Learning to be Credible

In-Koo Cho[1] and Thomas J. Sargent[2]

January 1997

This paper was prepared for presentation at the conference to celebrate the Bank of Portugal's 150^{th} birthday.

[1] Brown University

[2] University of Chicago and Hoover Institution, Stanford University

Learning to be Credible

1. Credibility as Conforming

The birthday of a Central Bank provides a suitable occasion to describe our recent attempts to analyze the *acquisition* of macroeconomic credibility. Central banks care about how credibility can be sustained and how it can be earned. Recent work on credibility has deepened our appreciation for how difficult it can be for a well motivated central bank to *earn* a good reputation for pursuing low inflation.

The theory of credible economic policy demonstrates the possibility of attaining good macroeconomic outcomes if a system of expectations prevails that provides policy makers with incentives not to use inflation to ameliorate unemployment. But this theory describes multiple equilibrium systems of expectations, many with bad outcomes. The multiplicity of expectations systems is integral to the theory, a product of insisting, via backward induction, on a complete (i.e., perfect) theory of rational behavior. Backward induction requires the theory to describe the self-interested behavior of the government under all possible observed histories and under various systems of expectations.

The theory of credible policy describes the behavior of policy makers who find it in their interests to conform to the public's expectations about their behavior. Within the theory, it is impossible to distinguish between the government's strategy as an object reflecting the government's choices, or as a descriptor of the public's expectations about the government's behavior. The pure theory is silent in terms of advice about how a government might go about acquiring a good reputation (i.e., by somehow manipulating the public's views about it).

To coax from the theory a prediction about which equilibrium system of expectations is likely to be observed requires appealing to something that would make us attach more likelihood to some of the equilibria. We need to step outside the model, into the domain of 'stability theory', to find a method of equilibrium selection.

A purpose of studying *learning* in macroeconomics has been to select among equilibria. Woodford (1990) and Marcet and Sargent (1989b) have used adaptive least squares learning schemes to select equilibria in macroeconomic models with large numbers of rational expectations equilibria. Here we describe some of our efforts to apply broadly similar learning algorithms to contexts in which the multiplicity of equilibria resides in the credibility aspect. Briefly, our research program is to adapt the least squares learning method so that it applies to repeated economies, where the object being learned about is a reputation. An impediment to developing a theory of learning for reputations is the high dimension of the *state* usually used to describe a 'reputation'. The key object in the theory is a repeated-game or repeated-economy strategy, defined as a sequence of functions mapping histories of outcomes into current outcomes. The space of strategies is large because the space of possible histories is so large.

We compose our theory of learning in two broad steps. First, we restrict strategies to a space of manageable dimension, without throwing out 'too much.' This we achieve by encoding strategies using simple neural networks fed by summary statistics of historical outcomes. By not throwing out 'too much', we refer to Cho's (1995) demonstration that this set of strategies is large enough to recover the many payoffs supportable in the folk theorem. Second, we adapt the least squares algorithms of Woodford (1990) and Marcet and Sargent (1989a) to apply to the neural networks. We tamper with details of the algorithms to build in aspects of experimentation,[1] not needed in the settings of Woodford and Marcet and Sargent, but needed here.

We describe our ideas in the context of a classic macroeconomic example, Kydland and Prescott's 'Phillips curve' example. This example has been a laboratory for studying credibility in macroeconomics, and it serves our purposes well. We shall first describe Kydland's and Prescott's one-period economy, then do our work with a repeated version of it.

The remainder of the paper is organized as follows. Section 2 describes the example one-period economy. By exhibiting the inferiority of a Nash equilibrium outcome *vis a vis* a Ramsey outcome, we show the value of a commitment mechanism. Section 3 describes the infinitely repeated economy. It describes 'linear strategies' as economical devices for encoding 'reputations'. We adapt to our economy Cho's (1995) 'folk theorem', expressed in terms of these linear strategies. Section 3 is about equilibrium representation, a necessary prologue to formulating an adaptive theory of learning. Section 4 formulates the learning problem in terms of a (2×2) version of the economy, using linear strategies to parameterize peoples' beliefs. We state a theorem about the possible limit points of the learning algorithm. Relative to the 'folk theorem', our theorem narrows the set of possible outcomes: it states that under the learning algorithm, only the Ramsey outcome and the Nash outcome eventually occur with positive probability. This is an encouraging result in terms of equilibrium selection. But simulations of the section 4 model are less encouraging in terms of macroeconomics because the Nash outcome occurs most of the time. This reaffirms the value of a commitment mechanism. Section 5 describes our incomplete efforts to extend the analysis to the case of continuous action spaces. A reader not interested in these extensions may skip this section and move directly to section 6, which states our conclusions.

[1] See Fudenberg and Kreps (1994).

2. One-period Economy

The one-period model is like one used by Kydland and Prescott (1977), stated in terms of concepts used by Stokey (1989, 1990). Let U_t, y_t, x_t be the unemployment rate, the government's policy action, and the public's (average) expectation of the government's policy action at t, respectively. Kydland and Prescott took y to be the rate of inflation and x to be the public's expectation of the rate of inflation. The government's one-period payoff is

$$-.5(U_t^2 + y_t^2). \tag{1}$$

Unemployment is determined by an 'expectational Phillips curve'

$$U_t = U^* - \theta(y_t - x_t), \quad \theta > 0. \tag{2}$$

Substituting (2) into (1) gives the one-period payoff of the government as the function $v(x, y)$ defined by

$$v(x_t, y_t) = -.5[(U^* - \theta(y_t - x_t))^2 + y_t^2]. \tag{3}$$

There is a continuum of private agents, each of whose choice is to set ξ_t, its expectation about y_t.[2] The average over all households' settings of ξ_t is x_t. To capture that each private agent solves a forecasting problem, we assume that the one period payoff function of a private agent is $u(\xi, x, y)$ defined by

$$u(\xi, x, y) = -.5[(y - \xi)^2 + y^2]. \tag{4}$$

Given y, each agent maximizes its payoff (solves its forecasting problem) by setting $\xi = y$. Since all private agents face the same problem, $x = \xi$.

For some of this paper, we specify that $y_t \in Y$, and $x_t \in X \equiv Y$. We assume that Y is compact subset of the real line, $Y = [0, y^{\#}]$, where $y^{\#} > 0$. However, to simplify the analysis, we shall also first work in the spirit of Stokey (1989, 1991) and analyze a numerical example in which $Y = X = \{C, D\}$, a two element set, where C denotes low inflation and D denotes high inflation.

We work with the following objects.

Rational Expectations Equilibrium: A *rational expectations equilibrium* is a triple (U, x, y) satisfying (2) and $y = x$.

Government Best Response: Given the public's expectation x, a government *best response* satisfies $y = \arg\max_y w(x, y)$. Let $y = B(x)$ be the best response function.

[2] We follow Chari and Kehoe (1989) and Stokey (1989) in distinguishing between the ξ chosen by the representative agent and the average of ξ over all such agents, x.

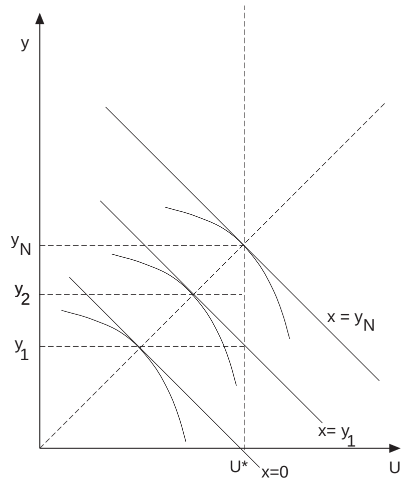

Figure 1. Nash equilibrium, Ramsey outcome, and 'best response dynamics'. The straight lines depict a family of Phillips curves for different levels of expected inflation x, with slope $-\theta = -1$; curves are drawn for $x = 0, y_1, y_N$. The Nash equilibrium outcome is U^*, y_N. The Ramsey outcome is $U^*, 0$. The government's best response setting for y, given x, occurs at the tangency of an indifference equation induced by (1) with the Phillips curve indexed by x.

