Working Paper Series

Optimal Policy with Probabilistic Equilibrium Selection

WP 01-03

Huberto M. Ennis
Federal Reserve Bank of Richmond

Todd Keister
Department of Economics and Centro de Investigación Económica, ITAM

THE FEDERAL RESERVE BANK OF RICHMOND
RICHMOND ▪ BALTIMORE ▪ CHARLOTTE

Optimal Policy with Probabilistic Equilibrium Selection*

Huberto Ennis
Research Department, Federal Reserve Bank of Richmond
huberto.ennis@rich.frb.org

Todd Keister
Department of Economics and Centro de Investigación Económica, ITAM
keister@itam.mx

Federal Reserve Bank of Richmond Working Paper No. 01-03
June 2001

Abstract

This paper introduces an approach to the study of optimal government policy in economies characterized by a coordination problem and multiple equilibria. Such models are often criticized as not being useful for policy analysis because they fail to assign a unique prediction to each possible policy choice. We employ a selection mechanism that assigns, *ex ante*, a probability to each equilibrium indicating how likely it is to obtain. With this, the optimal policy is well defined. We show how such a mechanism can be derived as the natural result of an adaptive learning process. This approach generates a theory of how government policy affects the process of equilibrium selection; we illustrate this theory by applying it to problems related to the choice of technology and the optimal sales tax on Internet commerce.

JEL Classification Nos.: E61, D83, H21
Keywords: Coordination failure, government policy, learning

* We thank seminar participants at Cornell, the Federal Reserve Bank of Richmond, ITAM, and Purdue. We are also grateful to Preston McAfee for helpful comments and especially to Karl Shell for emphasizing the importance of studying how government policy can affect the process of equilibrium selection. The views expressed herein are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Richmond nor those of the Federal Reserve System.

1 Introduction

The extent to which purchases made over the Internet should be subject to sales tax is currently the subject of a lively debate in the United States (and elsewhere). The Internet Tax Freedom Act of 1998 was essentially an agreement to postpone any major decisions for at least three years, meaning that the issue is still very much undecided. The arguments both for and against special tax treatment of Internet transactions are numerous, but one of the arguments in favor of special treatment strikes us as particularly interesting. This argument[1] relies on the idea that there is a natural network externality: as more people engage in e-commerce, the efficiency of e-commerce increases for all users. The future size of the electronic market is currently unclear, and the cost advantages of e-commerce over traditional retailing methods make a large electronic marketplace a better outcome (in terms of social welfare). Lower e-commerce taxes are believed to substantially increase the number of people who use this medium for commerce.[2] Preferential tax treatment, therefore, is aimed at trying to generate a large e-commerce market.

Putting this argument into the language of formal economic models, the claim is that there are multiple equilibria with differing levels of e-commerce activity. To keep things simple, suppose there are two equilibria. In one, the e-commerce market is small and engaging in e-commerce is not very profitable, leading few people to do so. In the other, the market is large and very efficient, leading many people to engage in e-commerce. Because e-commerce is claimed to be more efficient than traditional distribution methods, the latter equilibrium socially dominates the former. In this language, the argument is that by giving e-commerce transactions preferential tax treatment the government can encourage the economy to settle into the better equilibrium. In other words, the proposed policy is an attempt to affect the process by which an equilibrium is selected.

This argument raises a fundamental issue that reaches far beyond the specific example: whenever there are multiple equilibria, actions taken by the government may affect which equilibrium is selected. This means that, in such situations, simply formulating the optimal policy problem requires a theory of how the policy choice affects the equilibrium selection process. In the present paper, we provide such a theory for exactly the type of situation described above: coordination-problem economies with multiple equilibria. Our approach is to assign a *probability* to each equilibrium indicating how likely it is to obtain. These probabilities then generate a well-defined optimal policy problem. Solving this problem requires taking into account the effect that the policy choice has on the probabilities. We show how an adaptive learning process naturally generates such probabilities

[1] See Zittrain and Nesson (2000) for one statement of this argument in the popular press.

[2] Indeed, Goolsbee (2000) estimates that applying existing sales taxes to Internet commerce would reduce the number of online buyers by about 24 percent.

and that they are affected by policy in intuitive and interesting ways.

We apply our methods to both the Internet sales tax issue and to the general problem of the choice between technologies in the presence of network externalities. In such situations, there are often two (symmetric, pure-strategy) equilibria – one where the good technology is adopted and one where the bad technology is used. These are typically both strict equilibria, neither of which is easily refined away. There is ample evidence from the historical and experimental literatures that both Pareto-dominant and Pareto-dominated equilibria can arise in such settings.[3] Hence rules that select only Pareto-optimal equilibria do not seem to describe such situations very well. On the other hand, it is often intuitively clear that one equilibrium seems more likely to obtain than the other. As an example, consider the problem of an agent deciding whether to engage in e-commerce or traditional commerce. To simplify the story, suppose she believes the e-commerce market will be either "thick" or "thin."[4] Suppose further that the gain from choosing e-commerce is large if the market is thick and that the loss from choosing e-commerce is small (near zero) if the market is thin. If all agents are identical to this one, then there are two strict equilibria. However, in a world with uncertainty and imperfect information, it seems intuitively more likely that the economy would settle on the thick-market equilibrium since the large payoff of the e-commerce equilibrium will lead expected-utility-maximizing agents to choose e-commerce for a large range of beliefs about the market thickness. In the words of Harsanyi and Selten (1988), the e-commerce equilibrium is risk dominant. Hence, it seems intuitively plausible that a risk-dominant equilibrium might be expected to obtain more often than a risk-dominated one. How much more often would seem to depend on the strength of its risk dominance.

Experimental evidence lends strong support to the probabilistic view of equilibrium selection. Van Huyck, Battalio, and Beil (1990, 1991) report that, in a series of coordination-game experiments, the frequency with which the subjects converged to each equilibrium varied systematically with the treatment variables. In other words, the *ex ante* probability distribution across equilibria was non-degenerate and was affected by small changes in the structure of the game.[5] We build on the model of adaptive learning in Howitt and McAfee (1992) and show that it generates precisely this type of behavior. Under this learning process, the economy can converge to either equilibrium. Where it converges depends in part on the particular realizations of uncertainty along the learning path. In the context of our leading example, suppose that the efficiency level of e-commerce is not known with

[3] See Cooper (1999) for an excellent review of this literature.

[4] Another type of model where expectations about market thickness can generate multiple equilibria is the market game. See Peck, Shell, and Spear (1992).

[5] See Crawford (1997) for an interesting discussion of these results and for an estimation of a general learning model using the experimental data.

certainty. If the first few observations make e-commerce seem attractive, agents will begin to coordinate on the high e-commerce decision, making this equilibrium more likely. We demonstrate that the probability of converging to a particular equilibrium is negatively related to the *risk factor* of the equilibrium. (A low risk factor corresponds to a strongly risk-dominant equilibrium.) The risk factor, in turn, depends on the government's policy choice. This is the mechanism through which policy affects equilibrium selection – by making a particular choice less risky for agents in the stochastic environment, the government can make that equilibrium more likely to obtain.

This mechanism assigns an objective probability to each equilibrium and thereby allows us to address the question of optimal government policy. We take as the government's objective the maximization of expected utility across equilibria. The optimal policy derived under our approach typically differs from that derived using any deterministic selection criterion. We show, for example, that it is generally not optimal to choose the policy that maximizes the (utility) value of the good equilibrium. The reason is that, by deviating in some direction, there is a first-order gain in the likelihood of attaining the equilibrium, but no first-order loss in the value of that equilibrium. We also show that even when it is feasible to eliminate the bad equilibrium, it may not be optimal to do so. It may be optimal to allow the bad outcome to occur with low probability in exchange for a higher (utility) value if the good equilibrium occurs.

Our approach stands in contrast to the recent work of Morris and Shin (1998), which uses informational imperfections to eliminate the multiplicity of equilibria in a related coordination game. In that approach, economic fundamentals (and the information structure) determine how agents will coordinate; there is no role for chance. In our approach, chance plays an important role, with the economic fundamentals determining the probabilities. Our approach is therefore more in the spirit of Cole and Kehoe (2000). That paper studies a coordination game embedded in a dynamic general equilibrium model and calculates (for given fundamentals) the maximum probability of a bad outcome that is consistent with equilibrium. Hence chance plays a role and economic fundamentals determine the possible probabilities. However, our approach assigns a unique probability to each outcome, eliminating the need to focus on the worst case scenario.

When there are multiple equilibria, policy choices can affect equilibrium selection by acting as a sunspot variable and guiding the coordination of agents' actions.[6] For example, agents might expect the high e-commerce equilibrium to obtain if and only if e-commerce is completely tax free. This would then be a rational expectations equilibrium. It is extremely difficult, however, to formally

[6] Manuelli and Peck (1992) show how intrinsically important economic variables can also act as sunspot variables with respect to agents' expectations.

model agents' expectations regarding such effects.[7] A variety of papers have focused instead on the "worst" equilibria and looked for policies that eliminate these as possible outcomes.[8] Our approach differs from these in that it assigns an objective probability to each equilibrium and hence generates a unique probability distribution over the set of equilibria. This can be thought of as endogenizing agents' expectations, although doing so requires a departure from rational expectations in the learning process. The distribution tells us how much weight the government should give to each of the possible outcomes. This allows us to ask not only *can* the government eliminate a particular bad equilibrium, but also *should* the government eliminate it.

The remainder of the paper is organized as follows. In the next section, we lay out our approach to optimal policy analysis and equilibrium selection in a general setting. The focus is on illustrating the central ideas rather than proving specific results. In Section 3, we specialize to a class of coordination-problem economies that contains the examples discussed above, and we provide a detailed analysis of both the learning process and the optimal policy problem in that setting. In Section 4, we then apply our approach to simple models of technology choice. Finally, in Section 5 we offer some concluding remarks.

2 The General Approach

In this section, we outline our approach in a general setting that allows us to highlight the important features with minimal complications. We consider economies with a continuum of identical agents. We focus on symmetric equilibria, where the fact that we have identical agents means that social welfare is the same as individual welfare. Studying symmetric equilibria facilitates the analysis of the optimal policy, as there is no need to impose a social welfare function.

2.1 The Model and Equilibrium

Each agent must choose an action a from a set $A \subset \mathbb{R}^N$. There is a (benevolent) government that chooses a policy τ from a set $T \subset \mathbb{R}$. In addition, there is aggregate uncertainty represented by the variable $c \in C \subset \mathbb{R}$. The distribution of c is given by f, a probability measure on $\mathbb{B}(C)$, the Borel subsets of C.[9] The utility of each agent depends on each of these variables *and* on the average action in the economy $\overline{a}$; this external effect is what will generate multiple equilibria. Since there is a continuum of agents, we have a truly competitive economy and each agent correctly perceives his

[7] See Ghiglino and Shell (2000), section 7, for an interesting discussion of this issue.

[8] Contributions along these lines in the sunspots literature include Grandmont (1986), Woodford (1986), Smith (1994), Goenka (1994), and Keister (1998). This approach is also common in the literature on financial fragility.

[9] We are assuming that both c and τ are one-dimensional, but this does not seem to be important for the results.

own action to have no effect on $\overline{a}$. Each agent's utility is given by the function

$$V : A \times \overline{A} \times T \times C \to \mathbb{R},$$

where $\overline{A}$ is the convex hull of A. We assume that the set T is convex and that V is (jointly) strictly concave and twice continuously differentiable in all of its arguments. These assumptions are made so that we can use a simple first-order condition to characterize optimal government policy.

Each agent has a belief about the values of $\overline{a}$ and c that is represented by a joint probability measure F on $\overline{A} \times C$. The government policy is commonly known. Each agent therefore maximizes expected utility by solving

$$\max_{a \in A} \int V(a, \overline{a}, \tau, c)\, dF(\overline{a}, c). \tag{1}$$

Rational expectations equilibrium is a natural solution concept in this environment; we focus on symmetric outcomes.

Definition: A *symmetric rational-expectations equilibrium* is an $a^* \in A$ such that

(i) Agents have rational expectations, that is, we have

$$F(\overline{a}, \mathbb{C}) = \left\{ \begin{array}{c} f(\mathbb{C}) \\ 0 \end{array} \right\} \text{ for } \overline{a} \left\{ \begin{array}{c} = \\ \neq \end{array} \right\} a^* \text{ for all } \mathbb{C} \in \mathbb{B}(C) \text{, and}$$

(ii) Given these beliefs, agents are optimizing (that is, a^* solves (1)).

We denote the set of equilibria for a given government policy by $\mathcal{E}(\tau)$. Our interest is in situations where there is more than one such equilibrium for at least some values of τ, so that equilibrium selection is an issue.[10] It is well known in the literature on coordination problems that some degree of payoff complementarity is necessary for the existence of multiple equilibria (see Cooper and John [1988]), so we are assuming that the V function has this property. In Section 4, we work through explicit examples where this is the case.

