

Discussion Paper 100

Institute for Empirical Macroeconomics  
Federal Reserve Bank of Minneapolis  
250 Marquette Avenue  
Minneapolis, Minnesota 55480-0291

May 1995

## **On the Dynamic Selection of Mechanisms for Provision of Public Projects**

Roger Lagunoff\*

University of Pennsylvania

### **ABSTRACT**

This paper describes a dynamic model in which the provision mechanism for a public project is itself the object of locational choice of individuals. Individuals in an ongoing society must choose between a Majority Rule mechanism and a Voluntary Contribution mechanism. Each mechanism determines a funding decision for a local public project which is repeated over time. Generations of individuals asynchronously supercede their "parents," creating an entry/exit process that allows individuals with possibly different beliefs to enter society. A *self confirming equilibrium (SCE) belief process* describes an evolution of beliefs in this society consistent with a self confirming equilibrium of the repeated location/provision game. Due to Fudenberg and Levine (1993), SCE is weaker than Nash as it requires correct forecasts of an individual only along the realized path during the individual's lifetime. Since individuals' beliefs on out-of-equilibrium behavior may vary, an SCE belief process may admit random and heterogenous forecasts in the form of mutations of beliefs across generations as newborn individuals enter the system. It is shown that the process with belief mutation results in a globally absorbing state in which the Majority Rule mechanism is the *unique survivor* of the two.

Corresponding address: Department of Economics, University of Philadelphia, Philadelphia, PA 19104, rlagunof@pennsas.upenn.edu.

\*I wish to thank Prajit Dutta, Ed Green, Jim Jordan, Peter Linhart, George Mailath, Stephen Morris, Patrick Kehoe, Akihiko Matsui, Roy Radner, and Alberto Trejos for very helpful comments and discussions. The suggestions of an associate editor are gratefully acknowledged. Any errors, of course, remain my own. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

# 1 Introduction

There is a variety of collective choice mechanisms that are used to allocate the local public goods of an organization or a community. How do these mechanisms arise? This paper examines the possibility that certain evolutionary forces could be responsible for selecting one mechanism over another.

Most models of mechanism selection in collective choice problems characterize the optimal arrangement relative to information held by the participants. However, the selection of the mechanism itself is left in the hands of an exogenous “social planner.”<sup>1</sup> On the other hand, the scant literature in which the mechanism is chosen by the participants themselves considers only a single one-time choice problem among a fixed group of individuals.<sup>2</sup>

This paper offers an alternative notion of mechanism selection based on the operation of a dynamic on decentralized locational decisions of individuals over alternative organizations. By associating to each location a mechanism for the provision of a public project, individuals select the mechanism by their choice of location. The aggregation of these choices determine which mechanisms are used. We examine the nature of the “selection pressures” that are created by individuals’ locational decisions over time and ask whether these selection pressures systematically favor certain mechanisms over others.

For concreteness, we study the special case in which there are only two locations. Taking the mechanism to be the sole basis for selection, it is assumed that each location has a distinct internal provision mechanism which determines both the level and distribution of funding for a local public project. The project is identical in each location and is produced if there is a sufficient number of discrete contributions collected from the membership. To simplify further we focus on two extremes: a voluntary contribution mechanism is used in one location, while a majority voting mechanism is used in the other. Individuals can choose both which location to join, and, given the provision mechanism of their chosen location, whether or not to support the project in that location. Three types of Nash equilibria of this location/provision game emerge: (1) the project is funded only in the location with the Voluntary Contribution mechanism; (2) it is only funded in the location with Majority Rule; (3) it is funded by neither. Generally, all three types of equilibria coexist.

As a result of this multiplicity, the static model has very little to say about mechanism selection. As well as modeling the mechanics of intertemporal movement between mechanisms, the dynamic process in this model is helpful as an equilibrium selection device. Not all static equilibria will necessarily survive with the passage of time.

---

<sup>1</sup>See Groves and Ledyard (1987) for a survey on a part of the vast literature.

<sup>2</sup>See Crawford (1985), Holmström and Myerson (1983), Lagunoff (1992), and Vassilakis (1992).

The dynamic selection process is modeled here as a repeated game with asynchronous moves in which the population is subject to periodic turnover. Individuals “die” at stochastic points in time and are replaced by “newborn” individuals who inherit the same physical characteristics as their “parents.” Because the environment is stationary, each of the static equilibrium outcomes is also an equilibrium outcome of the repeated game.

Here, however, the possibility exists that over long periods of time, behavior evolves due to accumulation of heterogeneous beliefs across generations. The possibility of heterogeneous forecasts of incoming individuals creates a natural tension with standard Nash equilibrium analysis of the repeated game which admits no such heterogeneity since all individuals’ forecasts in a Nash equilibrium are assumed to be everywhere correct. Heterogeneity in this model is possible by considering a weaker behavioral concept than Nash equilibrium. Instead, it is assumed that the strategies chosen by all individuals constitute a *self confirming equilibrium*, a concept introduced by Fudenberg and Levine (FL) (1993a,b). The self confirming concept weakens the Nash concept by requiring only that all individuals make correct predictions about the realized path of play while they remain in society. Unlike in a Nash equilibrium, beliefs about *out-of-equilibrium* behavior need not be correct.<sup>3</sup> Unlike in a Nash equilibrium, an individual who may outlive his contemporaries may forecast long run changes unforeseen by his contemporaries while they were alive. This possibility allows a newborn to enter society with possibly different beliefs than his parents. The slow but steady accumulation of heterogeneous beliefs of this type may induce a drift away from certain static equilibria and toward others.

Under this kind of dynamic, the specifics upon which we elaborate below, the main result is that the Majority Rule location is the unique survivor in the long-run. Specifically, it is shown that there is a random stopping time at which societal behavior settles into the set of Majority Rule funded profiles and departs thereafter with probability zero. That is, the set of behavior profiles in which only the Majority Rule mechanism funds the project constitutes a globally absorbent set.<sup>4</sup>

The process governing the belief arrivals as new individuals enter society is referred to as a *self confirming equilibrium (SCE) belief arrival process*. Three conditions on this process are sufficient to guarantee that funding of the public good exclusively by Majority Rule is

---

<sup>3</sup>The particular equilibrium concept used here is due to Fudenberg and Levine. However, equilibrium concepts with forecasts unfalsified by observation are not new. Hahn’s (1973) *conjectural equilibrium* has this feature. Applications of conjectural equilibrium to game theory include Battigalli and Guatoli (1988), and Rubinstein and Wolinsky (1990). Another cousin of SCE is Kalai and Lehrer’s (1993) *subjective equilibrium*, the main difference being that, unlike SCE, subjective equilibrium requires that beliefs admit no strategic correlation.

<sup>4</sup>Matsui (1992, 1993), and Gilboa and Matsui (1991) refer to this property in related contexts as “social stability.”

globally absorbing in the sense described above. First, we assume that the belief arrival process is stochastic and, in particular, is a stationary Markov process. Second, we assume that beliefs of incoming individuals mutate across generations, and the distribution on these mutations are independent across individuals and across time. Third, the stochastic process is assumed to have maximal support with respect to the first two properties. This third requirement allows the maximal variation in behavior to occur through time.

