Government policy and the probability of coordination failures

Huberto M. Ennis[a,*], Todd Keister[b]

[a] *Research Department, Federal Reserve Bank of Richmond, P.O. Box 27622, Richmond, VA 23261, USA*
[b] *Centro de Investigación Económica, Instituto Tecnológico Autónomo de México (ITAM), Av. Camino Santa Teresa 930, México, D.F. 10700, Mexico*

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. We show how such a mechanism can be derived as the natural result of an adaptive learning process. This approach leads to a well-defined optimal policy problem, and has important implications for the conduct of government policy. We illustrate these implications using a simple model of technology adoption under network externalities.
© 2003 Elsevier B.V. All rights reserved.

1. Introduction

In a wide variety of situations, a government or policymaker must decide on a policy before private agents choose their actions. For example, a government that wants to subsidize investment must set the subsidy rate before agents choose their investment plans. In making this decision, the policymaker anticipates that the final outcome will be an equilibrium allocation of the economy with the chosen policy. Together with a measure of social welfare, this leads to a well-defined optimal policy problem

---

* Corresponding author. Tel.: +1-804-697-8988; fax: +1-804-697-8255.
  E-mail addresses: huberto.ennis@rich.frb.org (H.M. Ennis), keister@itam.mx (T. Keister).

if the economy has a unique equilibrium for each possible policy choice. However, many economic models possess multiple equilibria. Formulating the optimal policy problem in this case is problematic: at which of the equilibrium allocations should welfare be evaluated? In addition, it seems possible that actions taken by the policy-maker may in some way affect which equilibrium will obtain. In such situations, therefore, formulating the optimal policy problem requires a theory of how the policy choice affects the equilibrium selection process. In this paper, we provide such a theory for coordination-problem economies with identical agents and multiple, Pareto-ranked equilibria. Our approach is to assign a *probability* to each equilibrium indicating how likely it is to obtain. We show how an adaptive learning process naturally generates such probabilities. This approach leads to a well-defined optimal policy problem that takes into account the effects of the policy on the selection process. We demonstrate through examples that these effects are intuitively appealing and can change the policy prescription of a model in interesting ways.

Our examples are related to problems of technology adoption and technology choice in the presence of network externalities. In such situations, there are often two (symmetric, pure-strategy) equilibria – one where the good technology is adopted by all agents and one where a bad technology is used. These are typically both strict equilibria, neither of which is easily refined away. Interesting policy questions arise in such environments. For example, consider the issue of the taxation of Internet access and purchases made over the Internet. The arguments both for and against special tax treatment of Internet transactions are numerous, but one of the arguments strikes us as being of particular theoretical interest. The argument is based on the idea that there are large network externalities which generate multiple equilibria with differing levels of e-commerce activity.[1] 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. E-commerce is claimed to be more efficient than traditional distribution methods, so that the latter equilibrium socially dominates the former. By giving e-commerce transactions preferential tax treatment, the argument goes, 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 type of argument raises a variety of questions: How significant is the impact of government policy on the selection process? Will a large enough subsidy ensure that the high e-commerce equilibrium obtains? If so, is such a large subsidy optimal? Our goal is to provide a formal way of addressing these types of questions.

There is ample evidence from the historical and experimental literatures that both Pareto-dominant and Pareto-dominated equilibria are observed in coordination-problem environments (see Cooper, 1999, for a review of these literatures). Therefore, using a deterministic selection criterion such as assuming that the Pareto-best equilibrium will obtain does not seem to describe these settings well. If both equilibria are possible, however, we need to be able to assign probabilities to them in order for the government

---

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

to have a well-defined optimal policy problem. The obvious question is where such probabilities would come from. In some economies, it is intuitively clear that one equilibrium should be more likely to obtain than the other. As an example, consider the problem of an individual agent deciding whether to engage in e-commerce or traditional commerce. To simplify the story, suppose she believes that the e-commerce market will settle on either a "thick" or a "thin" equilibrium. 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. Then the large potential payoff will lead her to choose e-commerce for a wide range of beliefs about what other agents will do. If all agents are in a similar situation, then it seems intuitively more likely that the economy will coordinate to the thick-market equilibrium. In other words, the thick-market equilibrium is risk dominant in the terminology of Harsanyi and Selten (1988), and it seems reasonable to expect that a risk-dominant equilibrium is more likely to obtain than a risk-dominated one. How much more likely would seem to depend on the strength of its risk dominance.

We propose using an equilibrium selection mechanism that captures this view in a systematic way. To do so, we suppose that the probability assigned to each equilibrium is negatively related to the *risk factor* of the equilibrium. The risk factor of the thick-market equilibrium is the smallest probability that an individual agent can place on the e-commerce market being thick and still be willing to choose e-commerce. Our mechanism therefore assigns a high probability to the thick-market equilibrium when agents are willing to choose e-commerce for a wide range of beliefs. Because the government's policy choice changes the payoffs associated with agents' actions, it will also affect the risk factor and hence the probability of each equilibrium. In choosing a policy, the government must take this effect into account. 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. In general, our technology-adoption examples indicate that subsidies can have strong effects on equilibrium selection in environments with network externalities.

Addressing quantitative issues such as finding the optimal subsidy level obviously requires going beyond qualitative statements about how the probabilities depend on the risk factors. We show how an adaptive learning process can be used to assign an explicit probability to each equilibrium. In particular, we build on the learning process of Howitt and McAfee (1992) and show that it converges to one of the (symmetric, pure-strategy) equilibria of our model. To which equilibrium the economy converges depends in part on the particular realizations of uncertainty along the learning path. Hence convergence is stochastic and the learning process naturally generates a probability distribution over the set of equilibria. We demonstrate that the probability assigned to each equilibrium by the learning process is in fact negatively related to the risk

factor of the equilibrium. In the context of the e-commerce example, because giving Internet transactions preferential tax treatment lowers the risk factor of the high e-commerce equilibrium, it enlarges the set of beliefs for which agents are willing to choose e-commerce. During the learning process, therefore, agents are more likely to become "optimistic enough" to begin coordinating on high levels of e-commerce, which in turn makes convergence to the high e-commerce equilibrium more likely. Hence the learning process generates an equilibrium selection mechanism with precisely the intuitively-desirable properties discussed above.

The idea of using adaptive processes to determine whether or not a particular rational expectations equilibrium is a plausible outcome is, of course, not new.[2] The approach has received increasing attention in recent years, particularly in macroeconomics (see, for example, Sargent, 1993; Evans and Honkapohja, 2001; and the many references therein). Much of this literature focuses on the stability or instability (appropriately defined) of a particular equilibrium and uses this as a selection criterion – unstable equilibria can safely be ignored. Our problem differs from this literature in that both equilibria we focus on are stable in the sense that they can be approached with positive probability.[3] Hence we must look beyond the question of stability and assign a probability to each equilibrium in order to formulate an optimal policy problem. Experimental evidence lends strong support to this probabilistic view of equilibrium selection. Van Huyck et al. (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 nondegenerate and, in addition, was affected by small changes in the structure of the game, such as changes in the payoff associated with each action.[4] This is exactly the type of behavior displayed by the learning-based selection mechanism that we study.

Our approach also stands in contrast to that of Morris and Shin (1998), which uses informational imperfections to eliminate the multiplicity of equilibria in a related coordination game.[5] 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. In their approach, economic fundamentals (and the information structure) determine how agents will coordinate; equilibrium selection is deterministic. In our probabilistic approach, chance plays an important role, with the economic fundamentals determining the probabilities. Our approach is therefore closer

---

[2] See, for example, the survey by Blume et al. (1982) and the article by Lucas (1986).

[3] This issue also arises in other types of models. For example, Evans and Honkapohja (1994) show that multiple stable equilibria can arise in linear rational expectations models.

[4] See Crawford (2001) for an interesting discussion of these results and for an estimation of a general learning model using the experimental data.

[5] See also Morris and Shin (2003), which provides a detailed discussion of this global-games-based approach, along with a survey of applications.

in spirit to the sunspots-based approach used, for example, by Cole and Kehoe (2000) and Peck and Shell (2003). In those papers, a strategic agent (a debt-issuing government in the former case and a bank in the latter) must choose a policy which determines whether or not multiple equilibria exist. Eliminating the bad equilibrium (a debt crisis or a bank run) is costly. The probability of the sunspot state(s) in which agents choose the bad action is a parameter of the economy, and hence the probability that the bad equilibrium will occur (if it exists) is independent of the action chosen by the strategic agent. Our learning-based selection mechanism, in contrast, assigns a probability to each equilibrium that depends on the fundamentals of the economy, including the chosen policy. The approach therefore allows us to investigate how government policy can influence the likelihood of coordination to the good outcome.

The remainder of the paper is organized as follows. In the next section, we lay out a basic coordination-problem model, and we discuss the problem of determining an optimal policy when there may be multiple equilibria. In Section 3, we introduce a probabilistic equilibrium selection mechanism with certain desirable properties and derive the implications of such a mechanism for the optimal policy in a simple example. In Section 4, we analyze in detail the equilibrium selection mechanism generated by the learning process of Howitt and McAfee (1992), and in Section 5 we use this learning-based mechanism to compute the optimal policy in a more general model of technology choice under network externalities. Finally, in Section 6 we provide a summary of the results and offer some concluding remarks.

