left(\begin{pmatrix}\bar{\alpha}_0\\ \bar{\beta}_0\end{pmatrix}\Sigma_0\right) \quad (27)$$

and that the disturbance term is normally distributed as well. This has the advantage of making all the subsequent posterior distributions normal.

$$L_r = \sigma_{a0} + r^2\sigma_{a1} - 2r\sigma_{a0a1} + \sigma_e \quad (28)$$

$$L_m = \sigma_\alpha + m^2\sigma_\beta - 2m\sigma_{\alpha\beta} + \sigma_\varepsilon \quad (29)$$

Clearly, parametric uncertainty can reverse the choice of instruments, even if $\sigma_e > \sigma_\varepsilon$.

$$V(\mu_0) = -[w(\bar{x},\mu_0) + (\delta/(1-\delta))w(\bar{x},\mu_0)]$$
$$= -(1/(1-\delta))w(\bar{x},\mu_0) \quad (18)$$

Let $a^1(\mu_0) = \bar{x}$, and we see that for $\rho=\delta/(1-\delta)$, they are the same problem.∎

Next, we want to show the convexity of the value function. This will be instrumental in giving the policy authority the incentive to pursue actions that will yield uncertainty over future beliefs. This means a highly variable policy. We begin with

Lemma 2: The one–period value function is convex on $\Lambda(\Theta)$.

Proof: Our proof follows Marschak and Miyasawa(1968) and Easley and Kiefer(1988). Choose any two elements from the set of beliefs $\mu_a,\mu_b \in \Lambda(\Theta)$, and let μ_c be a convex combination: $\mu_c = \gamma\mu_a + (1-\gamma)\mu_b$. Let $\bar{x}(\mu)$ be the maximum of $-w(\bar{x},\mu)$ for beliefs μ.

$$\gamma V^1(\mu_a) + (1-\gamma)V^1(\mu_b) = \gamma V^1(\bar{x}(\mu_a),\mu_a) + (1-\gamma)V^1(\bar{x}(\mu_b),\mu_b)$$
$$\geq \gamma V^1(\bar{x}(\mu_c),\mu_c) + (1-\gamma)V^1(\bar{x}(\mu_c),\mu_c)$$
$$= V^1(\mu_c) \quad (19)$$

with the inequality based on revealed preference for a feasible alternative, and the equality resulting from the linearity of rewards in beliefs.∎

We depart from the literature on experimentation at this point. Prescott(1972) and Grossman, Kihlstrom and Mirman(1977) show that larger actions $x_1 > x_2$ are more informative[5]. This is sufficient to show that the policy authority will choose an action that exceeds the optimal one–period action. Our goal here, though, given the recent results of McLennan(1987) and Easley and Kiefer(1988), is to characterize conditions under which the policy authority will pursue actions that lead to beliefs converging to θ^*.

Proposition II: If $V^1(\mu)$ is convex, then there exists a $\delta^* < 1$, such that if $\delta \geq \delta^*$, $\mu_\infty = \theta^*$, almost surely.

Proof: We condense several propositions from Easley and Kiefer(1988). Let $S \subset \Theta$ be the set of beliefs invariant for any action $a^1(\mu)$. By the martingale convergence theorem, $\mu \to \mu_\infty$, and straightforward arguments give us that the support of μ_∞ will be S. Convexity of $V^1(\mu)$ along with (A2) give us that S is finite. We need then to consider δ's corresponding to each element in S.
Consider now some $\bar{\mu} \in S - \{\theta^*\}$. Let $a^1(\bar{\mu}) = \bar{x}$. From Lemma 2, we know that $\bar{x}$ does not solve $E[-w(\bar{x},\bar{\mu})]$. There then exist δ's: $0 \leq \delta^* \leq \delta < 1$, where δ^* is a critical discount factor for which learning is incomplete, such that $\bar{x}$ does not solve (17). By Lemma 1 then, it cannot solve the infinite horizon problem.∎

The optimal policy choice turns on the discount factor. It is similarly easy to show that for small δ (you discount the future highly), incomplete learning can be an optimal policy. This is just Rothschild's(1974) two–armed bandit example, where incomplete learning of the demand curve by an "experimenting" monopolist is optimal. Kiefer and Nyarko and McLennan(1984) also have examples of incomplete learning.

