ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

LEARNING WITHIN RATIONAL-EXPECTATIONS EQUIILIBRIUM

Thomas W.L. Norman

Number 591
January 2012

Manor Road Building, Oxford OX1 3UQ

Learning Within Rational-Expectations Equilibrium

Thomas W. L. Norman*

Magdalen College, Oxford

January 18, 2012

Abstract

Models of macroeconomic learning are populated by agents who possess a great deal of knowledge of the "true" structure of the economy, and yet ignore the impact of their own learning on that structure; they may learn *about* an equilibrium, but they do not learn *within* it. An alternative learning model is presented where agents' decisions are informed by hypotheses they hold regarding the economy. They periodically test these hypotheses against observed data, and replace them if they fail. It is shown that agents who learn in this way spend almost all of the time approximating rational-expectations equilibria. *Journal of Economic Literature* Classification Number: E00, D83, D84.

Key Words: rational-expectations equilibrium; learning; hypothesis testing.

I.

Ever since rational-expectations equilibrium took its place at the core of macroeconomic analysis, learning has received significant attention as a foundation for this behaviorally demanding concept. The field of "adaptive learning," in particular, has considered whether an economy composed of agents with parametric uncertainty over the "true model" of the economy, might learn their way to rational-expectations equilibrium through econometric estimation (Bray 1982, Radner 1982, Marcet and Sargent 1989, Woodford 1990, Evans and Honkapohja 1999, Evans and Honkapohja 2001). The area of "least squares learning," for example, has produced the concept of "expectational stability" (Evans and Honkapohja 2001), which has been used both to justify and to select amongst rational-expectations equilibria.

Such models are, however, subject to a misspecification problem captured well by Thomas Sargent in his recent Presidential Address to the American Economic Association:

*I am grateful for helpful discussions with Heinrich Nax, Chris Wallace and Peyton Young. Email thomas.norman@magd.ox.ac.uk.

> "We cannot appeal to the same econometrics that lets a rational expectations econometrician learn an equilibrium because an econometrician is *outside* the model and his learning is a sideshow that does not affect the data generating mechanism. It is different when people learning about an equilibrium are inside the model. Their learning affects decisions and alters the distribution of endogenous variables over time, making them aim at moving targets."
>
> Sargent (2008), pp. 13–14.

Using Bray and Kreps' (1987) terminology, agents in these models may learn *about* an equilibrium, but they do not learn *within* it.

A similar nonstationarity is encountered in the related game-theoretic literature on learning to play Nash equilibrium. There, players try to learn their opponents' behavior, but in adapting their responses may in turn alter that very behavior. The question of whether this process settles down on a Nash equilibrium over time is a nontrivial one. Kalai and Lehrer's (1993) model of Bayesian learning delivers convergence to Nash equilibrium under the so-called "grain of truth" assumption that players' prior beliefs place positive probability on their opponents' equilibrium strategies, but this assumption has been criticised as too demanding (Nachbar 1997, 2001, 2005, Foster and Young 2001).[1] Absent any prior information on opponents' strategies or payoffs, the possibility of learning Nash equilibrium is more delicate (Hart and Mas-Colell 2003, 2005). Foster and Young (2003) provide a stochastic model of learning in a repeated game where players respond almost optimally to hypotheses about their opponents' behavior, and periodically test these hypotheses against observed play.[2] They show that such a learning process spends almost all of the time approximating Nash equilibria.

Sargent (2008) highlights the promise of this approach in a macroeconomic setting. Might such a learning process, when applied to a macroeconomic model, spend almost all of the time approximating rational-expectations equilibria? This paper answers this question in the affirmative. Specifically, it models an economy composed of individuals and a government, who repeatedly make decisions (e.g. wage demands and monetary policy, respectively) that collectively determine payoff-relevant economic data. They make those decisions in (almost optimal) response to hypotheses they hold regarding the generation of the data, which they periodically test against observation. When they reject a hypothesis, they adopt a new one at random, and the hypothesis tests are "powerful" in the sense that the probability of type-I and -II errors declines exponentially in the quantity of observed data. The main result shows that, if all agents care about accurately predicting the data, such a learning process spends almost all of the time approximating rational-expectations equilibria.

[1] Blume and Easley (1982) have analyzed an analogous model in the macroeconomic learning literature, where economic agents try to learn which of finitely many models of the economy is true by Bayesian updating from observed data.

[2] This model has some of the characteristics of Kreps' (1998) favored "anticipated utility" approach.

This result is noteworthy in the context of the macroeconomic learning literature for three main reasons. First, under this process, there need be *no* knowledge of either the "true model" of the economy or of other agents' payoffs. Nonetheless, these hypothesis-testing strategies come close to equilibrium. Second, the learning in the model does not treat the environment as stationary. Instead, agents continually test their hypotheses against the data that they have produced; hypotheses that once performed well may cease to do so and be rejected. Third, provided all agents have an incentive to predict the data accurately, the equilibrium concept is rational-expectations equilibrium, rather than the self-confirming equilibrium highlighted by Sargent (2008). Agents play optimally with high probability, but randomly with low probability, so that off-equilibrium-path beliefs cannot go untested in perpetuity, and hence cannot be wrong in long-run equilibrium.

II.

We now present an example, based on one in Sargent (2008), to illustrate both the stability of rational-expectations equilibria and the instability of self-confirming equilibria under our learning process. Letting U be the unemployment rate, π the inflation rate, v the systematic part of the inflation rate chosen by the government, z_1 the average wage, z_2 the output gap, $\theta_0 > 0$ the natural rate of unemployment, $\theta_1 > 0$ the slope of the Phillips curve, and θ_2, θ_3 volatility parameters, U and π in a given period t are given by

$$\begin{aligned} U &= \theta_0 - \theta_1\theta_3 z_2 + \theta_2 z_1 \\ \pi &= v + \theta_3 z_2. \end{aligned}$$

Suppose that $z = (z_1, z_2)$ is determined by the wage bargaining and output decisions $w = (w_1, \ldots, w_n)$ of the n individuals in the economy. Letting $\tilde{z}$ be a 2×1 Gaussian random vector with mean zero and identity covariance, if the government's model in period t is

$$\begin{aligned} U &= \theta_0 - \theta_1\theta_3 \tilde{z}_2 + \theta_2 \tilde{z}_1 \\ \pi &= v^* + \theta_3 \tilde{z}_2, \end{aligned} \tag{1}$$

and its one-period loss function is $\mathrm{E}(U^2 + \pi^2)$, then $v^* = 0$ is its optimal policy response.[3] If, furthermore, individuals share the same model, and their optimal responses $w^* = (w_1^*, \ldots, w_n^*)$ generate the random vector $\tilde{z}$, then the economy is in rational-expectations equilibrium.