2. The model

In this section, we describe an abstract coordination-problem model and the resulting equilibria. We consider economies with a $[0, 1]$ continuum of identical agents and with two symmetric, pure-strategy equilibria. Because we focus on symmetric equilibria with identical agents, social welfare is the same as individual welfare and hence is well defined.

2.1. The environment and equilibrium

Each agent must choose an action a from the set A. We assume that agents face a binary decision, in addition to other payoff-relevant choices. Therefore, we model the agents' choice set as

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

We will later interpret g as choosing the "good" technology and b as choosing the "bad" one, with $x \in X$ representing the amount of investment in the chosen technology. 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. [6]

There is a (benevolent) government that chooses a policy τ from a set $T \subset \mathbb{R}$. We assume that the set T includes zero, which has the natural interpretation of a

---

[6] See, for example, Harsanyi and Selten (1988), Kandori et al. (1993), and Matsui and Matsuyama (1995).

"passive" or *laissez–faire* policy. In addition, there is aggregate uncertainty represented by the random variable c, which has support $C \subset \mathbb{R}$ and distribution f. The utility of each agent depends on each of these variables *and* on the average action in the economy $\bar{a}$; this external effect is what will generate multiple equilibria. Because there is a continuum of agents, we have a truly competitive economy and each agent correctly perceives her own action to have no effect on $\bar{a}$. Each agent's utility is given by the function

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

where $\bar{A}$ is the convex hull of A. We assume that agents have rational expectations: each agent correctly anticipates the value of $\bar{a}$ and knows the distribution f of c. The government policy is also commonly known. Each agent therefore maximizes expected utility by solving

$$\max_{a \in A} \int V(a, \bar{a}, \tau, c) \, \mathrm{d}f(c). \tag{1}$$

We now define a rational expectations equilibrium; we focus on symmetric outcomes in pure strategies.

Definition. A *symmetric rational-expectations equilibrium* is an $a^* \in A$ such that a^* solves problem (1) when the value of $\bar{a}$ is held constant at a^*.

Our interest is in situations where there is more than one such equilibrium for at least some values of τ. 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 Sections 3 and 5, we work through explicit examples where this is the case.

It is worth pointing out here that under rational expectations, the timing of the choices of the elements of the vector a does not matter (as long as all agents make each choice simultaneously). For example, suppose that all agents first make the binary choice (g or b), and then observe the binary choices of others before making the continuous choice x. The set of rational expectations equilibria in this environment is exactly the same as in an environment where all choices are made simultaneously. We have left the timing unspecified because we view our rational-expectations model as applying to any of these environments.

Define x_g^* as the solution to

$$x_g^* = \arg\max_x \int V((g,x),(g,x_g^*),\tau,c) \, \mathrm{d}f(c).$$

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

$$V_{gg}(\tau) \equiv \int V((g,x_g^*),(g,x_g^*),\tau,c) \, \mathrm{d}f(c). \tag{2}$$

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

$$V_{bg}(\tau) \equiv \max_x \int V((b,x),(g,x_g^*),\tau,c)\,\mathrm{d}f(c).$$

We assume that $V_{gg}(\tau) > V_{bg}(\tau)$ holds for all $\tau \in T$, meaning that an equilibrium where all agents choose g always (strictly) exists. We define x_b^*, $V_{bb}(\tau)$, and $V_{gb}(\tau)$ in an analogous way. We assume that $V_{bb}(0) > V_{gb}(0)$ holds, so that an equilibrium where all agents choose b (strictly) exists when the government is passive. Finally, we assume that $V_{gg}(\tau) > V_{bb}(\tau)$ holds for all $\tau \in T$, so that the equilibria are clearly Pareto ranked: g is always the "good" outcome and b the "bad" one.

2.2. The optimal policy problem

We now turn our attention to the problem of determining the optimal policy. The traditional approach is to focus on a particular equilibrium, which might be selected by a formal rule or might be the focus of attention solely because it has some desirable properties. The government recognizes, of course, that the equilibrium action x_j^* is a function of τ. The benevolent government chooses its policy to maximize the expected utility of agents in this equilibrium. The traditional optimal policy problem is therefore

$$\max_{\tau \in T} V_{jj}(\tau), \tag{3}$$

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 subscript indicates that this is the optimal policy when the policymaker knows that equilibrium j will obtain. A problem with this approach is that the resulting optimal policy typically depends on which equilibrium was selected; that is, we typically have $\tau_g^* \neq \tau_b^*$. 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 allow for an equilibrium selection mechanism that assigns a probability to each equilibrium.

Definition. An *equilibrium selection mechanism* is a function $\pi : T \rightarrow [0,1]$ that assigns, for each policy choice τ, a probability $\pi(\tau)$ to the good equilibrium and $(1-\pi(\tau))$ to the bad equilibrium. We require that $\pi(\tau) = 1$ hold if the bad equilibrium does not exist for that policy choice.

One standard approach to equilibrium selection is to assume that agents coordinate their actions based on the realization of a sunspot variable. Suppose that after the government policy is set but before private agents make their choices, an extrinsic, publicly observed state of nature s is drawn from the uniform distribution on $[0,1]$. Suppose further that if τ is such that both the good and the bad equilibrium exist, all agents follow the strategy "choose (g,x_g^*) if $s \leqslant \pi$; otherwise choose (b,x_b^*)" for some number π. Then in the resulting sunspot equilibrium, π is the ex ante probability

that the good outcome will obtain. We work directly with the probabilities of the equilibria, but one can view our equilibrium selection mechanism as a function that gives the cutoff value π in a fully specified sunspots model for each possible policy choice. The standard sunspots approach is to assume that π is independent of τ (as long as both equilibria exist). Our definition allows for this possibility, but also allows for the cutoff value to vary with the policy choice.

Once a probability is assigned to each equilibrium, the natural goal of the government is to maximize the expected utility of agents across equilibria. The optimal policy problem is therefore

$$\max_{\tau \in T} \pi(\tau)V_{gg}(\tau) + (1 - \pi(\tau))V_{bb}(\tau). \tag{4}$$

Recalling the definitions of V_{gg} and V_{bb} above (see expression (2)), we see that τ influences the value of the objective function in three ways: (i) by directly affecting the value V, (ii) by affecting the equilibrium action x_j^*, and (iii) by affecting the equilibrium selection mechanism π. Deriving the optimal policy this way leads to interesting and sometimes surprising results, as we demonstrate through examples below. To gain some intuition here, suppose that π and V are differentiable in τ. Then the first-order condition for this problem is

$$\pi(\tau)V'_{gg}(\tau) + (1 - \pi(\tau))V'_{bb}(\tau) + \pi'(\tau)(V_{gg}(\tau) - V_{bb}(\tau)) = 0. \tag{5}$$

This is in many ways the central equation of the paper; its solution is the optimal policy τ^*. It says that in deriving the optimal policy, 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. Note that the first-order condition under the standard sunspots approach, where π is independent of τ, would consist of the first two terms on the left-hand side of Eq. (5), but not the third.[7] Including the effect of the policy on the probability of reaching each equilibrium is the distinguishing characteristic of our approach. It is also important to note that the solution to the traditional optimal policy problem (3) is misleading even in the case where $\tau_g^* = \tau_b^*$ happens to hold. That is, even when the same policy would maximize welfare in both equilibria, this will *not* be the optimal policy if a small deviation can increase the probability of reaching the good equilibrium. We now present an example to illustrate the implications of Eq. (5) for the optimal subsidy level in a simple model of network externalities.[8]

---

[7] It is worth pointing out here that even under the standard sunspots approach, the outcome of the model with probabilistic equilibrium selection is not a (mere) randomization over the two certainty equilibrium allocations. The values of x_g^* and x_b^* will be affected by the sunspot variable because the policymaker's choice of τ will depend on π. In this sense, the policymaker in our model plays a role similar to that of the unrestricted consumers in Cass and Shell (1983).

[8] While our examples below deal with network externalities in investment, we should emphasize that this methodology is applicable to a wide range of environments. For example, in Ennis and Keister (2003a) we study the implications of Eq. (5) in a model of bank runs.

3. A leading example

In this section we consider a simple model of network externalities that captures some of the crucial features of the Internet sales tax debate. In this model, each agent gains utility from "transacting" with other agents. These transactions occur through one of two technologies: g (the "good" one) represents Internet transactions and 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 transacting depends on the productivity of the transactions technology, which is characterized by a network externality. An agent's utility level is given by

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

The term $\bar{x}_g$ is the total amount of effort employed by agents using technology g, that is, $\bar{x}_g = \int_\Gamma x_i \, di$, where Γ is the set of agents using technology g. The function h is increasing and strictly concave; higher levels of total effort in technology g make each agent using the technology more productive. In other words, Internet-based transacting becomes more efficient as more people use it. We assume that $h(0) > 0$ holds, so that the good technology is productive even if no one else is using it. We also assume the Inada-type condition that $h'(x)$ goes to zero as x goes to infinity. The fixed cost c of operating technology g is stochastic. However, because the cost enters utility additively in this example, only the expected cost $\bar{c}$ will matter in a rational expectations equilibrium.