2.2 Optimal Policy

We now turn our attention to the problem of determining the optimal policy. The traditional approach is to focus on a particular equilibrium $a^* \in \mathcal{E}(\tau)$. This equilibrium might be selected by a formal rule, or it might be the focus of attention solely because it has some desirable properties. The government recognizes, of course, that the equilibrium action a^* is a function of τ. The benevolent government chooses its policy to maximize the expected utility of each agent in this symmetric equilibrium. The

[10] The multiplicity here opens the door to a richer set of possible rational expectations equilibria (related to the correlated equilibria of the coordination game). We choose to restrict our definition exclusively for simplicity (see the remarks in Section 3.1).

traditional optimal policy problem is therefore

$$\max_{\tau \in T} \int V\left(a^*\left(\tau\right), a^*\left(\tau\right), \tau, c\right) df\left(c\right).$$

The shortcoming of this approach is that the resulting optimal policy typically depends on which equilibrium was selected. As we mentioned above, the evidence on coordination games indicates that in many cases a unique equilibrium simply cannot be singled out as the prediction of the model. As a result, this approach does not yield a clear policy prescription.

Suppose instead that we allowed for a probabilistic equilibrium selection mechanism. Such a mechanism assigns, for each value of τ, a probability distribution over the set of equilibria $\mathcal{E}\left(\tau\right)$. We denote the probability of equilibrium a^* by $\pi\left(a^*, \tau\right)$. As the notation indicates, this distribution will typically depend on the government's choice of policy. In the next subsection, we provide a foundation for this type of mechanism by showing that it can be derived as the limiting outcome of an adaptive learning process in the model. First, however, we discuss the implications of such a mechanism for the optimal policy problem. Once a probability is assigned to each equilibrium, the natural goal of the government is to maximize the expected utility of agents across equilibria.[11] The optimal policy problem is therefore

$$\max_{\tau \in T} \int \left(\int V\left(a^*\left(\tau\right), a^*\left(\tau\right), \tau, c\right) df\left(c\right) \right) d\pi\left(a^*\left(\tau\right), \tau\right). \tag{2}$$

Notice how τ enters the objective function by (i) directly affecting the value V, (ii) affecting the equilibrium action a^*, and (iii) affecting the equilibrium selection mechanism π.

Deriving the optimal policy this way leads to interesting and sometimes surprising results as we demonstrate in Section 4. To gain some intuition here, suppose that there are two equilibria, one "good" and the other "bad." Let $V_g\left(\tau\right)$ and $V_b\left(\tau\right)$ represent the utility level of agents in each of these equilibria. Also, let $\pi\left(\tau\right)$ represent the probability of converging to the good equilibrium (that is, $\pi\left(g; \tau\right)$). Then, the optimal policy problem can be written as

$$\max_{\tau \in T} \pi\left(\tau\right) V_g\left(\tau\right) + \left(1 - \pi\left(\tau\right)\right) V_b\left(\tau\right). \tag{3}$$

Suppose (again for illustration) that π is differentiable in τ.Then the first-order condition for this problem is

$$\pi\left(\tau\right) V_g'\left(\tau\right) + \left(1 - \pi\left(\tau\right)\right) V_b'\left(\tau\right) + \pi'\left(\tau\right)\left(V_g\left(\tau\right) - V_b\left(\tau\right)\right) = 0.$$

This is in many ways the central equation of the paper; its solution is the optimal policy τ^*. It says

[11] Agents now face a compound lottery – first an equilibrium is selected and then a state of nature is realized. Maximizing expected utility across equilibria amounts to a reduction of this into a single lottery.

that the optimal policy is the result of the balancing of three forces. One must consider not only the effect of the policy on the utility of agents in each of the two equilibria, but also the effect of the policy on the equilibrium selection mechanism.

One way to think about this approach is to contrast it with a situation where agents use a sunspot variable to coordinate their actions. In that case, with probability π there is sunspot activity and with probability $(1 - \pi)$ there is not, and π is independent of government policy (and everything else in the economy). There is then an equilibrium where agents choose $\left(g, x_g^*\right)$ when sunspots are present and (b, x_b^*) when they are absent.[12] If the government must choose τ before the realization of the sunspot variable, it faces exactly problem (3) except that π does not depend on τ. One can think of our approach, therefore, as endogenizing the properties of the extrinsic variable on which agents coordinate their actions. This allows us to study the impact that the government has on this coordination device and the resulting changes in the optimal policy problem.

The crucial question at this point is obviously where the function π comes from. In the next subsection, we show how it can be obtained as the natural result of an adaptive learning process within this environment.

2.3 A Probabilistic Equilibrium Selection Mechanism

In this subsection, we describe how a general learning process induces a probabilistic equilibrium selection mechanism. In section 3, we provide a more detailed analysis of the approach in a specific class of models. Studying learning within this model requires changing the information possessed by agents so that there is something for them to learn about. In addition to the endogenous uncertainty about each other's actions, we introduce uncertainty about the distribution f. This means that agents need to learn about the fundamentals of the economy while they learn about market conditions, and the coevolution of their beliefs about the two objects determines where the economy converges.

Our interpretation of the learning process is similar to that in Lucas (1986), which advocates using learning to investigate the plausibility of different equilibria. Using adaptive behavior to predict the actual performance of the economy along the learning transition does not constitute, for Lucas, a "serious hypothesis." We also do not think of our learning process as an accurate description of the short-run behavior of the economy. Rather, we view it as a mechanism that accurately reflects the likelihood with which the economy will end up in each equilibrium. Our primary interest is in how these likelihoods are affected by the government's actions.

We model the learning process as taking place over an infinite sequence of discrete (artificial)

[12] In general, however, sunspot equilibria need not be mere randomizations over equilibria of the non-sunspots economy, as shown by Cass and Shell (1983).

time periods $t = 0, 1, 2, \ldots$. In period t, each agent must choose an action a_t, and then observe a state c_t. The state c_t is drawn from the true distribution f over C and is drawn independently across periods. As a result, agents will asymptotically learn the true distribution f. The action a_t is chosen to maximize current-period expected utility, given the agent's time t beliefs (the beliefs at $t = 0$ are the initial priors). The agent then observes the average action $\overline{a}_t$ and updates his beliefs, and the process repeats itself.

Formally, let Y_t denote the random vector that describes the possible outcomes of the aggregate economy in period t of the learning process. A realization of Y_t is denoted $y_t = (\overline{a}_t, c_t)$, and therefore we have $y_t \in Y = \overline{A} \times C$. Let Φ represent the set of distributions over Y that agents consider possible for Y_t. We may choose to restrict this set to, say, distributions of a particular parametric form. Agents' beliefs at time t are given by a probability measure over Φ that we denote by μ_t.

Agents begin the learning process with a common prior belief μ_0. We consider a recursive learning process with the following structure

$$\mu_{t+1} = \Gamma_t(\mu_t, y_t). \tag{4}$$

Note that Y_t is, in general, an endogenous random variable. Given the current state of beliefs and the structure of the economy, described by a mapping Ψ, we have

$$Y_t = \Psi(\mu_t; \tau). \tag{5}$$

Together, equations (4) and (5) and the initial prior beliefs μ_0 fully describe the learning system. The fact that the laws of evolution of the endogenous variables are determined in part by the learning process make this system self-referential. (Agents are learning about a system that is being influenced by the learning processes of people like themselves.) This property of the system implies that agents are not learning about a *fixed* data-generating process (see Marcet and Sargent [1989]). The limiting behavior of beliefs is especially complicated due to this fact. We are interested in adaptive rules Γ that satisfy certain "natural" requirements (see Woodford [1990]). For example, we would like to consider learning algorithms that satisfy asymptotic consistency: beliefs converge almost surely to the "true" distribution. In self-referential systems there is actually no "true" (fixed) distribution. Although we could still check this requirement for the non-self-referential part of the system, we will only consider adaptive rules that satisfy an even stronger requirement: convergence of beliefs to the rational expectations equilibrium beliefs. Using the notation above, we can define a rational expectations equilibrium as a random vector $\widehat{Y}$ such that, together with the belief $\widehat{\mu}$ that puts full mass on $\widehat{Y}$, we have

$$\widehat{Y} = \Psi(\widehat{\mu}; \tau). \tag{6}$$

In other words, an equilibrium random vector is such that if agents' beliefs put probability one on that vector, then agents' actions will generate that same vector. We are interested in economies with multiple rational expectations equilibria, that is, with more than one random vector solving (6). Let $\widehat{\Phi}(\tau) \subset \Phi$ be the set of solutions to equation (6) and $\widehat{\Theta}(\tau)$ the corresponding set of belief distributions that put full mass on a particular element of $\widehat{\Phi}(\tau)$. To study the asymptotic behavior of system (4) it is convenient to consider the set Y^{∞} of infinite sequences $\{y_t\}_{t=1}^{\infty}$ and the induced probability measure P over that space (which depends on the prior distribution μ_0 and, through (5), the government policy τ). We then focus on adaptive rules that for all sequences $y \in Y^{\infty}$ satisfy

$$\lim_{t \to \infty} \mu_t = \mu_{\infty} \in \widehat{\Theta}(\tau).$$

Note that this implies two things. First, it implies that the sequence $\{\mu_t\}$ converges and second, that it converges to a distribution that puts full mass on only one element of the set $\widehat{\Phi}$. Finally, note that if the adaptive learning process satisfies this condition, then it will induce a probability distribution over the set $\widehat{\Phi}(\tau)$ that is given by

$$\Pr(\widehat{Y}) = P(y \in Y^{\infty} : \mu_{\infty} = \widehat{\mu} \in \widehat{\Theta}(\tau)\ ;\ \tau).$$

This is the foundation for our proposed equilibrium selection mechanism (ESM). The probability that we assign to each equilibrium is the probability of the set of sequences $\{y_t\}$ that converge to it.

It is important to note that both the endogenous variables $\overline{a}_t$ and the exogenous random variables c_t are essential to our problem. If vector Y_t contained only endogenous variables, then given the prior distribution μ_0, the economy would always follow the same path during the learning process. Then, there would be no chance of convergence with positive probability to more than one rational expectations equilibrium. By the same token, if vector Y_t contained only exogenous (random) variables then the set $\widehat{\Phi}(\tau)$ would be a singleton (containing only the measure that puts full mass on the true distribution of Y_t) and hence there would be no equilibrium selection problem.

It is implicit in the formulation of problem (2) is that the policy maker knows f, the true distribution of random variable c. Recall that our interest is in determining the optimal policy in the static, rational-expectations model of Section 2.1. In that context, the policy maker must make a single choice τ, knowing f. This may seem somewhat at odds with our learning story – the policy maker chooses τ while knowing f, and then agents begin to learn about f. However, this is an unavoidable aspect of using learning to generate an equilibrium selection mechanism. A rational expectations equilibrium of the model in Section 2.1 – our model of interest – requires that everyone, including the government, know the true distribution f.[13] However, if agents began the learning process knowing f, their initial beliefs would uniquely determine the outcome and learning would not be selecting the equilibrium. In other words, if an equilibrium could be selected using methods that are entirely consistent with the original model, there would not have been multiple equilibria to begin with.

In a sense, this is the same type of point made by Morris and Shin (1998) in the context of a model of self-fulfilling currency attacks. However, it is also where our approach differs fundamentally from theirs. We perform equilibrium selection for a given economy, whereas they change the economy so that the equilibrium is unique. In particular, they change the informational structure so that agents receive different signals about economic fundamentals and must act on the basis of this (incomplete) information. The result is, in general, not an equilibrium of the original economy. We keep the original economy as our fundamental object of interest and study an adaptive learning process that converges to an equilibrium of this economy. We think of our approach in the following way. The model in section 2.1 abstracts from the uncertainty that is undoubtedly present in any economic situation. This is a standard approach and, as long as the uncertainty is not too large, seems reasonable in the context of the static model. We then add this uncertainty in for the learning process, where it interacts with agents' beliefs about each other's actions in a crucial way.

Another implication of this approach is that the policy maker cannot change τ during the learning process or after the economy has converged to an equilibrium. This seems a natural requirement in any event for two reasons. First, the policymaking process may be such that changing policies is costly and/or involves time lags. Second, and more importantly, a policy change could easily cause a "jump" in agents' beliefs, which would reset the learning process to some new initial condition. The strategic interplay between policy changes and changes in beliefs is a very difficult problem, and we defer this to future work.

3 Coordination Problems

In order to simplify the analysis, we specialize the model to economies with a coordination problem and two symmetric, pure-strategy equilibria. This is most easily done by assuming that agents face a binary choice, such as the choice between two competing technologies.