NASH EQUILIBRIUM: A *Nash equilibrium* is a pair (x, y) satisfying (i.) $x = y$, and (ii.) $y = B(x)$.

RAMSEY PROBLEM (OUTCOME): The *Ramsey problem* is $\max_y w(y, y)$. The *Ramsey outcome* is the value of y that attains the maximum of the Ramsey problem.

MIN-MAX ACTIONS: The government's min-max strategy attains $\underline{v} \equiv \min_x \max_y v(x, y)$. The public's min-max strategy attains $\underline{u} \equiv \min_y \max_x u(x, y)$. Let $\underline{y}$ be the government's min-max strategy and $\underline{x}$ be the public's.

'SECURITY' AND 'INDIVIDUALLY RATIONAL' PAYOFF LEVELS: The *security level* payoffs of the government and the public, respectively, are $v^{\#} = v(\underline{y}, B(\underline{y}))$, and $u^{\#} = u(\underline{y}, \underline{y})$. Any payoff vector (v, u) exceeding $(v^{\#}, u^{\#})$ is called *individually rational.*

For the Kydland-Prescott economy, the government's best response function is evidently

$$y = B(x) = \frac{\theta}{\theta^2 + 1} U^* + \frac{\theta^2}{\theta^2 + 1} x. \tag{5}$$

The Nash equilibrium is $y = x = \theta U, U = U^*$. The government's min-max choice is $y^{\#}$, and so is the public's.[3] The Ramsey outcome is $y = x = 0, U = U^*$. Figure 1 illustrates the situation when $\theta = 1$. Evidently, $v(x, y)$ is higher in the Ramsey outcome than in the Nash, a difference referred to in macroeconomics as the 'time inconsistency' problem. It measures the value of giving the government access to a 'commitment technology.' The literature on credibility investigates how 'reputation' might substitute for such a technology.

3. The Repeated Economy

To represent reputation, we study infinite repetition of the one-period economy, for $t \geq 1$. Given (x_t, y_t), $u_t = u(x_t, y_t)$ is the time t payoff of the private sector, and $v_t = v(x_t, y_t)$ is the time t payoff of the government. Each agent is infinitely patient and so ranks outcome paths by their long run average payoffs:

$$\begin{aligned} u &= \liminf_{T \to \infty} T^{-1} \sum_{t=1}^{T} u_t \\ v &= \liminf_{T \to \infty} T^{-1} \sum_{t=1}^{T} v_t. \end{aligned} \tag{6}$$

At time t, the public and the government both observe the history $h_t = (x^{t-1}, y^{t-1})$, where a superscript denotes a record of observations from 1 to $t-1$; h_o is the 'null history' (nothing has been observed). A *strategy* for the repeated economy is a sequence of functions, one for each $t \geq 1$, mapping time t histories into choices. A *strategy profile* for the repeated economy is a pair of strategies, one for the government, another for the public. Given a strategy for the government (public), a repeated economy strategy for the public (or the government) is a *best response* if it maximizes u (or v) of (6). A strategy

[3] The government's min-max problem comes reduces to $\min_x v(x, B(x))$. The maximization piece of the public's problem yields $x = y$; the 'outer' minimization piece then yields $y = x = y^{\#}$.

profile is a *subgame perfect equilibrium* for the repeated economy if each strategy is a best response for each $t \geq 1$ and for each history h_t.

Linear strategies

The immense size of the space of histories makes the number of repeated economy strategies and equilibria very large and unmanageable for the purposes of studying learning. For this reason, we follow Cho (1995) and focus on a small subset of strategies, called linear strategies. These strategies summarize histories by a pair of simple statistics (averages); then use a hyperplane to divide the space of histories into two categories, those calling for one choice, and those calling for another choice. The role played by the hyperplane gives these strategies their name, *linear* strategies.

An easy way to define the strategies is in terms of a simple 'neuron' known as a perceptron (see Cho and Sargent (1996)). Define $\overline{u}_t = \sum_{s=1}^t u_s/t$ and $\overline{v}_t = \sum_{s=1}^t v_s/t$. Let Y denote the Heaviside function

$$Y(z) = \begin{cases} 1 & \text{if } z \geq 0; \\ 0 & \text{if } z < 0. \end{cases}$$

Let $i = p$ refer to the public, and $i = g$ refer to the government. Let $B_i(Y)$ be the action by agent $i \in \{p, g\}$, as determined by the value of the Heaviside function. We refer to the argument of $Y(\cdot)$ as an 'internal state', and to the composition $B_i(Y(z))$ as a 'threshold rule.' These 'two action, two condition' strategies are simple, but not so simple as to loose any values supportable as subgame perfect equilibria without our restrictions on strategies.

Representation of equilibria

We use $\bar{u}_t, \bar{v}_t$ as a pair of state variables in terms of which to cast strategies for the repeated economy. Any individually rational payoff vectors can be represented by a Nash equilibrium payoff where each agent uses a linear strategy.

Let $(u^*, v^*) = (u(x^*, y^*), v(x^*, y^*))$ be an individually rational payoff vector. We will use (u^*, v^*) as a 'target' which we want to support as a subgame perfect equilibrium of the infinitely repeated economy. Recall that $\underline{x}$ and $\underline{y}$ are the minmax strategies of the private sector and the government; that the set of feasible actions for the private sector and the government is $[0, y^{\#}]$; and that $\underline{x} = \underline{y} = y^{\#}$.

Folk theorem for linear strategies: Let the argument of the private sector's perceptron be $z_p = v^* - \overline{v}_t$. Let the argument of the government's perceptron be $z_g = u^* - \overline{u}_t$. The strategies $B_p(1) = x^*, B_p(0) = \underline{x}$, $B_g(1) = y^*, B_g(0) = \underline{y}$ form a subgame perfect equilibrium with long run average payoff vector (u^*, v^*).

A proof can be constructed by adapting arguments in Cho (1995).

In the next two sections, we use linear strategies to parameterize peoples' beliefs in order to formulate a theory of learning. Readers of Woodford (1990) and Marcet and Sargent (1989a) will recognize how in selecting *some* parameterization of beliefs, we are mimicking the steps they used to create a least squares theory of learning.

4. Two-by-two Economy

To introduce the problem of learning in the repeated economy, we begin with a simplified version of the infinitely repeated economy, in which the government and the public are each restricted to choose between two actions. When the government chooses one of two rows from $\{C, D\}$, while the private sector selects one of the two columns in each period, then the payoff pairs $(v(x,y), u(x,x,y))$ are as recorded in the following matrix:

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} 3,3 & 0,0 \\ 4,0 & 1,1 \end{pmatrix}. \end{array}$$

We interpret the first row C as the policy of low inflation and D as the high inflation policy. Similarly, the first row C is the action (actually, the *expectation*) of the private sector that induces low unemployment, and D is the expectation of high inflation. It is easy to verify that (D, D) is the unique Nash equilibrium, and that (C, C) is the Ramsey outcome.