Due to the externalities, it is 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 Θ 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. The subsidy would be paid for by all agents through either higher taxes on other activities or reduced government expenditures. Since the subsidy in our model is financed by a lump-sum tax (which affects neither the efficiency nor the cost of effort of technology b), we are effectively assuming that the government does not raise the tax rate on traditional commerce to pay for the e-commerce subsidy.

3.1. Equilibrium

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

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

The solution to this problem is given by $x_i = 1/a$. If instead the agent chooses technology g, her problem is

$$\max_{x_i} x_i h(\bar{x}_g) - (1 - \tau) \frac{a}{2} x_i^2 - \bar{c} - \Theta, \tag{6}$$

which is solved by

$$x_i = \frac{h(\bar{x}_g)}{(1 - \tau)a}. \tag{7}$$

The agent then compares the expected utility given by each technology and chooses the more promising one. This choice clearly depends on the agent's beliefs about $\bar{x}_g$. We first look for an equilibrium where all agents choose technology g and $\bar{x}_g$ is positive. From Eq. (7), the equilibrium effort level would be given by the unique solution to

$$x_g^* = \frac{h(x_g^*)}{(1 - \tau)a}. \tag{8}$$

This generates utility level

$$V_{gg}(\tau) = \frac{h(x_g^*)^2}{2(1 - \tau)a} - \bar{c} - \Theta,$$

where the value of the lump-sum tax Θ is given by the government's budget constraint

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

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

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

to hold. We assume that the parameter values are such that we have $V_{gg}(0) > V_{bg}(0) = V_{bb}$, so that this good equilibrium exists and Pareto dominates the bad outcome when the government is passive. Straightforward calculations 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). In other words, we require that V_{bb} be greater than $V_{gb}(0)$, or that

$$\frac{1}{2a} > \frac{h(0)^2}{2a} - \bar{c}$$

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

$$\tau > \bar{\tau} \equiv 1 - \frac{h(0)^2}{1 + 2a\bar{c}},$$

the government can eliminate the bad equilibrium (and thus be sure that the good equilibrium will obtain).

3.2. Equilibrium selection and the optimal policy

In the introduction we argued that, intuitively, it seems that the probability assigned to each of the equilibria should be related to the notion of risk dominance. We now make this relationship precise by using the risk factors of the equilibria. Following Young (1998), we define the *risk factor of equilibrium g* to be the smallest probability ρ_g such that if an agent believes $\bar{a}$ will be (g, x_g^*) with probability strictly greater than ρ_g (and (b, x_b^*) otherwise), then she strictly prefers choosing (g, x_g^*) to choosing (b, x_b^*).[9] The risk factor of equilibrium b is then equal to $(1 - \rho_g)$. The risk factor measures how willing an agent is to choose an equilibrium action when she is unsure of what the market conditions will be. When an equilibrium has a low risk factor, agents are willing to choose that action for a wide range of beliefs, and therefore we should expect that the economy is very likely to coordinate on that outcome. In other words, the probability of each equilibrium should be a decreasing function of its risk factor. For the example in this section, the risk factor is given by

$$\rho_g = \frac{(1 - \tau)(1 + 2\bar{c}a) - h(0)^2}{h(x_g^*(\tau))^2 - h(0)^2}.$$

Notice that ρ_g is strictly decreasing in the subsidy rate τ; a higher subsidy makes an agent willing to adopt the good technology for a wider range of beliefs about market conditions. Therefore, if the probability of the good equilibrium π is decreasing in its risk factor, then this probability will be increasing in the subsidy rate. We now state this property as an assumption, and derive the resulting policy implications. We make the assumption in terms of derivatives only to keep the exposition simple. In the next section, we derive the properties of the function π generated by an adaptive learning mechanism.

Assumption. $\pi'(\tau) > 0$ for all $\tau < \bar{\tau}$.

The simplicity of the present example allows 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.) This implies that the same policy choice is (weakly) optimal for both equilibria, and therefore the traditional optimal policy problem (3) gives a clear policy prescription. In other words, one might be tempted to think that the issue of equilibrium selection can be safely ignored here, because τ_g^* is an optimal choice regardless of which equilibrium is selected. However, as we have argued above, the correct problem to solve is (4), which for this example simplifies to

$$\max_{\tau \in T} \pi(\tau)(V_{gg}(\tau) - V_{bb}). \tag{9}$$

---

[9] Young's definition applies to 2×2 games, but the extension to our setting is straightforward. Morris et al. (1995) introduce the related notion of p-dominance for two-player, multi-action games. In the binary-choice environment, the risk factor of equilibrium j is equal to the smallest number p such that the action profile where both agents choose j is p-dominant.
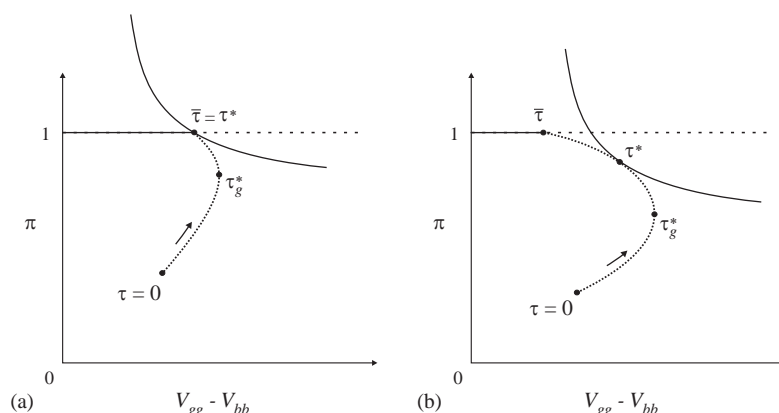
Fig. 1. Optimal policy.

There are two cases to consider. If we have

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

then the policy that maximizes the utility 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 eliminates 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

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

In this case, the government faces a tradeoff. By increasing the subsidy level above τ_g^*, it can increase the probability that the good equilibrium obtains. However, such high subsidy levels make the good equilibrium less attractive. This tradeoff is illustrated in the two panels of Fig. 1. For each value of τ, we plot the pair $(V_{gg} - V_{bb}, \pi)$ generated by the 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 $\bar{\tau}$. Increasing τ beyond $\bar{\tau}$ is clearly inefficient as π cannot increase further and V_{gg} continues to decrease.

The level curves of the objective function in (9) are hyperbolas. If eliminating the bad equilibrium is not very costly, as in part (a) of the figure, then $\bar{\tau}$ is the optimal policy. If the situation is as in part (b) of the figure, however, the optimal policy will fall somewhere strictly between τ_g^* and $\bar{\tau}$. 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. We summarize these results in the following proposition.

Proposition 1. *If $\tau_g^* \geqslant \bar{\tau}$ holds, the optimal subsidy rate is τ_g^*. If $\tau_g^* < \bar{\tau}$ holds, the optimal subsidy rate is greater than τ_g^*. In the latter case, we can have either* (i) $\tau^* = \bar{\tau}$, *in which case the optimal policy eliminates the bad equilibrium, or* (ii) $\tau^* < \bar{\tau}$, *in which case the optimal policy allows the bad equilibrium to occur with positive probability.*

For the case where the solution is interior (as in part (b) of the figure), the result can also be seen from Eq. (5). For this example, $V_{bb}'(\tau)$ is always zero. At τ_g^*, the derivative of V_{gg} is zero, but because $\pi'(\tau)$ is positive the left-hand side of the equation is positive; expected utility is still increasing in τ. Only by increasing τ further and making V_{gg}' negative can the first-order condition be made to hold.

What does this analysis 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, unless correcting the externality also eliminates the coordination problem, the optimal subsidy level is necessarily *higher* than the externality-correcting level, 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 Fig. 1 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 for the economy to face some risk over which equilibrium is selected because eliminating the bad equilibrium is too costly.

So far we have taken the equilibrium selection mechanism π as given and assumed that it possesses some desirable properties (such as being increasing in the subsidy rate). In the next section, we turn our attention to the important question of where this function might come from.

4. Deriving a probabilistic equilibrium selection mechanism

In this section, we describe how the learning process in Howitt and McAfee (1992) can be used to generate an equilibrium selection mechanism that has the intuitively appealing properties discussed above. Studying learning in our 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 of the exogenous random variable c. 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 will determine where the process 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. We do not think of the learning process as an accurate description of the short-run behavior of the economy. Rather, we view it as a mechanism that provides a good approximation of the likelihood with which the agents in the economy will coordinate on each of the equilibria. We therefore think of learning as occurring in "fast time" or artificial time, and we evaluate welfare only at the limiting points (the equilibria of the static model). [10]

The learning process takes place over an infinite sequence of discrete (artificial) time periods $t = 0, 1, 2, \dots$ Each period is like the static model of Section 2, except that agents do not necessarily have rational expectations. An agent faces two sources of uncertainty: the cost variable c and the "market conditions" variable $\bar{a}$. We assume that the cost variable can only take on two values, c_H and c_L, with $c_H > c_L$. Letting $\bar{p}$ denote the probability of c_L, f is then a Bernoulli distribution with parameter $\bar{p}$. During the learning process, agents do not know the value of $\bar{p}$. Instead, they begin with a common prior that is diffuse, i.e., a uniform distribution over $[0, 1]$. [11] In each period, the actual cost c_t (which is common to all agents) is an independent draw from the true distribution f, and agents use this information to update their belief about $\bar{p}$. Because all agents begin with the same prior and observe the same information, they hold identical beliefs about $\bar{p}$ at every point in time.