3.1 A Binary-Choice Model

We now restrict the choice of agents to include a binary decision, in addition to other payoff-relevant choices (such as how much to produce with the chosen technology). Therefore, we model the agents'

[13] Rewriting the model in Section 2.1 to include uncertainty about f would not resolve the issue; it would only move it to the next level: beliefs about the distribution of the true distribution.

choice set as

$$A = \{g, b\} \times X, \qquad \text{where } X \subset \mathbb{R}^N.$$

(We will later interpret g as choosing the "good" technology and b as choosing the "bad" one.) The analog of this binary choice in game-theoretic analysis is the 2×2 game that has received so much attention in the literature on equilibrium selection.[14]

Define x_g^* as the solution to

$$x_g^* = \arg\max_x \int V\left((g, x), \left(g, x_g^*\right), \tau, c\right) df(c).$$

Then x_g^* will be the value taken by x in an equilibrium where all agents choose g (if such an equilibrium exists). We assume that this equation has a unique solution. The utility value to each agent of being in this equilibrium would then be

$$V_{gg} \equiv \int V\left(\left(g, x_g^*\right), \left(g, x_g^*\right), \tau, c\right) df(c).$$

We similarly define the utility that an agent would receive from choosing b (and then choosing x optimally) when (almost) every other agent is choosing g by

$$V_{bg} \equiv \max_x \int V\left((b, x), \left(g, x_g^*\right), \tau, c\right) df(c).$$

We assume that $V_{gg} > V_{bg}$ holds, meaning that there is indeed an equilibrium where all agents choose g. We define x_b^*, V_{bb}, and V_{gb} in an analogous way, and we assume that $V_{bb} > V_{gb}$ holds. Then V_{bb} is the utility agents receive in the equilibrium where all agents choose b, and V_{bg} is the utility value of deviating unilaterally from this equilibrium. Finally, we assume that $V_{gg} > V_{bb}$ holds, so that g is the "good" equilibrium and b the "bad" one.

There is an important point to be mention here. The definition of rational expectations equilibrium that we will be using (see Section 2.1) is somewhat restrictive, even for this specific kind of coordination game. It is well known that coordination games often admit a large number of correlated equilibria. For example, we could have grouped agents within a finite number of different coordinated groups and then constructed correlated equilibria for which the equilibrium probability distribution over the aggregate state $\overline{a}$ is non-degenerate. This is important because, during the learning process, agents hold prior probabilities over a set of possible nondegenerate distributions of the aggregate state $\overline{a}$ (as opposed to holding prior distributions directly over the possible states $\overline{a}$ or equivalently only over degenerate distributions).[15] Agents are trying to learn about equilibrium

[14] See, for example, Harsanyi and Selten (1988), Kandori, Mailath, and Rob (1993), and Matsui and Matsuyama (1995).

[15] A subset of these correlated equilibria is the set of (symetric information) sunspot equilibria. Our definition rules out

behavior and hence define prior beliefs broadly enough to include all possible types of equilibrium. As it turns out, without the existence of a generally accepted correlating device (see Fudenberg and Tirole [1991]), the final equilibria are always of the type described by our restrictive definition. In addition, much of the analysis in this paper can be extended to include such correlation with only minor modifications.

3.2 Optimal Policy

The traditional optimal policy problem is to maximize $V_{jj}(\tau)$, where j is either g or b depending on which of the two equilibria was selected. We denote the solution to this problem by τ_j^*, where the j subscript indicates that this is the optimal policy when the government knows that equilibrium j will obtain.

In contrast, our approach uses a probability distribution π over the set of equilibria $\{g, b\}$. The optimal policy problem is then given by (3). Without placing further structure on the vector of utilities V, little can be said about the optimal policy τ^*. It is typically not equal to either τ_g^* or τ_b^*. It may fall between these two values, or it may be higher than both. We now look at a specific learning model and the properties of the π that it generates.

3.3 Bayesian Learning and Equilibrium Selection

In this section, we specialize to a particular learning process, that of Howitt and McAfee (1992). We use this process primarily because it has a simple graphical representation that allows us to illustrate how policy affects the equilibrium selection process. We should emphasize, however, that the details of the learning process are not critical for our story. As Section 2.3 indicates, any of a broad class of adaptive processes could be used. The only real requirement is that the process converges to each of the two equilibria with positive probability.

We assume that the exogenous random variable c takes on only two values, c_H and c_L. We will later interpret c as a random fixed cost of operating technology g. In that case, c_L will correspond to a low cost and c_H to a high cost. This assumption implies that f is a Bernoulli distribution; let $\overline{p}$ be the (true) probability of c_L. We use V_{gg}^L (resp. V_{gg}^H) to denote the (*ex post*) utility value of equilibrium g when c takes the value c_L (resp. c_H). Therefore we have

$$V_{gg} = \overline{p} V_{gg}^L + (1 - \overline{p}) V_{gg}^H .$$

Subscripts added to other variables are defined similarly. We use V to denote the vector whose elements are V_{ij}^k for $i = g, b$; $j = g, b$; $k = H, L$.

those as well. For the relationship between sunspots and correlated equilibria, see Peck and Shell (1991).

We also assume that each agent knows that $\overline{a}$ will be either $\left(g, x_g^*\right)$ or $\left(b, x_b^*\right)$. This is a restriction on the support Φ of the set of beliefs. The agent believes that which of these two events occurs is the result of an i.i.d. Bernoulli random variable, with the probability that $\left(g, x_g^*\right)$ is chosen given by $\overline{q}$. Each agent is uncertain about the values of $\overline{p}$ and $\overline{q}$. She has (independent) beliefs about $\overline{p}$ and $\overline{q}$ which are given by beta distributions over $[0,1]$ with means p and q, respectively. The agent begins at $t = 0$ with diffuse (uniform) priors for each variable and updates these beliefs using Bayes' rule.[16] Because all agents begin with the same priors and observe the same information, they hold identical beliefs at every point in time.

We place the following restrictions on the utility values:

$$V_{gb}^L > V_{bb}^L \quad \text{and} \quad V_{bg}^H > V_{gg}^H. \tag{7}$$

The first of these inequalities implies that if p were equal to one, then g would be the optimal choice for the agent regardless of q. In other words, if the agent is optimistic enough about the variable c, she will choose g regardless of her beliefs about market conditions. The second inequality is the reverse; it implies that if p were zero, then b would be the optimal choice regardless of q. This implies that agents believe it is possible that either one of the equilibria exists or that both exist. We will see below that these assumptions guarantee that the economy has positive probability of converging to each equilibrium from any current set of beliefs. This prevents initial beliefs from having too strong of an effect on the final outcome.

Intuitively, the learning process works as follows. The agent knows that either the good technology will catch on or the bad one will. In other words, one of the two symmetric equilibria will occur. What she does not know is what probability to assign to the good outcome. In the same way, she knows that the cost of operating the good technology will be either low or high, but she does not know what probability to assign to the low cost. She begins the learning process with beliefs about each of these probabilities. In each period, after making a decision, she receives two signals that she uses to update these beliefs for the next period. One signal is a random draw from the cost distribution, either c_L or c_H. With this she updates her belief about the likelihood that c_L will be the true state. The other signal is about the market. She observes either $\left(g, x_g^*\right)$ or $\left(b, x_b^*\right)$ and uses this to update her belief about the likelihood that the good technology will catch on. There is no "true" distribution for this signal to be drawn from; instead, the signal is equal to the optimal response given the agents'

[16] Diffuse priors are a standard way of representing "minimal" prior knowledge about a parameter. See, for example, Zellner (1971).

current (shared) beliefs. The expected utility given belief (p_t, q_t) of each choice is given by

$$\begin{aligned} g &: \; p_t q_t V_{gg}^L + (1-p_t) q_t V_{gg}^H + p_t (1-q_t) V_{gb}^L + (1-p_t)(1-q_t) V_{gb}^H \\ b &: \; p_t q_t V_{bg}^L + (1-p_t) q_t V_{bg}^H + p_t (1-q_t) V_{bb}^L + (1-p_t)(1-q_t) V_{bb}^H . \end{aligned}$$

Straightforward algebra shows that the agent chooses g if we have[17]

$$q_t \geq \frac{V_{bb}^H - V_{gb}^H - \left(V_{gb}^L - V_{gb}^H - V_{bb}^L + V_{bb}^H\right) p_t}{\left(V_{gg}^H - V_{gb}^H - V_{bg}^H + V_{bb}^H\right)(1-p_t) + \left(V_{gg}^L - V_{gb}^L - V_{bg}^L + V_{bb}^L\right) p_t}. \tag{8}$$

For a given level of p_t, the agent will prefer technology g if she thinks it is likely enough that other agents will be choosing technology g. Following Howitt and McAfee (1992), we can represent this graphically as in Figure 1.

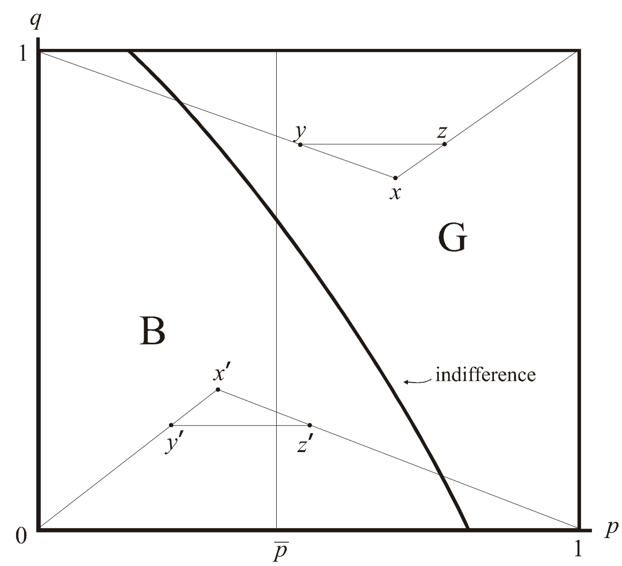


Figure 1: The dynamics of beliefs

The box in Figure 1 represents the set of all possible (p, q) pairs. We define the set G to be the set of points (p_t, q_t) such that (8) holds. The set B is the complement of G. It is straightforward to show that (7) implies that the curve separating these regions must begin to the right of $(0, 1)$ and end to the left of $(1, 0)$. One can show that the curve is smooth, strictly decreasing, and that it can be either convex or concave depending on the functional forms chosen (in Howitt and McAfee [1992] and in the examples in Section 4, it is linear).

Let (p_t, q_t) denote the expected value of (p, q) according to the beliefs held by agents at date t. In particular, diffuse priors imply that we have $p_0 = q_0 = \frac{1}{2}$. In each iteration of the learning algorithm,

[17] Here we are imposing the tie-breaking rule that agents choose g when they are indifferent. This is not important for the results; the only real requirement is that all agents take the same action.

the observation of $(\overline{a}_t, c_t)$ provides useful information for updating (p_t, q_t). Bayesian updating of beliefs allows us to write the expected value of the parameter of the distributions after t observations as

$$p_{t+1} = \left\{ \begin{array}{l} \eta_t p_t \\ \eta_t p_t + (1 - \eta_t) \end{array} \quad \text{if} \;\; c_t = \begin{array}{c} c_H \\ c_L \end{array} \right\} \tag{9}$$

and

$$q_{t+1} = \left\{ \begin{array}{l} \eta_t q_t \\ \eta_t q_t + (1 - \eta_t) \end{array} \quad \text{if} \;\; (p_t, q_t) \in \begin{array}{c} B \\ G \end{array} \right\}, \tag{10}$$

where $\eta_t = (t+2)/(t+3)$, for $t = 0, 1, 2, \ldots$. Such Bayesian learning has a nice representation in Figure 1. If the time t beliefs fall in G, all agents will choose $\left(g, x_g^*\right)$ and hence the value of q will increase. As Howitt and McAfee (1992) point out, the posterior beliefs always lie on the line segment connecting the prior beliefs with one of the corners of the box. From point x, for example, we would move to y if c_H is observed and to z if c_L is observed. Similarly, if the original point is x' (in B), $\overline{a} = (b, x_b^*)$ will be observed, and q will decrease. We move to point y' if c_H is observed and to z' if c_L is observed.

Suppose that in period t, agent's beliefs are represented by the point x. To which equilibrium will the economy converge? The answer depends crucially on the sequence of realizations $\{c_t\}$, especially the next few observations. Because x falls in G, we know that agents will be choosing g and hence q_t will be rising. Suppose, however, that the economy is "unlucky" and receives a string of realizations of c_H. Then p_t will be falling and eventually beliefs will cross into region B. At this point, agents will begin to choose b and q_t will start to fall.

Bayesian updating consistently estimates the value of $\overline{p}$, that is, we have $p_t \to \overline{p}$ almost surely as $t \to \infty$. Whether q_t converges to zero or one depends (very roughly speaking) on whether beliefs are in G or B when p_t settles down. A sufficiently unlucky sequence of realizations of c_t will lead the economy into region B and therefore make convergence to the bad equilibrium likely. Conversely, a sufficiently lucky economy will be driven into G, making convergence to the good equilibrium likely. Howitt and McAfee (1992) formalize this argument showing that the probability of converging to each equilibrium is positive. We now show that the learning process must converge.[18] We then use monte carlo simulation to confirm that the system only converges (with positive probability) to the two rational expectations equilibria. These two results combine to show that the learning process generates a valid equilibrium selection mechanism. The proof of the first result and the description of the computations for the second are contained in the appendix.