To study learning, we take two steps. First, we adopt a parametric specification of the public's and the government's beliefs about each other. These parameterizations are rich enough to support the 'folk theorem' described above. Second, we describe an adaptive learning algorithm for updating the parameters in these belief functions as the outcomes of the economy unfold over time.

Parameterizing beliefs

The private sector forms its expectations by assuming that the government's policy is determined by a linear strategy that selects D if and only if

$$\alpha_o + \sum_{s \in S} \alpha_1(s) G_t(s) \geq 0$$

where

$$G_t(s) = \# \left\{ t' \leq t \middle| s_t = s \right\}.$$

Let g_t be the empirical frequency of outcome $s \in \{C, D\}^2$, where elements $s \in S$ are ordered $(C, C), (C, D), (D, C), (D, D)$. To simplify notation, let

$$\alpha = (\alpha_1, \alpha_o)'$$

be a column vector in $\mathbb{R}^5$, and let f_t be the 1×5 'data' vector

$$f_t = (G_t, 1)/t.$$

The vector f_t records the empirical frequencies of the four possible outcomes observed historically, and unity. The private sector thinks the government will choose D in period t if

$$f_{t-1}\alpha \geq 0$$

and C if

$$f_{t-1}\alpha < 0.$$

That is, the public believes $B_g(1) = D, B_g(0) = C$.

Estimation of beliefs

Let $\hat{\alpha}_t$ be the "estimator" for α based on information available at the end of period $t-1$. We assume that the private sector updates the estimator according to the stochastic gradient algorithm:

$$\hat{\alpha}^*_{t+1} = \hat{\alpha}_t + \eta_t \left[(y_t - B_g(Y(f_{t-1}\hat{\alpha}_t))) \mathcal{K}_g f'_{t-1} + \epsilon_t \right]$$

where

$$Y(z) = \begin{cases} 1 & \text{if } z \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

and where $\epsilon_t \sim N(0,1)$. Since the sign of $\hat{\alpha}_t f_{t-1}$ determines the value of Y, it makes sense to project $\hat{\alpha}^*_{t+1}$ into a compact set if the norm of the estimator becomes too large. Let $\overline{M}_g > 0$. Define

$$\hat{\alpha}_{t+1} = \pi_g(\hat{\alpha}^*_{t+1}) = \begin{cases} \dfrac{\overline{M}_g \hat{\alpha}^*_{t+1}}{|\hat{\alpha}^*_{t+1}|} & \text{if } |\hat{\alpha}^*_{t+1}| > \overline{M}_g \\ \hat{\alpha}^*_{t+1} & \text{if } |\hat{\alpha}^*_{t+1}| \leq \overline{M}_g. \end{cases}$$

Combining the recursive formula and the projection facility, we have a learning scheme of the private sector about the government's strategy.

The sequence $\{\eta_t\}$ is called the 'gain' sequence. It is a decreasing positive sequence that eventually approaches zero at a rate that influences whether the learning system converges. In contrast to Woodford [1990] and many existing adaptive learning models, we assume that

$$\eta_t \sqrt{t} \to \infty$$

and

$$\eta_t \log t \to 0.$$

Since η_t is asymptotically bounded from below by $1/\sqrt{t}$, η_t converges to 0 at a rate much slower than $1/t$ as assumed in Woodford [1990]. In this way, we can ensure that the estimator is updated quickly in response to the recent observations.

Except for the appearance of the weighting matrix $\mathcal{K}_g$ and the stochastic experimentation term ϵ_t, this is a standard stochastic gradient algorithm for estimating α; $\mathcal{K}$ is a coefficient-specific, history dependent adjustment to the 'gain' in the adaptive algorithm. We let $\mathcal{K}_g$ be a diagonal matrix, and use it to represent the private sector's perception about the government's intention to choose C:

$$\mathcal{K}_g = \text{diag}\left[\kappa_g^{(C,C)}, \kappa_g^{(C,D)}, \kappa_g^{(D,C)}, \kappa_g^{(D,D)}, \kappa_g^{\alpha_o}\right].$$

In principle, one can make κ_g^s depend on the entire history. But to keep faith with the hypothesis of boundedly rational players, we focus on κ_g^s's that depend upon a summary statistic of history, in particular the sign of $f_{t-1}\hat{\alpha}_t$. We will use κ^g to parameterize the government's 'prejudices' in processing the data via its recursive algorithm for updating beliefs about the private sector's expectations about its behavior.

Let $\mathbf{K}_g^c$ be the collection of $\mathcal{K}_g$ that satisfies the following properties:

$$f\mathcal{K}_g f'|_{\hat{\alpha}f>0} - f\mathcal{K}_g f'|_{\hat{\alpha}f\leq 0} \neq 0 \quad \forall f, \forall j; \tag{7}$$

and

$$\frac{\kappa_g(D,s_i)|_{\hat{\alpha}f<0}}{\kappa_g(D,s_i)|_{\hat{\alpha}f\geq 0}} > \frac{f\mathcal{K}_g f'|_{\hat{\alpha}f<0}}{f\mathcal{K}_g f'|_{\hat{\alpha}f\geq 0}} > \frac{\kappa_g(C,s_i)|_{\hat{\alpha}f<0}}{\kappa_g(C,s_i)|_{\hat{\alpha}f\geq 0}} \quad \forall f, \forall s_i, \forall i \neq g. \tag{8}$$

As a typical example, consider

$$\mathcal{K}_g = \begin{cases} \text{diag}[\theta, \theta, \theta(1+\delta), \theta(1+\delta), 0] & \text{if } \hat{\alpha}f < 0 \\ \text{diag}[1+\delta, 1+\delta, 1, 1, 0] & \text{if } \hat{\alpha}f \geq 0. \end{cases} \tag{9}$$

where

$$\theta > 1 + \delta > 1.$$

By choosing $\theta > 1+\delta$, we force the private sector to update its estimator of the government's policy faster when $\hat{\alpha}_t f_{t-1} < 0$ than when $\hat{\alpha}_t f_{t-1} \geq 0$. Whenever the private sector's hope for a good policy from the government is dashed, the private sector updates its estimator quickly, so that in the following round, it is more likely that the private sector is ready for the bad policy. On the other hand, even after the private sector is encountered with a good policy unexpectedly, it updates its estimator rather slowly. In this sense, (9) shows the private sector to be suspicious about the government's intention to carry out a good policy. Hence, (7) is satisfied. As (9) satisfies also (8), the private sector updates $\hat{\alpha}_t(D,s_i)$ $[\hat{\alpha}_t(C,s_i)]$ more quickly if the government chooses D $[C]$ unexpectedly than if the government chooses D $[C]$ as predicted by the private sector.

We construct the learning scheme of the government about the private sector's strategy in the same manner. Assume that the government conjectures that the private sector's behavior is driven by a linear strategy that selects D if

$$f_{t-1}\beta > 0$$

and C if

$$f_{t-1}\beta \leq 0$$

where

$$\beta = (\beta_1(s), \beta_o)_{s\in\{C,D\}^2}.$$

Let $\hat{\beta}_t$ be the estimator for β based on the information available at the end of period $t-1$. We assume that the government updates its estimator according to

$$\hat{\beta}^*_{t+1} = \hat{\beta}_t + \eta_t \left[\left(x_t - B_p(Y(f_{t-1}\hat{\beta}_t)) \right) \mathcal{K}_p f'_{t-1} + \epsilon_t \right]$$

where x_t is the actual action by the private sector in period t. Given $\overline{M}_p$, define the projection facility as

$$\hat{\beta}_{t+1} = \pi_p(\hat{\beta}^*_{t+1}) = \begin{cases} \dfrac{\overline{M}_p \hat{\beta}^*_{t+1}}{|\hat{\beta}^*_{t+1}|} & \text{if } |\hat{\beta}^*_{t+1}| > \overline{M}_p \\ \hat{\beta}^*_{t+1} & \text{if } |\hat{\beta}^*_{t+1}| \leq \overline{M}_p \end{cases}$$

that pushes the estimator $\hat{\beta}^*_{t+1}$ into a compact set, once its norm becomes too large.