Agents treat the uncertainty about market conditions $\bar{a}$ in the same way. In particular, it is assumed that in each period agents know that $\bar{a}_t$ will be one of the equilibrium values from the static model, either (g, x_g^*) or (b, x_b^*). Agents believe that which of these two events occurs is also the result of an exogenous i.i.d. Bernoulli random variable, with the probability of (g, x_g^*) given by $\bar{q}$. In this sense, an agent in the Howitt–McAfee learning model is boundedly rational; she does not realize that the market conditions are determined by the actions of other agents who hold the same beliefs as she does. All agents begin with a diffuse prior over $[0, 1]$ for $\bar{q}$, and hence their beliefs about $\bar{p}$ and $\bar{q}$ are independent.

In addition to the above restrictions on beliefs, the Howitt–McAfee framework imposes some restrictions on the utility functions and on the timing of decisions during the learning process. First, the random variable c must be a fixed cost that enters utility in an additively separable fashion (as it does in the example in the previous section). We assume that the cost of choosing technology g is given by $\gamma_g c$ and the cost of choosing technology b by $\gamma_b c$. The examples in Sections 3 and 5 both have γ_g equal to one and γ_b equal to zero, but the only important thing is that the cost affect the two

---

[10] Abstracting from the transition also allows us to avoid the question of how agents' beliefs might be affected by changes in government policy during the learning process. A change in policy could, for example, reset beliefs to the initial prior distribution. If the model were dynamic but the government could not change policy during the transition, these issues do not arise; in this case including welfare along the transition would not have much effect on the optimal policy. This is because economies that are highly likely to converge to the good equilibrium also spend a large fraction of the transition "near" the good equilibrium.

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

technologies differently, that is, that $\gamma_g \neq \gamma_b$ hold. This separability assumption allows us to write $V_{gg}(\tau)$ as

$$V_{gg}(\tau) = \hat{V}_{gg}(\tau) - \gamma_g c,$$

where the term $\hat{V}_{gg}(\tau)$ does not depend on c. We define $\hat{V}_{gb}(\tau)$, $\hat{V}_{bg}(\tau)$, and $\hat{V}_{bb}(\tau)$ similarly. Next, the restriction on the timing of decisions is as follows. At the beginning of period t, an agent first chooses a technology (g or b). She then observes which technology has been adopted by the "market" before deciding on her level of investment. If, for example, she observes that the market has adopted g, then from the form of her beliefs about $\bar{a}_t$ described above she knows that the market level of investment will be x_g^*. If she has also chosen technology g, then because of the separability assumption she will choose $x_i = x_g^*$ regardless of her beliefs about the distribution of the cost variable. Similarly, if both she and the market have chosen technology b, she will choose the investment level $x_i = x_b^*$. Because all agents hold the same beliefs at all points in time, they will all take the same actions and therefore the value of $\bar{a}_t$ will be either (g, x_g^*) or (b, x_b^*), validating the beliefs described above. After investing, agents observe the realization of the cost variable c_t. They then use the observations of c_t and $\bar{a}_t$ to update their beliefs about $\bar{p}$ and $\bar{q}$ using Bayes' rule, and the next period begins. As discussed above, the timing of decisions does not matter in a rational expectations equilibrium, since each agent correctly anticipates market conditions. During the learning process, however, agents' beliefs are less precise and therefore this timing is important. We choose this particular structure so that we can use the Howitt–McAfee learning process with a minimum of modifications. The Howitt–McAfee process has a simple graphical representation that allows us to clearly 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. Any of a broad class of adaptive processes could be used (see Guesnerie and Woodford, 1992; Woodford, 1990). The only real requirements are that the process (i) converges to a rational expectations equilibrium with probability one and (ii) can converge to each equilibrium with positive probability.

Because the information an agent receives is independent of the actions she takes, she will always maximize current-period expected utility. Expected utility is linear in probabilities and therefore depends only on the means of the time-t beliefs about $\bar{p}$ and $\bar{q}$, which we denote (p_t, q_t). The expected utility of each choice at time t is given by

$$g : q_t \hat{V}_{gg} + (1 - q_t)\hat{V}_{gb} - \gamma_g[p_t c_L + (1 - p_t)c_H],$$

$$b : q_t \hat{V}_{bg} + (1 - q_t)\hat{V}_{bb} - \gamma_b[p_t c_L + (1 - p_t)c_H].$$

Straightforward algebra shows that an agent will choose g if we have [12]

$$q_t \geqslant \frac{\hat{V}_{bb} - \hat{V}_{gb} + (\gamma_g - \gamma_b)c^H - (\gamma_g - \gamma_b)(c^H - c^L)p_t}{\hat{V}_{gg} - \hat{V}_{gb} - \hat{V}_{bg} + \hat{V}_{bb}}. \tag{10}$$

---

[12] 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.
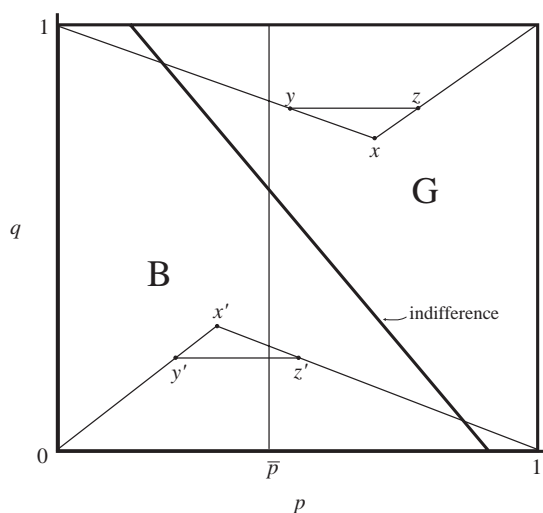
Fig. 2. The dynamics of beliefs.

For a given level of p_t, the agent will prefer technology g if q_t is high enough, i.e., if she thinks it is likely enough that other agents will be choosing technology g. Following Howitt and McAfee (1992), we can represent the optimal choice graphically as in Fig. 2, in which the box represents the set of all possible (p_t, q_t) pairs. We define G to be the set of points (p_t, q_t) such that condition (10) holds. The set B is the complement of G.

We need to place some additional restrictions on the utility values in order for this process to generate a well-behaved equilibrium selection mechanism. The statement of these restrictions depends on the relative magnitudes of γ_g and γ_b, so we assume that $\gamma_g > \gamma_b \geqslant 0$ holds, meaning that the cost variable has a larger impact on the good technology (as is the case in our examples). We then require that

$$(\gamma_g - \gamma_b)c_L < \hat{V}_{gb} - \hat{V}_{bb} < (\gamma_g - \gamma_b)\bar{c} < \hat{V}_{gg} - \hat{V}_{bg} < (\gamma_g - \gamma_b)c_H \qquad (11)$$

hold. The two inequalities in the center of this expression simply state that both the good and the bad equilibrium (strictly) exist. The leftmost inequality 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 cost variable c, she will choose g regardless of her beliefs about market conditions. The rightmost inequality is the reverse; it implies that if p were zero, then b would be the optimal choice regardless of q. Together, these last two conditions imply that during the learning process agents believe it is possible that either one of the equilibria exists or that both exist. Howitt and McAfee (1992) show 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. It is straightforward to show that the conditions given above imply that

the line separating regions B and G in the figure is downward sloping, begins to the right of $(0, 1)$, ends to the left of $(1, 0)$, and crosses $\bar{p}$ in the interior of the box.

Bayesian updating allows us to write the means of the belief distributions after t observations as functions of the previous-period means and the new observations,

$$
p_{t+1} = \left\{ \begin{array}{l} \eta_t p_t \\ \eta_t p_t + (1 - \eta_t) \end{array} \right. \text{if } c_t = \left. \begin{array}{l} c_{\mathrm{H}} \\ c_{\mathrm{L}} \end{array} \right\}
$$

and

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

where $\eta_t = (t + 2)/(t + 3)$, for $t = 0, 1, 2, \ldots$. This process has a nice representation in Fig. 2. If the time t beliefs fall in G, all agents will choose (g, x_g^*) 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 region B), $\bar{a} = (b, x_b^*)$ 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 some period t, agents' beliefs are represented by the point x in the figure. What will the long-run behavior of the learning process be? The fact that the laws of evolution of the endogenous variable $\bar{a}_t$ 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 others like themselves. In particular, agents are not learning about a *fixed* data-generating process (see Marcet and Sargent, 1989). The limiting behavior of beliefs is especially complicated in such cases. We show below that Howitt and McAfee's learning rule satisfies an important (and strong) condition: convergence of beliefs to the self-confirming equilibrium beliefs (see Sargent, 1999). In other words, the learning process will converge to a rational expectations equilibrium of the static model. To which equilibrium will the economy converge? The answer depends crucially on the sequence of realizations $\{c_t\}$. Because x is in region G, we know that agents will choose 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 $\bar{p}$; that is, we have $p_t \to \bar{p}$ almost surely as $t \to \infty$. Whether q_t converges to zero or one depends (roughly speaking) on whether beliefs are in region G or in region 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 stay in region G and therefore converge to the good equilibrium. Howitt and McAfee (1992) formalize this argument, showing that the probability of converging to each equilibrium is positive. However, they do not show that the learning process necessarily converges. We

now show that the process does indeed converge with probability one.[13] 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 in accordance with our definition in Section 2.2. The proof of the first result and the description of the computations for the second are contained in the appendix.