[18] Convergence is guaranteed in purely Bayesian models because beliefs form a martingale. This result does not apply to our setting, however, because agents learn using a misspecified model. It is well known that in such an environment, learning need not converge (hence the importance of this first proposition). See Blume and Easley (1998) and Nyarko (1991).

Proposition 1 *The learning process* $\{p_t, q_t\}$ *converges with probability one.*

Proposition 2 *Let* π *be the empirical probability of the set of sequences* $\{c_t\}$ *such that* $\{p_t, q_t\} \to (\overline{p}, 1)$ *. Then the empirical probability of the set of sequences* $\{c_t\}$ *such that* $\{p_t, q_t\} \to (\overline{p}, 0)$ *is equal to* $(1 - \pi)$ *. In other words, the simulation of the learning process always converges to a symmetric rational expectations equilibrium.*

The value of π clearly depends on the position of the curve in Figure 1 and will change when this curve is shifted by, say, a policy change. The curve can be interpreted in a way that relates it to the risk factor of each equilibrium. Following Young (1998), we define the *risk factor of equilibrium* $j \in \{g, b\}$ to be the smallest probability ρ such that if an agent believes $\overline{a}$ will be $\left(j, x_j^*\right)$ with probability strictly greater than ρ, then $\left(j, x_j^*\right)$ is her unique optimal action. In Figure 1, define $\overline{q}$ to be the level of q where the curve separating the two regions crosses the $\overline{p}$ line. Then the risk factor of equilibrium g is given by $\overline{q}$, while that of equilibrium b is given by $(1 - \overline{q})$. It is clear from the diagram that this alone is not enough to determine π – the position of the entire curve matters. Because agents do not know $\overline{p}$ during the learning process, it matters how risky each strategy seems for each possible belief p. This leads us to give the following definition.

Definition: The *risk factor* of action j *given belief* p is the smallest probability ρ such that if an agent's beliefs about $\overline{p}$ have mean p and the agent believes $\overline{a}$ will be $\left(j, x_j^*\right)$ with probability strictly greater than ρ, then $\left(j, x_j^*\right)$ is her unique optimal action.

The risk factor of action g given belief p is equal to the height of the line at p in Figure 1. (Note that this is equal to one for low enough values of p and zero for high enough values). In the next proposition, we show that enlarging the region G (by shifting the curve down in some way) strictly increases the probability of attaining the good equilibrium. Hence, a change in the economy that uniformly lowers the risk factor of an equilibrium will increase the probability that the economy reaches that equilibrium.

Proposition 3 *If the risk factor of action* $j \in \{g, b\}$ *given belief* p *decreases for some* p *and does not increase for any* p, *then the probability of equilibrium* j *strictly increases.*

Even though the statement of this proposition is rather intuitive, the proof is fairly complex. It involves establishing that when the curve shifts, (i) the area between the new and old curves is visited with positive probability and (ii) the asymptotic behavior of a trajectory that visits this area is changed with positive probability. The proof is contained in the appendix.

The condition that the risk factor not increase for any belief p may seem strong, but changes in the individual elements of vector V have exactly this effect. This can be seen by working with equation (8). For every value of p_t such that the curve is in the interior of the box, q_t is either strictly increasing or strictly decreasing in each element of V. Using Proposition 3, this implies that π is monotone in each of the elements of vector V. This is important because it is through these elements that the policy parameter τ affects π. We state this result as a corollary.

Corollary 1 *The value of π is strictly increasing in V_{gg}^H, V_{gg}^L, V_{gb}^H, and V_{gb}^L. It is strictly decreasing in V_{bb}^H, V_{bb}^L, V_{bg}^H, and V_{bg}^L.*

We now apply this approach to two simple examples and show how it generates interesting insights into the corresponding optimal policy problems.

4 Applications

In this section, we apply our approach to the type of technology-choice problems that we discussed in the introduction. We give two examples, one that captures the central features of the debate over the Internet sales tax and another that applies to problems of technology choice more generally. Each example puts a different structure on the utility functions V and how they are affected by government policy; this allows us to highlight different features of the optimal policy that may emerge under our approach.

4.1 Taxing Internet Transactions

We first present a simple model that captures some of the crucial features of the Internet sales tax debate. In line with the analysis above, there is a continuum of identical agents. Each agent gains utility from "transacting" with other agents. These transactions occur through one of two technologies, which we label g and b. Technology g (the "good" one) represents Internet transactions and technology b (the "bad" one) traditional store-based methods. Transacting requires effort and we denote the agent's choice of effort level by $x_i \in [0, \infty)$. The utility cost of this effort is quadratic. The utility derived from this effort depends on the productivity of the transactions technology, which is characterized by a network externality. The agent's utility level is given by

$$u_i = \left\{ \begin{array}{c} x_i f(\overline{x}_g) - \frac{a}{2}x_i^2 - c \\ x_i - \frac{a}{2}x_i^2 \end{array} \right\} \text{ if technology } \left\{ \begin{array}{c} g \\ b \end{array} \right\} \text{ is used.}$$

The term $\overline{x}_g$ is the total amount of effort employed by agents using technology g, that is, $\overline{x}_g = \int_\Gamma x_i di$, where Γ is the set of agents using technology g. The function f is increasing and strictly concave; higher levels of total effort in technology g make each agent using the technology more productive (this is the external effect that will generate multiple equilibria). In other words, Internet-based transacting becomes more efficient as more people use it. We assume that $f(0) > 0$ holds, as does the Inada-type condition

$$\lim_{x\to\infty} f'(x) = 0. \tag{11}$$

In addition, the fixed utility cost c of operating technology g is stochastic. Keeping with the analysis above, we assume $c \in \{c_L, c_H\}$, with $c_L < c_H$ and the probability of c_L given by $\overline{p}$. We denote the expected value of c by $\overline{c}$. Everything is measured in terms of utility, so the agent's objective function will be linear in probabilities and only the expected cost $\overline{c}$ will matter in a rational expectations equilibrium.

This is obviously a very stylized model of exchange. A detailed analysis of government policy and equilibrium selection in an economy with explicit search frictions and decentralized exchange can be found in Ennis and Keister (2000).

Because of the externalities, it seems natural for the government to consider intervening to encourage effort in technology g. We assume that the government does this by subsidizing such effort. Let $\tau \in [0, 1)$ be the rate of *ad valorem* subsidy and let T be the lump-sum tax (on all agents) that finances this subsidy. This subsidy can be thought of as the discount from the standard sales tax that Internet transactions receive. A passive government would charge the same tax rate on both types of transactions. An active government would lower the rate on Internet commerce, effectively subsidizing electronic transactions. This subsidy would be paid for by all agents through either higher tax rates on other activities or reduced government expenditures. For simplicity, we assume that this policy does not affect the efficiency or the cost of effort for agents using technology b (that is, that the government does not raise the tax rate on traditional commerce to pay for the e-commerce subsidy).

4.1.1 Equilibrium

We first examine the set of rational expectations equilibria for a given government policy. If the agent chooses technology b, he will choose his effort level to solve

$$\max_{x_i}\ x_i - \frac{a}{2}x_i^2 - T.$$

The solution to this problem is given by $x_i = \frac{1}{a}$. If instead the agent chooses technology g, his problem is

$$\max_{x_i}\ x_i f(\overline{x}_g) - (1-\tau)\frac{a}{2}x_i^2 - \overline{c} - T, \tag{12}$$

which is solved by

$$x_i = \frac{f(\overline{x}_g)}{(1-\tau)\,a}. \tag{13}$$

The agent then compares the expected utility given by each technology

$$\begin{aligned} g &: \frac{f(\overline{x}_g)^2}{2(1-\tau)a} - \overline{c} - T \\ b &: \frac{1}{2a} - T \end{aligned}$$

and chooses the more promising one. This choice clearly depends on the agent's beliefs about $\overline{x}_g$. In our rational expectations equilibria, this number is known with certainty. We only look at symmetric equilibria, where all agents act identically. We look first for the good equilibrium, where $\overline{x}_g$ is positive. From (13), the equilibrium effort level is given by the unique solution to

$$x_g^* = \frac{f(x_g^*)}{(1-\tau)a}. \tag{14}$$

This generates utility level

$$V_{gg}(\tau) = \frac{f(x_g^*)^2}{2(1-\tau)a} - \overline{c} - T,$$

where the value of T is given by the government's budget constraint

$$T = \frac{\tau a}{2}(x_g^*)^2.$$

This is an equilibrium as long as no individual agent is made better off by choosing technology b. That is, we need

$$V_{gg}(\tau) > V_{bg}(\tau) = \frac{1}{2a} - T$$

to hold. We assume that the parameter values are such that we have $V_{gg}(0) > V_{bg}(0) = V_{bb}$, so that the good equilibrium exists when the government is passive and Pareto dominates the bad equilibrium. It is straightforward to show that the difference $(V_{gg} - V_{bg})$ is strictly increasing in τ, and therefore the good equilibrium exists for all subsidy levels.

We also assume that the parameter values are such that the bad equilibrium exists when the government is passive (otherwise the problem of equilibrium selection does not arise). For this, we require

$$V_{bb} > V_{gb}(0)$$

or

$$\frac{1}{2a} > \frac{f(0)^2}{2a} - \overline{c}.$$

This does not, however, mean that the bad equilibrium exists for all values of τ. By setting

$$\tau > \overline{\tau} = 1 - \frac{f(0)^2}{1 + 2a\overline{c}},$$

the government can eliminate the bad equilibrium (and thus be sure that the good equilibrium will obtain). Finally, we assume that the conditions in (7) hold for all $\tau < \overline{\tau}$. Notice that, while all of the previous assumptions depend on $\overline{c}$, these two depend on c_L and c_H, respectively. As a result, they can be satisfied for any $\overline{c}$ by making the spread between c_L and c_H large enough.

4.1.2 Learning and Optimal Policy

Learning in this example proceeds exactly as described in Section 3.3. From Proposition 1, we know that the equilibrium selection mechanism π is well-defined for this example. We can use Proposition 3 to determine how the policy variable τ affects π.

Proposition 4 *In this example, π is strictly increasing in τ for $\tau < \overline{\tau}$.*

Proving this proposition entails showing that increasing the subsidy rate strictly expands the region G in Figure 1. The details are contained in the appendix, but the intuition is straightforward. Increasing the subsidy to technology g always makes it more attractive relative to b (since the lump-sum tax is paid by the agent regardless of her choice). This leads agents to choose g for a wider range of beliefs, which makes convergence to the good equilibrium more likely. This does *not*, however, imply that a higher level of subsidy is always better from a welfare standpoint, as we show below.

The simplifying assumptions of this example allow us to gain a fair amount of insight into the nature of the optimal policy problem. The utility value of the bad equilibrium is independent of the policy chosen (if no one engages in Internet transactions, the sales tax rate on such transactions is irrelevant). Hence the only relevant factors for determining the expected utility of agents are the utility value of the good equilibrium V_{gg} and the probability of reaching that equilibrium π. The optimal policy problem can be written as

$$\max_{\tau} \ \pi\left(\tau\right)\left(V_{gg}\left(\tau\right) - V_{bb}\right). \tag{15}$$

There are two cases to consider. If we have

$$\tau_g^* > \overline{\tau},$$

then the policy that maximizes the value of the good equilibrium also eliminates the bad equilibrium, ensuring that the good equilibrium will obtain. In this case, correcting the externality present in the good technology is sufficient to make that technology a dominant choice and therefore to eliminate the coordination problem. When this happens, τ_g^* is clearly the optimal policy.

The more interesting case is where the coordination problem remains even when the externality is being corrected, or when we have

$$\tau_g^* < \overline{\tau}.$$

In this case, the government faces a tradeoff. By increasing the subsidy level above τ_g^*, the good equilibrium becomes less attractive. However, deviating from the good equilibrium also becomes less attractive and thereby makes the good equilibrium more likely. This tradeoff is illustrated in the two panels of Figure 2. For each value of τ, we plot the pair $(V_{gg} - V_{bb}, \pi)$ generated by the policy.