Since the dynamics of this model “select” some types of equilibria over others, this work is similar in spirit to the literature in evolutionary games with perpetual randomness (“mutation”) and adjustment dynamics.<sup>5</sup> Specifically, the dynamics of this paper are closest to Matsui and Rob (1992) and Lagunoff and Matsui (1995). These papers differ from much of the previous work by considering forward looking agents rather than myopic or short-sighted ones.

A distinctive feature of this paper is that it considers an evolution which is consistent with individuals having foresight. The forecasts of individuals are consistent with equilibrium of the fully repeated game. Yet, the system tolerates random perturbations of beliefs as new individuals enter the system. The role of the perturbations in this model is to allow an accumulation of the “right” sequence of beliefs in order to escape profiles other than those in which Majority Rule funding prevails. By contrast the restriction to equilibrium prevents transitions away from the Majority Rule funded profiles. The two forces are only compatible when there is *more latitude* in beliefs than is permitted under the Nash equilibrium assumption. The self confirming assumption permits such latitude.<sup>6</sup>

Putting the dynamics aside, a main feature of the stage game in this model is the Tiebout-like migration between locations, which has antecedents in Tiebout (1956) and Buchanan (1967). More recently, Greenberg (1983), Scotchmer (1985), Wooders (1978, 1988), Scotchmer and Wooders (1987), and others have described models in which both the composition of membership and/or the nature of the local public good technology determine the spatial distribution of voters. Unlike these models, the object of locational choice here is the mechanism *itself*. Caplin and Nalebuff (1992, 1993) describe a locational model with some similarities to our stage game. The policies of each location or jurisdiction in their model are determined interactively with location decisions of individuals. The main difference between their model and our stage game is that they allow the possibility an organizational response to migration which is not considered here, and the individuals in their model are

---

<sup>5</sup>See Mailath (1992) for a survey and references contained therein. See also, Blume (1993) and Ellison (1993).

<sup>6</sup>Fudenberg and Levine (1993a) give examples of games in which inferences “off the play path” may well generate deviations from some types of behavior and toward others. Nöldeke and Samuelson (1992) give further examples of this kind of behavior under the self confirming hypothesis in a model with random matching learning dynamics.

small, having no strategic effects, while strategic interactions between individuals here play an important role. None of the aforementioned models have considered explicit dynamics of locational movement.

Section 2 describes the static model, while in Section 3 the dynamic extension of the model is laid out. There, the assumptions on the belief process are described. An example is described to demonstrate the difference between Nash and self confirming equilibrium dynamics. Section 4 states the main result and provides some intuition for the proof of the result. Section 5 gives concluding remarks. Some limitations of the present model and potential extensions are considered. For some environments selection of this type seems to work well. For others it does not. Section 6 is an appendix with a proof of the main result.

## 2 The Static Model

A society of agents  $I = \{1, \dots, \ell\}$  divide themselves into two disjoint locations or “jurisdictions” denoted by the sets  $V$  and  $M$  where  $V \cup M = I$ , with  $|V| = n$ ,  $|M| = m$  so that  $n + m = \ell$ . There is a local public project or a fixed local activity that can be produced in each location and which yields a simultaneous benefit to all members of the location. Each location member is assumed to be able to make a fixed, unitary contribution of a single consumption good toward the production of the project. This contribution is best viewed as a physical effort such as military or civic service as it is assumed to be nonrefundable. In order to contribute one unit, agent  $i$  incurs a cost or disutility of  $c_i > 0$ . The project in each location is funded if and only if the number of contributors from the location’s membership is no less than some integer  $x^*$ . An agent  $i \in I$  residing in one of the two locations derives a benefit  $v_i$  from the project if it is produced in that location.

Each location is differentiated by the mechanism used to fund the project and determine each member’s contribution. In location  $V$  the provision mechanism is completely decentralized. Contributions toward the project are voluntary. We refer to this mechanism as the Voluntary Contribution mechanism. In the second location,  $M$ , provision is more centralized, being determined by the median voter. We refer to this as the Majority Rule mechanism. Both are defined below.

### Location $V$ : The Voluntary Contribution Mechanism

Let  $x_i \in \{0, 1\}$  denote the contribution from person  $i \in V$ , and let  $x^V = (x_i)_{i \in V}$  denote the *action profile* of members of  $V$ . Let  $\bar{n} = \sum_{i \in V} x_i$ , the number of contributors. An agent

$i \in V$  has a payoff given by

$$w_i(x_i, x^V) = \begin{cases} v_i - c_i x_i & \text{if } \bar{n} \geq x^*, \\ -c_i x_i & \text{if } \bar{n} < x^*. \end{cases} \quad (1)$$

In (1), agent  $i$  gets  $v_i - c_i x_i$  if the project is funded, while he gets  $-c_i x_i$  if the project is not funded. We assume that no side payment of private goods is possible. Therefore, agent  $i$ 's utility depends only on his own contribution and the number of contributors,  $\bar{n}$ , in the location. We assume that  $x^* > 1$  so that no one has a unilateral incentive to contribute on his own. If  $n > x^*$ , some individuals may “free ride” on the efforts of others, and so the number of potential noncontributors (*free riders*) in location  $V$ , given by  $n - x^*$ , is technologically determined.

### Location $M$ : The Majority Rule Mechanism

Let  $x_i \in \{0, 1\}$ , as before, but here it has the interpretation of being the “vote” or “opinion” of agent  $i \in M$ . The vote  $x_i = 1$  is a vote in favor of the project, while  $x_i = 0$  is a vote against it. Let  $x^M = (x_i)_{i \in M}$  denote the profile of these votes/opinions. Letting  $\bar{m} = \sum_{i \in M} x_i$ , i.e., the number of affirmative votes, we define the payoff of an agent residing in location  $M$  by

$$w_i(x_i, x^M) = \begin{cases} v_i - c_i \frac{x^*}{\bar{m}} & \text{if } \bar{m} > m/2 \text{ and } m \geq x^*, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

The Majority Rule mechanism in location  $M$  funds the public project if and only if a numerical majority support the project. The 51% threshold was chosen for plausibility but is not essential to the analysis. The Unanimity rule, for example, generates similar results. If a majority is in favor of the project then the costs are distributed evenly and anonymously through a lottery that selects  $x^*$  contributors with equal probability at each potential revision time.<sup>7</sup> Each individual is therefore chosen with probability  $x^*/m$ . Unlike the Voluntary Contribution mechanism, Majority Rule requires some degree of coercion. Individuals that oppose the project incur a positive expected contribution if sufficiently many others vote differently.

### The Location Model:

The choice of mechanism is modeled as a Tiebout-style choice of location. An agent's choice of jurisdiction/ location coincides with his choice of mechanism. As with, say, religious

---

<sup>7</sup>The reason for the lottery is simple. Without it, the Majority Rule location usually yields an inefficient outcome since more than  $x^*$  individuals must pay if the project is funded. I therefore prefer a comparison of mechanisms such that the equilibria that support cooperation are Pareto noncomparable. Ledyard and Palfrey (1994) show that the Majority Voting (“Lottery Draft”) mechanism is interim efficient in a class of environments of incomplete information with two player types.

or political affiliations, locations might be metaphors for distinct organizational characteristics, rather than spatially separated communities.