[5] The motivation for this word comes from Blackwell's(1953) theorem on the comparison of experiments. Blackwell defines an experiment x_1 to be more informative than x_2 if every loss vector $w(x_2,\mu)$ is also attainable with x_1. Clearly $x_1 > x_2$ is sufficient for x_1 to be more informative.

For large enough δ in which complete learning of θ^* is optimal, we have established that continually variable policy may be necessary for the policy authority to maximize welfare. We show below in two popular macroeconomic models, the implications of these results.

STOCHASTIC CONTROL EXAMPLES

We consider two examples. The first re–examines the conclusions for optimal policy in a stochastic IS/LM model, and the other seeks to do the same in a rational expectations model.

Optimal Policy in the IS/LM Model

Our first example is motivated by Poole's(1970) paper on the choice of instruments in an IS/LM model. We demonstrate that taking account of parametric uncertainty can reverse the standard conclusions of that paper. Following Poole, we pose a linear stochastic version of Hick's IS/LM model

$$y = a_0 + a_1 r + e_t \quad a_1 < 0 \quad e_t \sim N(0,\sigma_e) \quad (20)$$
$$m = b_0 + b_1 Y + b_2 r + v_t$$
$$b_1 > 0,\ b_2 < 0 \quad v_t \sim N(0,\sigma_v) \quad (21)$$

where y is output, r is the interest rate, m is the money supply (we will not even purport to distinguish between real and nominal magnitudes), e_t and v_t are disturbance terms, not assumed independent, and a's and b's are parameters. Solving out for the reduced forms for income we get

$$y = a_0 + a_1 r + e_t \quad (22)$$

$$y = \alpha + \beta m_t + \varepsilon_t \quad (23)$$

where $\alpha = (a_0b_2 - a_1b_0)/(a_1b_1 + b_2)$, $\beta = a_1/(a_1b_1 + b_2)$, and $\varepsilon_t = (b_2e - a_1v)/(a_1b_1 + b_2)$.

The monetary authority's function here is to minimize a quadratic loss around some optimal y^*:

$$L(m,y) = (y^* - y)^2 \quad (24)$$

Given no parametric uncertainty, the expected losses with r and m as the instruments are

$$L_r = \sigma_e \quad (25)$$

$$L_m = \sigma_\varepsilon = (a^2{}_1\sigma_v - 2a_1b_2\sigma_{ev} + b^2{}_2)/(a_1b_1 + b_2)^2 \quad (26)$$

The standard conclusion in the static model is if $\sigma_e > \sigma_\varepsilon$ (which implies that $\sigma_v/\sigma_e < b_1$) then a money stock policy is to be preferred.

Consider now the case where a_0, a_1, α, β are unknowns. Assume that the agent holds a bi–variate normal prior,

$$\mu(\alpha,\beta|\bar{\alpha}_0, \bar{\beta}_0, \Sigma_0) = N\left(\begin{pmatrix}\bar{\alpha}_0\\ \bar{\beta}_0\end{pmatrix}\Sigma_0\right) \quad (27)$$

and that the disturbance term is normally distributed as well. This has the advantage of making all the subsequent posterior distributions normal.

$$L_r = \sigma_{a0} + r^2\sigma_{a1} - 2r\sigma_{a0a1} + \sigma_e \quad (28)$$

$$L_m = \sigma_\alpha + m^2\sigma_\beta - 2m\sigma_{\alpha\beta} + \sigma_\varepsilon \quad (29)$$

Clearly, parametric uncertainty can reverse the choice of instruments, even if $\sigma_e > \sigma_\varepsilon$.

Now turn to the multi-period problem. We want to show that the single period results can be reversed for policy "experimentation" reasons in the multi-per od case. We assume that, given prior beliefs, $L_m > L_r$, but $\sigma_e > \sigma_\varepsilon$, so that in the absence of parametric uncertain y, the money stock policy would be preferred. We show now it may be optimal to use the money stock policy, even thoug i its one period loss for initial beliefs is higher.