Let $\mathbf{K}^c_p$ be the collection of $\mathcal{K}_j$ that satisfies the following properties:

$$f\mathcal{K}_p f'|_{\hat{\beta}f>0} - f\mathcal{K}_p f'|_{\hat{\beta}f\leq 0} \neq 0 \quad \forall f, \forall j; \tag{10}$$

and

$$\frac{\kappa_p(D,s_i)|_{\hat{\beta}f<0}}{\kappa_p(D,s_i)|_{\hat{\beta}f\geq 0}} > \frac{f\mathcal{K}_p f'|_{\hat{\beta}f<0}}{f\mathcal{K}_p f'|_{\hat{\beta}f\geq 0}} > \frac{\kappa_p(C,s_i)|_{\hat{\beta}f<0}}{\kappa_p(C,s_i)|_{\hat{\beta}f\geq 0}} \quad \forall f, \forall s_i, \forall i \neq p. \tag{11}$$

As a typical example, consider

$$\mathcal{K}_p = \begin{cases} \text{diag}[\theta, \theta(1+\delta), \theta, \theta(1+\delta), 0] & \text{if } \hat{\beta}f < 0 \\ \text{diag}[1+\delta, 1, 1+\delta, 1, 0] & \text{if } \hat{\beta}f \geq 0. \end{cases} \tag{12}$$

where

$$\theta > 1 + \delta > 1.$$

Theorem 1. If $\mathcal{K}_j \in \mathbf{K}_j$ for each $j \in \{g, p\}$, then $\forall \delta > 0 \ \exists T(\delta) \geq 1$ such that

$$Pr \begin{pmatrix} x_t - B^p(Y(\hat{\beta}_t f_{t-1})) =0 & \forall t \geq T(\delta) \\ y_t - B^g(Y(\hat{\alpha}_t f_{t-1})) =0 & \end{pmatrix} \geq 1 - \delta$$

and only (C, C) (the Ramsey outcome) and (D, D) (the one-period component economy Nash equilibrium) is played with positive frequency in the limit.

Although one can show that the two parties learn rational expectations in the limit, and that they coordinate between the Ramsey outcome (C, C) and (D, D), it is difficult to calculate the limit frequency of (C, C) and (D, D). As a result, we rely on computer simulations.

We choose $\mathcal{K}_g$ and $\mathcal{K}_p$ according to (9) and (12) with $\delta = 0.2$ and $\theta = 1.3$ so that each party is quite suspicious about the other side's intention of playing a good action (C). We choose $\eta_t = 0.1/[\log(t+1)]^2$ and $\epsilon_t \sim N(0, 1)$. Table 1.1 reports the evolution of empirical frequency of outcomes and the average forecasting errors of each party. Each sample path is as long as 100,000 periods. Average forecasting error is calculated as the proportion of periods when the forecast is not accurate. Notice that the average forecasting error diminishes.

Table 1.2 shows an experiment for which $\theta = .6, \delta = .2$, which makes each party more optimistic about the others performance. The simulations indicate that (C, C) eventually occurs more often than under the more pessimistic $\mathcal{K}_g, \mathcal{K}_p$.

Table 1.1

$\delta = .2, \theta = 1.3$, 'pessimism'

Periods	Empirical Frequencies				Average Forecasting Error	
$\times$10,000	(C,C)	(C,D)	(D,C)	(D,D)	Government	Public
1	0.1681	0.1207	0.1222	0.5890	0.2429	0.2461
2	0.0993	0.0706	0.0780	0.7520	0.1486	0.1572
3	0.0662	0.0471	0.0527	0.8340	0.0998	0.1079
4	0.0497	0.0354	0.0414	0.8736	0.0768	0.0838
5	0.0397	0.0283	0.0343	0.8977	0.0626	0.0713
6	0.0331	0.0236	0.0300	0.9133	0.0536	0.0626
7	0.0284	0.0202	0.0263	0.9251	0.0465	0.0553
8	0.0248	0.0177	0.0232	0.9343	0.0409	0.0486
9	0.0221	0.0157	0.0206	0.9416	0.0363	0.0432
10	0.0199	0.0141	0.0186	0.9474	0.0327	0.0389

Table 2.1

$\delta = .2, \theta = .6$, less 'pessimism'.

Periods	Empirical Frequencies				Average Forecasting Error	
$\times$10,000	(C,C)	(C,D)	(D,C)	(D,D)	Govt	Public
1	0.1910	0.0925	0.1552	0.5614	0.2476	0.2873
2	0.1323	0.0834	0.1411	0.6432	0.2245	0.2623
3	0.1081	0.0760	0.1312	0.6847	0.2072	0.2410
4	0.1099	0.0694	0.1202	0.7006	0.1895	0.2185
5	0.1298	0.0627	0.1096	0.6978	0.1724	0.1973
6	0.1412	0.0605	0.1035	0.6948	0.1640	0.1874
7	0.1513	0.0567	0.0948	0.6973	0.1515	0.1734
8	0.1551	0.0533	0.0914	0.7002	0.1447	0.1635
9	0.1492	0.0498	0.0857	0.7153	0.1355	0.1538
10	0.1452	0.0479	0.0822	0.7247	0.1301	0.1473

5. General Action Spaces

Representing Beliefs

We extend our analysis from a 2×2 economy to an economy with many actions. In most economic models, the action space is assumed to be a closed interval of the real line. However, to avoid technical problems, we consider a "discretized" action space. Let $\mathbf{N}$ be the set of all positive integers. Let $X = [\underline{x}, \overline{x}]$ be a closed interval and let its discretized counterpart be

$$X_h = \left\{ \underline{x} + kh \middle| \; \exists k \in \mathbf{N}, \underline{x} + kh \le \overline{x} \right\}.$$

Similarly, let $Y = [\underline{y}, \overline{y}]$ be the action space of the government with the discretized counterpart

$$Y_h = \left\{ \underline{y} + kh \middle| \; \exists k \in \mathbf{N}, \underline{y} + kh \le \overline{y} \right\}.$$

To simplify notation, however, we shall describe the learning model in terms of the continuous action spaces X and Y by suppressing subscript h from X_h and Y_h.

The general action space necessarily complicates the specifications of these aspects of the problem: (i.) parameterization of the decision rule, (ii.) choice of 'best response' to one's belief about the rule used by the other actor(s); and (iii.) composition of the recursive rule for updating beliefs.

Let g_t be the empirical distribution over $X \times Y$, and

$$f_t = (g_t, 1/t)$$

as before. Define

$$\alpha : X \times Y \cup \{z^o\} \to \mathbb{R}$$

and

$$\beta : X \times Y \cup \{z^o\} \to \mathbb{R}$$

where $\alpha_o = \alpha(z^o)$ and $\beta_o = \beta(z^o)$ are the coefficients assigned to the thresholds as in the 2×2 economy. Define

$$\alpha f_t = \sum_{(x,y) \in X \times Y} \alpha(x,y) g_t(x,y) + \frac{\alpha_o}{t}$$

and

$$\beta f_t = \sum_{(x,y) \in X \times Y} \beta(x,y) g_t(x,y) + \frac{\beta_o}{t}$$

The private sector believes that the government acts according to

$$B_g \left[Y \left(\alpha f_t \right) \right].$$

Let y_+ and y_- be the action by the government when the value of the Heaviside function is 1 and 0, respectively: $y_+ = B_g(1), y_- = B_g(0)$. Because each agent in the private sector has no strategic influence over the outcome, his best response is to "copy" the government's strategy in order to achieve perfect prediction and 'coordination':

$$x_t = B_g(Y(\alpha f_{t-1})).$$

Because private agents can use information already available from the estimated strategy of the government, they can economize the cost of building a best response by just imitating that strategy.