To simplify, we ignore “transportation” costs between locations. Let  $s_i \in \{V, M\}$  denote agent  $i$ ’s location decision, and  $i$ ’s payoff in the location game is given by  $w_i(x_i, x^M)$  if  $s_i = M$ , and  $w_i(x_i, x^V)$  if  $s_i = V$ .  $m$  denotes the number of agents who choose location  $M$ . Likewise  $n$  for  $V$ . The aggregate profile,  $(x_i, s_i)_{i \in I}$ , is assumed to be unobservable, however, the aggregate distribution of actions,  $z = (n, \bar{n}, m, \bar{m})$  is assumed to be publically observable. This *anonymity* assumption plays an important role in the analysis. This location/provision game has three types of (pure strategy) Nash equilibria which are characterized in the table below.

**Nash Equilibria of Location/Provision Model**

Equilibrium	location V	location M
$E_N$ : no funding	$\bar{n} = 0$	$\bar{m} < \frac{m}{2}$ or $m < x^*$
$E_M$ : $M$ -funded	$\bar{n} = 0$ with $\frac{v_i}{c_i} < \frac{x^*}{m+1}, i \in V$	$\bar{m} > \frac{m}{2}$ and $m \geq x^*$ with $\frac{v_i}{c_i} \geq \frac{x^*}{m}, i \in M$
$E_V$ : $V$ -funded	$\bar{n} = x^*$	$m = 0$

In the type of equilibrium displayed in the first row, neither jurisdiction takes up the project. We refer to this as the “no funding” equilibrium, the set of which is denoted by  $E_N$ . The second row displays an equilibrium in which location  $M$  funds the project; location  $V$  does not since it consists only of individuals who would rather not pay the tax imposed in  $M$ . We refer to the set of equilibria of this type as the Majority Rule funded equilibrium, labelled  $E_M$ . The last row displays the set of equilibria, denoted by  $E_V$ , in which  $V$  funds the project and no one locates in location  $M$ . No one person’s defection from an equilibrium,  $E_V$ , can fund the project under Majority Rule, and each noncontributing member of  $V$  receives a free-riding benefit by staying in  $V$ . Clearly, the sets  $E_V$ ,  $E_M$  and  $E_N$  are disjoint.

**Proposition** *If  $(x_i, s_i)_{i \in I}$  is a pure strategy Nash equilibrium of the static location/provision game then  $(\bar{x}_i, \bar{s}_i)_{i \in I} \in E_V \cup E_M \cup E_N$ . Furthermore,  $E_N \neq \emptyset$ , while  $E_M \neq \emptyset$  iff there is  $m \geq x^*$  with  $|\{i \in I : \frac{v_i}{c_i} \geq \frac{x^*}{m}\}| = m$ , and  $E_V \neq \emptyset$  iff  $|\{i \in I : \frac{v_i}{c_i} \geq 1\}| \geq x^*$ .*

By the Proposition there is no equilibrium in which the project is undertaken in both jurisdictions. The reason is simple. If it were otherwise, then potential free riders in  $M$  would prefer to move to  $V$ , while contributors of  $V$  would generally prefer to move to  $M$ . The remainder of the argument is straightforward and therefore omitted. The Proposition

also shows that the parameter sets that support each of the three sets of equilibria are nested. No-funding equilibria always exist.  $V$ -funded equilibria require the most stringent conditions of the three. For the three types of equilibria to coexist requires minimally that  $v_i \geq c_i$ , for at least  $x^*$  of the  $\ell$  individuals. This will be assumed in the sequel. We ignore mixed strategies as they will be unimportant in the dynamic extension of the model.

A few remarks concerning the comparison of equilibria are in order. The choice and limitation of the model to the particular two mechanisms might seem arbitrary. At first glance, it may seem that the “deck is stacked” against the Voluntary mechanism since agents cannot share the burden of the costs as they can in the Majority Rule mechanism. Despite this, we observe that any pair of equilibria,  $E_V$  and  $E_M$ , are Pareto noncomparable. Though one can verify that  $E_M$  comprises the set of Strong Nash equilibria, it is only by assuming cooperative coalition formation at the outset that profiles in  $E_V$  are ruled out. Moreover, those who expect it possible to free ride or those with  $v_i < c_i$  will always prefer the Voluntary mechanism, giving it a base of support that does not exist in location  $M$ . To conclude, despite some apparent advantages, *there is no formal argument in the noncooperative (Nash) framework that allows one to predict that the Majority Rule equilibria will prevail*. The multiplicity problem only worsens in the repeated game.

There are also many other mechanisms one could consider in a model locational choice. Nevertheless, these two represent important benchmarks. Both are commonly observed, and are differentiated by the degree of centralization or coercion required of the participants.

Another, quite reasonable, objection to this location story is that it precludes any kind of organizational response. Why can't the location that can no longer provides the public good change simply change its rules? Some types of internal responses have been studied.<sup>8</sup> However, since individuals, not locations, make decisions, it is no longer obvious that any full blown, political economic model of local decision making will result in immediate repudiation of an inefficient mechanism. Our goal here is to isolate the dynamic effects of locational decisions as a selective force on the mechanisms. The type of locational dynamics modeled here are able, as it turns out, to account for mechanism selection in the long run.

Observe, finally, that assumptions of form of contribution technology rule out the possibility of fractional contributions in some common unit of account. Military service, charity work, and community service are all examples in which the physical presence of an agent is required, the physical effort is nonrefundable, the contribution decision may be lumpy, and the resulting outcome infeasible below some threshold. If the model were generalized to allow individual contributions from a continuum, the analysis would complicate considerably since it creates an infinite state space in the dynamic model. In the summary, we discuss

---

<sup>8</sup>See, for example, Caplin and Nalebuff (1992), (1993).



the changes in the model by considering fractional contributions.

## 3 The Dynamic Model

### 3.1 Asynchronous Choice and Replacement

In this section a certain dynamic that operates on the static game is examined.<sup>9</sup> We now consider a society with an infinite horizon. Time is continuous. The provision problem is described by a continuous repetition of the static location game.

Unlike standard repeated games, however, all agents' decisions in this model are made asynchronously, occurring at stochastic points in time, and follow a Poisson process with parameter  $\epsilon > 0$ , which is constant and independent across time and across individuals. Moreover, at every revision opportunity for  $i \in I$ , with probability  $\gamma > 0$  he "dies" and is replaced with a "newborn" individual with the same index. A newborn is assumed to inherit the same physical characteristics and information about the past history from his predecessor. However, the newborn individual is distinguished by his beliefs which may be different than those of his predecessor. We refer to the index " $i$ " as the progeny rather than the individual person, and let  $ij$  index the  $j$ th newborn individual in the  $i$ th progeny. The replacement process is therefore also a Poisson process with the rate of replacement of an agent given by  $\epsilon\gamma$ . It is assumed that  $\epsilon$  and  $\gamma$  are known by each individual. As before, we assume that each individual does not directly observe other progenies' behavior, nor their revisions and replacements dates. However, he does observe the aggregate behavioral distribution at any point in time.