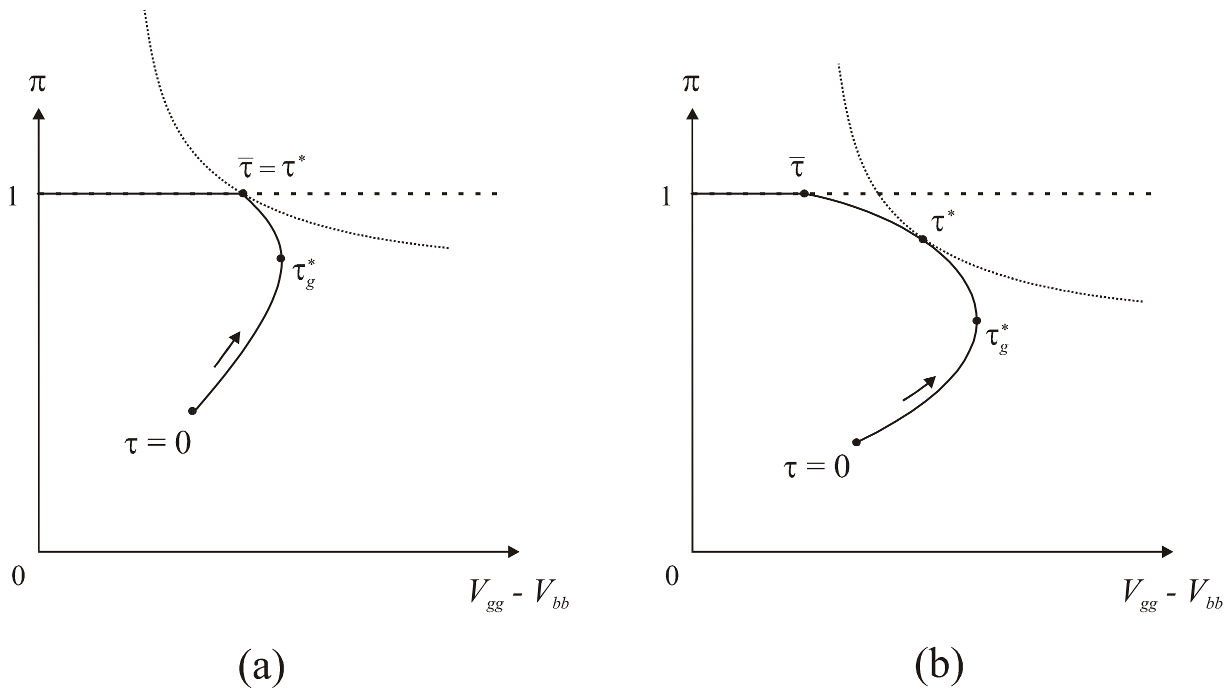


Figure 2: Optimal policy

The point generated by $\tau = 0$, for example, has a value of π strictly between zero and one, since we assumed that both equilibria (strictly) exist when the government is passive. As we increase τ, we trace out a curve of feasible points. The arrows in the figure indicate the direction of movement along the curve as τ increases. Initially, both V_{gg} and π are increasing in τ. When we reach the policy τ_g^*, we know that V_{gg} is at its maximum level. For higher subsidies, V_{gg} starts to fall, but π continues to increase until it reaches unity at subsidy level $\overline{\tau}$. Increasing τ beyond this is clearly inefficient as π cannot increase further and V_{gg} continues to decrease.

The level curves of the objective function in (15) are hyperbolas. If eliminating the bad equilibrium is not very costly, as in part (a) of the figure, then $\overline{\tau}$ is the optimal policy. If, however, the situation is as in part (b) of the figure, the optimal policy will fall somewhere in between τ_g^* and $\overline{\tau}$. In this case, the optimal policy under our approach is different than that which would be derived under any deterministic selection criterion. In particular, it should be clear from the diagram that the optimal policy will never be τ_g^*. This is because increasing τ a little past this point causes a small (second-order) loss in the value of the good equilibrium and brings a larger (first-order) increase in the probability of reaching that equilibrium.

What does this tell us about the debate over taxing Internet transactions? The model tells us that Internet commerce *should* be subsidized for two reasons. The first is straightforward – we assumed that there is a network externality that should be corrected. However, the optimal subsidy level is necessarily *higher* than the level that would just correct the externality because higher subsidies make the good equilibrium more likely to obtain. This second reason for the subsidy is new to our approach and can only be seen in a model where the subsidy can affect the equilibrium selection process. In addition, as part (b) of Figure 2 shows, it may not be optimal to subsidize e-commerce so much that the good equilibrium is certain to obtain. Instead, it may be optimal to face some risk over which equilibrium is selected because eliminating the bad equilibrium is too costly.[19]

4.2 The Choice of Technology

We can modify the above example to study situations where there are two available technologies, both of which are subject to network externalities.[20] There is again a continuum of identical agents whom we now think of as producing a single commodity and consuming their own output. Each agent has available two production technologies, g and b, and can operate only one of them. Both technologies again require (costly) effort as an input, and the agent chooses an effort level $x_i \in [0, \infty)$. Utility is linear in output and is given by

$$u_i = \left\{ \begin{array}{c} g x_i f\left(\overline{x}_g\right) - \frac{a}{2} x_i^2 - c \\ b x_i f\left(\overline{x}_b\right) - \frac{a}{2} x_i^2 \end{array} \right\} \text{ if technology } \left\{ \begin{array}{c} g \\ b \end{array} \right\} \text{ is used.}$$

As before, the function f represents the network externality which now applies to both technologies. We maintain all of our assumptions about f from the previous section, including $f(0) > 0$. (This last condition will help guarantee that there are no zero-effort equilibria in this setting.) We assume $g > b$, so that technology g has a higher marginal product than technology b for a given amount of total effort. As before, there is a stochastic fixed utility cost c of operating technology g.

We again assume that the government subsidizes effort. Let τ be the rate of *ad valorem* subsidy and let T be the lump-sum tax that finances this subsidy. We now assume that the same subsidy level must apply to all effort (the government cannot distinguish effort devoted to technology b from effort devoted to technology g) and likewise, the same tax is paid by all agents. Hence, we are not allowing

[19] Indeed, we have abstracted from the distortions caused by increasing other tax rates, which would increase the cost of eliminating the bad equilibrium further.

[20] A well-known example of such a situation was the adoption of video cassette recorders with the competing Beta and VHS technologies. See Katz and Shapiro (1986) and the references therein for a discussion of this and other examples.

the government to pick the "winning" technology by subsidizing it and taxing the other. Instead, the government can only encourage (or discourage) the entire industry. Because the two technologies differ, the chosen level of encouragement will affect the equilibrium selection process.

4.2.1 Equilibrium

Regardless of the technology $j \in \{g, b\}$ chosen by the agent, his optimization problem will resemble (12) above. The solution is therefore of the same form as (13), or

$$x_i = \frac{jf(\overline{x}_j)}{(1-\tau)a}. \tag{16}$$

The agent compares the expected utility generated by each technology

$$\begin{aligned} g &: \frac{(gf(\overline{x}_g))^2}{2(1-\tau)a} - \overline{c} - T \\ b &: \frac{(bf(\overline{x}_b))^2}{2(1-\tau)a} - T \end{aligned}$$

and chooses the more promising one. We again look at symmetric equilibria, where one of the two numbers $\overline{x}_j$ will be positive and the other zero. We look first for the good equilibrium, where $\overline{x}_g$ is positive. As above, the equilibrium effort level is given by the unique solution to (14). We then have

$$V_{gg}(\tau) = \frac{\left(gf\left(x_g^*\right)\right)^2}{2(1-\tau)a} - \overline{c} - T_g$$

and

$$V_{bg}(\tau) = \frac{(bf(0))^2}{2(1-\tau)a} - T_g,$$

where, from the government's budget constraint we have

$$T_g = \frac{\tau a}{2}\left(x_g^*\right)^2.$$

We assume that the parameter values are such that we have $V_{gg}(0) > V_{bg}(0)$ and hence the good equilibrium exists when the government is completely passive. It is straightforward to show that this implies that the good equilibrium exists for all values of τ.

Next, we look at an equilibrium where $\overline{x}_b$ is positive. Such an equilibrium is a solution to

$$x_b^* = \frac{bf(x_b^*)}{(1-\tau)a}.$$

This leads to

$$V_{bb}(\tau) = \frac{(bf(x_b^*))^2}{2(1-\tau)a} - T_b$$

and

$$V_{gb}(\tau) = \frac{(gf(0))^2}{2(1-\tau)a} - \overline{c} - T_{b,}$$

where T_b is defined similarly to T_g above. We also assume that we have $V_{bb}(0) > V_{gb}(0)$, so that the bad equilibrium exists when there is no government intervention. In this example, the difference $(V_{bb} - V_{gb})$ is strictly increasing in τ, which implies that the bad equilibrium exists for all values of τ. Hence, unlike in the previous example, policy cannot be used here to eliminate the bad equilibrium. We also assume that (7) holds and that we have $V_{gg}(0) > V_{bb}(0)$, so that g is indeed the good equilibrium.

4.2.2 Learning and Optimal Policy

Appealing to Proposition 1 again gives us an equilibrium selection mechanism π. Because of the additional complexity of this example (all of the elements of the vector V now depend directly on τ), drawing diagrams such as those in Figure 2 is not possible. For this reason, we move directly to numerical analysis. We use the following functional form and parameter values

$$f(x) = (x+1)^{\frac{1}{2}},\ a = 2\ , b = 1\ , g = 1.4\ , \overline{c} = 0.5$$

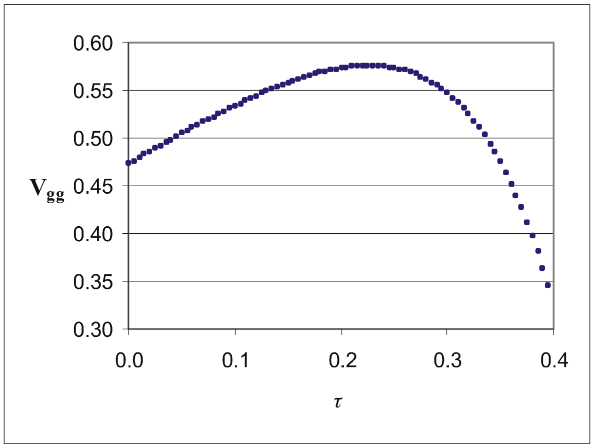
Furthermore, we take $c_H = 2$, $c_L = -1$, and $\overline{p} = \frac{1}{2}$.

We use a grid of values for $\tau \in [0, 0.4]$, with a step size of 0.005. For each value of τ, we simulate the learning process $500{,}000$ times and set $\pi(\tau)$ equal to the fraction of times the process converges to the good equilibrium.[21] The results of this exercise are presented in a series of graphs in Figure 3. The first four graphs plot the elements of V as functions of the policy τ. (These are just the graphs of the expressions given in the previous section.) The last two graphs are the results of the simulations; they give π and expected utility as functions of τ. The former shows π to be increasing, as expected.[22] The latter shows that the optimal policy is $\tau^* = 0.24$, larger than both $\tau_b^* = 0.185$ and $\tau_g^* = 0.225$.

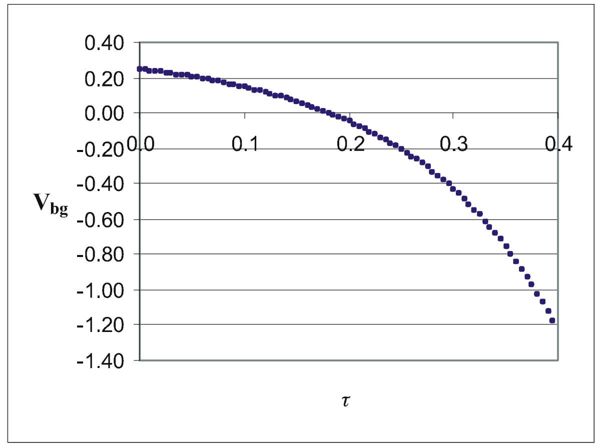
This last result clearly demonstrates the importance of considering equilibrium-selection effects in determining the optimal policy. In this example, τ_b^* and τ_g^* are the obvious candidates for the optimal policy. A policy maker who faces uncertainty about which equilibrium will obtain might be tempted to choose something in between these two values. This would be the correct approach if π did not depend on τ. Our analysis shows, however, that such a choice is *not* correct for the chosen parameter values. As the subsidy level is increased beyond τ_g^*, the utility value is decreasing for both equilibria.

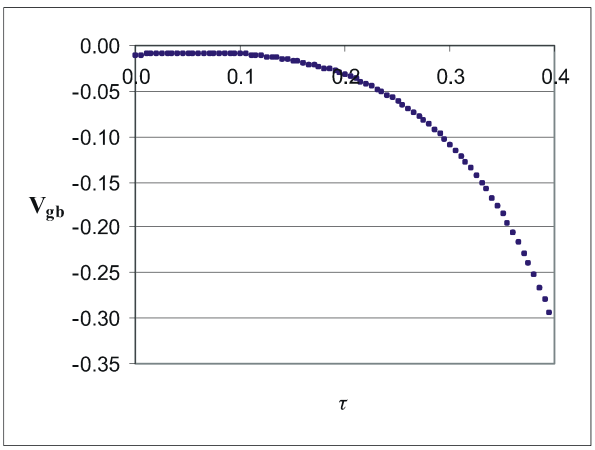
[21] The computations are done in FORTRAN, using the RAN2 algorithm for generating random numbers. The source code is available from the authors upon request.