Similarly, the government believes that the private sector chooses its action according to the threshold rule:

$$B_p\left[Y\left(\beta f_t\right)\right]$$

where $x_+ = B_p(1)$ and $x_- = B_p(0)$. In contrast to the agents in the private sector, the government can manipulate the sequence of outcomes in order to manipulate the beliefs of the private sector to its advantage. As a result, calculating the government's optimal strategy is more complicated.

Given its belief about the private sector's strategy, many different strategies can attain the maximum payoff by the government. To proceed with an analysis of learning, we must adopt some principle for selecting among these optimal strategies. We select a particular strategy which is very simple to calculate; simple because the government uses information only about the single-period economy to identify its optimal strategy for the infinitely repeated economy, and fully exploits information already available from estimation of the private sector's strategy.

First calculate y_+ and y_- as follows. Draw a hyperplane

$$\mathbf{H}_p = \left\{ f \middle| \beta f = 0 \right\},$$

which divides the set of empirical frequencies $\Delta(X \times Y)$ into two half spaces. We use this hyperplane to help us identify four numbers: $y_{+,c}$, $y_{+,v}$, $y_{-,c}$ and $y_{-,v}$. By subscript c, we mean "confirming to the sign constraint" and by v, we mean "violating the sign constraint," where the 'sign constraint' will be determined with respect to the hyperplane defined above.

Define $y_{+,c}$ as a solution of

$$\max_y v(x_+, y)$$

subject to

$$\alpha(x_+, y) \geq 0$$

if a solution exists. Suppose that the private sector's action is indeed driven by the linear strategy assumed by the government. Then given that the state is such that private sector is choosing x_+, a long run best response of the government is to choose the one-period best response against x_+. In fact, if the above inequality holds, then the government can attain $\max_y v(x_+, y)$ in the long run. Similarly, define $y_{-,c}$ as a solution of

$$\max_y v(x_-, y)$$

subject to

$$\alpha(x_-, y) \le 0$$

with the same convention when no $\alpha(x, y)$ satisfies the constraint.

Define $y_{+,v}$ as a solution of

$$\max_y v(x_+, y)$$

subject to

$$\alpha(x_+, y) < 0.$$

One can interpret $y_{+,v}$ as a strategy of the government designed to manipulate the private sector's strategy. Notice that in order to let the private sector choose (really, *expect*) x_+, the above inequality must be reversed. When the government is choosing $y_{+,v}$ against x_+, the internal state of the private sector is decreasing, and soon the private sector is expecting x_-. Such strategic behavior is necessary if the government wants the private sector to expect x_- in order to achieve its long run maximum payoff. Similarly, define $y_{-,v}$ as a solution of

$$\max_y v(x_-, y)$$

subject to

$$\alpha(x_-, y) > 0.$$

Let $v_{+,c}$, $v_{+,v}$, $v_{-,c}$ and $v_{-,v}$ be the payoff of the government corresponding to each maximization problem. If the constrained maximization problem has no solution, set the corresponding payoff as $\min_{(x,y)} v(x, y) - 1$. Calculate the intersection between line segment

$$[(u(x_+, y_{+,v}), v(x_+, y_{+,v})), (u(x_-, y_{-,v}), v(x_-, y_{-,v}))]$$

and the payoff vectors induced by $f \in \mathbf{H}_g$:

$$\left\{ \left(\sum_{(x,y)} u(x, y) f(x, y), \sum_{(x,y)} v(x, y) f(x, y) \right) \;\middle|\; f \in \mathbf{H}_g \right\}.$$

Denote the intersection as $(\hat{u}, \hat{v})$. Then $\hat{v}$ is the long run average payoff of the government that can achieved by alternating between $(x_+, y_{+,v})$ and $(x_-, y_{-,v})$.

Given its view of the private sector's strategy, the government has three different sorts of strategies. The first is to confirm $v(x_+, y_{+,c})$ by keeping the internal state of the private sector positive. The second is to confirm $v(x_-, y_{-,c})$ by keeping the private sector's internal state negative. The final alternative is to alternate between $(x_+, y_{+,v})$ and $(x_-, y_{-,v})$. The choice of the government's long run strategy is determined by comparing the long run payoffs from these three strategies.

Suppose that

$$\hat{v} > \max(v_{+,c}, v_{-,c})$$

so that alternating between the two outcomes achieves the maximum payoff. Then set $y_+ = y_{+,v}$ and $y_- = y_{-,v}$ to implement $\hat{v}$ as the long run average payoff. Suppose that

$$\hat{v} \leq \max(v_{+,c}, v_{-,c}).$$

Then the government should implement

$$\max(v_{+,c}, v_{-,c}).$$

In particular, if $v_{+,c} \geq v_{-,c}$, then set $y_+ = y_{+,c}$ and $y_- = y_{-,v}$. In that way, the government can make the internal state of the private sector positive after a finite number of rounds, which then forces the private sector to expect x_+. Similarly, if $v_{+,c} \leq v_{-,c}$, set $y_+ = y_{+,v}$ and $y_- = y_{-,c}$.

When the government is trying to confirm $v_{+,c}$, y_- is used only to manipulate what the government believes is the public's internal state (i.e., the argument of its perceptron) to be positive in finite rounds. Thus, in this case, the government can choose any value for y_- that satisfies

$$\alpha(x_-, y_-) > 0$$

without influencing the long run convergence property of the learning dynamics to be described below. We opt for $y_- = y_{-,v}$ because in this way, for any possible configuration of $\hat{\alpha}$'s, the government has to calculate at most 4 numbers. The same argument applies for the choice of y_+ when the government is trying to enforce $v_{-,c}$ in the long run.

If the government chooses its optimal strategy in this 'most economical' way, then the private sector's conjecture becomes consistent in the sense that the government's best response is indeed a linear strategy. Similarly, given the way the private sector selects its best response, the government's model of the private sector's strategy is also consistent.

Representation of Learning

Because α, β, x_+, x_-, y_+ and y_- are not directly observable, they must be estimated from the actual outcomes. Let $\hat{\alpha}_t$, $\hat{\beta}_t$, $\hat{x}_{+,t}$, $\hat{x}_{-,t}$, $\hat{y}_{+,t}$ and $\hat{y}_{-,t}$ be the respective estimators based on information available at the end of period $t-1$. Define

$$\hat{B}_{g,t}(Y(\hat{\alpha}_t f_{t-1})) = \begin{cases} \hat{y}_{+,t} & \text{if } \hat{\alpha}_t f_{t-1} > 0 \\ \hat{y}_{-,t} & \text{if } \hat{\alpha}_t f_{t-1} \leq 0 \end{cases}$$

and

$$\hat{B}_{p,t}(Y(\hat{\beta}_t f_{t-1})) = \begin{cases} \hat{x}_{+,t} & \text{if } \hat{\beta}_t f_{t-1} > 0 \\ \hat{x}_{-,t} & \text{if } \hat{\beta}_t f_{t-1} \leq 0 \end{cases}$$

as the estimated behavior rule of the government and the private sector, respectively. Without loss of generality, we assume that $\hat{B}_{g,t}$ and $\hat{B}_{p,t}$ satisfy monotonicity properties:

$$\hat{B}_{g,t}(1) \geq \hat{B}_{g,t}(0)$$

and

$$\hat{B}_{p,t}(1) \geq \hat{B}_{p,t}(0).$$

In the 2×2 economy, once the sign of the internal state is determined, the action is naturally assigned, because only 2 actions are available. But in the general case being studied here, each party has also to choose an action to be taken conditioned on the sign of the internal state, and therefore, the action itself must be estimated.