Denote an individual's *effective discount rate* by  $r = \tilde{r} + \epsilon\gamma$ , where  $\tilde{r}$  is the physical discount rate and  $\gamma\epsilon$  is the death rate. The effective rate  $r$  relative to the revision rate  $\epsilon$ , summarized by the parameter  $\lambda \equiv \frac{r}{\epsilon}$ , gives the discounted average delay between revision opportunities. The purpose of the revision process is merely to ensure that individual's choices remain asynchronous. The artificial friction imposed by the revision process is will be eliminated in the limit as we let adjustment delays go to zero, that is, we let  $\lambda \rightarrow 0$ .

We argue that the dynamic structure assumed here is a natural one to study the provision of a public project. The stochastic replacement process has a natural demographic interpretation — birth and death (but it could also correspond to entry and exit among firms). The

---

<sup>9</sup>The reader is forewarned that, to understand the details of the dynamic properties, the explanations may appear divorced from the applied problem. The relevance for the specific mechanism selection problem will be re-emphasized in Sections 3.4 and 4.

constant flow of individuals both in and out of locations is common in most locations where the organizational structure is designed to outlast any particular generation.

A *feasible sample path*  $p$  is a function defined on  $[0, \infty)$  that gives, at each instant  $t$ , the realization  $p(t)$  of each agent's action,  $(x_{it}, s_{it})$ ,  $i \in I$ , and all agents' revision and replacement opportunities that are consistent with the dynamic structure of the model. A *continuation path from  $t$*  is some  $t$ -length shifted path  $p_t$  defined by  $p_t(\tau) = p'(t + \tau)$  for some feasible path  $p'$ . For a path  $p$  let  $p^t = \{p(\tau)\}_{0 \leq \tau < t}$  denote the history up to  $t$ . Let  $z^t = \{(n_\tau, \bar{n}_\tau, m_\tau, \bar{m}_\tau)\}_{\tau < t}$  denote the aggregate history up to  $t$ . Finally, let  $\mathcal{A}$  denote the set of Poisson arrival realizations with generic element  $\alpha = (\alpha_1, \dots, \alpha_\ell) \in \mathcal{A}$  where  $\alpha_i(t)$  is the number of births/deaths in progeny  $i$  by time  $t$ .

An individual who has a revision opportunity at time  $t$  must make a location and contribution/voting choice. To do this he chooses a strategy to maximize his expected discounted utility stream given his beliefs about continuation paths from  $t$ . The following description of this utility stream will suffice. Given a path  $p$ , define each  $i$ 's flow utility by  $u_i(p(t)) \equiv w_i(x_{it}, x^{st})$  for each  $t$ . A *strategy*,  $f_{ij}$ , for individual  $ij$  is a mapping that gives for each  $t$  and each possible history  $p^t$  a location/opinion-contribution in  $\{M, V\} \times \{0, 1\}$ . Strategy  $f_{ij}$  can only vary with aggregate histories  $z^t$  and his own behavioral history. Given a strategy  $f_{ij}$  for  $ij$ , a *belief* for individual  $ij$  at time  $t$  is given by  $\phi_{ij}(f_{ij}, p^t)$ , which is a probability distribution over all paths given history  $p^t$  and his own strategy  $f_{ij}$ .

Since individuals within a progeny are identical in every respect other than their beliefs, the newborn  $ij$  can be identified by his *belief*  $\phi_{ij}$  which is defined technically at "date 0" for the entire path. Formally,  $\phi_{ij}(f_{ij}, p^t)$  will denote the time  $t$  Bayesian update of  $\phi_{ij}(f_{ij}, p^0)$  which is a measure that assigns positive probability only to those sets of paths consistent with initial state  $p^0$  and with  $ij$ 's strategy  $f_{ij}$ . (Since beliefs are defined directly on paths rather than on others' behavior strategies, an individual must place full support only on those paths that are consistent with his own actions, thus the formal dependence of  $ij$ 's belief on his own strategy.) Individual  $ij$ 's belief must, of course, also be consistent with his information.

A complete description of the model requires a specification of consistency properties of beliefs and strategies. We defer the formal statement of these properties to Section 6 where they will become important for the proof of the main result. It suffices for now to observe that the physical restrictions of the model imply that a strategy  $f_{ij}$  prescribes a behavior that is constant everywhere except possibly at Poisson revision points. Also, an individual's belief over sample paths must be consistent with some belief over others' behavior strategies. This ensures the standard restriction that, given  $ij$ 's conditional belief  $\phi_{ij}(f_{ij}, p^t)$ , by varying his own strategy  $ij$  does not affect his own beliefs about others' choice of behavior strategy

(but may effect others' behavior!). Denote by  $\Phi_i$  the set of possible beliefs for an individual of progeny  $i$ . The formal construction of  $\Phi_i$  is contained in Section 6.

Given history  $p^t$ , a newborn individual  $ij$  with beliefs  $\phi_{ij}(\cdot, p^t)$  has expected utility given by

$$E_{\phi_{ij}(f_{ij}, p^t)} \left[ r \int_t^\infty e^{-r(\tau-t)} u_i(p_t(\tau)) d\tau \right] \quad (3)$$

where  $E_{\phi_{ij}(f_{ij}, p^t)}$  is the expectation given probability  $\phi_{ij}(f_{ij}, p^t)$  over sample paths.

### 3.2 Belief Arrival Process

The demographic structure of the replacement process is meant to capture that of an ongoing society in which new individuals enter with possibly different beliefs about future locational and and provisional decisions than their predecessors. These new beliefs accumulate in society as old individuals “die off” sequentially and are replaced. The evolution of behavior is determined by the evolution of beliefs.

The effect that beliefs have on behavior is summarized by a probability distribution on the sample paths which is induced by a realization of the belief of every individual who will, at some point, take part in the game. Let  $\Phi = \times_{i=1}^n \times_{j=1}^\infty \Phi_i$ , denote the product space of all individuals' beliefs. Let  $\phi$  be element of  $\Phi$ . We refer to  $\phi$  as a *belief realization*. A *belief arrival process* (or simply *belief process*), denoted by  $\pi$ , is defined as a probability distribution on  $\Phi \times \mathcal{A}$  that jointly determines the beliefs and arrival times of each newborn.

One useful interpretation of the belief arrival process is that “nature draws” a belief realization  $\phi$  and an arrival realization from the “urn”  $\Phi \times \mathcal{A}$  at time  $t = 0$ . This realization simultaneously determines the beliefs and arrival times of all individuals, including those not yet “alive.” An individual arrives at date  $t$  with a belief that is the Bayesian posterior of the belief drawn at time  $t = 0$ . Of course, this interpretation is a fiction since prior beliefs “arrive” sequentially as newborns enter the game.