Let's map this problem into the environment of Section IV. The value function is given by (15) with maximization for m $\in [0,\overline{m}]$ and with

$$w(m,\mu) = -\Sigma_\Theta \int \gamma(y^* - \alpha - \beta m_t - \varepsilon_t)^2 p(y_t | m_t, \theta_i)\, dP\lambda(\theta_i) \qquad (30)$$

The minimum one-period loss would be to choose $m^* = (y^* - \hat{\alpha})/\overline{\beta}$. PII implies that this cannot solve the infinite horizon problem for high enough discount factors. Further more, no constant money stock policy $m_t = \overline{m}$ can be optimal either.

Some Surprises for the Lucas Supply Function?

Our structure for this problem posits that output is given solely by a Lucas "surprise" supply function:

$$y_t = \alpha + \beta(p_t - p^e_t) + \varepsilon_t \quad \varepsilon_t \sim N(0,\sigma_\varepsilon) \qquad (31)$$

y is output, p are prices and p^e is their expected level. We assume that prices react to the money stock, m, and real disturbances, ξ:

$$p_t = m_t + \xi_t \qquad \xi_t \sim \text{i.i.d. } N(0,\sigma_\xi) \qquad (32)$$

Given our i.i.d. assumption on the real shocks,

$$p^e_t = m^e_t \qquad (33)$$

The monetary authority's role is assumed to be minimizing the variance of output. In the case where monetary policy is perfectly anticipated but α and β are unknown, we linearize $\beta\xi$ through a Taylor expansion around $\beta = \overline{\beta}$, $\xi = \overline{\xi} = 0$: $\beta\xi \approx \overline{\beta}\overline{\xi} + \overline{\beta}(\xi - \overline{\xi}) + \overline{\xi}(\beta - \overline{\beta}) = \overline{\beta}\xi$. The variance of output is

$$\text{var}(y_t) = \sigma_\alpha + \overline{\beta}^2\sigma_\xi - 2\overline{\beta}\sigma_{\alpha\xi} + \sigma_\varepsilon \qquad (34)$$

We want to show that it might be optimal for the monetary authority to undertake unpredictable policy. To keep the problem simple, we will assume that the expected value for m is some target value for the money supply, say $\overline{m}_t$. In this case, the variance of output is

$$\text{var}(y_t) = \sigma_\alpha + \overline{\beta}^2\sigma_\xi + (m^2_t - \overline{m}^2_t)\sigma_\beta - 2\overline{\beta}\sigma_{\alpha\xi} - 2(m_t - \overline{m}_t)\sigma_{\alpha\beta} - 2(m_t - \overline{m}_t)\sigma_{\beta\xi}. \qquad (35)$$

which is larger than (34), except in the unlikely instance that the monetary authority's beliefs about the parameters were correlated with the real shocks. If we took α and β to be structural parameters, we could safely assume these covariances were zero.

The monetary authority pays one quadratic price for its deception, equal to $(m^2_t - \overline{m}^2_t)\sigma_\beta$. Suppose though that σ_α were very large (given that it proxies the natural rate, it is not farfetched), then we would be in the situation of the first example. We might want to trade some current period losses for future gains.

The Fed could come to learn α and β solely through the effects of the real shocks ξ, $\Sigma_{t=0}{}^T(\xi_t - 0)^2 \rightarrow \infty$ as $T \rightarrow \infty$, so we are assured of convergence. Without these shocks, the Fed would have to pursue "experimentation." Only an infinite series of surprises from the monetary authority would maximize the value function. Any predictable monetary policy would be sub-optimal.

CONCLUSIONS

Learning must be studied explicitly we think for two reasons. The first is that rational expectations cannot properly be regarded as the limitation of all learning processes. Beliefs do always settle down (there is a point at which beliefs no longer change), but as Section II revealed, beliefs may not converge to the true structural model.