But if the government's model is misspecified, denying the natural rate hypothesis

[3](1) is the Lucas (1973) aggregate supply curve.

with, say,

$$U = b_0 - b_1(v + b_3\tilde{z}_2) + b_2\tilde{z}_1$$
$$\pi = v^* + b_3\tilde{z}_2,$$

then its optimal policy response is

$$v^* = h(b) = \frac{-b_1 b_0}{1 + b_1^2}.$$

There then exists a unique self-confirming equilibrium in which

$$b_0 = \theta_0 + \theta_1 h(\theta)$$
$$b_1 = -\theta_1$$
$$b_2 = \theta_2$$
$$b_3 = \theta_3,$$

and the systematic part of inflation is $h(b)$.[4] This is Kydland and Prescott's (1977) time-consistent equilibrium. In an adaptive learning model, the government's parameter estimates converge to this b, from which there is no long-run escape.

However, in the model of learning by hypothesis testing presented here, individuals come close to optimizing, but can make mistakes with positive probability. Therefore, enough mistakes will eventually accumulate for the government's misspecified model to be rejected, and hence for the self-confirming equilibrium to be exited, whereupon there is some minimum probability of agents adopting the true model with the Lucas aggregate supply curve (1). But this rational-expectations equilibrium is also vulnerable to mistakes, with type-I errors leading to disequilibrium with positive probability. Long-run analysis thus involves quantifying the relative time spent in the different equilibria, and here the self-confirming equilibrium (or more generally, rational-expectations disequilibrium) has a disadvantage: failing to reject a model that is incorrect off the equilibrium path is a type-II error, the probability of which declines exponentially in the quantity of observed data under "powerful" hypothesis tests. If enough data is collected then, leaving the self-confirming equilibrium becomes relatively likely. And since leaving the rational-expectations equilibrium requires a type-I error in hypothesis testing, this becomes relatively unlikely. These observations can be used to derive conditions under which almost all of the time is spent approximating rational-expectations equilibrium.

[4] In a self-confirming equilibrium, all agents have correct forecasting distributions for events observed often along the equilibrium path, but possibly incorrect views about events that are rarely observed.

III.

Suppose that there is a set I of *individuals* in the economy, indexed by $i = 1, \ldots, n$, each member of which makes a vector of *decisions* w_i from a finite set W_i, resulting in a von Neumann–Morgenstern utility $u_i(w_i, x)$, where x is a finite set of *variables*. Let $w = (w_1, \ldots, w_n) \in W := \prod_{j=1}^{n} W_i$ be a typical vector of individuals' decisions. Partition $x = [v\ y]$ into a vector v of decisions taken by the *government* G from a finite set V, and a vector $y \in Y$ of all other variables. Let $j = 1, \ldots, n, G$ index the $n+1$ *agents* (individuals plus government) in the economy, and define $m := \max\{|W_1|, \ldots, |W_n|, |V|\}$.

The agents' decisions are repeated through infinite discrete time, $t = 0, 1, 2, \ldots$. Let $x^t = [v^t\ y^t] \in V \times Y = X$ denote the vector of values taken by the variables in period t, i.e. the *data*, which is perfectly monitored by the agents. A *time series* records all of the data over time, and is denoted $\gamma \in \Gamma$; $\overline{\gamma}^t = (x^1, x^2, \ldots, x^t) \in \overline{\Gamma}$ then denotes the *initial time series* in periods 1 through t inclusive, and $\Gamma(\overline{\gamma}^t) = \{\zeta \in \Gamma \mid \overline{\zeta}^t = \overline{\gamma}^t\}$ the set of all *continuations* of the initial time series $\overline{\gamma}^t$. Individual i's *history* $\omega_i \in \Omega_i$ records all of his decisions over time, with $\overline{\omega}_i^t$ denoting an *initial history* and $\Omega_i(\overline{\omega}_i^t)$ its set of continuations.

Each agent j has a *model* of the data-generating process, specifying a conditional probability distribution $p_j^t(x \mid x^{t-1}, b_j)$ on this period's data, given last period's data x^{t-1} and a vector $b_j \in \mathcal{B}$ of at most countably many *parameters*. The space $\mathcal{B}$ of possible parameter vectors is assumed to be a compact metric space with metric d. We also assume that, given any $x^{t-1} \in X$, each $p_j^t(\cdot \mid x^{t-1}, b_j)$ is *Lipschitz continuous* in b_j with Lipschitz constant K. The government's model will obviously specify v^t with certainty, but for convenience we treat it in the same manner as the individuals' models. Let $b = (b_1, \ldots, b_n, b_G) \in \mathcal{B}^{n+1}$ denote a typical vector of models.