If an individual  $ij$  with belief  $\phi_{ij}$  chooses a strategy  $f_{ij}^*$  that maximizes expected utility in (3), then a belief realization  $\phi \in \Phi$ , and arrival realization  $\alpha$  therefore determine a collection of strategies  $f = (f_{ij})$ ,  $i \in I$ ,  $j = 1, \dots, \infty$  which maximize (3). For  $\pi$ -almost every  $(\phi, \alpha)$  pair, the asynchronous choice aspect of the model uniquely determines the collection  $f$  of best responses. In turn,  $f$  determines a probability measure,  $\mu_f$ , on path space. This measure satisfies certain physical restrictions implied by the Poisson revision processes. Integrating over the  $(\phi, \alpha)$  realizations determines a measure  $\mu_\pi$  on the location/provision paths induced by process  $\pi$ . We refer to the measure  $\mu_\pi$  as the *path process*. Our aim will be to characterize

the asymptotic properties of the induced  $\mu_\pi$  as a function of the parameters  $x^*$ ,  $\lambda$ ,  $(v_i, c_i)_{i \in I}$ , and  $\pi$ .

### 3.3 Self Confirming Belief Arrival Processes

The standard requirements of repeated game Nash equilibria dictate that the prior beliefs of all individuals are everywhere correct. This means that beliefs are correct not only the realized path of societal behavior, but also over successive layers counterfactual possibilities as well. These requirements seem excessive, but more importantly for our purposes they allow no latitude for mutation in beliefs to change behavior.

Instead, we consider beliefs that give rise to *self confirming equilibrium (SCE)* behavior. Introduced by Fudenberg and Levine (1993a), in a SCE individuals need not have mutually consistent and correct forecasts about counterfactual information.

**Definition 1** A *self confirming equilibrium (SCE)* is a collection of strategies  $f^* = (f_{ij}^*)$ ,  $i \in I$ ,  $j = 1, \dots, \infty$ , and a belief and arrival realization  $(\phi, \alpha)$  such that each individual  $ij$  has belief  $\phi_{ij} = \text{proj}_{ij}\phi$  and chooses a strategy  $f_{ij}^*$  such that

- (i)  $f_{ij}^*$  maximizes the payoff in (3) given  $\phi_{ij}$ ;
- (ii) for  $\mu_{f^*}$ -almost every history  $p^t$  and for any public knowledge event  $B$  that occurs during  $ij$ 's lifetime,<sup>10</sup>

$$\phi_{ij}(B | f_{ij}^*, p^t) = \mu_{f^*}(B | p^t). \quad (4)$$

A self confirming equilibrium (SCE) is a collection of strategies and beliefs, one for each individual, with the property that an individual  $ij$  makes the correct prediction about observable behavior of society, but only on those paths that are in the support of the realized distribution while that individual remains alive. The definition of Nash equilibrium is almost identical to Definition 1 with an important qualification that in a Nash equilibrium, equation (4) holds for every history  $p^t$ . Hence, every Nash equilibrium of the repeated game is a SCE, however, the converse is not true.

**Definition 2** A *self confirming equilibrium (SCE) belief arrival process* is a belief process  $\pi$  which satisfies: for  $\pi$ -almost every  $(\phi, \alpha)$  pair, there is a collection of strategies  $(f_{ij}^*)$  that, together with  $(\phi, \alpha)$ , constitute a self confirming equilibrium.

---

<sup>10</sup>Letting  $\mathcal{F}_i$  denote the  $\sigma$ -algebra that contain events observable to individuals in progeny  $i \in I$ , a public knowledge event  $B$  is one in which  $B \in \bigwedge_i \mathcal{F}_i$ .

Definition 2 defines a larger class of belief processes than if “SCE” were replaced with “Nash” in the definition. Every Nash belief process is also an SCE process. In a Nash equilibrium, beliefs are perfectly correlated both on realized paths and on those that arise with  $\mu_\pi$  probability zero. Hence, Nash belief processes are typically measures that involve a public randomization (possibly degenerate) over belief realizations  $\phi$ , each consistent with some Nash equilibrium. Though self confirming beliefs also require a high degree of correlation, the correlation need not be perfect. In either SCE or Nash equilibrium, the belief process simply describes the evolution of beliefs in this society that are consistent with equilibrium behavior.

### 3.4 An Example

To demonstrate the nature of self confirming belief dynamics of mechanism selection, suppose that there are three progenies with  $v_i = v > 1$ , and  $c_i = 1$  for each  $i = 1, 2, 3$ . Assume that two of three contributions are required to fund the project. Given these parameters, either type of funding equilibrium set,  $E_V$  or  $E_M$ , has the property that all three agents locate at the same mechanism. An individual receives  $v - 1$  or  $v$  in a  $V$ -funded profile depending on whether he and another contribute or whether he “free rides” while the other two contribute. He receives  $v - \frac{2}{3}$  in mechanism  $M$  if two of the three vote for the project in  $M$ .

Fix time  $t_0$  such that  $((V, 0), (V, 1), (V, 1)) \in E_V$  is the profile. The three individuals living at time  $t_0$  are located in  $V$ ; the first two contribute while the third is a free rider. Let  $1', 2'$ , and  $3'$  denote the three individuals alive at time  $t_0$ . Let  $1'', 2''$ , and  $3''$  denote their successors.

We will sketch a scenario in which there is a self confirming equilibrium transition from  $((V, 0), (V, 1), (V, 1)) \in E_V$  to  $((M, 1), (M, 1), (M, 1)) \in E_M$  which is not possible with Nash as the solution concept. To do this we suppose initially that  $1', 2'$  and  $3'$  all believe that  $((V, 0), (V, 1), (V, 1))$  is permanent. That is, each believes that the other two will keep their current actions forever. With this initial condition, a Nash equilibrium must prescribe that the three progenies *necessarily* remain in  $((V, 0), (V, 1), (V, 1))$  forever, even after the initial cohort of  $1', 2'$  and  $3'$  have died. The reason is that with these beliefs, no subsequent individual can anticipate a departure from  $((V, 0), (V, 1), (V, 1))$  since such a belief must contradict that of the initial cohort. In a Nash equilibrium, the beliefs of all individuals must be everywhere correct.

Suppose, however, that  $1'$  is replaced by  $1''$  at time  $t_1$  who adopts the belief that one of the other two lineages will depart for  $M$  once  $2'$  and  $3'$  have been replaced. Upon observing this switch, the other player will follow. Individual  $1''$  believes that if he attempts to pre-

empt this switch by leaving early for location  $M$ , he will induce a “punishment phase” in  $E_N$ . Individual  $1''$ 's best response under this belief is to keep his current action until he sees a departure toward  $M$ . Notice that the beliefs of  $1''$ ,  $2'$ , and  $3'$  are perfectly congruent on those publically observed trajectories on which  $1''$ ,  $2'$ , and  $3'$  *all coexist*. Individual  $1''$  will keep his action  $(V, 1)$  as long as  $2'$  and  $3'$  remain alive. Hence, each existing individual's prediction that the other two existing individuals keep their current actions is correct. Resulting behavior is therefore self confirming. Now consider that  $2'$  is replaced by  $2''$ . Given the self confirming assumption and the fact that  $1''$  remains alive at any given future time with positive probability, individual  $2''$ 's beliefs must be consistent with the continuation of the SCE. Individual  $2''$  will then assign the same probability that there is a departure toward  $M$  as  $1''$  conditional on he and  $1''$  both remaining alive. Given this belief, individual  $2''$ 's best response is also leave for location  $M$  once he sees the first departure.