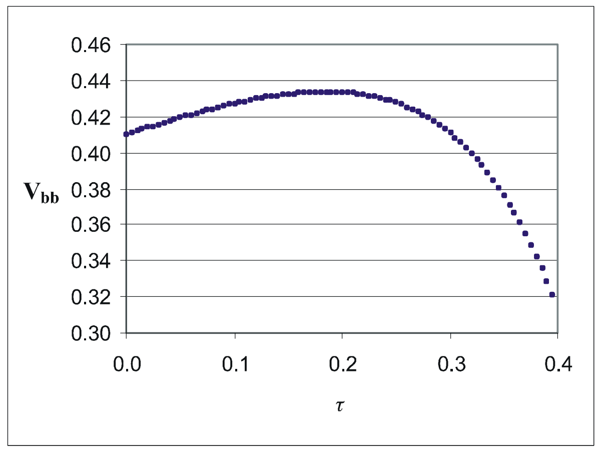
[22] The discontinuities in π come from the discrete nature of the learning process. Some of the points in the box in Figure 1 are visited much more often than others. The value of τ that first makes such an "important" point fall in region G is accompanied by a discrete jump in π.



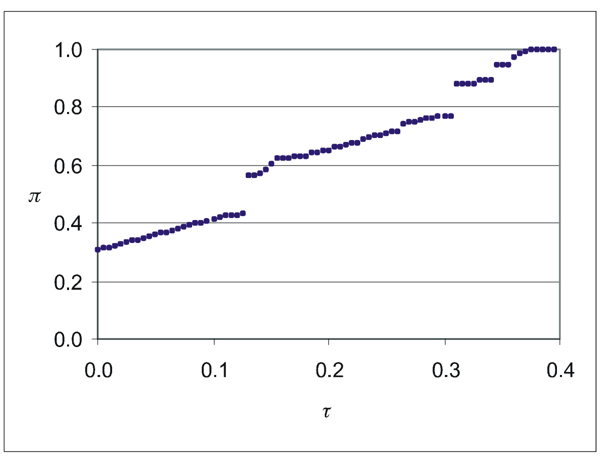
$\tau_g^* = 0.225$

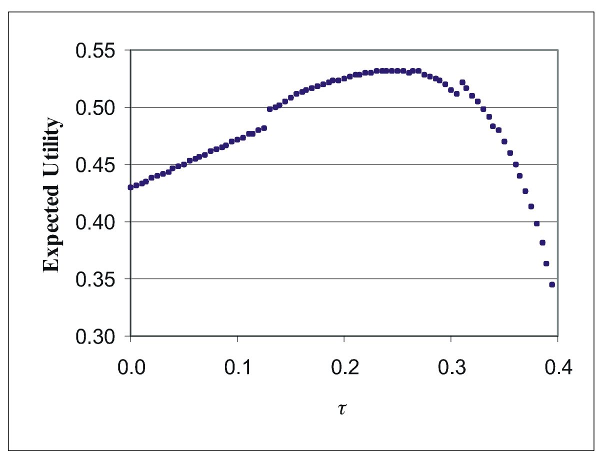






$\tau_b^* = 0.185$





$\tau^* = 0.240$

Figure 3: Computational results

The good technology suffers less from this increase, though, and hence becomes a relatively more attractive option. As a result, the probability of the good equilibrium is still increasing. As long as this effect is large enough (relative to the decreases in the utility values), expected utility continues to increase.

Notice that the government in this example does not need to know which technology is good and which is bad since the same subsidy level is applied to both. The point is that increasing the subsidy to the entire industry will have a positive effect of equilibrium selection, simply because the good technology always benefits more (or suffers less) from such an increase. This example shows that policies designed to correct for network externalities can be more powerful (and therefore important) than is currently recognized. As a result, calculations that ignore the process by which an equilibrium is selected can substantially underestimate the optimal subsidy level.

5 Concluding Remarks

The main point of this paper is that using a probabilistic equilibrium selection mechanism can bring models of multiple equilibria to bear on policy questions in interesting and informative ways. We believe that the probabilistic view of equilibrium selection is both appealing and plausible, and we have shown how adaptive learning naturally generates such a mechanism. We have also shown through examples that taking the equilibrium-selection effect into account can reveal that policy may be more potent than is commonly recognized. We have illustrated our approach using a very simple model: a static model with identical agents making a binary choice and a focus on symmetric equilibria. This allowed us to present the issues involved and to discuss the workings of the learning process as transparently as possible, but it certainly does not mean that the general approach applies only in such models. In other work (Ennis and Keister [2000]), we use this approach in a model with explicit search frictions and decentralized exchange. In that context we discuss optimal aggregate demand management policies, taking into account the effect that policy can have on the equilibrium selection process. This is only one of many possible applications. We conclude by mentioning two other areas that we find promising.

Financial Crises: Models of financial crises are often of the coordination-problem type that we have studied here. We discussed above the relationship between our approach and that of Morris and Shin (1998), who deal with currency crises. Our approach could easily be applied to their model and would allow one to answer questions such as: under what conditions is it optimal for the government to allow currency crises to occur with positive probability? Both Cole and Kehoe (2000) and Chari and Kehoe (2000) consider similar issues, and our approach could be brought to bear on their questions. Cooper and Corbae (2000) study financial collapses (such as occurred in the Great Depression) as a form of coordination problem. Again, our approach could be used to study the question of optimal policy and whether or not that entails completely eliminating the possibility of a collapse.

Monetary Economics: It is well known that monetary models typically exhibit multiplicity of equilibrium, with at least one equilibrium where money has no value. This is true in overlapping-generations as well as search-theoretic environments. Suppose an economy is in a monetary steady state and the government is considering a substantial change in policy. Extensions of our approach to these settings would allow one to address questions such as: What is the probability that the policy change will cause the economy to switch to another (possibly hyperinflationary) equilibrium? Taking this into account, what is the optimal policy? Although extending our analysis to these dynamic settings is far from trivial, the potential payoff seems very high.

Appendix A. Proofs

Proposition 1: The learning process $\{p_t, q_t\}$ converges with probability one.

Proof: The dynamics of p_t are independent of q_t and represent a standard statistical learning process. Let Ω be the set of all possible sequences $\{c_t\}_{t=1}^{\infty}$ and ω be an element of this set. Let $A \subset \Omega$ be the set of ω such that $p_t(\omega) \to \overline{p}$; by the strong law of large numbers we know that the probability of the set A is one. We will show that for each ω in A, the learning process converges.

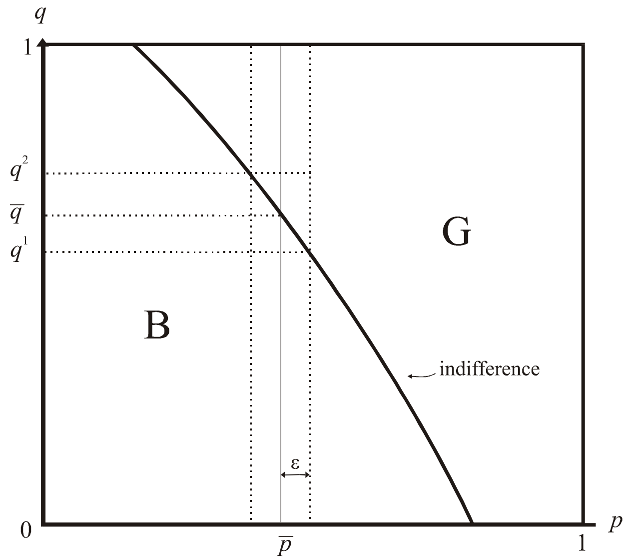


Figure 4: Convergence of beliefs

Let the function $q = f(p)$ represent the curve separating the two regions in Figure 4. Consider a sequence of small numbers $\{\varepsilon_n\}$ converging to zero. For each n, define

$$\begin{aligned} q_n^1 &= f(\overline{p} + \varepsilon_n) \\ q_n^2 &= f(\overline{p} - \varepsilon_n) . \end{aligned}$$

Notice that we have $q_n^1 < \overline{q} < q_n^2$ for all n. Fix a particular ω in A, so that we know $p_t(\omega)$ converges to $\overline{p}$. Then for each $\varepsilon_n > 0$, there exists a t_n such that $t > t_n$ implies

$$|p_t(\omega) - \overline{p}| < \varepsilon_n.$$

That is, for any small band around $\overline{p}$, the sequence $p_t(\omega)$ will eventually enter the band and never leave. If $q_t(\omega)$ is ever sufficiently low after this happens, that is, if we have

$$q_t(\omega) < q_n^1 \qquad \text{for any } t \geq t_n, \tag{A-1}$$

then the trajectory will never switch regions again (because doing so would require leaving the ε-band). Hence all future observations will involve all agents choosing b, and therefore q_t will converge to zero. Similarly, if we have

$$q_t(\omega) > q_n^2 \qquad \text{for any } t \geq t_n, \tag{A-2}$$

all future observations will involve all agents choosing g, and $q_t(\omega)$ will converge to one. Therefore if, for any n, either (A-1) or (A-2) is satisfied, the learning process converges to one of the two equilibria. The conclusion of the proposition is therefore established unless we have

$$q_t(\omega) \in \left[q_n^1, q_n^2\right] \qquad \text{for all } t \geq t_n, \text{ for all } n.$$

The continuity of f implies that q_n^1 and q_n^2 both converge to $\overline{q}$ as n goes to infinity, so in this case $q_t(\omega)$ must converge to $\overline{q}$. This establishes our claim. ■

Proposition 2: Let π be the empirical probability of the set of sequences $\{c_t\}$ such that $\{p_t, q_t\} \rightarrow (\overline{p}, 1)$. Then the empirical probability of the set of sequences $\{c_t\}$ such that $\{p_t, q_t\} \rightarrow (\overline{p}, 0)$ is equal to $(1-\pi)$. In other words, the simulation of the learning process always converges to a symmetric rational expectations equilibrium.

Proof: Note that, because of proposition 1, we only need to verify that the system does not converge to the point $(\overline{p}, \overline{q})$. The law of motion for q_t has a discontinuity along the curve in Figure 4, which passes through $(\overline{p}, \overline{q})$. This makes the asymptotic behavior around that point very difficult to study analytically. For this reason, we turn to a numerical methodology. From (8), we have that the curve separating the two regions is of the form

$$q = \frac{\alpha - p}{\beta(1-p) + \gamma p},$$

where α, β, γ satisfy

$$0.5 > \alpha > 1,$$
$$\beta < \alpha,$$

and

$$\beta + \gamma > 2\alpha - 1 > 0.$$

We fix $\overline{p} = 0.5$ and construct a grid over the three-dimensional parameter space containing a total of 20 points in each dimension. Points in this grid that satisfy the previous inequalities correspond to different shapes of the curve in Figure 4.[23] For each such point, we simulate $1,000$ runs of the

[23] To cover uniformly all possible shapes of the curve we construct the grid using an auxiliary variable, $\xi = \gamma - \beta$. The restrictions over parameters imply that $\xi > -1$. When $\xi < 0$ the curve is concave and when $\xi > 0$ the curve is convex. Hence, we consider a grid that concentrates half of the points in the negative range of ξ and half in the positive range.

learning process.

We set convergence bounds for q_t in the following manner. First we compute $\overline{q}$. Then we define the variables $bound_G = \min\{(1-\overline{q})/5, 0.1\}$ and $bound_B = \min\{\overline{q}/5, 0.1\}$. When q_t goes beyond $bound_G$ and the system has not switched zones for the last $2,000$ iterations, we say the economy has converged to the good equilibrium. For the bad equilibrium we use a similar procedure when q_t goes below $bound_B$. If neither of these events has occurred after $300,000$ iterations, we say that the economy did not converge to one of the two equilibria. (This would be the case, for example, if q_t were to converge to $\overline{q}$.)

The convergence bounds may seem somewhat large, but it should be kept in mind that the step size of a Bayesian learning process decreases fairly rapidly. As an example, a process that reaches $q_t = 0.1$ after $2,000$ steps and that continues monotonically approaching the bad equilibrium will take over $18,000$ more steps to reach $q_t = 0.01$.[24] As a result, tightening the convergence bounds is computationally very expensive. However, this small step size also means that the probability of a sequence switching regions after not having switched in the previous 2000 iterations is likely to be minuscule.

In every case, the system converged to one of the two rational expectation equilibria. In fact, the process never switches regions after the first quarter of the total number of possible iterations. Based on this, we claim that the empirical support of the limit of the learning process is the two rational expectations equilibria. The Fortran code is available from the authors upon request. ■

Proposition 3: If the risk factor of action $j \in \{g, b\}$ given belief p decreases for some p and does not increase for any p, then the probability of equilibrium j strictly increases.

The proof of this proposition is fairly long so we first offer a brief discussion. It is fairly straightforward to see that the probability of equilibrium j cannot decrease. Suppose the curve in Figure 4 shifts in such a way that the new region G strictly contains the old one. Pick any ω in Ω and let (p_t, q_t) be the sequence generated by ω before the shift, and $(\widetilde{p}_t, \widetilde{q}_t)$ the sequence afterwards. Then we have

$$\widetilde{p}_t = p_t \quad \text{and} \quad \widetilde{q}_t \geq q_t \quad \text{for all } t.$$

This implies that if (p_t, q_t) converged to $(\overline{p}, 1)$ before the change, it will still do so after the change and therefore π cannot decrease.