Example 1 In the example of Section II, the variables are the systematic part of inflation v and $y = (U, \pi, z)$, the individuals' decisions w are over wage bargaining and output, and the parameter vector b (or θ) is four-dimensional. ■

Example 2: Linear Models Suppose that we allow only linear models with parameters drawn from compact real intervals. Then the model space $\mathcal{B}$ may be identified with a compact subspace of finite-dimensional Euclidean space, and equipped with a standard metric such as that induced by the Euclidean norm, $d(z, z') = \|z - z'\|$. ■

Example 3: Infinite Series Suppose that we allow any model that can be expressed as the sum of a certain infinite series (e.g. power series, Fourier series) with parameters drawn from the unit interval $[0, 1]$. The model space $\mathcal{B}$ may then be identified with the

product of countably many copies of the unit interval, $[0,1]^\infty$, and equipped with the metric

$$d(z,z') = \sum_{k=1}^{\infty} \frac{|z_k - z'_k|}{2^k(1+|z_k - z'_k|)},$$

restricted to $[0,1]^\infty \times [0,1]^\infty$. This metrizes the standard product topology on $[0,1]^\infty$ (Bessaga and Pelczynski 1975); hence, $\mathcal{B} = [0,1]^\infty$ is compact by Tychonoff's Theorem.[5] ■

Each individual i also has a *(behavioral) response* that defines the conditional probability $q_i^t(w_i \mid x^{t-1}, a_i)$ that he chooses w_i in period t, given last period's data x^{t-1} and a parameter vector $a_i \in \mathcal{A}$. Similarly, the government has a response that defines the conditional probability $q_G^t(v \mid x^{t-1}, a_G)$ that it chooses v in period t, given last period's data x^{t-1} and a parameter vector $a_G \in \mathcal{A}$. The space $\mathcal{A}$ of possible parameter vectors is again assumed to be a compact metric space with metric d.[6] We also assume that, given any $x^{t-1} \in X$, each $q_j^t(\cdot \mid x^{t-1}, a_j)$ is *Lipschitz continuous* in a_j with Lipschitz constant $(n+1)/m$. Let $a = (a_1, \ldots, a_n, a_G) \in \mathcal{A}^{n+1}$ denote a typical vector of agents' responses, and let $d(a,a') \leq \sum_{j=1}^{n+1} d(a_j, a'_j)$, $a, a' \in \mathcal{A}^{n+1}$.

Letting $\rho_i < 1$ be individual i's discount factor, and given an initial time series $\overline{\gamma}^{t-1}$, an initial history $\overline{\omega}_i^{t-1}$, and continuations $\gamma \in \Gamma(\overline{\gamma}^{t-1})$ and $\omega_i \in \Omega_i(\overline{\omega}_i^{t-1})$, the future stream of utilities discounted to period t is given by

$$U_i^t(\omega_i, \gamma) = (1-\rho_i) \sum_{t'=t}^{\infty} \rho_i^{t'-t} u_i(w_i^{t'}, x^{t'}).$$

Utilities are normalized so that, for each individual, the maximum utility in any period is one and the minimum utility is zero. Letting μ_{a_i,b_i} be the conditional probability measure over time series and histories induced by the model b_i and the response a_i, individual i's expected utility at time t over all continuations is

$$\mathrm{E}(U_i^t(\omega_i,\gamma) \mid a_i, b_i, x^{t-1}) = \frac{\int_{\Omega_i(\overline{\omega}_i^{t-1}) \times \Gamma(\overline{\gamma}^{t-1})} U_i^t(\omega_i,\gamma) d\mu_{a_i,b_i}}{\int_{\Omega_i(\overline{\omega}_i^{t-1}) \times \Gamma(\overline{\gamma}^{t-1})} d\mu_{a_i,b_i}},$$

which we will call $U_i^t(a_i, b_i)$ to simplify notation. Given a small $\sigma_i > 0$, a_i is a *σ_i-optimal response to b_i* if

$$U_i^t(a_i, b_i) \geq U_i^t(a'_i, b_i) - \sigma_i, \quad \forall t, \forall a'_i.$$

Each individual i is assumed to have a *σ_i-optimal response function* $A_i^{\sigma_i} : \mathcal{B} \to \mathcal{A}$ (mapping from models into σ_i-optimal responses), which in turn is assumed to be *continuous* in b_i and in each payoff $u_i(w_i, x)$, and *diffuse* in the sense that each possible decision

[5] A concise reference on the topological notions and results used in this paper is Sutherland (1975).

[6] This is an abuse of notation, as $\mathcal{A}$'s metric could be different from $\mathcal{B}$'s, but no confusion should result.

$w_i \in W_i$ is made with positive probability.[7] Such an $A_i^{\sigma_i}$ will be called a *σ_i-smoothed best response function*. Similarly, equipping the government with a von Neumann–Morgenstern utility function $u_G(x)$ and discount factor $\rho_G < 1$, we can define $U_G^t(\gamma) = (1-\rho_G)\sum_{t'=t}^{\infty} \rho_G^{t'-t} u_G(x^{t'})$, and the analogous objects $U_G^t(a_G, b_G) = \mathrm{E}(U_G^t(\gamma) \mid a_G, b_G, x^{t-1})$ and $A_G^{\sigma_G} : \mathcal{B} \to \mathcal{A}$. Let $A^\sigma : \mathcal{B}^{n+1} \to \mathcal{A}^{n+1}$ give a typical vector of σ_j-smoothed best response functions.

Suppose that y is determined in each period t via an economic *structure* $y = f_\theta(v^t, w^t)$, where $f_\theta : V \times W \to Y$ given a parameter vector $\theta \in \Theta$.[8] Suppose also that there exists a continuous function $B : \mathcal{A}^{n+1} \to \mathcal{B}$ giving the *correct model* for any given response vector a.[9] Let B^{n+1} denote a vector of $n+1$ copies of B. Agent j, seeking $B(a)$, subjects his model to a hypothesis test at random intervals. A *null hypothesis* H_0 is a statement of the form "the real process generating the data from time t on is described by b_j^t." An alternative hypothesis H_1 replaces b_j^t with another parameter vector in $\mathcal{B}$. If agent j is not conducting a test at the beginning of period t_0, he begins a new test with probability $1/s_j$. Having started a *test phase*, he collects data on how the process evolves over the next s_j periods—a *data sample*—all the while continuing to play $A_j^{\sigma_j}(b_j^{t_0})$. At the end of period $t_0 + s_j$, he conducts the test: if the null hypothesis is accepted, he retains his previous model, i.e. $b_j^{t_0+s_j+1} = b_j^{t_0}$; if it is rejected, he chooses a new model $b_j^{t_0+s_j+1}$ according to a probability measure $h_j^{t_0+s_j+1}(b_j|\overline{\gamma}^{t_0+s_j})$. The chosen new model remains at least until the next hypothesis test completed by j.[10] The h_j-measure of every open τ_0-ball of models is assumed to be *flexible* in the sense of being bounded below by some $h_*(\tau_0)$ that is positive for all $\tau_0 > 0$. The following result entitles us to make this assumption.