Finally, at time  $t_3$ ,  $3'$  is replaced by  $3''$ . Given the continuation of the SCE individual  $3''$ , believing that his departure for  $M$  will induce the others to move immediately to  $M$  and support the project, individual  $3''$  will switch if revision opportunities are sufficiently rapid. This entire scenario is sketched in the table below.

Individual			Time	Newborn's belief
$1'$	$2'$	$3'$	$t_0$	$((V, 0), (V, 1), (V, 1))$ “forever”
$1''$	$2'$	$3'$	$t_1$	$((V, 0), (V, 1), (V, 1)) \rightarrow ((M, 1), (M, 1), (M, 1))$ only after $2'$ , $3'$ replaced”
$1''$	$2''$	$3'$	$t_2$	$((V, 0), (V, 1), (V, 1)) \rightarrow ((M, 1), (M, 1), (M, 1))$ only after $3'$ replaced”
$1''$	$2''$	$3''$	$t_3$	$((V, 0), (V, 1), (V, 1)) \rightarrow ((M, 1), (M, 1), (M, 1))$ immediately upon revision”

Observe, that the beliefs of, say  $3'$ , while consistent with SCE, is inconsistent with Nash equilibrium. The reason is that if the old individual  $3'$  were to defect to location  $M$ , then individuals  $1''$  and  $2''$  would then infer that the new individual  $3''$  had arrived and would subsequently depart for location  $M$ . But if this switch were to be anticipated by  $3'$  then he would most certainly defect to location  $M$  since he receives the public good at a lower cost to himself. This means that  $3'$ 's original belief that the others remain in location  $V$  forever is misforecast. The incompatibility of  $3'$ 's belief with the Nash assumption lies in the fact

that his forecasts are incongruent with those of 1'' and 2'' on the out-of-equilibrium event that 3' himself departs for  $M$ .

Notice that neither 1'' nor 2'' observe the birth and death times in other lineages. They do not need to. All that is required is that when nature pulls from the urn a belief for, say, individual 2'' this belief must be consistent with those of the other existing players along the realized path. It may seem strange at first that player 2'' can have beliefs that depend on the order of arrival in other lineages whose replacement time he cannot observe. The reason this is not as absurd as from first glance is that individuals do not "choose" their beliefs. That is, replacement times are conditioning events for the process  $\pi$ , but not for the individuals themselves.

### 3.5 The Assumptions on Beliefs

The following assumptions allow random heterogeneity in the beliefs of newborns consistent with SCE belief arrivals. First, observe that the Poisson processes order the newborns by their arrival times. Let  $\phi_k$  denote the realized belief of the  $k$ th newborn. (We consider only the case of  $k > \ell$  since the first  $\ell$  individuals are placed in the world at time 0.) Given an arrival realization  $\alpha$ , let  $Y(k)$  denote the interval of time that the latest newborn is  $k$ . Let  $I(k)$  denote the  $\ell$  individuals, one from each progeny, who are alive at the entry time  $\min_t Y(k)$  of the  $k$ th newborn (the set  $I(k)$  includes the  $k$ th individual). Denote  $\phi^k = (\phi_{ij})_{ij \in I(k)}$  to be the  $n$ -tuple of realized beliefs of individuals alive at the time of the  $k$ th newborn.

(A1) (Stationary Markov Belief Process) The belief process  $\pi$  defines a stationary Markov process on the belief realizations  $\phi$ , with transition probabilities expressed as

$$\pi(\{\phi^k \in A\} \mid \phi^{k-1})$$

for any measurable set  $A \subseteq \Phi^\ell$ .

(A2) (Independent Increments) Let  $t(k)$  denote the random variable representing the time that the first  $k$  individuals over all progenies have died. Then for each  $k$ ,

$$\phi_{k+1}(\cdot, p(t(k))) \text{ and } \phi^k \text{ are mutually independent.} \quad (5)$$

(A3) (Maximal Support) There is no other SCE belief process that has strictly larger support than  $\pi$ .

Assumption (A1) is made to simplify the result and could likely be relaxed. Assumption (A2), an “independent increments”-like property, allows beliefs in a progeny to mutate over time.<sup>11</sup> The idea is that, conditional on the position of the path process at the time of death of the first  $k - 1$  individuals, the  $k$ th individual can take a random draw among priors that are consistent with the path process realized thus far. This is possible despite the obvious correlation of beliefs in an SCE process by allowing random changes in “out-of-equilibrium” beliefs to precipitate changes in the set of self-confirming expectations across generations. Finally we assume that the SCE belief process has maximal support. This gives the maximal variation in beliefs across time which remain self confirming. In the previous example, assumptions (A1)-(A3) allows that individual 1’’s belief over events beyond 2’’s and 3’’s lifetime is drawn independently from 2’ and 3’.

## 4 Selection of a Mechanism

**Theorem** There exists a  $\hat{\lambda}$  such that if  $\lambda \leq \hat{\lambda}$  then for any SCE belief process  $\pi$  that satisfies (A1)-(A3) given  $\lambda$ , for any initial state  $p^0$  and for  $\mu_\pi$ -almost every  $p$ , there is a finite random stopping time  $\tau$  such that

1.  $p(\tau) \in E_M$
2.  $\mu_\pi(\{p(t) \notin E_M, \text{ some } t > \tau\}) = 0$ .

The proof is in the Appendix. The Theorem states that in the “fully rational limit” as  $\lambda \rightarrow 0$ , there is an eventual transition from any profile to a Majority-Rule funded profile such that the probability of departure thereafter from  $E_M$  is zero. It is in this sense that *funding in the Majority Rule location is globally absorbing with respect to  $\mu_\pi$* .

The Theorem does not, of course, imply that under “rational expectations” only location  $M$  and its associated Majority Rule mechanism prevails. Static equilibria in the sets  $E_V$  and  $E_N$  remain Nash, hence, self confirming equilibria of the repeated game without belief mutations.

**Some Intuition behind the Proof** The result may be broken down into two parts. First, it can be shown that there is an infinite, unbounded number of stopping times at which

---

<sup>11</sup>It has been shown by Lagunoff and Matsui (1993) in a non-equilibrium model with “belief mutations,” where mutations can be represented as independent marginal distributions  $\pi_{ij}$  over each  $ij$ ’s beliefs, that if  $\lambda$  is sufficiently small then the path distribution is ergodic and therefore has no absorbing states.



societal behavior will end up in  $E_M$ . The intuition for this is that assumptions (A1)-(A3) guarantee that the “right” sequence of beliefs to admit a transition will occur infinitely often. This sequence is roughly the sequence described in the example in Section 3. Recall that this sequence of beliefs is consistent with SCE, but not Nash equilibrium.

Second, it may be shown that there is one such stopping time from which no departure from the set  $E_M$  is possible. To see why societal behavior may be absorbed in  $E_M$ , observe that the self confirming hypothesis must identify as public knowledge the long run average time an average (anonymous) agent spends as a contributor in  $V$ . This knowledge must be acquired at some point in time before these contributors depart for  $E_V$ . Since no agent can beat the current odds of contributing fraction  $\frac{x^*}{m}$  of the time on average, and since there is a small delay in reaching an equilibrium  $E_V$  from  $E_M$ , each individual strictly prefers to remain in equilibrium  $E_M$ .