Because each side of the economy observes the action rather than the sign of the internal state of the other side's strategy, we use

$$Y(y_t - \hat{y}_{e,t})$$

where

$$\hat{y}_{e,t} = \frac{\hat{y}_{+,t} + \hat{y}_{-,t}}{2}$$

as the proxy for the sign of the *actual* state of the government's strategy. In 2×2 economy, $\hat{\alpha}_t$ is updated if the estimated sign of the internal state $\hat{\alpha}_t f_{t-1}$ does not match the observed sign of the internal state $\alpha_t f_{t-1}$ represented in terms of the action of the government. Following the same idea, we can write down the updating scheme of the private sector of the government's linear classifier as

$$\hat{\alpha}_{t+1} = \hat{\alpha}_t + \eta_t \left[\left(Y\left(y_t - \hat{y}_{e,t}\right) - Y(\hat{\alpha}_t f_{t-1}) \right) \mathcal{K}_g f_{t-1} + \epsilon_{p,t} \right]$$

where $\mathcal{K}_g$ is a positive definite diagonal matrix. To define the updating rule for $\hat{y}_{\cdot,t}$, we use the counters:

$$\tau^g_{+,t} = \tau^g_{+,t-1} + Y(y_t - \hat{y}_{e,t})$$
$$\tau^g_{-,t} = \tau^g_{-,t-1} + (1 - Y(y_t - \hat{y}_{e,t})).$$

These count the number of periods according to the internal states of the government's linear strategy. Define the updating rule of $\hat{y}_{\cdot,t}$ as

$$\begin{aligned}\hat{y}_{+,t+1} &= \hat{y}_{+,t} + \eta_{\tau^g_{+,t}}\left[y_t - \hat{y}_{+,t}\right]Y(y_t - \hat{y}_{e,t})\\ \hat{y}_{-,t+1} &= \hat{y}_{-,t} + \eta_{\tau^g_{-,t}}\left[y_t - \hat{y}_{-,t}\right](1 - Y(y_t - \hat{y}_{e,t})).\end{aligned}$$

These form $\hat{y}_{+,t}$, $\hat{y}_{-,t}$ as (weighted) averages of past observed values of the setting for y associated with the two types of histories that the perceptron discriminates between.

Similarly, we can define

$$\hat{x}_{e,t} = \frac{\hat{x}_{+,t} + \hat{x}_{-,t}}{2},$$

$$\hat{\beta}_{t+1} = \hat{\beta}_t + \eta_t\left[\left(Y(x_t - \hat{x}_{e,t}) - Y(\hat{\beta}_t f_{t-1})\right)\mathcal{K}_p f_{t-1} + \epsilon_{g,t}\right]$$

and

$$\begin{aligned}\hat{x}_{+,t+1} &= \hat{x}_{+,t} + \eta_{\tau^p_{+,t}}\left[x_t - \hat{x}_{+,t}\right]Y(x_t - \hat{x}_{e,t})\\ \hat{x}_{-,t+1} &= \hat{x}_{-,t} + \eta_{\tau^p_{-,t}}\left[x_t - \hat{x}_{-,t}\right](1 - Y(x_t - \hat{x}_{e,t}))\end{aligned}$$

where

$$\begin{aligned}\tau^p_{+,t} &= \tau^p_{+,t-1} + Y(x_t - \hat{x}_{e,t})\\ \tau^p_{-,t} &= \tau^p_{-,t-1} + (1 - Y(x_t - \hat{x}_{e,t}))\end{aligned}$$

and $\mathcal{K}_p$ is a positive definite diagonal matrix.

First examine the mean dynamics of the linear classifiers:

$$\frac{d\hat{\alpha}}{dt} = \left[Y(y - \hat{y}_e) - Y(\hat{\alpha} f)\right]\mathcal{K}_g f$$

and

$$\frac{d\hat{\beta}}{dt} = \left[Y(x - \hat{x}_e) - Y(\hat{\beta} f)\right]\mathcal{K}_p f$$

where all variables without time subscript are obtained by interpolating the discrete counterpart with time subscript, and setting time very large. The time scale is determined according to η_t instead of $1/t$. Since $\eta_t t \geq \eta_t\sqrt{t} \to \infty$, we can treat the empirical frequency as a constant in the limit. Then we can represent the mean dynamics of the estimator of the actions as follows:

$$\begin{aligned}\frac{d\hat{y}_+}{dt} &= (y - \hat{y}_+)Y(y - \hat{y}_e)\\ \frac{d\hat{y}_-}{dt} &= (y - \hat{y}_-)\left[1 - Y(y - \hat{y}_e)\right],\end{aligned}$$

and

$$\begin{aligned}\frac{d\hat{x}_+}{dt} =&(x - \hat{x}_+)Y(x - \hat{x}_e)\\ \frac{d\hat{x}_-}{dt} =&(x - \hat{x}_-)\left[1 - Y(x - \hat{x}_e)\right].\end{aligned}$$

Before writing down the appropriate Lyapounov function for the mean dynamics, it is instructive to see what should happen in the limit by setting the right hand sides of the above differential equations equal to 0:

$$\begin{aligned}\left[Y(x - \hat{x}_e) - Y(\hat{\beta} f)\right] \mathcal{K}_p f =&0\\ (x - \hat{x}_+)Y(x - \hat{x}_e) =&0\\ (x - \hat{x}_-)\left[1 - Y(x - \hat{x}_e)\right] =&0\end{aligned}$$

and

$$\begin{aligned}\left[Y(y - \hat{y}_e) - Y(\hat{\alpha} f)\right] \mathcal{K}_g f =&0\\ (y - \hat{y}_+)Y(y - \hat{y}_e) =&0\\ (y - \hat{y}_-)\left[1 - Y(y - \hat{y}_e)\right] =&0.\end{aligned}$$

A careful examination reveals that

$$\hat{\beta} f \geq 0 \Leftrightarrow x - \frac{\hat{x}_+ + \hat{x}_-}{2} \geq 0 \Leftrightarrow x = \hat{x}_+ \geq \hat{x}_-$$

and

$$\hat{\beta} f < 0 \Leftrightarrow x - \frac{\hat{x}_+ + \hat{x}_-}{2} < 0 \Leftrightarrow x = \hat{x}_- < \hat{x}_+.$$

The third term implies that the government entertains perfect foresight of the private sector's action. Since the individual agent in the private sector has no strategic power, the private sector's optimal response is to set its expectation equal to the estimated inflation rate of the government:

$$x = \begin{cases} \hat{y}_+ & \text{if } \hat{\alpha} f \geq 0 \\ \hat{y}_- & \text{if } \hat{\alpha} f < 0. \end{cases}$$

Hence, the dynamics of $\hat{x}$ in combination with the behavior rule of the private sector requires that

$$[\hat{\alpha} f][\hat{\beta} f] > 0$$

at any stable solution in the limit.