Lemma 1 *There exists a flexible probability measure h_j.*

Proof. Consider a cover $\mathcal{U}$ of $\mathcal{B}$ composed of infinitely many open τ_0-balls. By compactness of $\mathcal{B}$, $\mathcal{U}$ has a finite subcover $\mathcal{V}$. Let h_j allocate measure 1 uniformly over the elements of $\mathcal{V}$. ∎

Given the initial time series $\overline{\gamma}^{t_0-1}$, let $\alpha_{j,s_j}(b_j^{t_0}, \overline{\gamma}^{t_0-1})$ be the probability of making a type-I error by rejecting a correct $b_j^{t_0}$; and given an alternative model $b_j \neq b_j^{t_0}$, let $\beta_{j,s_j}(b_j, b_j^{t_0}, \overline{\gamma}^{t_0-1})$ be the probability of making a type-II error by failing to reject $b_j^{t_0}$ in favour of a correct b_j. Now, given a small *tolerance level* $\tau > 0$, define

$$\beta_{j,s_j,\tau} = \sup_{\overline{\gamma},t,b_0} \sup_{b_j:\, d(b_j,b_0)>\tau} \beta_{j,s_j}(b_j, b_0, \overline{\gamma}^{t-1}),$$

[7] Given the Lipschitz continuity of models and responses in their parameters, continuity of $A_i^{\sigma_i}$ can always be satisfied.

[8] There is, however, no assumed common knowledge of f_θ, and n is assumed to be large enough that each individual ignores the effect of his own decisions on y.

[9] This assumption is more reasonable the richer the model space.

[10] This new model can—though need not—be viewed as the updating of j's belief given his new information. Indeed, Foster and Young (2003) show, under an additional assumption, that hypothesis testing is almost rational.

and

$$\alpha_{j,s_j} = \sup_{\overline{\gamma},t,b_0} \alpha_{j,s_j}(b_0, \overline{\gamma}^{t-1}).$$

A *powerful* family of tests is assumed to be employed by each agent j, in the Foster and Young (2003) sense that there exist functions $k_j(\tau) > 0$ and $r_j(\tau) > 0$ such that, for every tolerance $\tau > 0$, there is a test in the family such that[11]

$$\alpha_{j,s_j} \leq k_j(\tau)\mathrm{e}^{-r_j(\tau)s_j} \quad \text{and} \quad \beta_{j,s_j,\tau} \leq k_j(\tau)\mathrm{e}^{-r_j(\tau)s_j}.$$

Define, analogously with $\beta_{j,s_j,\tau}$, the least upper bound $\alpha_{j,s_j,\tau}$ on rejecting the null when the true distribution is within τ of the null. The following result adapts Foster and Young's (2003) Lemma 1 to the present setting.

Lemma 2 *Suppose that under the null hypothesis the conditional probability of observing any given combination of decisions in any period is at least ξ. Then for every $\tau > 0$,*

$$\alpha_{j,s_j,\tau} \leq \left(1 + \frac{K\tau}{\xi}\right)^{s_j} \alpha_{j,s_j}.$$

Proof. For notational convenience, consider the situation where the test is applied to the data $\overline{\gamma}^{s_j}$ generated in the first s_j periods. Denote the null distribution on the initial time series by $p(\overline{\gamma}^{s_j} \mid b_0)$, and define the *rejection set* R to be the set of $\overline{\gamma}^{s_j}$ whose observation leads to the rejection of H_0. Let $p(\overline{\gamma}^{s_j} \mid b_j)$ be the distribution resulting from any $b_j \neq b_0$ lying within τ of b_0. Then, by Lipschitz continuity of each p_j^t,

$$\left|p(x^t|x^{t-1}, b_j) - p(x^t|x^{t-1}, b_0)\right| \leq K\tau \quad \text{for all } t \leq s_j.$$

We wish to show that if $p(R \mid b_0) \leq \alpha$, then

$$p(R \mid b_j) \leq \left(1 + \frac{K\tau}{\xi}\right)^{s_j} \alpha.$$

Since $p((v^t, y^t)|x^{t-1}, b_0) = p_G(v^t \mid x^{t-1}, b_G) \prod_{i=1}^n p_i(w_i^t \mid x^{t-1}, b_0) \geq \xi$ for each $t \leq s_j$, where $(v^t, w^t) = f^{-1}(y^t)$, it follows that

$$\frac{p(x^t|x^{t-1}, b_j)}{p(x^t|x^{t-1}, b_0)} \leq 1 + \frac{K\tau}{\xi}.$$

[11] This assumption is more demanding the richer the model space.

So,

$$\begin{aligned} p(R \mid b_j) = \int_R p(\overline{\gamma}^{s_j} \mid b_j)\, d\overline{\gamma}^{s_j} &= \int_R \prod_{t=1}^{s_j} p(x^t \mid x^{t-1}, b_j)\, d\overline{\gamma}^{s_j} \\ &\leq \int_R \prod_{t=1}^{s_j} \left(1 + \frac{K\tau}{\xi}\right) p(x^t \mid x^{t-1}, b_0)\, d\overline{\gamma}^{s_j} \\ &= \left(1 + \frac{K\tau}{\xi}\right)^{s_j} \int_R \prod_{t=1}^{s_j} p(x^t \mid x^{t-1}, b_0)\, d\overline{\gamma}^{s_j} \\ &= \left(1 + \frac{K\tau}{\xi}\right)^{s_j} p(R \mid b_0), \end{aligned}$$

from which the desired conclusion follows. ■

Define $s^* = \max_j s_j$ and $s_* = \min_j s_j$.