In the three person example in Section 3, an individual from , say, progeny 1 will only depart from  $M$  if he believes that he will be a free rider more than  $1/3$  of the time. This threshold comes from the fact that he is currently a contributor  $2/3$  of the time on average in the Majority Rule mechanism. He must then believe that his defection will induce one of the other agents, say agent 2, to split time contributing toward the good in  $V$ . Such a strategy could be accomplished through time averaged mixed strategy.<sup>12</sup> If the time split between the two agents is 50/50, for example, then both agents will defect from  $(M, 1)$  since each would obtain a long run average payoff of  $v - \frac{1}{2}$  in  $V$  instead of the  $v - \frac{2}{3}$  obtained in the Majority Rule mechanism. Finally, if both agents 1 and 2 defect to  $V$ , then agent 3 is better off being a full time contributor in  $V$  than staying alone in  $M$ .

This scenario, however, involves beliefs that violate the stated assumptions. The reason is that since beliefs are self confirming, agent 3 also anticipates this defection. Furthermore, since only the aggregate distribution,  $z = (n, \bar{n}, m, \bar{m})$ , not the profile  $(x_i, s_i)_{i \in I}$ , is observed (admittedly this is farfetched in a 3 person example), agents 1 and 2 have no way of monitoring their jointly coordinated time averaging strategy. In particular, once all the agents are located in  $V$ , agent 3 has an incentive to stop contributing at whichever time there is a changeover between agents 1 and 2. Since the fact that a deviation has occurred is not common knowledge, the continuation of the public knowledge path must be a SCE continuation. Hence, the lottery induced by time averaging in the Voluntary mechanism is unenforceable. This means that no one can do better than  $v - \frac{2}{3}$ .

---

<sup>12</sup>Agent 1 contributes fully some fraction of the time, then defects to  $(V, 0)$ , while agent 2 jumps in for the remainder of the time. This ignores the delays due to the stochastic revision times. These delays can be made arbitrarily small, however, as  $\lambda \rightarrow 0$ .

Therefore, an individual will defect to  $E_V$  only if he believes that a punishment ensues if he fails to defect. Such a punishment can only be conditioned on the aggregate distribution  $z$  and on calendar time. Since  $\lambda > 0$ , this punishment must therefore be a positive probability event, and each individual must correctly predict its likelihood as long as he remains alive. If at some time some initial cohort of arrivals in  $E_M$  all believe that  $E_M$  will prevail indefinitely, then no individual who overlaps with them will ever initiate such a punishment. Moreover, no individual who overlaps with individuals who overlap ... with this initial cohort will ever initiate a punishment. Therefore, all individuals will assign probability zero to its occurrence. Hence, no contributor will defect to a profile in  $E_V$ , and so behavior is absorbed into  $E_M$ .

The reason that escape from  $E_M$  is precluded only at certain times is that the initial cohort to arrive in  $E_M$  must have the “right” beliefs — a condition which, again, is guaranteed infinitely often by (A1)-(A3).

Finally, observe that if individuals’ beliefs were not self confirming, then a transition to some profile in  $E_V$  could occur if each individual expected that he himself would be a permanent free rider in  $E_V$ . Of course, to achieve a profile in  $E_V$ , someone having such beliefs must be wrong.

## 5 Concluding Remarks

This model provides a setting in which agents choose a mechanism by choosing the location of their choice. The comparison was made between a voluntary contribution mechanism and a majority rule mechanism. The mechanism that emerges is the one that prevails in the long run when entry and exit over time infuses society with new and heterogeneous beliefs, even though societal beliefs are self confirming. The dynamics of entry and exit with new beliefs help to create transitions from one mechanism to another. The results suggest that the Majority Rule mechanism is the unique survivor of such a process in the two-mechanism example.

An upshot of the SCE dynamic formulated here is that individuals acquire beliefs one-at-a-time. Change occurs only after sufficient accumulation of “like minded” individuals. The accumulation of these beliefs may occur over many generations and over many progenies. An unfortunate aspect is that mechanism selection in this model is an asymptotic property. One would like to say something more about the realized path. The problem is that the path properties depend critically on details of the belief process  $\pi$ .

A second limitation is that the class of environments studied here is narrow. Obvious

extensions are possible. One such extension is to consider a smooth provision technology and/or fractional contributions. In any SCE process that satisfies the hypothesis of Theorem 2, only the static equilibria are candidates for steady states. Hence, examining “smooth” public goods may not add much to the analysis, except possibly to eliminate the extreme nature of the static equilibria (i.e., either full funding or no funding). Furthermore, transitions from states that are not equilibria of the static game always exist, and so any dynamics that introduce all the possible sets of self confirming expectations will lead society away from these states. As for contributions from a continuum, technical difficulties arise since the argument that certain profiles are visited by the process infinitely often breaks down with an infinite state space. It might usefully be argued nevertheless that certain *neighborhoods* of the profile space are recurrent. If so, the SCE dynamics formulated here may extend to quite general public goods environments.

Another possible extension is to examine location in a general class of mechanisms. The class of *binary decision mechanisms* studied in Lagunoff and Matsui (1995) is a good candidate as it includes many commonly observed mechanisms, including both Voluntary Contribution and Majority Rule. That paper examined the stability properties under completely heterogeneous beliefs about behavior in each mechanism. The performance of each mechanism was examined in isolation. A locational setting provides a natural endogenous selection “mechanism” over provision mechanisms. In the general case, it is likely that multiple jurisdictions survive. In a multi-jurisdictional structure, it is conceivable that the presence of one mechanism in the structure may actually *enhance* the long run viability of another, since it might attract the type of agent who is likely to destroy cooperation in the other mechanism. Hence, the framework may allow the possibility of complementarities between institutions that serve the same purpose.

## 6 Appendix: Proof of Theorem

Before proceeding with the argument, it is first necessary to define the notion of an open set in the support of  $\pi$ . Fix an initial profile  $p^0 \equiv (x_0, s_0)$  and let  $\Omega$  denote the space of all sample paths. Each individual’s information is described by the  $\sigma$ -algebra  $\mathcal{F}_i$ . For now we drop the subscript  $j$  for  $ij$ . First, fix a strategy  $f_i$  for an individual and let  $F_i$  denote the set of all such strategies. To be consistent with the physical restrictions of the model requires that for each  $f_i \in F_i$ ,  $f_i(p^t) = (x_{it}, s_{it})$  only if  $t$  is a revision opportunity for  $i$ . To construct the space of beliefs, let  $D_i$  denote the set of probability measures on  $(\Omega, \mathcal{F}_i)$ , and  $B_{f_i} \in \mathcal{F}_i$  denote the set of paths  $p$  with  $s_i(t) = f_i(p^t)$  on the revision points of  $i$ . We define the set of possible beliefs,  $\Phi_i$ , as the space of  $\mathcal{F}_i$ -measurable mappings,  $\phi_i : (B_{f_i}, p^0) \mapsto \phi_i(f_i, p^0)$  for

each  $i$  which satisfy two conditions. First,  $\phi_i(f_i, p^0) \in D_i$  puts full support on  $B_{f_i}$ . Second, for any two strategies  $f_i$  and  $\hat{f}_i$ , and for any  $\mathcal{F}_i$ -measurable set  $B \subset B_{f_i} \cap B_{\hat{f}_i}$ ,

$$\phi_i(B \mid f_i, p^0, B_{f_i} \cap B_{\hat{f}_i}) = \phi_i(B \mid \hat{f}_i, p^0, B_{f_i} \cap B_{\hat{f}_i}) \quad (6)$$

Equation (6) assures that  $i$ 's beliefs are consistent with some belief over others' behavior strategies and does not vary with  $i$ 's own strategy.