The second reason concerns questions of policy. The idea that anticipated monetary policy is neutral has become conventional wisdom in macroeconomics. We have demonstrated that this is only true in a world in which the only uncertainty is over policy. Parametric uncertainty is sufficient to make this statement quite inaccurate.

We have also given a new perspective on volatility. Usually this is regarded as a bad thing, since most loss or utility functions are quadratic. In the multi-period control problem though, variability is informative about the future. Weighing short-term losses against long-term gains is particularly important for policy authorities with long horizons.

A series of counter-examples were presented in the same spirit as Sargent and Wallace(1975). Their work motivated macroeconomists to look fundamentally at the question of expectations. This paper has entirely the same objective. Recent advances in the learning literature though have enabled us to move a step closer towards bringing expectations formation into the optimizing paradigm. We feel as though the implications for policy are as challenging to the conventional wisdom as the original contributions by the rational expectations school.

REFERENCES

Amemiya, T. (1985) Advanced Econometrics, Cambridge: Harvard U. Press.

Billingsley, P. (1979) Probability and Measure, New York: John Wiley.

Blackwell, D. (1953) "Equivalent Comparisons of Experiments" Annals of Mathematical Statistics 24, p.265–73.

Blackwell, D. (1965) "Discounted Dynamic Programming," Annals of Mathematical Statistics 36, p.226–35.

Cagan, P. (1956) "The Monetary Dynamics of Hyperinflation," in Studies in the Quantity Theory of Money, (M.Friedman ed.), Chicago: U. of Chicago Press.

Chow, G. (1981) Econometric Analysis by Control Methods, New York: John Wiley.

Easley, D. and N. Kiefer (1988) "Controlling A Stochastic Process with Unknown Parameters," Econometrica 56, p.1045–64.

I would like to thank Yaw Nyarko for several helpful discussions. Richard Arnott and Chris Maxwell made useful comments on an earlier draft. The usual disclaimer applies.

Now turn to the multi-period problem. We want to show that the single period results can be reversed for policy "experimentation" reasons in the multi-period case. We assume that, given prior beliefs, $L_m > L_r$, but $\sigma_e > \sigma_\varepsilon$, so that in the absence of parametric uncertainty, the money stock policy would be preferred. We show now it may be optimal to use the money stock policy, even though its one period loss for initial beliefs is higher.

Let's map this problem into the environment of Section IV. The value function is given by (15) with maximization for m $\in [0,\overline{m}]$ and with

$$w(m,\mu) = -\Sigma_\Theta \int_Y (y^* - \alpha - \beta m_t - \varepsilon_t)^2 p(y_t | m_t, \theta_i)\, dP\lambda(\theta_i) \quad (30)$$

The minimum one-period loss would be to choose $m^* = (y^* - \hat{\alpha})/\bar{\beta}$. PII implies that this cannot solve the infinite horizon problem for high enough discount factors. Further more, no constant money stock policy $m_t = \overline{m}$ can be optimal either.

<u>Some Surprises for the Lucas Supply Function?</u>

Our structure for this problem posits that output is given solely by a Lucas "surprise" supply function:

$$y_t = \alpha + \beta(p_t - p^e_t) + \varepsilon_t \quad \varepsilon_t \sim N(0,\sigma_\varepsilon) \quad (31)$$

y is output, p are prices and p^e is their expected level. We assume that prices react to the money stock, m, and real disturbances, ξ:

$$p_t = m_t + \xi_t \qquad \xi_t \sim \text{i.i.d. } N(0,\sigma_\xi) \quad (32)$$

Given our i.i.d. assumption on the real shocks,

$$p^e_t = m^e_t \quad (33)$$

The monetary authority's role is assumed to be minimizing the variance of output. In the case where monetary policy is perfectly anticipated but α and β are unknown, we linearize $\beta\xi$ through a Taylor expansion around $\beta = \bar{\beta}$, $\xi = \bar{\xi} = 0$: $\beta\xi \approx \bar{\beta}\bar{\xi} + \bar{\beta}(\xi - \bar{\xi}) + \bar{\xi}(\beta - \bar{\beta}) = \bar{\beta}\xi$. The variance of output is

$$\text{var}(y_t) = \sigma_\alpha + \bar{\beta}^2\sigma_\xi - 2\bar{\beta}\sigma_{\alpha\xi} + \sigma_\varepsilon \quad (34)$$