Lemma 3 *For each tolerance τ, there is a value $c(\tau) \leq \tau$ such that*

$$\forall j, \quad \alpha_{j,s_j,c(\tau)}, \beta_{j,s_j,\tau} \leq k(\tau) e^{-r(\tau)s_*}.$$

Proof. Let $\xi_j(\sigma_j)$ denote the minimum probability with which agent j chooses each decision in each period; diffuse response functions imply that this is strictly positive. Every combination of decisions occurs with probability at least $\xi = \prod_{j=1}^{n+1} \xi_j(\sigma_j)$. Letting

$$c_j(\tau) = \frac{r_j(\tau)\xi}{2K},$$

Lemma 2 implies that

$$\alpha_{j,s_j,c_j(\tau)}, \beta_{j,s_j,\tau} \leq \left(1 + \frac{Kc_j(\tau)}{\xi}\right)^{s_j} \alpha_{j,s_j}.$$

Since j employs a powerful family of tests, $\alpha_{j,s_j} \leq k_j(\tau)\mathrm{e}^{-r_j(\tau)s_j}$, so that

$$\begin{aligned} \alpha_{j,s_j,c_j(\tau)}, \beta_{j,s_j,\tau} &\leq \left(1 + \frac{Kc_j(\tau)}{\xi}\right)^{s_j} k_j(\tau)\mathrm{e}^{-r_j(\tau)s_j} \\ &= \left(1 + \frac{r_j(\tau)}{2}\right)^{s_j} k_j(\tau)\mathrm{e}^{-r_j(\tau)s_j} \\ &\leq \mathrm{e}^{r_j(\tau)s_j/2} k_j(\tau)\mathrm{e}^{-r_j(\tau)s_j} = k_j(\tau)\mathrm{e}^{-r_j(\tau)s_j/2}. \end{aligned}$$

Moreover, under powerful tests,

$$\beta_{j,s_j,\tau} \leq k_j(\tau)\mathrm{e}^{-r_j(\tau)s_j} \leq k_j(\tau)\mathrm{e}^{-r_j(\tau)s_j/2}.$$

Finally, define $c(\tau) = \min_j c_j(\tau)$, $r(\tau) = \min_j r_j(\tau)/2$ and $k(\tau) = \max_j k_j(\tau)$. ■

Agents are assumed to *use comparable amounts of data*, in the sense that there exists a $p \in (1,2)$ such that $s_j \leq s_{j'}^p$ for every j and j'. We say that *prediction matters by at most* ϵ for agent j if he has a response that is ϵ-optimal for every b_j. We say that he is an *ϵ-good predictor* if

$$\lim_{T\to\infty} \sup \frac{1}{T} \sum_{t=1}^{T} d(b_j^t, B(a^t))^2 \leq \epsilon \quad \text{a.s.},$$

i.e. if the mean square error of his predictions is almost surely bounded above by ϵ as $t \to \infty$. a^t is an *ϵ-self-confirming equilibrium* at time t if

$$U_j^t(a_j^t, B(a^t)) \geq \max_{a_j'} U_j^t(a_j', B(a^t)) - \epsilon, \quad \forall j.$$

If a^t is an ϵ-self-confirming equilibrium and all agents are ϵ-good predictors, then a^t is called an *ϵ-rational-expectations equilibrium*.

Theorem 1 *Given any $\epsilon > 0$, if the σ_j are small (given ϵ), if the test tolerances τ_j are sufficiently fine (given ϵ and σ_j) and if the amounts of data collected, s_j, are sufficiently large (given ϵ, σ_j and τ) then:*

1. *the responses are ϵ-self-confirming equilibria at least $1-\epsilon$ of the time; and*
2. *all agents for whom prediction matters by at least ϵ are ϵ-good predictors.*

Corollary 1 *If prediction matters by at least ϵ to all agents, then the responses are ϵ-rational-expectations equilibria at least $1-\epsilon$ of the time.*

The theorem is implied by the following lemma, whose proof shares some of the structure of Foster and Young's (2003) proof of their Lemma 2.

Lemma 4 *Given any $\epsilon > 0$, there exist functions $\sigma(\epsilon)$, $\tau(\epsilon, \sigma)$, $s(\epsilon, \sigma, \tau)$ such that if these functions bound the parameters with the same names, then at least $1-\epsilon$ of the time t,*

1. *$d(a^t, A^\sigma(B^{n+1}(a^t))) \leq \epsilon/2$.*
2. *$|U_j^t(a_j^t, B(a^t)) - \max_{a_j'} U_j^t(a_j', B(a^t)| \leq \epsilon$ for all j.*
3. *$d(b_j^t, B(A^\sigma(b^t))) \leq \epsilon$ for every agent i for whom prediction matters by at least ϵ.*

Proof. Let a^* be an arbitrary fixed point of the mapping $A^\sigma \circ B^{n+1}$ (which must exist as A^σ is a continuous function from a compact convex space into itself); it follows that a vector b^* of $n+1$ copies of $B(a^*)$ is a fixed point of $B^{n+1} \circ A^\sigma$. Choose $\sigma_j \leq \epsilon/2$, $\forall j$. Since

each $A_j^{\sigma_j}(\cdot)$ is continuous in b_j and in u_j, both of which lie in compact domains, there is a $\delta > 0$ such that

$$\forall u, t, j, \quad \forall b_j, b'_j, \quad d(b_j, b'_j) \leq \delta \quad \Rightarrow \quad d\left(A_j^{\sigma_j}(b_j), A_j^{\sigma_j}(b'_j)\right) \leq \frac{\epsilon}{2(n+1)}, \tag{2}$$