By following the same logic, we have

$$\hat{\alpha} f \geq 0 \Leftrightarrow y - \frac{\hat{y}_+ + \hat{y}_-}{2} \geq 0 \Leftrightarrow y = \hat{y}_+ \geq \hat{y}_-$$

and

$$\hat{\alpha} f < 0 \Leftrightarrow y - \frac{\hat{y}_+ + \hat{y}_-}{2} < 0 \Leftrightarrow y = \hat{y}_- < \hat{y}_+.$$

Again, the third term implies that the private sector learns to perfectly foresee the government's action. Depending upon how the government responds to $\hat{\beta} f$, either

$$[\hat{\alpha} f][\hat{\beta} f] \geq 0$$

or

$$[\hat{\alpha} f][\hat{\beta} f] \leq 0$$

must hold in the limit.

Combining the mean dynamics of the private sector and the government, we conclude that if the mean dynamics has a stable solution, then

$$x_t - \hat{x}_t \to 0, \quad y_t - \hat{y}_t \to 0$$

and

$$(\hat{\alpha}, \hat{\beta}, f) \to Q^* = Q_1^* \cup Q_2^*$$

where

$$Q_1^* = \left\{ (\hat{\alpha}, \hat{\beta}, f) \middle| [\hat{\alpha} f][\hat{\beta} f] \geq 0, \quad B_g(0) \leq B_g(1) \right\}$$

and

$$Q_2^* = \left\{ (\hat{\alpha}, \hat{\beta}, f) \middle| [\hat{\alpha} f][\hat{\beta} f] = 0, \quad B_g(0) > B_g(1) \right\}.$$

Indeed, one can choose

$$\mathcal{L} = (\hat{x} - x)^2 + (\hat{y} - y)^2 + \min_{(\hat{\alpha}', \hat{\beta}', f') \in Q^*} \left| (\hat{\alpha}, \hat{\beta}, f) - (\hat{\alpha}', \hat{\beta}', f') \right|^2$$

as the Lyapounov function to verify that

$$(\hat{\alpha}, \hat{\beta}, f) \to Q^*$$

according to the mean dynamics.

However, this convergence result *per se* does not imply that the two parties eventually perfectly forecast the action by the other side. Notice that the mean dynamics of $\hat{\alpha}$ and $\hat{\beta}$ are discontinuous along the boundary of Q^*. Therefore, if $(\hat{\alpha}, \hat{\beta}, f)$ converges to the boundary of Q^*, then either party may sustain persistent forecasting errors. To obtain perfect foresight, $(\hat{\alpha}, \hat{\beta}, f)$ needs to remain in the interior of Q^* almost surely.

In order to strengthen the convergence result above, we need to solve two problems. First because Q_2^* has an empty interior, the mean dynamics do not lead $(\hat{\alpha}, \hat{\beta}, f)$ to Q_2^*. Then we need to ensure that $(\hat{\alpha}, \hat{\beta}, f)$ remains in the interior of Q_1^*,

First we need to impose some monotonic conditions on $\mathcal{K}_p$ and $\mathcal{K}_g$ in order to eliminate Q_2^* from the set of solutions of the mean dynamics. A straightforward calculation shows that

$$(\hat{\alpha}, \hat{\beta}, f) \to Q_2^*$$

only if

$$\begin{aligned}
&\lim_{s\to 0}\lim_{K\to\infty} \frac{\hat{\beta}_{t_K+m(t_K+s)} - \hat{\beta}_{t_K}}{s} \propto \Lambda \\
&= \lim_{s\to 0}\lim_{K\to\infty} \frac{1}{s}\left[\sum_{t=t_K}^{m(t_K+s)} \eta_t \left[\frac{\mathcal{K}_p|_{\hat{\beta}_t f_{t-1}<0}}{f_{t-1}\mathcal{K}_p f'_{t-1}|_{\hat{\beta}_t f_{t-1}<0}} - \frac{\mathcal{K}_p|_{\hat{\beta}_t f_{t-1}\geq 0}}{f_{t-1}\mathcal{K}_p f'_{t-1}|_{\hat{\beta}_t f_{t-1}\geq 0}}\right] f'_{t-1}\right].
\end{aligned}$$

Recall that in order to have $(\hat{\alpha}, \hat{\beta}, f) \to Q_2^*$, it is necessary that the government's policy selection satisfy

$$y_+ < y_-.$$

Suppose that Λ is the estimator for β, and against Λ, the government's policy selection is such that

$$y_+ > y_-.$$

Then we can conclude that $(\hat{\alpha}, \hat{\beta}, f)$ no longer stays in the neighborhood of Q_2^*. Rather, it converges to Q_1^* as desired.

We need some restrictions on $\mathcal{K}_p$ such that Λ has the desired properties. Suppose that

$$\text{sgn}\Lambda(x,y) = \begin{cases} + & \text{if } y \geq \gamma(x) \\ - & \text{if } y < \gamma(x) \end{cases}$$

where γ is an increasing function of x and $\text{sgn}\Lambda(x,y)$ is the sign of $\Lambda(x,y)$. One can easily show that against such Λ, the best response of the government must be such that

$$y_- < y_+.$$

Therefore, if one can find a condition for $\mathcal{K}_p$ such that the associated Λ induces the best response of the government so that $y_- < y_+$, then $(\hat{\alpha}, \hat{\beta}, f)$ converges to Q_1^* instead of Q_2^* according to the mean dynamics.

Let $\kappa_p(x,y)$ be a diagonal element in $\mathcal{K}_p$ associated with outcome (x,y). Define

$$\kappa_p(x,y) = \begin{cases} \theta & \text{if } y > \gamma(x) \\ \theta(1+\delta) & \text{if } y \leq \gamma(x) \end{cases}$$

if $\hat{\beta} f < 0$; and

$$\kappa_p(x,y) = \begin{cases} 1+\delta & \text{if } y > \gamma(x) \\ 1 & \text{if } y \leq \gamma(x) \end{cases}$$

if $\hat{\beta} f \geq 0$. Let

$$\theta > 1 + \delta > 1.$$

This specification of $\mathcal{K}_p$ captures the government's pessimism about its ability to sustain the private sector's expectations of low inflation. If $\hat{\beta} f < 0$, the government expects that the private sector will expect low inflation rate x_-. But if the private sector actually expects high inflation (i.e., $x \geq x^e$), then the government updates $\hat{\beta}$ so that it becomes more likely that the government expects that the private sector will forecast a high inflation rate in the following round. Similarly, if $\hat{\beta} f \geq 0$, but the private sector actually expects a low inflation, then the government will adjust $\hat{\beta}$ so that in the following round, the government forecasts that the private sector expects low inflation. Because $\theta > 1 + \delta$, the government responds more rapidly when its forecast of low expected inflation rate is dashed than when its forecast of high expected inflation rate is proved incorrect.

Similarly, let $\mathcal{K}_g = \mathrm{diag}[\kappa_g(x,y)]_{(x,y)\in X \times Y}$ be a diagonal matrix such that

$$\kappa_g(x,y)\Big|_{\hat{\alpha} f < 0} > \kappa_g(x,y)\Big|_{\hat{\alpha} f \geq 0}$$

to represent the private sector's pessimism about the government's intention to maintain low inflation rate.