We want to show that it might be optimal for the monetary authority to undertake unpredictable policy. To keep the problem simple, we will assume that the expected value for m is some target value for the money supply, say $\overline{m}_t$. In this case, the variance of output is

$$\text{var}(y_t) = \sigma_\alpha + \bar{\beta}^2\sigma_\xi + (m^2_t - \overline{m}^2_t)\sigma_\beta - 2\bar{\beta}\sigma_{\alpha\xi} - 2(m_t - \overline{m}_t)\sigma_{\alpha\beta} - 2(m_t - \overline{m}_t)\sigma_{\beta\xi}. \quad (35)$$

which is larger than (34), except in the unlikely instance that the monetary authority's beliefs about the parameters were correlated with the real shocks. If we took α and β to be structural parameters, we could safely assume these covariances were zero.

The monetary authority pays one quadratic price for its deception, equal to $(m^2_t - \overline{m}^2_t)\sigma_\beta$. Suppose though that σ_α were very large (given that it proxies the natural rate, it is not farfetched), then we would be in the situation of the first example. We might want to trade some current period losses for future gains.

The Fed could come to learn α and β solely through the effects of the real shocks ξ, $\Sigma_{t=0}^T(\xi_t - 0)^2 \to \infty$ as $T \to \infty$, so we are assured of convergence. Without these shocks, the Fed would have to pursue "experimentation." Only an infinite series of surprises from the monetary authority would maximize the value function. Any predictable monetary policy would be sub-optimal.

CONCLUSIONS

Learning must be studied explicitly we think for two reasons. The first is that rational expectations cannot properly be regarded as the limitation of all learning processes. Beliefs do always settle down (there is a point at which beliefs no longer change), but as Section II revealed, beliefs may not converge to the true structural model.

The second reason concerns questions of policy. The idea that anticipated monetary policy is neutral has become conventional wisdom in macroeconomics. We have demonstrated that this is only true in a world in which the only uncertainty is over policy. Parametric uncertainty is sufficient to make this statement quite inaccurate.

We have also given a new perspective on volatility. Usually this is regarded as a bad thing, since most loss or utility functions are quadratic. In the multi-period control problem though, variability is informative about the future. Weighing short-term losses against long-term gains is particularly important for policy authorities with long horizons.

A series of counter-examples were presented in the same spirit as Sargent and Wallace(1975). Their work motivated macroeconomists to look fundamentally at the question of expectations. This paper has entirely the same objective. Recent advances in the learning literature though have enabled us to move a step closer towards bringing expectations formation into the optimizing paradigm. We feel as though the implications for policy are as challenging to the conventional wisdom as the original contributions by the rational expectations school.

REFERENCES

Amemiya, T. (1985) <u>Advanced Econometrics</u>, Cambridge: Harvard U. Press.

Billingsley, P. (1979) <u>Probability and Measure</u>, New York: John Wiley.

Blackwell, D. (1953) "Equivalent Comparisons of Experiments" <u>Annals of Mathematical Statistics</u> 24, p.265–73.

Blackwell, D. (1965) "Discounted Dynamic Programming," <u>Annals of Mathematical Statistics</u> 36, p.226–35.

Cagan, P. (1956) "The Monetary Dynamics of Hyperinflation," in <u>Studies in the Quantity Theory of Money</u>, (M.Friedman ed.), Chicago: U. of Chicago Press.

Chow, G. (1981) <u>Econometric Analysis by Control Methods</u>, New York: John Wiley.

Easley, D. and N. Kiefer (1988) "Controlling A Stochastic Process with Unknown Parameters," <u>Econometrica</u> 56, p.1045–64.

I would like to thank Yaw Nyarko for several helpful discussions. Richard Arnott and Chris Maxwell made useful comments on an earlier draft. The usual disclaimer applies.