Proposition 2. *The learning process $\{p_t, q_t\}$ converges with probability one.*

Proposition 3. *Let π be the empirical probability of the set of sequences $\{c_t\}$ such that $\{p_t, q_t\} \rightarrow (\bar{p}, 1)$. Then the empirical probability of the set of sequences $\{c_t\}$ such that $\{p_t, q_t\} \rightarrow (\bar{p}, 0)$ is equal to $(1 - \pi)$. In other words, the simulation of the learning process always converges to a symmetric rational expectations equilibrium of the static model.*

We define our equilibrium selection mechanism to be the function $\pi(\tau)$ that: (i) takes on the value of π delivered by Proposition 3 when τ is such that both equilibria strictly exist, and (ii) is equal to unity when the bad equilibrium does not strictly exist.[14] Note that when τ is changed, the dividing line in Fig. 2 shifts and therefore the value π delivered by the learning process will typically change.

It is important to note that both the endogenous variables $\bar{a}_t$ and the exogenous random variables c_t are essential to our story. If the observation used for updating beliefs in each period contained only endogenous variables, then given the initial prior distribution, the economy would always follow the same path during the learning process. There would be no chance of convergence with positive probability to more than one rational expectations equilibrium.[15] By the same token, if the observation used for updating beliefs contained only exogenous (random) variables, then there would be only one possible convergence point for agents' beliefs (the "true" distribution) and hence there would be no equilibrium selection problem.

Going back to the formulation of the optimal policy problem (4), notice that it is implicit there that the policymaker knows f, the true distribution of the random variable c. Recall that our interest is in determining the optimal policy in the static,

---

[13] 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 proposition). See Blume and Easley (1998) and Nyarko (1991).

[14] When an equilibrium exists but is not strict, it requires agents to choose a weakly dominated strategy. Such equilibria are typically not considered to be plausible outcomes. In any event, a non-strict equilibrium would be assigned a probability of zero by the learning process if we weakened condition (11) to allow for this case.

[15] In other words, if agents did not have to learn about the exogenous random variables, we could only look at the local stability of each equilibrium with respect to agents' beliefs about the endogenous variables. This would be analogous to studying the expectational stability of the equilibria, as introduced in Evans (1989) and Evans and Honkapohja (1992) (see also Evans and Honkapohja, 2001). In our cases of interest, all equilibria will be locally stable in this sense, and thus local methods will not enable us to formulate an optimal policy problem.

rational-expectations model of Section 2. In that context, the policymaker must make a single choice τ, knowing f. This may seem somewhat at odds with our learning story – the policymaker chooses τ while knowing f, and then agents begin to learn about f. We could instead change the original economy so that there is uncertainty about f, and impose that the policymaker choose τ while having the same belief with which agents will begin the learning process. The analysis of the optimal policy would be somewhat messier, but the effects that we highlight would still be present and important. [16] The reason we do not do this is that it changes the original problem. Our goal is to find the optimal policy in the model of Section 2, not in a noisy variation of it. In an equilibrium of our model of interest, everyone (including the government) knows the true distribution f. 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. We also assume that the policymaker cannot change τ during the learning process or after the economy has converged to an equilibrium. This again follows from the nature of the problem we begin with. Our interest is in settings where the policy must be chosen before private agents choose their actions, so that the question of equilibrium selection arises in the optimal policy problem. The learning process occurs quickly, and then the equilibrium actions of private agents are observed. At this point it is too late for the policy to be changed. [17] If this were not the case, the policymaker could follow an equilibrium-dependent strategy, which is a very different environment from the one we consider.

The dividing line in Fig. 2 can be interpreted in a way that relates it to the risk factor of each equilibrium. Define $\bar{q}$ to be the level of q where the line separating the two regions crosses $\bar{p}$. Then the risk factor of equilibrium g is given by $\bar{q}$, while that of equilibrium b is given by $(1 - \bar{q})$. It is clear from the diagram that this alone is not enough to determine the value of π; the position of the entire dividing line matters. Because agents do not know $\bar{p}$ during the learning process, it matters how risky each strategy seems for each possible belief p. This leads us to provide the following definition.

Definition. The *risk factor of action g given belief p* is the smallest probability $\rho_{g,p}$ such that if an agent's beliefs about $\bar{p}$ have mean p and the agent believes $\bar{a}$ will be (g, x_g^*) with probability strictly greater than $\rho_{g,p}$, then she strictly prefers choosing (g, x_g^*) to choosing (b, x_b^*). The *risk factor of action b given belief p* is given by $\rho_{b,p} = 1 - \rho_{g,p}$.

The risk factor of action g given belief p is equal to the height of the dividing line at p in Fig. 2. (Note that this is equal to one for low enough values of p and to zero for high enough values.) In the next proposition, we show that enlarging the set G (by shifting the line down) strictly increases the probability of attaining the good

[17] Imagine, for example, trying to collect additional sales tax on past transactions. Alternatively, one could think that changing policies is costly or involves some time lags.

equilibrium. Hence, a change in the economy that uniformly lowers the risk factor of an equilibrium action will increase the probability that the economy reaches that equilibrium.

Proposition 4. *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 convergence to 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 dividing line in Fig. 2 shifts, (i) the area between the new and old lines 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 change uniformly for all beliefs p may seem strong, but changes in the values of the $\hat{V}$ variables have exactly this effect. This follows from the form of condition (10), which defines the line separating the two regions in Fig. 2. Moving any single $\hat{V}$ term causes the line to shift either uniformly up or uniformly down, which (using Proposition 4) makes the direction of the resulting change in π unambiguous. This is important because it is through these terms that the policy parameter τ affects π. We state this result as a corollary.

Corollary 1. *The value of* π *is strictly increasing in* $\hat{V}_{gg}$ *and* $\hat{V}_{gb}$. *It is strictly decreasing in* $\hat{V}_{bb}$ *and* $\hat{V}_{bg}$.

As a first step in illustrating the implications of these results, we return to the example of Section 3. In the next proposition, we show that when the Howitt–McAfee learning process is applied to the example, the resulting equilibrium selection mechanism π is strictly increasing in τ. In other words, the learning process generates the basic property of the function π that we claimed intuitively should hold (see Section 3.2).

Proposition 5. *When the Howitt–McAfee learning process is applied to the example in Section 3, the resulting* π *is strictly increasing in* τ *for* $\tau < \bar{\tau}$.

The proof of the proposition is contained in the appendix, but the intuition should be fairly clear. An increase in the subsidy will make agents willing to choose the good technology for a strictly larger set of beliefs. In other words, the set G in Fig. 2 becomes strictly larger, and hence the learning process is more likely to converge to the good outcome.

Finding the actual optimal policy under this learning process obviously requires finding the actual values of the function π, which can only be done by simulating the learning process numerically. If we are going to simulate the model, however, we can do away with many of the simplifying assumptions of the example in Section 3 (which were imposed so that the optimal policy problem could be presented graphically in Fig. 1). In the next section, we show how additional insights can be generated

by computing the equilibrium selection mechanism and the optimal policy in a more general example.

5. A general technology-choice problem

We now modify the example of Section 3 to study situations where there are two available technologies, both of which are subject to network externalities. [18] 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 each agent chooses an effort level $x_i \in [0, \infty)$. Utility is linear in output and is given by

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

The function h again represents the network externality, which now applies to both technologies. We maintain all of our assumptions about h from the previous example, including $h(0) > 0$. We assume that $g > b$ holds, so that technology g has a higher marginal product than technology b for a given amount of total effort. Also as before, the stochastic fixed cost applies only to technology g (in the notation of Section 4, γ_g is equal to unity and γ_b to zero).

We again assume that the government subsidizes effort. Let τ be the rate of *ad valorem* subsidy and let Θ 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. Likewise, the same tax is paid by all agents. In this way, we are not allowing the government to pick the "winning" technology by subsidizing one and taxing the other. Instead, the government can only encourage (or discourage) the entire industry.

5.1. Equilibrium

Regardless of the technology $j \in \{g, b\}$ chosen by an agent, her optimization problem will resemble (6). The solution is therefore of the same form as Eq. (7), or

$$
x_i = \frac{jh(\bar{x}_j)}{(1 - \tau)a}.
$$

An agent compares the expected utility generated by each technology and chooses the more promising one. We again look at symmetric equilibria, where all agents choose the same technology and therefore one of the two numbers $\bar{x}_j$ is positive and the other

---

[18] 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.

zero. We look first for an equilibrium where $\bar{x}_g$ is positive. As above, the equilibrium effort level is given by the unique solution to Eq. (8). We then have

$$V_{gg}(\tau) = \frac{(gh(x_g^*))^2}{2(1 - \tau)a} - \bar{c} - \Theta_g$$

and

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

where the value of the lump-sum tax is given by the government's budget constraint

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

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

Next, we look for an equilibrium where $\bar{x}_b$ is positive. The effort level in such an equilibrium is given by the solution to

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

This leads to

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

and

$$V_{gb}(\tau) = \frac{(gh(0))^2}{2(1 - \tau)a} - \bar{c} - \Theta_b,$$

where Θ_b is defined similarly to Θ_g above. We also assume that $V_{bb}(0) > V_{gb}(0)$ holds, 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 the conditions in (11) hold, and that we have $V_{gg}(0) > V_{bb}(0)$, so that g is indeed the good equilibrium.