and

$$\delta < \frac{\epsilon}{2(n+1)}. \tag{3}$$

Letting $r_j > 0$ be the diameter of the image of $A_j^{\sigma_j}(\mathcal{B})$ in $\mathcal{A}$ (i.e. the maximal distance between two points belonging to the image), agent j is *unresponsive* if $r_j \leq \delta$ (and *responsive* otherwise). Since, for any given $x^{t-1} \in X$, $r_j \leq \delta$ implies that $\sup_{b_j, b'_j \in \mathcal{B}} |q_j^t(w_j|x, A_j^{\sigma_j}(b_j)) - q_j^t(w_j|x, A_j^{\sigma_j}(b'_j))| \leq (n+1)\delta/m < \epsilon/2m$ by Lipschitz continuity of each q_j^t, and hence that $\sup_{b_j, b'_j \in \mathcal{B}} \sum_{w_j \in W_j} |q_j^t(w_j|x, A_j^{\sigma_j}(b_j)) - q_j^t(w_j|x, A_j^{\sigma_j}(b'_j))| < \epsilon/2$, agent j must be responsive if prediction matters to j by more than ϵ. We will assume that prediction matters to the government by more than ϵ (otherwise policymaking is a trivial problem).

Fix a small tolerance level

$$\tau \leq \delta. \tag{4}$$

We say that a model vector b is *good* for j if

$$d(b_j, B(A^\sigma(b))) \leq \tau; \tag{5}$$

otherwise it is *bad* for j. b is *all good* if it is good for all agents; it is *fairly good* if it is good for all responsive agents, and it is *bad* if it is not fairly good.

Claim 1 *If the model vector b^t is fairly good at least $1 - \epsilon$ of the time, then all three statements of Lemma 4 follow.*

Proof. By equations (4) and (3), $\tau < \epsilon$. If b^t is fairly good then, it follows that $d(b_j, B(A^\sigma(b))) \leq \epsilon$ for responsive agents, and hence all agents for whom prediction matters by at least ϵ (since $\epsilon > \delta$). This establishes statement 3 of the lemma.

Next we establish statement 1. For every responsive agent j, (5) and (4) imply that $d(b_j, B(A^\sigma(b))) \leq \tau \leq \delta$. Hence, by (2), $d(A_j^{\sigma_j}(b_j), A_j^{\sigma_j}(B(A^\sigma(b)))) \leq \epsilon/2(n+1)$ for each responsive agent, i.e. $d(a_j, A_j^{\sigma_j}(B(A^\sigma(b)))) \leq \epsilon/2(n+1)$. For each unresponsive agent, meanwhile, $d(a_j, A_j^{\sigma_j}(B(A^\sigma(b)))) \leq \delta < \epsilon/2(n+1)$. Putting these together gives $d(a, A^\sigma(B^{n+1}(a))) \leq \epsilon/2$.

It remains to show statement 2. For each j, the response $A_j^{\sigma_j}(B(a^t))$ involves a utility loss of at most $\sigma_j \leq \epsilon/2$ as compared to a best response to $B(a^t)$, say $\hat{a}_j$. From above, we know that $d(a_j^t, A_j^{\sigma_j}(B(a^t))) \leq \epsilon/2(n+1)$, so that for any given $x^{t-1} \in X$, $|q_j^t(w_j|x^{t-1}, a_j^t) - q_j^t(w_j|x^{t-1}, A_j^{\sigma_j}(B(a^t)))| \leq \epsilon/2m$ by Lipschitz continuity of each q_j^t, and

$\sum_{w_j \in W_j} |q_j^t(w_j|x^{t-1}, a_j^t) - q_j^t(w_j|x^{t-1}, A_j^{\sigma_j}(B(a^t)))| \leq \epsilon/2$. Hence, and because the utility functions are bounded between 0 and 1, the payoff difference between a_j^t and $A_j^{\sigma_j}(B(a^t))$ is at most $\epsilon/2$. Thus, $|U_j^t(a_j^t, B(a^t)) - U_j^t(\hat{a}_j, B(a^t))| \leq \epsilon/2 + \epsilon/2 = \epsilon$. ■

Claim 2 *The model vector b^t is fairly good at least $1-\epsilon$ of the time.*

Proof. Recall from Lemma 3 that, for each τ, there is a real number $0 < c(\tau) < \tau$ such that whenever agents' models are within $c(\tau)$ of the truth they reject with exponentially small probability, and when their models are at least τ away from the truth they reject with probability close to one. In particular, there exist functions $k(\tau)$ and $r(\tau)$ such that whenever an agent's model is within $c(\tau)$ of the correct model, he rejects with probability at most

$$k(\tau)\mathrm{e}^{-r(\tau)s_*}.$$

Next, we claim that there is a value $\psi > 0$ such that, for our chosen model fixed point b^*,

$$\forall j, \quad d(b_j, b_j^*) < \psi \quad \Rightarrow \quad d(b_j, B(A^\sigma(b))) < c(\tau). \tag{6}$$

Indeed, B is continuous on the compact domain $\mathcal{A}^{n+1}$, giving uniform continuity. Hence, there exists $\chi > 0$ such that

$$\forall a, a', \quad d(a, a') < \chi \quad \Rightarrow \quad d(B(a), B(a')) < \frac{c(\tau)}{2}.$$

Similarly, since $A_j^{\sigma_j}$ is continuous on the compact domain $\mathcal{B}$ for each j, there exists $\psi > 0$ such that

$$\forall j, \quad \forall b_j, b_j', \quad d(b_j, b_j') < \psi \quad \Rightarrow \quad d\left(A_j^{\sigma_j}(b_j), A_j^{\sigma_j}(b_j')\right) < \chi/(n+1).$$

Choosing $\psi < c(\tau)/2$, it follows that

$$\begin{aligned} d(B(A^\sigma(b)), b_j) &\leq d(B(A^\sigma(b)), b_j^*) + d(b_j^*, b_j) \\ &= d(B(A^\sigma(b)), B(A^\sigma(b^*))) + d(b_j^*, b_j) \\ &< \frac{c(\tau)}{2} + \frac{c(\tau)}{2}. \end{aligned}$$

This establishes (6). The values of δ, τ and ψ defined in (3), (4) and (6) will remain fixed for the remainder of the proof.