Showing that the probability actually increases is much more difficult because it requires establishing specific properties of the (probabilistic) behavior of trajectories in the box. We break this task

[24] Using the equation $q_{t+1} = \eta_t q_t$, it can be shown that $q_{t+n} = [(t+2)/(t+2+n)]q_t$ and hence that $n = (t+2)[(q_t/q_{t+n}) - 1]$ holds. For the numbers above, this gives us $n = 2,002(10-1) = 18,018$.

into parts. We present and prove three lemmas and then use these results to prove the proposition. Our first lemma applies for a fixed curve f and shows that any open set near enough to the center of the box is visited with positive probability. Let p^1 and p^2 denote the values of p at which f intersects the top and the bottom of the box, respectively (see Figure 5). We then have the following.

Lemma 1: Fix any $t_0 \geq 0$, any starting point (p_{t_0}, q_{t_0}), any target point $(\widehat{p}, \widehat{q})$ with $p_1 < \widehat{p} < p_2$, and any $\varepsilon > 0$. Then there exists a finite number $T \geq t_0$ and a sequence $\{c_t\}_{t=t_0}^{T}$ such that the trajectory from (p_{t_0}, q_{t_0}) is within ε of $(\widehat{p}, \widehat{q})$ at time T.

Proof of lemma 1: Suppose $(\widehat{p}, \widehat{q})$ is below the curve f, as depicted in Figure 5. (The reverse case is completely symmetric.) Draw the line segment starting at the origin, running through $(\widehat{p}, \widehat{q})$ and ending on f. Let $(\widetilde{p}, \widetilde{q})$ denote the endpoint of this segment on f, and let x denote the entire segment. Consider a band around this segment with width $\delta = \sqrt{2\varepsilon}$ (so that a δ-square around $(\widehat{p}, \widehat{q})$ falls both inside this band and inside the ε-ball). Notice that if a trajectory enters this band between $(\widehat{p}, \widehat{q})$ and $(\widetilde{p}, \widetilde{q})$ when the maximum step size is less than δ, a long enough sequence of consecutive $c_t = H$ realizations will lead the trajectory to land in the ε-ball, as desired.

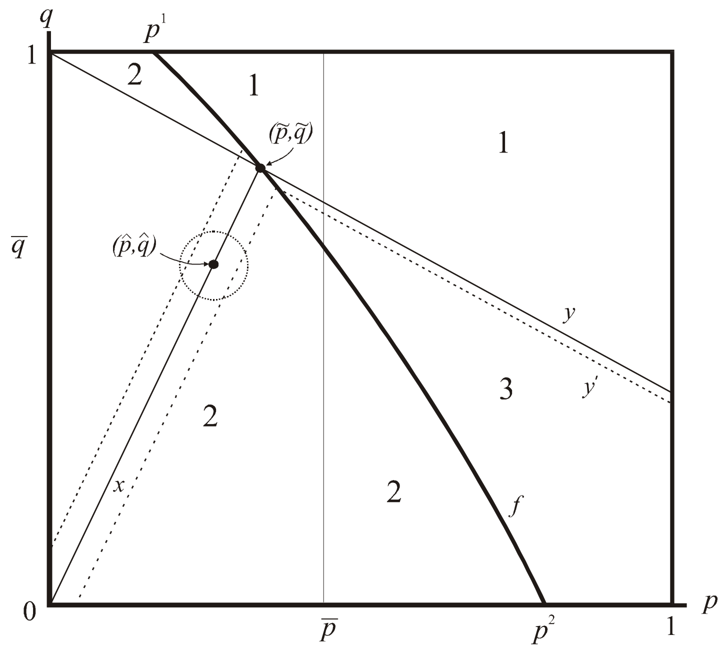


Figure 5: Visiting a neighborhood of $(\widehat{p}, \widehat{q})$

Next, draw the line from $(0, 1)$ passing through $(\widetilde{p}, \widetilde{q})$; denote this line y. Also draw the parallel line segment that intersects f at the same point as the lower bound of the δ-band and continues to the right. Denote this segment by y'. Suppose that a trajectory lands in the strip between y and y' when the maximum step size is less than δ. Then, a sufficiently long sequence of $c_t = H$ realizations will lead the trajectory to first cross f into the δ-band around x, and to then land in the ε-ball around $(\widehat{p}, \widehat{q})$, as desired. All that remains, therefore, is to show that for an arbitrary starting point and time, there exists a finite sequence of realizations that will lead the trajectory to land in this strip at a time when the maximum step size is less than δ.

To do this, we use the line y and the curve f to divide the box into three regions as labelled in Figure 5. First, suppose (p_{t_0}, q_{t_0}) is in region 1. Then a long enough sequence of consecutive $c_t = H$ realizations will bring the trajectory into region 2. From any point in region 2, a long enough sequence of $c_t = H$ realizations will make $q_t < \widetilde{q}$, and then a long enough sequence of $c_t = L$ realizations will take the trajectory into region 3.[25]

From any point in region 3, a long enough sequence of $c_t = H$ realizations will lead the trajectory to either (i) land in the strip between y and y' or (ii) step across this strip and land in region 1. Notice that if t is large enough (so that the step size is small enough), the former will necessarily occur. If (ii) occurs, the above process can be repeated to construct a (long but finite) sequence of realizations that leads the trajectory to cycle until (i) occurs. Because the maximum step size is converging to zero, (i) must occur with a maximum step size of less than δ in finite time. ■

Note that all the arguments in the proof of Lemma 1 involve finite sequences of specific realizations of c_t and hence involve events that would occur with (perhaps very low but) positive probability. In other words, starting from any point, the probability of entering any open set containing values of p between p^1 and p^2 is positive. The next lemma shows that once a target neighborhood is reached, p_t can stay in that neighborhood for arbitrarily long periods of time.

Lemma 2: Pick any $p_T \in (0, 1)$, any $\varepsilon > 0$, and any $N \geq 1$. Let T be large enough that the maximum step size of p_t is less than ε. Then there exists a sequence of realizations $\{c_t\}_{t=T}^{T+N}$ such that p_t remains in the interval $(p_T - \varepsilon, p_T + \varepsilon)$ for all t satisfying $T \leq t \leq T + N$.

Proof of Lemma 2: If $p_t \leq p_T$, then a realization of $c_t = L$ will ensure that p_{t+1} is in the desired interval. If $p_t \geq p_T$, then a realization of $c_t = H$ will do the same. This allows one to construct a sequence of realizations of arbitrary length that keeps p_t within ε of p_T. ■

Lemma 2 shows that the behavior of p_t can be "controlled" for finite periods of time using events of positive probability. The next lemma provides an infinite-period counterpart, showing that the probability of staying in any neighborhood of $\overline{p}$ is positive in the long run.

Lemma 3: Fix any $\varepsilon > 0$ and any T large enough that the maximum step size of p at T is less than ε.

[25] It is also possible that the trajectory will enter the δ-band between $(\widehat{p}, \widehat{q})$ and $(\widetilde{p}, \widetilde{q})$ before getting to region 3, at which point switching to $c_t = H$ will lead to the desired result if the maximum step size is less than δ.

Suppose we have a partial history of realizations $\omega^T = \{c_t\}_{t=1}^T$ such that $p_T \in (\overline{p} - \varepsilon, \overline{p} + \varepsilon)$. Then we have

$$\Pr\left[p_t \in (\overline{p} - \varepsilon, \overline{p} + \varepsilon) \text{ for all } t \geq T \mid \omega^T\right] > 0.$$

Proof of Lemma 3: Suppose this is not true. Then there exists an $\varepsilon > 0$, a $T \geq 1$ (where the maximum step size of p is less than ε), and a partial history ω^T with $p_T \in (\overline{p} - \varepsilon, \overline{p} + \varepsilon)$ such that

$$\Pr\left[p_t \notin (\overline{p} - \varepsilon, \overline{p} + \varepsilon) \text{ for some } t > T \mid \omega^T\right] = 1. \tag{A-3}$$

Pick an arbitrary partial history $\omega^{\widehat{T}}$ such that the generated partial trajectory $\{\widehat{p}_t\}_{t=1}^{\widehat{T}}$ (which does not necessarily pass through the point p_T above) has $\widehat{p}_{\widehat{T}} \in (\overline{p} - \varepsilon, \overline{p} + \varepsilon)$.

Returning to our original partial history ω^T, we can construct a finite sequence of N realizations that will lead the trajectory $\{p_t\}_{t=T}^{T+N}$ to (i) stay in the interval $(\overline{p} - \varepsilon, \overline{p} + \varepsilon)$ and (ii) land on the point $\widehat{p}_{\widehat{T}}$ at time $T + N$. To show this, we define

$$x_t = \left\{ \begin{array}{c} 1 \\ 0 \end{array} \right\} \text{ if } c_t = \left\{ \begin{array}{c} L \\ H \end{array} \right\}.$$

We can then write the period t belief as

$$p_t = \frac{1 + \sum_{i=1}^t x_i}{1 + t - \sum_{i=1}^t x_i}.$$

Notice that $\widehat{p}_{\widehat{T}}$ must be a rational number and can therefore be written as the ratio of two integers I and J. For the partial history ω^T, define

$$N_L = I - \left(1 + \sum_{i=1}^T x_i\right)$$

and

$$N_H = J - \left(1 + T - \sum_{i=1}^T x_i\right).$$

The integers I and J can be chosen large enough that N_L and N_H are both non-negative and that $T + N \geq \widehat{T}$, where $N = N_H + N_L$. Then appending a sequence of N realizations, N_L of which are $c_t = L$, to the partial history ω^T will lead the trajectory to land on $\widehat{p}_{\widehat{T}}$ at time $T + N$, satisfying (ii). These realizations can be ordered as in Lemma 2 to keep p_t in the ε-band around $\overline{p}$ so that (i) is also satisfied. (As long as positive numbers of both types of realizations remain, move in the direction of $\overline{p}$. Then the last string of (identical) realizations will lead monotonically to $\widehat{p}_{\widehat{T}}$.)

Because it is of finite length, this string of N realizations follows ω^T with positive probability. By (i), the trajectory has not exited the ε-band around $\overline{p}$ between periods T and $T+N$. Therefore, by (A-3) it must do so after time $T+N$ with probability one. In other words, we have

$$\Pr\left[p_t \notin (\overline{p}-\varepsilon, \overline{p}+\varepsilon) \text{ for some } t > T+N \mid \omega^{T+N}\right] = 1.$$

However, any continuation history that, when appended to ω^{T+N}, causes p_t to exit the ε-band for some t will also cause $\widehat{p}_t$ to exit the ε-band when appended to $\omega^{\widehat{T}}$. This is because the step size at $T+N$ is smaller than at $\widehat{T}$, so that p_t will be closer to their common starting point $\left(\widehat{p}_{\widehat{T}} = p_{T+N}\right)$ than is $\widehat{p}_t$ for every t. Recall that (by independence) the set of continuation histories and their probabilities is the same after every partial history. Therefore the probability of exiting the ε-band following $\omega^{\widehat{T}}$ is at least as great as that following ω^{T+N}, which is unity. Therefore we have

$$\Pr\left[p_t \notin (\overline{p}-\varepsilon, \overline{p}+\varepsilon) \text{ for some } t > \widehat{T} \mid \omega^{\widehat{T}}\right] = 1.$$

This is true for any partial history $\omega^{\widehat{T}}$ at any time that it enters the ε-band – it must exit the band with probability one. This implies that if we look at complete histories ω, we must have

$$\Pr\left[p_t \notin (\overline{p}-\varepsilon, \overline{p}+\varepsilon) \text{ infinitely often}\right] = 1,$$

which contradicts the strong law of large numbers. ■

In other words, this lemma shows that if the trajectories following some partial history were driven away from $\overline{p}$ with probability one, the same would be true for the trajectories following every partial history. This would imply that convergence to $\overline{p}$ is a zero-probability event, which we know is false.

We now use these lemmas to prove Proposition 3.

Proof of Proposition 3: We will focus on the case where the curve shifts in such a way that the risk factor of action g decreases for some $p < \overline{p}$, as depicted in Figure 6. The other cases are completely symmetric and therefore omitted. Let f_1 denote the curve before the change and f_2 the curve afterwards.[26] Let Δ denote the set of points lying between the two curves and between the vertical lines at p^1 and p^2 (where the original curve intersects the top and the bottom of the box). Note that the continuity of the curves implies that Δ has a non-empty interior.

Pick an arbitrary open ball in the set Δ, and let 2ε be the radius of this ball. Consider the ε-ball centered at the same point. Pick t_0 large enough so that the maximum step size for p_t is less than

[26] The continuity and decreasing nature of the curves are the only properties that we use here. Equation (8) actually puts much more structure on the nature of the possible changes. However, this structure does not seem to simplify the proofs in any way, and we therefore do not use it.