5.2. Optimal policy

Appealing to Propositions 2 and 3 gives us an equilibrium selection mechanism π. Because of the additional complexity of this example (all of the V terms now depend directly on τ), drawing diagrams such as those in Fig. 1 is not possible; therefore we move directly to numerical analysis. We use the following functional form and parameter values:

$$h(x) = (x + 1)^{1/2}, \quad a = 2, \quad b = 1, \quad g = 1.4, \quad \bar{c} = 0.5.$$
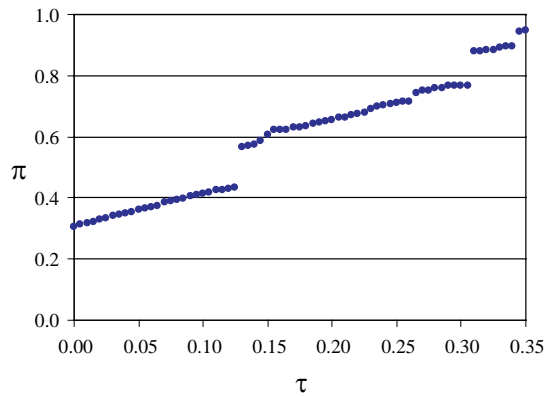
Furthermore, we take $c_{\mathrm{H}} = 2$, $c_{\mathrm{L}} = -1$, and $\bar{p} = 1/2$.

Fig. 3. The equilibrium selection mechanism $\pi(\tau)$.

Table 1

	τ	π	V_{gg}	V_{bb}	EU
$\tau = \tau_b^*$	0.185	0.642	0.5706	0.4339	0.5217
$\tau = \tau_g^*$	0.225	0.678	0.5764	0.4318	0.5298
$\tau = \tau^*$	0.240	0.703	0.5756	0.4297	0.5323

The resulting function π is presented in Fig. 3. The function is strictly increasing, as expected; higher subsidy rates benefit both technologies, but make g relatively more attractive. The discontinuities come from the discrete nature of the learning process. Some of the points in the box in Fig. 2 are visited with much higher probability than others, and the first value of τ that causes such an "important" point to switch regions will be accompanied by a discrete jump in π. For example, the point $(\frac{1}{2}, \frac{1}{2})$ represents the initial prior beliefs, and hence is the starting point for every "run" of the learning process. This point falls in region B if the government is passive. The first value of τ that enlarges region G enough to contain this point is $\tau = 0.13$. When τ is raised to this value, every possible trajectory begins with an increase in q rather than a decrease, which leads more trajectories to converge to $q = 1$ and thereby generates the first discrete jump shown in the figure. Similarly, a fraction $(1 - \bar{p})$ of the possible trajectories will begin with a high realization of the fixed cost. The first value of τ that enlarges the region G enough so that the economy stays in G after an initial high cost realization is given by $\tau = 0.31$, which corresponds to the second large jump in the figure. There are other, smaller points of discontinuity as well. As we discuss below, we have chosen the parameter values so that the discontinuities visible in the figure do *not* play a role in our results.

Table 1 presents the results of our computations for three different policy choices. The first row, $\tau = \tau_b^* \equiv 0.185$, shows the results when the subsidy level is chosen

to maximize the utility of agents in the bad equilibrium. The second row, $\tau = \tau_g^* \equiv$ 0.225, presents the same information when τ is chosen to maximize utility in the good equilibrium. The third row, $\tau = \tau^* \equiv 0.240$, corresponds to the optimal subsidy rate under probabilistic equilibrium selection. Notice that this optimal subsidy rate is *larger* than both τ_b^* and τ_g^*! This 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 government to choose. A policymaker who faces uncertainty about which equilibrium will obtain might be tempted to choose something in between these two values. Such an approach would be correct if π did not depend on τ, as in a standard sunspots-based model. Our analysis shows, however, that for the given parameter values such a choice is *not* correct when equilibrium selection effects are taken into account. As the subsidy level is increased beyond τ_g^*, the table shows that the utility value decreases for both of the equilibria. However, the higher subsidy rate still makes the good technology a relatively more attractive choice, and hence 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. We should emphasize that all three of these values of τ fall in between the first and second jumps in the π function in Fig. 3, and hence the presence of these discontinuities does not in any way affect our results.

   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 on equilibrium selection, simply because a subsidy has a larger effect on the good technology and therefore makes it a relatively more attractive choice. This example shows that policies designed to correct for network externalities can be substantially more powerful (and therefore more important) than an analysis that ignores the equilibrium selection process would indicate.

6. Concluding remarks

   The main point of this paper is that using a probabilistic equilibrium selection mechanism can bring models of coordination failure 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 some policies to be more potent than is commonly recognized.

   We use specific examples to make what we believe are more general points, applicable to a wide class of models with multiple equilibria. First, even when it is possible for the government to eliminate some bad outcome as an equilibrium, it may not be optimal to do so. It may be optimal to permit a bad outcome to occur with positive probability. Our approach takes into account both the costs and benefits of decreasing the probability of the bad equilibrium, and thereby can determine whether or not it is worthwhile to drive this probability all the way to zero.

Second, it is often not optimal to maximize the welfare of agents in any one equilibrium. Deviating slightly from such a policy will bring a small (second-order) loss in welfare if that equilibrium obtains, but can often shift probability from a bad to a good equilibrium in a way that more than compensates for this loss. This idea applies even in cases where the best policy from an equilibrium point of view does not depend on which equilibrium is selected (as in our example in Section 3). If the best policy is the same for each equilibrium, one might be tempted to think that the issue of equilibrium selection can be safely ignored and that this candidate policy should be chosen. However, we show this to be incorrect, because deviating from that policy can make a good equilibrium more likely to obtain and therefore raise *ex ante* welfare.

In situations where the best policy is different for each equilibrium, the same reasoning applies. In our example in Section 5, the government would like to give a low subsidy if agents coordinate on the bad technology, but would like to give a higher subsidy if agents coordinate on the good technology. Given that the policymaker does not know which technology will be used when the policy is set, she might be tempted to choose something in between the low and high levels. If equilibrium selection were independent of the policy (as in the standard sunspots approach), this would be the correct answer. However, taking the equilibrium-selection effect into account shows that the optimal subsidy is larger than this "averaging" approach would indicate. In fact, the example shows that the optimal subsidy can be even larger than the high level. Subsidizing the industry benefits users of the good technology more, and hence makes agents willing to choose the good technology for a wider range of beliefs. A government can therefore use policies of this sort to encourage the adoption of good technologies, even if it does not know which of the technologies is the better one.

Acknowledgements

Appendix A. Proofs

Proposition 2. *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
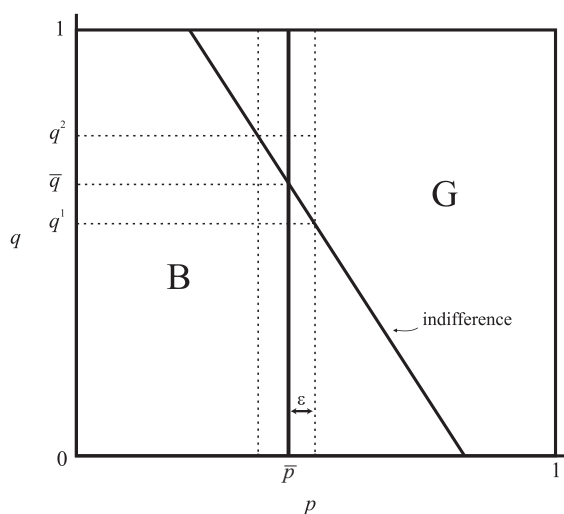
Fig. 4. Convergence of beliefs.

element of this set. Let $A \subset \Omega$ be the set of ω such that $p_t(\omega) \to \bar{p}$; 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.