Grossman, S., R. Kihlstrom, and L. Mirman (1977) "A Bayesian Approach to the Production of Information and Learning by Doing," Review of Economic Studies 44, p.533–547.

Jordan, J. (1985) "Learning Rational Expectations: The Finite State Case," Journal of Economic Theory 36, p.257–76.

Kiefer, N. and Y. Nyarko (1987) "Optimal Control of an Unknown Error Process with Learning," Cornell U. Working Paper #370.

Marschak, J. and H. Miyasawa (1968) "Economic Comparability of Information Systems," International Economic Review 9, p.137–74.

McLennan, A. (1984) "Price Dispersion and Incomplete Learning in the Long Run," Journal of Economic Dynamics and Control 7, p.331–47.

McLennan, A. (1987) "On the Optimality of Incomplete Learning," U. of Minnesota.

Mussa, M. (1975) "Adaptive and Regressive Expectations in a Rational Model of the Inflationary Process," Journal of Monetary Economics 1, p.432–42.

Nyarko, Y. (1988) "On the Convergence of Bayesian Posterior Processes in Linear Economic Models," Brown University, mimeographed.

Poole, W. (1970) "Optimal Choice of Monetary Policy Instruments in a Simple Stochastic Macro Model," Quarterly Journal of Economics, p.197–216.

Prescott, E. (1972) "The Multi-Period Control Problem Under Uncertainty," Econometrica 40, p.1043–58.

Rothschild, M. (1974) "A Two-Armed Bandit Theory of Market Pricing," Journal of Economic Theory 9, p.185–202.

Sargent, T. and N. Wallace (1975) "Rational Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply," Journal of Political Economy, 83, p.241–54.

Shiller, R. (1979) "The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure," Journal of Political Economy 87, p.1190–1219.

Taylor, J. (1974) "Asymptotic Properties of Multiperiod Control Rules in the Linear Regression Model," International Economic Review 15, p.472–84.

Zellner, A. (1971) An Introduction to Bayesian Inference in Econometrics, New York: John Wiley.

Grossman, S., R. Kihlstrom, and L. Mirman (1977) "A Bayesian Approach to the Production of Information and Learning by Doing," Review of Economic Studies 44, p.533–547.

Jordan, J. (1985) "Learning Rational Expectations: The Finite State Case," Journal of Economic Theory 36, p.257–76.

Kiefer, N. and Y. Nyarko (1987) "Optimal Control of an Unknown Error Process with Learning," Cornell U. Working Paper #370.

Marschak, J. and H. Miyasawa (1968) "Economic Comparability of Information Systems," International Economic Review 9, p.137–74.

McLennan, A. (1984) "Price Dispersion and Incomplete Learning in the Long Run," Journal of Economic Dynamics and Control 7, p.331–47.

McLennan, A. (1987) "On the Optimality of Incomplete Learning," U. of Minnesota.

Mussa, M. (1975) "Adaptive and Regressive Expectations in a Rational Model of the Inflationary Process," Journal of Monetary Economics 1, p.432–42.

Nyarko, Y. (1988) "On the Convergence of Bayesian Posterior Processes in Linear Economic Models," Brown University, mimeographed.

Poole, W. (1970) "Optimal Choice of Monetary Policy Instruments in a Simple Stochastic Macro Model," Quarterly Journal of Economics, p.197–216.

Prescott, E. (1972) "The Multi–Period Control Problem Under Uncertainty," Econometrica 40, p.1043–58.

Rothschild, M. (1974) "A Two–Armed Bandit Theory of Market Pricing," Journal of Economic Theory 9, p.185–202.

Sargent, T. and N. Wallace (1975) "Rational Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply," Journal of Political Economy, 83, p.241–54.

Shiller, R. (1979) "The Volatility of Long–Term Interest Rates and Expectations Models of the Term Structure," Journal of Political Economy 87, p.1190–1219.

Taylor, J. (1974) "Asymptotic Properties of Multiperiod Control Rules in the Linear Regression Model," International Economic Review 15, p.472–84.

Zellner, A. (1971) An Introduction to Bayesian Inference in Econometrics, New York: John Wiley.