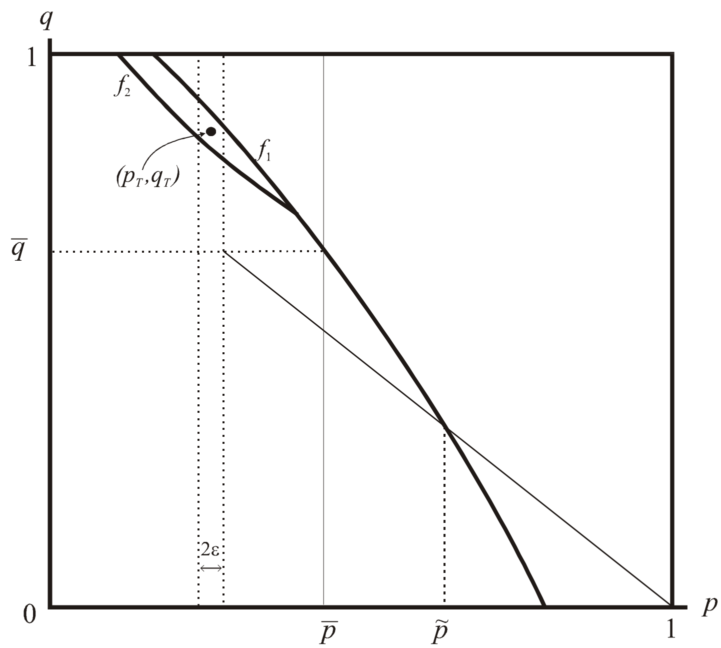


Figure 6: Asymptotic behavior depends on f

ε for all $t \geq t_0$. Then Lemma 1 tells us that there is a finite number $T \geq t_0$ and a sequence of realizations $\{c_t\}_{t=0}^{T}$ such that, under the dynamics generated by curve f_1, (p_T, q_T) falls in the ε-ball. Let ω^T denote this partial history. The dynamics generated by curve f_2 under ω^T will lead to the same p_T (since the behavior of p_t is independent of the curve) and to some $\widetilde{q}_T \geq q_T$.

Notice that the ε-ball around (p_T, q_T) is also contained in the set Δ (this was the reason for the 2ε radius of the original ball). Compute the number of steps N that would be required to move q from q_T to below $\overline{q}$. Then from Lemma 2 we know that there is a sequence of N realizations such that p_t stays in the interval $(p_T - \varepsilon, p_T + \varepsilon)$, and therefore the trajectory does not switch regions between periods T and $T + N$ (using either curve). Let ω^{T+N} denote the partial history in which these N realizations are appended to ω^T. This partial history leads the trajectory to a point like a in Figure 6 at time $T + N$ when the curve is f_1 and to a point like b when the curve is f_2.

Next, draw a line from $(p_T + \varepsilon, \overline{q})$ to $(0, 1)$. Let $\widetilde{p}$ be the value of p where this line crosses the curve f_1. Note that $\widetilde{p} > \overline{p}$ must hold. Also note that if p_t stays in the interval $(p_T - \varepsilon, \widetilde{p})$ for all future t, the trajectory from point a will never change regions and will therefore converge to $q = 0$. Likewise, under the same restriction, the trajectory from point b will converge to $q = 1$. Lemma 3 tells us that we have[27]

$$\Pr\left[p_t \in (p_T - \varepsilon, \widetilde{p}) \text{ for all } t > T + N \mid \omega^{T+N}\right] > 0.$$

[27] The asymmetry of the bounds around $\overline{p}$ is not important here. We could impose symmetric bounds as in Lemma 3 and then use a finite sequence of realizations to lead the trajectory into these bounds. We skip this simply to avoid introducing further notation.

Let D be the set of $\omega \in \Omega$ that begin with the partial history ω^{T+N} and then remain in the interval $(p_T - \varepsilon, \widetilde{p})$ for all $t \geq T + N$. By construction, the trajectory generated by any $\omega \in D$ would have converged to $(\overline{p}, 0)$ with the curve f_1, but converges to $(\overline{p}, 1)$ with the curve f_2. The probability of the set D is given by

$$\Pr[D] = \Pr\left[\omega^{T+N}\right] \cdot \Pr\left[p_t \in (p_T - \varepsilon, \widetilde{p}) \text{ for all } t > T + N \mid \omega^{T+N}\right],$$

which is positive because both terms on the right-hand side are positive. In addition, as discussed above, any ω that led the economy to converge to $(\overline{p}, 1)$ under the curve f_1 will do the same under the curve f_2. Therefore the probability of the set of sequences that lead the economy to converge to $(\overline{p}, 1)$ has strictly increased. ■

Proposition 4: In this example, π is strictly increasing in τ for $\tau < \overline{\tau}$.

Proof: In this example, using (14) allows us to write the curve dividing the regions G and B in Figure 1 as

$$q_t = \frac{(1-\tau)\left(\frac{1}{2a} + c_H\right) - \psi}{\frac{a}{2}(1-\tau)^2\left(x_g^*\right)^2 - \psi} - \frac{(1-\tau)(c_H - c_L)}{\frac{a}{2}(1-\tau)^2\left(x_g^*\right)^2 - \psi}\, p_t,$$

where the constant

$$\psi \equiv \frac{f(0)^2}{2a}$$

has been introduced to simplify the notation. The curve is a straight line in this case because of the linearity of utility in the cost c. Solving this equation for p_t yields

$$p_t = \frac{(1-\tau)\left(\frac{1}{2a} + c_H\right) - \psi}{(1-\tau)(c_H - c_L)} - \frac{\frac{a}{2}(1-\tau)^2\left(x_g^*\right)^2 - \psi}{(1-\tau)(c_H - c_L)} q_t.$$

We first look at the value of p_t when q_t is equal to unity (that is, the value of p_t such that agents would need to be certain that $\overline{x}_g = x_g^*$ in order to be willing to choose technology g). This is given by

$$p^1 = \frac{\frac{1}{2a} + c_H - \frac{a}{2}(1-\tau)\left(x_g^*\right)^2}{c_H - c_L}.$$

From (14), we see that x_g^* is strictly increasing in τ and also that $\frac{a}{2}(1-\tau)\left(x_g^*\right)^2$ can be replaced by $\frac{1}{2}x_g^* f\left(x_g^*\right)$. This latter term is strictly increasing in x_g^* hence in τ. This demonstrates that p^1 is strictly decreasing in τ. In Figure 1, this means that as τ increases, the intersection of the curve with the top of the box moves to the *left*.

We next examine the change in the intersection of the curve with the bottom of the box. This is the value of p_t when q_t is equal to zero, which is given by

$$p^2 = \frac{\frac{1}{2a} + c_H - \frac{\psi}{(1-\tau)}}{c_H - c_L}.$$

This is clearly also decreasing in τ, and hence this intersection moves to the left as well. Since in this example the curve is a straight line for all values of τ, this implies that a small increase in τ decreases the risk factor of action g given belief p for all p in (p^1, p^2). The risk factor for all other values is unchanged. Therefore, by Proposition 3, π strictly increases. ■

References

BLUME, L. and EASLEY, D. (1998), "Rational Expectations and Rational Learning", in M. Majumdar (ed.), *Organizations with Incomplete Information* (Cambridge: Cambridge University Press).

CASS, D. and SHELL, K. (1983), "Do Sunspots Matter?", *Journal of Political Economy,* **91,** 193-227.

CHARI, V.V. and KEHOE, P.J. (2000), "Financial Crises as Herds", (Working paper no. 600, Federal Reserve Bank of Minneapolis, March).

COLE, H. and KEHOE, T. (2000), "Self-fulfilling Debt Crises", *Review of Economic Studies* **67**, 91-116.

COOPER, R. (1999) *Coordination Games: Complementarities and Macroeconomics* (Cambridge: Cambridge University Press).

COOPER, R. and CORBAE, D. (2000), "Financial Fragility and Active Monetary Policy: A Lesson from the Great Depression" (mimeo., May).

COOPER, R. and JOHN, A. (1988), "Coordinating Coordination Failures in Keynesian Models", *Quarterly Journal of Economics* **103**, 441-463.

CRAWFORD, V. (1997), "Learning Dynamics, Lock-in, and Equilibrium Selection in Experimental Coordination Games", (Discussion paper 97-19, University of California, San Diego).

ENNIS, H. and KEISTER, T. (2000), "Aggregate Demand Management and Equilibrium Selection", (mimeo., December).

FUDENBERG, D. and TIROLE, J. (1991) *Game Theory* (Cambridge, MA: MIT Press).

GHIGLINO, C. and SHELL, K. (2000) “The Economic Effects of Restrictions on Government Budget Deficits”, *Journal of Economic Theory* **94**, 106-137.

GOENKA, A. (1994), “Fiscal Rules and Extrinsic Uncertainty”, *Economic Theory* **4**, 401-16.

GOOLSBEE, A. (2000), “In a World Without Borders: The Impact of Taxes on Internet Commerce”, *Quarterly Journal of Economics* **115**, 561-576.

GRANDMONT, J.M. (1986), “Stabilizing Competitive Business Cycles”, *Journal of Economic Theory* **40**, 57-76.

GUESNERIE, R. and WOODFORD, M. (1992), “Endogenous Fluctuations”, in J.-J. Laffont (ed.), *Advances in Economic Theory, Sixth World Congress*, Vol II (Cambridge: Cambridge University Press).

HARSANYI, J. and SELTEN, R. (1988) *A General Theory of Equilibrium Selection in Games* (Cambridge, MA: MIT Press).

HOWITT, P. and MCAFEE, R. P. (1992), “Animal Spirits”, *American Economic Review* **82**, 493-507.

KANDORI, M., MAILATH, G. J. and ROB, R. (1993) “Learning, Mutation, and Long Run Equilibria in Games”, *Econometrica* **61**, 29-56.

KATZ, M. L. and SHAPIRO, C. (1986) “Technology Adoption in the Presence of Network Externalities”, *Journal of Political Economy* **94**, 822-841.

KEISTER, T. (1998) “Money Taxes and Efficiency when Sunspots Matter”, *Journal of Economic Theory* **83**, 43-68.

LUCAS, R. E. (1986), “Adaptive Behavior and Economic Theory”, *Journal of Business* **59**, S401-S426.

MANUELLI, R. and PECK, J. (1992), “Sunspot-like Effects of Random Endowments”, *Journal of Economic Dynamics and Control* 16, 193-206.

MATSUI, A. and MATSUYAMA, K. (1995) “An Approach to Equilibrium Selection”, *Journal of Economic Theory* **65**, 415-434.

MORRIS, S. and SHIN, H.S. (1998) “Unique Equilibrium in a Model of Self-fulfilling Currency Attacks”, *American Economic Review* **88**, 587-597.

NYARKO, Y. (1991) “Learning in Mis-specified Models and the Possibility of Cycles”, *Journal of Economic Theory* **55**, 416-427.

PECK, J. and SHELL, K., “Market Uncertainty: Correlated and Sunspot Equilibria in Imperfectly Competitive Economies”, *Review of Economic Studies* **58**, 1011-1029.

PECK, J., SHELL, K. and SPEAR, S. (1992), “The Market Game: Existence and Structure of Equilibrium”, *Journal of Mathematical Economics* **21**, 271-299.

MARCET, A. and SARGENT, T. (1988), "The Fate of Systems with 'Adaptive' Expectations", *American Economic Review*, **78**, 168-172.

MARCET, A. and SARGENT, T. (1989), "Convergence of Least Squares Learning Mechanisms in Self-referential Linear Stochastic Models", *Journal of Economic Theory* **48**, 337-368.

SHELL, K. (1977), "Monnaie et Allocation Intertemporelle" (CNRS Séminaire Roy-Malinvaud, Paris, November. Title and abstract in French, text in English. Translation forthcoming in *Macroeconomic Dynamics*.)

SMITH, B. (1994), "Efficiency and Determinacy of Equilibrium under Inflation Targeting", *Economic Theory* **4**, 327-44.

VAN HUYCK, J., BATTALIO, R. and BEIL, R. (1990), "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure", *American Economic Review*, **80,** 234-248.

VAN HUYCK, J., BATTALIO, R. and BEIL, R. (1991), "Strategic Uncertainty, Equilibrium Selection Principles, and Coordination Failure in Average Opinion Games", *Quarterly Journal of Economics*, **106**, 885-910.

WOODFORD, M. (1986), "Stationary Sunspot Equilibria in a Finance Constrained Economy", *Journal of Economic Theory*, **40**, 128-137.

WOODFORD, M. (1990), "Learning to Believe in Sunspots", *Econometrica* **58**, 277-307.

YOUNG, H. P. (1998) *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions* (Princeton: Princeton University Press).

ZELLNER, A. (1971) *An Introduction to Bayesian Inference in Econometrics* (New York: John Wiley & Sons).

ZITTRAIN, J. and NESSON, R. (2000), "How to Think about the Internet Sales Tax Quandary", *The New Republic Online*, (May, available at http://thenewrepublic.com/online/jzrn050200.htmlfn.