Letting $g > 0$ be the diameter of the image of $B(a_1, \ldots, a_n, A_G^{\sigma_G}(\mathcal{B}))$ in $\mathcal{B}$, we will assume that the government is *potent* in the sense that $g > 2(n+1)\tau$. Next we establish the existence of a "wrong" model for a potent government, which invalidates all the individuals' models. We can choose $n+1$ points in the image set $B(a_1, \ldots, a_n, A_G^{\sigma_G}(\mathcal{B}))$ with no two points closer than 2τ, where τ is defined in (4). Hence there exist at least

two of these points, say b'_I and b''_I, such that

$$\forall i, \quad d(b_i, b'_I) \geq \tau \quad \text{and} \quad d(b_i, b''_I) \geq \tau.$$

Further, one or both of b'_I, b''_I is at least τ from b^*_i—say b'_I without loss of generality. Then we define $\zeta(b)$ to be any model b'_G such that $B(a_1, \ldots, a_n, A^{\sigma_G}_G(b'_G)) = b'_I$ and call this a *wrong model for the government given* b.

A *great state* is a state such that, for every agent j, j's model is within ψ of b^*_j and no agent is currently in a test phase. Consider the following sequence of events leading from a bad vector b^t to a great state.

Step 1. If the current model vector is good for G, an individual i with a bad model starts a new test phase as soon as all current tests are completed. After this test phase—during which no other agent starts a test—he rejects his current hypothesis and adopts a new model within ψ of $B(a^t)$. Repeat this step until (a great state is reached or) the model vector is bad for G, which occurs after a finite number m of tests with probability at least $g_m > 0$. (Duration: $(m+1)s^*$ periods.)

Step 2. The government starts a new test phase. During this test phase, no individual starts a test. Once its test phase is completed, the government rejects its current hypothesis and adopts a model within ψ of $\zeta(b^{t+(m+1)s^*})$. (Duration: at most s^* periods.)

Step 3. In succession, the individuals conduct tests on non-overlapping sets of data (whilst the government does not test). At the end of each of these n tests, the tester rejects and adopts a model within ψ of his part of b^*. (Duration: at most ns^* periods.)

Step 4. The government starts a test phase and no individuals test while it is in progress. It rejects this test and jumps to a point within ψ of b^*_G. (Duration: at most s^* periods.)

Step 5. If the number of periods in Steps 1–3 is $T < (n+m+3)s^*$, no player begins a test for the next $(n+m+3)s^* - T$ periods.

The duration of the whole sequence is exactly $(n+m+3)s^*$.

Each of the $n+m+2$ tests (following the initial partial test phase) ends with a rejection, and each such rejection occurs with probability at least $1/2$ under appropriately chosen test parameters. After rejection, a target of radius ψ must be hit within the rejector's model space. These $n+m+2$ events have probability at least $(h_*/2)^{n+m+2}$, where $h_* = h_*(\psi)$. In addition, each of the $n+m+2$ test phases must begin at a specific time and no other agents can be testing during another's phase; the probability of this is

at least

$$\left((1/s^*)(1-1/s_*)^{ns^*}\right)^{n+k+2}. \tag{7}$$

We can assume that all $s_i \geq 2$, so $(1-1/s_*)^{s_*} \geq 1/4$ and (7) is bounded below by

$$\left((1/s^*)(1/4)^{ns^*/s_*}\right)^{n+k+2}.$$

Finally, all agents must cease testing until $(n+m+3)s^*$ periods have elapsed from the start of the sequence, which occurs with probability at least

$$(1-1/s_*)^{(n+1)(n+m+3)s^*} \geq 1/4^{(n+1)(n+m+3)s^*/s_*}.$$

Putting all of this together, there is a positive integer N such that Steps 1–5 occur with probability at least

$$4^{-Ns^*/s_*} g_m (h_*/2s^*)^{n+m+2}.$$

By assumption $s^* \leq s_*^p$ for some $p \in (1,2)$, so $s^*/s_* \leq s_*^q$ where $q = p-1 > 0$. Hence, the probability of Steps 1–5 is at least

$$\begin{aligned} 4^{-Ns_*^q} g_m (h_*/2s_*^p)^{n+m+2} &= \exp(-\beta s_*^q + \ln\alpha - p(n+m+2)\ln s_*) \\ &\geq \exp(-\beta s_*^q + \ln\alpha - p(n+m+2)s_*^{q'}), \end{aligned}$$

where $\alpha, \beta > 0$ and $q' < 1$ are constants. This establishes the following fact.

If b^t is bad, the probability of being in a great state at time $t+(n+m+3)s^*$ is at least

$$\eta \equiv \exp(-\beta s_*^q + \ln\alpha - p(n+m+2)s_*^{q'}). \tag{8}$$

Suppose now that the process is in a great state at time t. Letting T be a large positive integer, we will compute the probability that no agent rejects a test over the next T periods. By construction, each null hypothesis is within $\tau_0 = c(\tau)$ of the truth, so the probability that a given agent j rejects a test is at most

$$\alpha_{j,s_j,\tau_0} \leq k_0 \mathrm{e}^{-r_0 s_i} \leq k_0 \mathrm{e}^{-r_0 s_*},$$

where $k_0 = k(\tau_0) > 0$ and $r_0 = r(\tau_0) > 0$.

At most nT/s_* tests are concluded over the course of the next T periods. Hence the probability that any agent rejects a hypothesis during periods $t+1, \ldots, t+T$ is bounded above by

$$(nT/s_*)k_0 \mathrm{e}^{-r_0 s_*} < T\mathrm{e}^{-4rs_*},$$

the inequality holding for all sufficiently large s_* and some $r > 0$.

We can now show that the process is in a bad state for a very small fraction of the

time. Letting $T = \mathrm{e}^{3rs_*}$, the probability of leaving a great state is at most

$$\mathrm{e}^{-rs_*}. \tag{9}$$

Starting from a given time t, let $\mathcal{E}$ denote the event "the realized states in at least ϵT of the periods $t+1, \ldots, t+T$ are bad." $\mathcal{E}'$ will then denote the sub-event of $\mathcal{E}$ in which no great state occurs before the last bad state, and $\mathcal{E}'' = \mathcal{E} - \mathcal{E}'$.