If both the private sector and the government are pessimistic about the opponent's intention of 'good will,' then the learning dynamics evolves in such a way that according to its associated mean dynamics

$$(\hat{\alpha}, \hat{\beta}, f) \to Q_1^*.$$

As long as

$$\eta_t \log t \to 0,$$

we can invoke the convergence theorem of Dupuis and Kushner [1981] to show that

$$(\hat{\alpha}_t, \hat{\beta}_t, f_t) \to Q_1^*$$

almost surely. Yet in order to ensure that

$$(\hat{\alpha}_t, \hat{\beta}_t, f_t) \to \partial Q_1^*$$

with probability 0, where ∂Q_1^* is the boundary of Q_1^*, we need to let η_t converge to 0 sufficiently slowly; namely,

$$\eta_t \sqrt{t} \to \infty.$$

Note that if $\eta_t\sqrt{t} \to 0$ as in Marcet and Sargent [1989] and Woodford [1990], then the learning dynamics can be approximated by the mean dynamics described above. But if $\eta_t\sqrt{t} \to \infty$, then the stochastic component of the learning dynamics does not vanish even in the limit, and the learning dynamics must be approximated by a stochastic differential equation.

Heuristically, one can assume that $(\hat{\alpha}, \hat{\beta}, f)$ stays in the closure of Q_1^* in the limit. In the interior of Q_1^*, the each estimator behaves like a random walk because each party makes a correct forecast about the opponent's action. However, along the boundary of Q_1^*, the estimator is pushed back to the interior according to the error correction process. In particular, as soon as $(\hat{\alpha}, \hat{\beta}, f)$ hits the boundary of Q_1^* at t, $\hat{\alpha}$ is pushed from the boundary according to

$$(Y(y - \hat{y}_e) - Y(\hat{\alpha} f))\,\mathcal{K}_g f$$

and $\hat{\beta}$ is pushed according to

$$\left(Y(x - \hat{x}_e) - Y(\hat{\beta} f)\right)\mathcal{K}_p f.$$

Hence, one can write the limit dynamics of $\hat{\alpha}$ and $\hat{\beta}$ as

$$\hat{\alpha}(t) = \hat{\alpha}(0) + \int_0^t dW_t + \int_0^t \mathcal{K}_g f d|\omega|$$

and

$$\hat{\beta}(t) = \hat{\beta}(0) + \int_o^t dW_t + \int_0^t \mathcal{K}_p f d|\omega|$$

where $|\omega|$ is implicitly defined as

$$|\omega(t)| = \int_0^t \mathbf{1}_{(\hat{\alpha},\hat{\beta},f)\in\partial Q_1^*} d|\omega(\tau)|.$$

One can easily verify that if both parties are 'pessimistic' about the opponent as defined above, then $(\hat{\alpha}, \hat{\beta}, f)$ are pushed back into the interior of Q_1^* along its boundary. Then one can invoke Kushner and Dupuis [1992] to show that

$$(\hat{\alpha}, \hat{\beta}, f) \in \partial Q_1^*$$

with probability 0, as desired.

If both the government and the private sector are pessimistic and $\mathcal{K}_p$ and $\mathcal{K}_g$ are constructed accordingly, the two parties eventually learn to forecast the other side's action perfectly. Since the private sector takes the same action as the estimated action of the government, this convergence result implies that the two parties end up taking the "same" action to achieve perfect coordination. Given our convention for identifying the action of the each party, this means that whenever the government is choosing a high inflation policy, the private sector is responding by expecting high inflation, and if the government is choosing the Ramsey policy (the low inflation policy), the private sector expects low inflation. To outsiders, the two parties behave as if each party can perfectly foresee the other side's action and respond appropriately.

6. Conclusions

The literature on credible economic policy raises the possibility that a good reputation can substitute for commitment. Systems of expectations exist that induce a policy maker to abstain forever from inflation. However, this finding provides little comfort because there are other systems of expectations within which a policy maker finds himself trapped, and which induce him to conform to his inflationary reputation. We interpret the literature on credible public policy as concluding that reputations alone cannot be relied on to substitute for a commitment mechanism, because there are so many reputations for bad outcomes.[4]

Our research asks whether a plausible theory of learning - or acquiring - reputations might narrow the range of outcomes to be expected from a reputational mechanism. We show how our theory of learning sharply reduces the range of outcomes to be expected, well below those described by the folk theorem. This is good news in terms of the predictive power of the theory. But our (2×2) analysis provides only very limited reasons to expect much from reputation. Our theorem raised our hopes, because it asserted that the limiting outcome would be better than the Nash, even if they might fall short of the Ramsey outcome. The simulations, especially the one with the 'suspicious' gain function, deflated our hopes, because the system produces the Nash outcome most of the time. The second simulation, with less suspicion in the gain, 'prompts' the system to learn to coordinate on a better mixture of Ramsey and Nash.

Our results thus serve to emphasize the importance of putting in place commitment mechanisms with more force than reputation. Maybe this is a useful message to convey to a conference celebrating an old central bank.

[4] See Chari, Christiano, and Eichenbaum's (1996) account of the dangers of an 'expectation trap' for monetary policy.

References

Bray, Margaret (1982). 'Learning, Estimation and Stability of Rational Expectations'. *Journal of Economic Theory*, 26: 318–339.

Chari, V. V., Lawrence Christiano, and Martin Eichenbaum (1996). 'Expectation Traps and Discretion'. Mimeo. Northwestern University.

Chari, V.V. and Patrick J. Kehoe (1990). 'Sustainable Plans'. *Journal of Political Economy*, Vol. 98, pp. 783–802.

Chari, V.V., Patrick J. Kehoe, and Edward C. Prescott. 'Time Consistency and Policy'. In Robert Barro (ed.), *Modern Business Cycle Theory* Harvard University Press, Cambridge, Massachusetts, pp. 265–305.

Cho, In-Koo (1995). 'Bounded Rationality, Neural Network and Folk Theorem in Repeated Games with Discounting'. *Economic Theory*, 4: 935–957.

Cho, In-Koo and Thomas J. Sargent (1996). 'Neural Networks for Encoding and Adapting in Dynamic Economies'. In Hans M. Amman, David A. Kendrick and John Rust (eds.), *Handbook of Computational Economics*. Elsevier Science, pp 443–470.

Dupuis, Paul and Harold J. Kushner (1989). 'Stochastic Approximation and Large Deviations: Upper Bounds and w.p.1 Convergence'. *SIAM Journal of Control and Optimization*, 27: 1108–1135.

Fudenberg, Drew and David M. Kreps (1994). 'Learning in Extensive Form Games, II: Experimentation and Nash Equilibrium'. Mimeo. Stanford University.

Kushner, Harold J. and Paul G. Dupuis (1992). *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer-Verlag.

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

Marcet, Albert and Juan Pablo Nicolini (1995). 'Recurrent Hyperinflations and Learning'. Mimeo. Universitat Pompeu Fabra.

Marcet, Albert and Thomas J. Sargent (1989a). 'Convergence of Least Squares Learning Mechanisms in Self Referential Linear Stochastic Models'. *Journal of Economic Theory*, 48: 337–368.

Marcet, Albert and Thomas J. Sargent (1989b). 'Least-Squares Learning and the Dynamics of Hyperinflation'. In William A. Barnett, John Geweke, and Karl Shell (eds.), *Economic Complexity: Chaos, Sunspots, Bubbles and Nonlinearity*. Cambridge University Press, pp. 119–137.

Marimon, Ramon (1995). 'Learning from Learning in Economics'. Mimeo. European University Institute.

Stokey, Nancy L (1989). 'Reputation and Time Consistency'. *American Economic Review*, Vol. 79, pp. 134–139.

Stokey, Nancy L (1991). 'Credible Public Policy'. *Journal of Economic Dynamics and Control*, Vol. 15, No. 4, pp. 627–656.

Weisbuch, Gérard (1990). *Complex System Dynamics: An Introduction to Automata Net-*

works. Lecture Note Volume II, Santa Fe Institute Studies in the Sciences of Complexity, Addison-Wesley.

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