Let the function $q = \phi(p)$ represent the line separating the regions G and B in Fig. 4. Consider a sequence of small numbers $\{\varepsilon_n\}$ converging to zero. For each n, define

$$q_n^1 = \phi(\bar{p} + \varepsilon_n),$$

$$q_n^2 = \phi(\bar{p} - \varepsilon_n).$$

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

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

That is, for any small band around $\bar{p}$, 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

$$q_t(\omega) < q_n^1 \quad \text{for any } t \geqslant t_n, \tag{A.1}$$

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

$$q_t(\omega) > q_n^2 \quad \text{for any } t \geqslant t_n, \tag{A.2}$$

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

$$q_t(\omega) \in [q_n^1, q_n^2] \quad \text{for all } t \geqslant t_n \text{ for all } n.$$

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

Proposition 3. *Let π be the empirical probability of the set of sequences $\{c_t\}$ such that $\{p_t, q_t\} \to (\bar{p}, 1)$. Then the empirical probability of the set of sequences $\{c_t\}$ such that $\{p_t, q_t\} \to (\bar{p}, 0)$ is equal to $(1 - \pi)$. In other words, the simulation of the learning process always converges to a symmetric rational expectations equilibrium of the static model.*

Proof. Note that, because of Proposition 2, we only need to verify that the system does not converge to the point $(\bar{p}, \bar{q})$. The law of motion for q_t has a discontinuity along the dividing line in Fig. 4, which passes through $(\bar{p}, \bar{q})$. This makes the asymptotic behavior around that point very difficult to study analytically. For this reason, we turn to a numerical methodology. From (10), we have that the curve separating regions G and B is linear, i.e., of the form

$$q = \alpha - \beta p,$$

Without loss of generality, we fix $\bar{p} = 0.5$. Then, the parameters α and β must satisfy the following two conditions:

$$1 \leqslant \alpha \leqslant 1 + \frac{\beta}{2},$$

and

$$\frac{\beta}{2} \leqslant \alpha \leqslant \beta.$$

We construct a grid of points over this two-dimensional parameter space to cover all the possible shapes of the dividing line. For this purpose we define two auxiliary variables x and y, with $x \in (0, 0.5)$ and $y \in (0.5, 1)$ and define $\beta \equiv 1/(y - x)$ and $\alpha \equiv y/(y - x)$. It is not hard to show that the values of α and β obtained in this manner satisfy the conditions above. We then construct a grid of values of x and y containing a total of 20 points in each dimension. For each such point, we compute α and β and we simulate 1,000 runs of the learning process.

We set convergence bounds for q_t in the following manner. First we compute $\bar{q}$. Then we define the variables $bound_G = \min\{(1-\bar{q})/5, 0.1\}$ and $bound_B = \min\{\bar{q}/5, 0.1\}$. 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 $\bar{q}$.)

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 $q_t = 0.01$.[19] 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 2,000 iterations is 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 these computations, 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 4. *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 convergence to 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 dividing line in Fig. 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 $(\tilde{p}_t, \tilde{q}_t)$ the sequence afterwards. Then we have

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

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

Showing that the probability actually increases requires us to show that there is a positive-probability subset $\Omega_1 \subset \Omega$ such that for all $\omega \in \Omega_1$ we have $q_t(\omega) \to 0$ but $\tilde{q}_t(\omega) \to 1$. This task is much more difficult because it requires establishing specific properties of the (probabilistic) behavior of trajectories in the box. We establish this result in two steps. First, we show that there is a positive-probability subset of Ω, call it Ω_2, such that the sequences (p_t, q_t) generated by all $\omega \in \Omega_2$ visit the area between the new and the old dividing line and hence their asymptotic behavior is *potentially* affected by the change in the line. This result is a consequence of our Lemma 1 below. Second, we show that there is a positive-probability subset $\Omega_1 \subset \Omega_2$ such that for all ω in this set, once the sequence (p_t, q_t) enters the area between the two dividing lines, it never changes regions again and hence its asymptotic behavior is *necessarily* affected by the change. This result follows from Lemmas 2 and 3 below.

We now formally state and prove the three lemmas and then use them to prove the proposition. Our first lemma applies for a fixed dividing line ϕ 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 ϕ intersects the top and the bottom of the box, respectively (see Fig. 5). Note that condition (11) implies that $p^1 < \bar{p} < p^2$ holds.

---

[19] 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 $n = 2{,}002(10 - 1) = 18{,}018$.
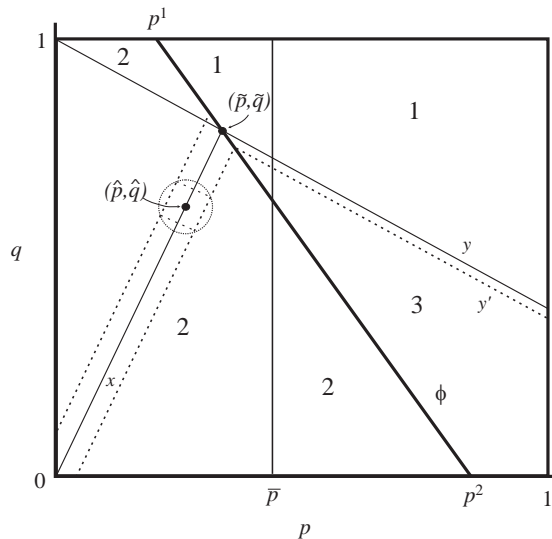
Fig. 5. Visiting a neighborhood of $(\hat{p}, \hat{q})$.

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

Proof of Lemma 1. Suppose $(\hat{p}, \hat{q})$ is below the line ϕ, as depicted in Fig. 5. (The reverse case is completely symmetric.) Draw the line segment starting at the origin, running through $(\hat{p}, \hat{q})$ and ending on ϕ. Let $(\tilde{p}, \tilde{q})$ denote the endpoint of this segment on ϕ, 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 $(\hat{p}, \hat{q})$ falls both inside this band and inside the ε-ball). Notice that if a trajectory enters this band between $(\hat{p}, \hat{q})$ and $(\tilde{p}, \tilde{q})$ when the maximum step size is less than δ, a long enough sequence of consecutive $c_t = c_H$ realizations will lead the trajectory to land in the ε-ball, as desired.

Next, draw the line from $(0, 1)$ passing through $(\tilde{p}, \tilde{q})$; denote this line y. Also draw the parallel line segment that intersects ϕ 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 $c_t = c_H$ realizations will lead the trajectory to first cross ϕ into the δ-band around x, and to then land in the ε-ball around $(\hat{p}, \hat{q})$, 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 lines y and ϕ to divide the box into three regions as labelled in Fig. 5. First, suppose (p_{t_0}, q_{t_0}) is in region 1. Then a long enough sequence of consecutive $c_t = c_H$ realizations will bring the trajectory into region 2. From any point

in region 2, a long enough sequence of $c_t = c_H$ realizations will make $q_t < \tilde{q}$, and then a long enough sequence of $c_t = c_L$ realizations will take the trajectory into region 3. [20]

From any point in region 3, a long enough sequence of $c_t = c_H$ 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 $N \geqslant 1$. 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 $T \leqslant t \leqslant T + N$.*

Proof of Lemma 2. If $p_t \leqslant p_T$, then a realization of $c_t = c_L$ will ensure that p_{t+1} is in the desired interval. If $p_t \geqslant p_T$, then a realization of $c_t = c_H$ 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 $\bar{p}$ 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. Suppose we have a partial history of realizations $\omega^T = \{c_t\}_{t=1}^T$ such that $p_T \in (\bar{p} - \varepsilon, \bar{p} + \varepsilon)$. Then we have*

$$\Pr[p_t \in (\bar{p} - \varepsilon, \bar{p} + \varepsilon) \text{ for all } t \geqslant T | \omega^T] > 0.$$

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

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

---

[20] It is also possible that the trajectory will enter the δ-band between $(\hat{p}, \hat{q})$ and $(\tilde{p}, \tilde{q})$ 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.

Pick an arbitrary partial history $\omega^{\hat{T}}$ such that the generated partial trajectory $\{\hat{p}_t\}_{t=1}^{\hat{T}}$ (which does not necessarily pass through the point p_T above) has $\hat{p}_{\hat{T}} \in (\bar{p} - \varepsilon, \bar{p} + \varepsilon)$. 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 $(\bar{p} - \varepsilon, \bar{p} + \varepsilon)$ and (ii) land on the point $\hat{p}_{\hat{T}}$ at time $T + N$. To show this, we define

$$x_t = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} \quad \text{if } c_t = \begin{Bmatrix} c_L \\ c_H \end{Bmatrix}.$$

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 $\hat{p}_{\hat{T}}$ 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 $T + N \geqslant T$ holds, where $N = N_H + N_L$. Then appending a sequence of N realizations, N_L of which are $c_t = c_L$, to the partial history ω^T will lead the trajectory to land on $\hat{p}_{\hat{T}}$ at time $T + N$, satisfying (ii) above. These realizations can be ordered as in Lemma 2 to keep p_t in the ε-band around $\bar{p}$ so that (i) is also satisfied. (As long as positive numbers of both types of realizations remain, move in the direction of $\bar{p}$. Then the last string of (identical) realizations will lead monotonically to $\hat{p}_{\hat{T}}$.)