If $\mathcal{E}'$ occurs, there are at least $\lfloor \epsilon T/(n+m+3)s^* \rfloor = k$ distinct times $t < t_1 < \cdots < t_k \leq t+T$ such that the following hold:

- $t_{l+1} - t_l \geq (n+m+3)s^*$ for $1 \leq l < k$,
- the state at time t_l is bad for $1 \leq l < k$,
- no great state occurs from t_1 to t_k.

From above, the probability of this event is at most $(1-\eta)^{k-1} < \mathrm{e}^{-\eta(k-1)}$.

Note that, since ϵ and n are fixed, we have $k - 1 > \mathrm{e}^{2rs_*}$ for all sufficiently large s_*.

Since $q, q' < 1$, (8) implies that $\eta > \mathrm{e}^{-rs_*}$ for all sufficiently large s_*; hence,

$$P(\mathcal{E}') \leq (1-\eta)^{k-1} \leq \mathrm{e}^{-\eta(k-1)} \leq \mathrm{e}^{-\mathrm{e}^{rs_*}},$$

which can be made as small as we wish—in particular, less than $\epsilon/2$—when s_* is large.

If $\mathcal{E}''$ occurs, the process does *not* remain in good states for at least T periods after a great state, so by (9)

$$P(\mathcal{E}'') \leq \mathrm{e}^{-rs_*},$$

which can also be made less than $\epsilon/2$ when s_* is sufficiently large. Putting all of this together it follows that, for all sufficiently large s_*,

$$P(\mathcal{E}) = P(\mathcal{E}') + P(\mathcal{E}'') \leq \epsilon.$$

Dividing all times t into disjoint blocks of length T, let Z_k be the fraction of bad times in the kth block. Having just shown that $P(Z_k \geq \epsilon) \leq \epsilon$ for all k, it follows that

$$\mathrm{E}(Z_k) \leq P(Z_k \geq \epsilon) \cdot 1 + P(Z_k < \epsilon) \cdot \epsilon \leq 2\epsilon.$$

Hence, the proportion of times that the process spends in bad states is almost surely less than 2ϵ. Rerunning the entire argument with $\epsilon/2$ establishes that b is fairly good at least $1-\epsilon$ of the time, establishing the claim. ■

The lemma follows from the two claims. ■

References

Bessaga, C., and A. Pelczynski (1975): *Selected Topics in Infinite-Dimensional Topology*. Polish Scientific Publishers, Warszawa.

Blume, L. E., and D. Easley (1982): "Learning to Be Rational," *Journal of Economic Theory*, 26, 318–339.

Bray, M. M. (1982): "Learning, Estimation, and the Stability of Rational Expectations," *Journal of Economic Theory*, 26, 340–351.

Bray, M. M., and D. M. Kreps (1987): "Rational Learning and Rational Expectations," in *Arrow and the Ascent of Modern Economic Theory*, ed. by G. R. Feiwel, pp. 597–625. New York University Press, New York.

Evans, G. W., and S. Honkapohja (1999): "Learning Dynamics," in *Equilibrium Analysis: Essays in Honor of Kenneth J. Arrow*, ed. by J. B. Taylor, and M. Woodford. Elsevier Science, North Holland, Amsterdam.

——— (2001): *Learning and Expectations in Macroeconomics*. Princeton University Press, Princeton, New Jersey.

Foster, D. P., and H. P. Young (2001): "On the Impossibility of Predicting the Behavior of Rational Agents," *Proceedings of the National Academy of Sciences of the USA*, 98(22), 12848–12853.

——— (2003): "Learning, Hypothesis Testing, and Nash Equilibrium," *Games and Economic Behavior*, 45, 73–96.

Hart, S., and A. Mas-Colell (2003): "Uncoupled Dynamics Do Not Lead to Nash Equilibrium," *American Economic Review*, 93, 1830–1836.

——— (2006): "Stochastic Uncoupled Dynamics and Nash Equilibrium," *Games and Economic Behavior*, 57, 286–303.

Kalai, E., and E. Lehrer (1993): "Rational Learning Leads to Nash Equilibrium," *Econometrica*, 61, 1019–1045.

Kreps, D. M. (1998): "Anticipated Utility and Dynamic Choice," in *Frontiers of Research in Economic Theory: The Nancy L. Schwartz Memorial Lectures, 1983–1997*, ed. by D. P. Jacobs, E. Kalai, and M. I. Kamien. Cambridge University Press, Cambridge, UK.

Kydland, F. E., and E. C. Prescott (1977): "Rules Rather Than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, 85, 473–491.

Lucas, R. E. (1973): "Some International Evidence on Output-Inflation Tradeoffs," *American Economic Review*, 63, 326–334.

Marcet, A., and T. J. Sargent (1989): "Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models," *Journal of Economic Theory*, 48, 337–368.

Nachbar, J. H. (1997): "Prediction, Optimization, and Learning in Repeated Games," *Econometrica*, 65, 275–309.

——— (2001): "Bayesian Learning in Repeated Games of Incomplete Information," *Social Choice and Welfare*, 18(2), 303–326.

——— (2005): "Beliefs in Repeated Games," *Econometrica*, 73, 459–480.

Radner, R. (1982): "Equilibrium Under Uncertainty," in *Handbook of Mathematical Economics*, ed. by K. J. Arrow, and M. D. Intriligator, vol. 2. North-Holland, Amsterdam.

Sargent, T. J. (2008): "Evolution and Intelligent Design," *American Economic Review*, 98, 5–37.

Sutherland, W. A. (1975): *Introduction to Metric and Topological Spaces*. Oxford University Press, Oxford.

Woodford, M. (1990): "Learning to Believe in Sunspots," *Econometrica*, 58, 277–307.