NON–CONVERGENCE TO RATIONAL EXPECTATIONS AND OPTIMAL MONETARY POLICY IN MODELS WITH LEARNING

Bruce Mizrach, Department of Economics, Boston College, Chestnut Hill, MA USA

Abstract: A principal argument in the rational expectations literature is the optimality of predictable policy. This paper illustrates that this claim does not hold in a world of parametric uncertainty for two reasons: (1) completely noiseless policy may lead to non–convergence to the true model parameters; (2) highly predictable policy is not very informative about the structure of the model. A series of examples illustrate the ramifications for macroeconomic policy.

Keywords: Learning systems; Bayes methods; stochastic control; dynamic programming;

INTRODUCTION

This paper examines the implications of learning for macroeconomic policy. In an environment in which the parameters of a structural model are unknown to economic agents, optimal policy differs strongly from the conventional wisdom. The leading argument in the rational expectations literature is policy ineffectiveness; that is, policy doesn't matter as long as its predictable. This translates into simple dictates for the monetary authority like k% growth rules.

Learning reverses the conventional wisdom in two ways. First, policy may be so predictable that it prevents economic agents from arriving at the true parameters of the structural model. Policy must continually vary for limiting beliefs to converge to rational expectations. Second, in dynamic models, optimal multi–period actions may differ greatly from the one–period maximizing policy. There may be incentives for the policy makers to use control variables in such a way as to ensure convergence of beliefs. Through a series of counter–examples in the spirit of Sargent and Wallace(1975), we present the dramatic implications of these conclusions for popular macroeconomic models.

We draw heavily on a wide–ranging literature. In studying the convergence of sequentially updated beliefs, we draw on the work of Taylor(1974), Jordan(1985) and Kiefer and Nyarko(1987). The problem of stochastic control has its origins in Blackwell's(1953) work on experimentation. He reformulated this question into the dynamic programming framework in a subsequent paper, Blackwell(1965). The finite parameter case we study here is closely related to the bandit problem. Rothschild(1974) first introduced this to the economic literature. McLennan(1984) presented an example of incomplete learning by a monopolist facing an unknown demand curve.

Prescott(1972) and Chow(1981) discuss applications of stochastic control to economic policy. Drawing on the recent work of Easley and Kiefer(1988) and McLennnan(1987) we extend this to a more general setting and provide an explicit characterization of the optimal policy.

Section II states sufficient conditions for the almost sure convergence of parameter estimates for a Bayes learner. Learning requires fluctuations in the independent variables, andwhen this is lacking, we have non–convergence to rational expectations.

We consider two examples in Section III in which policy is the critical factor in learning: the Cagan hyper–inflationary model of money demand and an expectations model of the term structure of interest rates. Both fixed rate growth rules and interest pegging are shown to be barriers to learning on the part of agents. The paper poses this paradox: while predictability in the policy environemnt is desirable in that it eliminates uncertainty over expectations of policy variables, in the absence of precise knowledge of the model's parameters, the agent cannot come to learn them over time in these examples.

We turn in Section IV to the issue of optimal policy in a stochastic control situation. We prove conditions that will give the monetary authority incentive to learn the model's paramters. As in the bandit problems, the policy authority trades off current reward for future gains. We offer two counter–examples in which the solution under certainty is very different from the case with parametric uncertainty. Poole's(1970) work on choice of monetary instruments in a stochastic IS/LM model is given a strikingly different interpretation. We show that the monetary authority may choose to use a money stock policy even when the interest pegging policy loss is initially higher. The second model is the Lucas–Sargent–Wallace model of aggregate supply. Here we show that if the monetary authority finds it optimal to learn the true parameters, the k% growth rule can never be optimal. We summarize the implications for rational expectations theory of these findings in a closing section.

NON–CONVERGENCE

The Environment

Let Ω be a complete and separable metric space, let F be its Borel field, and let P be the set of probability measures on F.[1] Consider a stochastic process $\{y_t\}^{\infty}_{t=1}$ defined on the

[1] A useful reference for this material is Billingsley(1979).