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 $\bar{p}$ 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

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

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 $\hat{p}_t$ to exit the ε-band when appended to $\omega^{\hat{T}}$. This is because the step size at $T + N$ is smaller than at $\hat{T}$, so that p_t will be closer to their common starting point ($\hat{p}_{\hat{T}} = p_{T+N}$) than is $\hat{p}_t$ 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 $\omega^{\hat{T}}$ is at least as great as that following ω^{T+N}, which is unity. Therefore we have

$$\Pr[p_t \notin (\bar{p} - \varepsilon, \bar{p} + \varepsilon) \text{ for some } t > \hat{T} \mid \omega^{\hat{T}}] = 1.$$
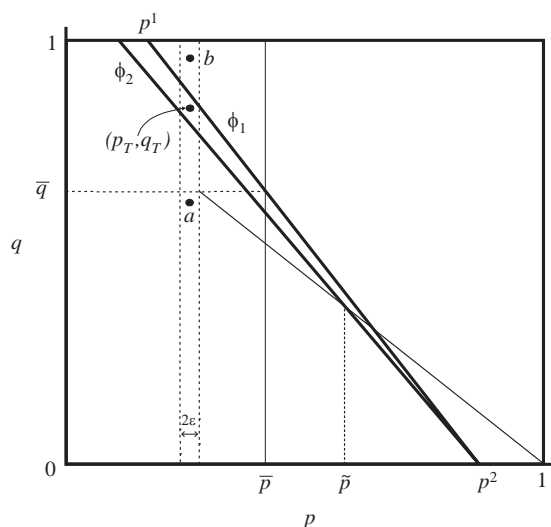
Fig. 6. Asymptotic behavior depends on ϕ.

This is true for any partial history $\omega^{\hat{T}}$ 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

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

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 $\bar{p}$ with probability one, the same would be true for the trajectories following every partial history. This would imply that convergence to $\bar{p}$ is a zero-probability event, which we know is false.

We now use these lemmas to prove Proposition 4.

Proof of Proposition 4. We will focus on the case where the dividing line shifts in such a way that the risk factor of action g decreases, as depicted in Fig. 6. The reverse case is completely symmetric. Let ϕ_1 denote the dividing line before the change and ϕ_2 the line afterwards. Let $\varDelta$ denote the set of points lying between the two dividing lines and between the vertical lines at p^1 and p^2 (where the original dividing line intersects the top and the bottom of the box). Note that $\varDelta$ has a non-empty interior and necessarily contains points with $p < \bar{p}$.

Pick an arbitrary open ball in the set $\varDelta$, 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 ε for all $t \geqslant t_0$. Then Lemma 1 tells us that there is a finite number $T \geqslant t_0$ and a sequence of realizations $\{c_t\}_{t=0}^{T}$ such that, under the dynamics generated by the dividing line ϕ_1, (p_T, q_T) falls in the ε-ball. Let ω^T denote this partial

history. The dynamics generated by the dividing line ϕ_2 under ω^T will lead to the same p_T (since the behavior of p_t is independent of the dividing line) and to some $\tilde{q}_T \geqslant q_T$.

   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 $\bar{q}$. 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 dividing line). 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 Fig. 6 at time $T + N$ when the line is ϕ_1 and to a point like b when the line is ϕ_2.

   Next, draw a line from $(p_T + \varepsilon, \bar{q})$ to $(0, 1)$. Let $\tilde{p}$ be the value of p where this line crosses ϕ_2. Note that $\tilde{p} > \bar{p}$ must hold. Also note that if p_t stays in the interval $(p_T - \varepsilon, \tilde{p})$ 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 [21]

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

Let D be the set of $\omega \in \Omega$ that begin with the partial history ω^{T+N} and then remain in the interval $(p_T - \varepsilon, \tilde{p})$ for all $t \geqslant T + N$. By construction, the trajectory generated by any $\omega \in D$ would have converged to $(\bar{p}, 0)$ with the dividing line ϕ_1, but converges to $(\bar{p}, 1)$ with the dividing line ϕ_2. The probability of the set D is given by

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

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 $(\bar{p}, 1)$ under the dividing line ϕ_1 will do the same under the dividing line ϕ_2. Therefore the probability of the set of sequences that lead the economy to converge to $(\bar{p}, 1)$ has strictly increased.   □

Proposition 5. *When the Howitt–McAfee learning process is applied to the example in Section 3, the resulting π is strictly increasing in τ for $\tau < \bar{\tau}$.*

Proof. In this example, using (8) allows us to write the line dividing the regions G and B in Fig. 2 as

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

where the constant

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

has been introduced to simplify the notation. Solving this equation for p_t yields

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

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 $\bar{x}_g = x_g^*$ in order to be willing to choose technology g). This is given by

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

From (8), we see that x_g^* is strictly increasing in τ and also that $(a/2)(1-\tau)(x_g^*)^2$ can be replaced by $(1/2)x_g^* h(x_g^*)$. This latter term is strictly increasing in x_g^* and hence in τ. This demonstrates that p^1 is strictly decreasing in τ. In Fig. 2, this means that as τ increases, the intersection of the dividing line with the top of the box moves to the *left*.

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

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

This is clearly also decreasing in τ, and hence this intersection moves to the left as well. Hence, 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 4, π strictly increases. $\square$

References

Blume, L., Easley, D., 1998. Rational expectations and rational learning. In: Majumdar, M. (Ed.), Organizations with Incomplete Information. Cambridge University Press, Cambridge, pp. 61–109.

Blume, L., Bray, M., Easley, D., 1982. Introduction to the stability of rational expectations equilibrium. Journal of Economic Theory 26, 313–317.

Cass, D., Shell, K., 1983. Do sunspots matter? Journal of Political Economy 91, 193–227.

Cole, H., Kehoe, T., 2000. Self-fulfilling debt crises. Review of Economic Studies 67, 91–116.

Cooper, R., 1999. Coordination Games: Complementarities and Macroeconomics. Cambridge University Press, Cambridge.

Cooper, R., John, A., 1988. Coordinating coordination failures in Keynesian models. Quarterly Journal of Economics 103, 441–463.

Crawford, V., 2001. Learning dynamics, lock-in, and equilibrium selection in experimental coordination games. In: Pagano, U., Nicita, A. (Eds.), The Evolution of Economic Diversity. Routledge, London and New York, pp. 133–163.

Ennis, H., Keister, T., 2003a. Economic growth, liquidity, and bank runs. Journal of Economic Theory 109, 220–245.

Ennis, H., Keister, T., 2003b. Aggregate demand management with multiple equilibria. Federal Reserve Bank of Richmond, Working Paper 03-04, June 2003.

Evans, G., 1989. The fragility of sunspots and bubbles. Journal of Monetary Economics 23, 297–317.

Evans, G., Honkapohja, S., 1992. On the robustness of bubbles in linear RE models. International Economic Review 33, 1–14.

Evans, G., Honkapohja, S., 1994. Learning, convergence, and stability with multiple rational expectations equilibria. European Economic Review 38, 1071–1098.

Evans, G., Honkapohja, S., 2001. Learning and Expectations in Macroeconomics. Princeton University Press, Princeton, NJ.

Guesnerie, R., Woodford, M., 1992. Endogenous fluctuations. In: Laffont, J.-J. (Ed.), Advances in Economic Theory, Sixth World Congress, Vol. II. Cambridge University Press, Cambridge, pp. 289–412.

Harsanyi, J., Selten, R., 1988. A General Theory of Equilibrium Selection in Games. MIT Press, Cambridge, MA.

Howitt, P., McAfee, R.P., 1992. Animal spirits. American Economic Review 82, 493–507.

Kandori, M., Mailath, G.J., Rob, R., 1993. Learning, mutation, and long run equilibria in games. Econometrica 61, 29–56.

Katz, M.L., Shapiro, C., 1986. Technology adoption in the presence of network externalities. Journal of Political Economy 94, 822–841.

Lucas, R.E., 1986. Adaptive behavior and economic theory. Journal of Business 59, S401–S426.

Marcet, A., Sargent, T., 1989. Convergence of least squares learning mechanisms in self-referential linear stochastic models. Journal of Economic Theory 48, 337–368.

Matsui, A., Matsuyama, K., 1995. An approach to equilibrium selection. Journal of Economic Theory 65, 415–434.

Morris, S., Shin, H.S., 1998. Unique equilibrium in a model of self-fulfilling currency attacks. American Economic Review 88, 587–597.

Morris, S., Shin, H.S., 2003. Global games: Theory and applications. In: Dewatripont, M., Hansen, L.P., Turnovsky, S. (Eds.), Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress, Vol. I. Cambridge University Press, Cambridge, pp. 56–114.

Morris, S., Rob, R., Shin, H.S., 1995. p-dominance and belief potential. Econometrica 63, 145–157.

Nyarko, Y., 1991. Learning in mis-specified models and the possibility of cycles. Journal of Economic Theory 55, 416–427.

Peck, J., Shell, K., 2003. Equilibrium bank runs. Journal of Political Economy 111, 103–123.

Sargent, T., 1993. Bounded Rationality in Macroeconomics. Oxford University Press, Oxford.

Sargent, T., 1999. The Conquest of American Inflation. Princeton University Press, Princeton, NJ.

Van Huyck, J., Battalio, R., Beil, R., 1990. Tacit coordination games, strategic uncertainty, and coordination failure. American Economic Review 80, 234–248.

Van Huyck, J., Battalio, R., Beil, R., 1991. Strategic uncertainty, equilibrium selection principles, and coordination failure in average opinion games. Quarterly Journal of Economics 106, 885–910.

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 University Press, Princeton, NJ.

Zellner, A., 1971. An Introduction to Bayesian Inference in Econometrics. Wiley, New York.

Zittrain, J., Nesson, R., 2000. How to think about the Internet sales tax quandary. The New Republic Online, available at http://thenewrepublic.com/online/jzrn050200.html