Now define the subset  $\Phi(f_i) \subseteq D_i$  as the set of measures  $\phi_i(f_i, p^0)$  which put probability one on the set  $B_{f_i}$ . The set  $\Phi(f_i)$  inherits subset topology of the weak\* topology on  $D_i$ . This topology on  $\Phi(f_i)$  can therefore be generated by open sets of the form:

$$N_\eta(\phi) \equiv \left\{ \hat{\phi} \in D_i : \left| E_{\hat{\phi}(f_i, p^0)}[U_i^m] - E_{\phi(f_i, p^0)}[U_i^m] \right| < \eta, \forall m = 1, \dots, M \right\}$$

where  $\eta > 0$  and  $U_i^m$ ,  $m = 1, \dots, M$  is any finite collection of bounded continuous functions of  $p$ . If we can show that  $\Phi(f_i)$  is closed then it is compact in this topology. But closedness of  $\Phi(f_i)$  must hold since the limit point of any sequence  $\{\phi^l(f_i, p^0)\}$  of measures that put full support on  $B_{f_i}$ , must also put full support on  $B_{f_i}$ . Hence, the set  $\Phi(f_i)$  is compact. We may now express the space of beliefs as the subset  $\Phi \subseteq \times_{f_i \in F_i} \Phi(f_i)$  which satisfies (6). The space  $\Phi_i$  inherits the product topology from  $\times_{f_i \in F_i} \Phi(f_i)$  and is compact.

Now suppose that  $\pi$  satisfies (A1)-(A3). We divide the argument in two parts. We show first:

**Claim 1** *There exists a  $\hat{\lambda} > 0$  such that for all  $\lambda \leq \hat{\lambda}$ , with  $\mu_\pi$ -probability one there is an unbounded sequence of random stopping times  $\{t_m(p)\}_{m=1}^\infty$  with  $t_1(p) < t_2(p) < \dots$  such that  $(x_{t_m(p)}, s_{t_m(p)}) \in E_M$  for all  $m$ .*

**proof of Claim 1** Fix a profile  $(x^*, s^*) \in E_M$ , and suppose that the profile at time  $t = \min Y(k)$  at which the  $k$ th newborn enters is  $(x_t, s_t) \in E_V$ . Let  $\{i_1, \dots, i_{x^*}\}$  denote the set of progenies that contribute in  $V$ . Relabel, if necessary, so that this set is given by  $\{1, \dots, x^*\}$ . We suppose without loss of generality that every individual would prefer to have the good even if he had to fully contribute, i.e.,  $|\{i : v_i - c_i > 0\}| = \ell$ .

Fix  $i \in \{1, \dots, x^*\}$ . Define  $t_i[t]$  to be  $i$ 's first decision point after time  $t$ . Now define the sets:

$$\begin{aligned} D_k^1 &= \{s_{1t_1[t(k)]} = M \text{ implies } s_{2t_2[t_1[t(k)]]} = M\} \\ D_k^2 &= \{s_{2t_2[t_1[t(k)]]} = M \text{ implies } s_{3t_3[t_2[t_1[t(k)]]]} = M\} \\ &\vdots \\ D_k^\ell &= \{s_{it_\ell[t_{\ell-1}, \dots, t_1[t(k)]]} = M \forall i\} \end{aligned}$$

The set  $D_k^1$ , for instance, is the event that  $i = 1$  locates in  $M$  without delay after  $t(k)$  implies that  $i = 2$  will also move to  $M$  without delay after  $t_1[t(k)]$ . Now define the sets  $\{C_k^i\}_{i=1}^\ell$  by

$$\begin{aligned} C_k^1 &= \cap_{i=1}^\ell D_k^i \\ C_k^2 &= \cap_{i=2}^\ell D_k^i \\ &\vdots \end{aligned} \tag{7}$$

so that  $C_k^1$  is the event that  $i = 1$ 's decision to move to location  $M$  from  $V$  induces an immediate chain reaction of all  $\ell - 1$  other progenies whatever their current location.

Now fix some  $\delta > 0$  and define the following sets in belief space by

$$B_\delta^i(p^{t(k)}) = \left\{ \phi_i : \max_{f_i} E_{\phi(f_i, p^{t(k)})} [U_i] > v_i - c_i \frac{x^*}{\ell} - \delta \right\} \tag{8}$$

In (8),  $B_\delta^i(p^{t(k)})$  is the neighborhood of beliefs that puts an individual from progeny  $i$  within  $\delta$  of his payoff  $v_i - c_i \frac{x^*}{\ell}$  of locating in  $M$  if he uses his optimal strategy starting from profile  $p^{t(k)}$ . Let  $\hat{\lambda}$  denote the supremum over all  $\lambda > 0$  such that there exists some  $\delta > 0$  for which  $\phi_i \in B_\delta^i(p^{t(k)})$  implies that  $\phi_i(C_k^i | \hat{f}_i, p^{t(k)}) = 1$  for maximizer  $\hat{f}_i$ . For all  $\lambda \leq \hat{\lambda}$  a belief in  $B_\delta^i(p^{t(k)})$  means that  $i$  anticipates that all individuals left in  $V$  after he departs immediately move to  $M$  and support the project at their first revision opportunity. Moreover, the delay before this is realized is sufficiently small so that having a belief  $\phi_i \in B_\delta^i(p^{t(k)})$  entails  $i$  choosing  $s_i = M$  without delay.

Hence, if  $\delta$  is small then if  $k + 1$  corresponds to agent  $i$ , then  $i$  moves/stays in  $M$  at  $t_i[t_{i-1}[\dots t_1[t(k)]\dots]]$ . Fix one such  $k$ . Since individual  $k+1$  is alive at time  $t_i[t_{i-1}[\dots t_1[t(k)]\dots]]$ ,  $i = 1 \dots, x^*$ , with positive probability, it follows that if event  $C_k^i$  occurs during  $(k + 1)$ 's lifetime then

$$\mu_\pi(C_k^i) = \gamma_i > 0, \quad \forall i = 1, \dots, \ell.$$

It therefore follows that there is some  $\gamma > 0$  for which

$$\mu_\pi(C_k \equiv \cap_{i=1}^\ell C_k^i) = \gamma.$$

Fix an arrival realization  $\alpha$ . Observe that for fixed  $\delta$ , by assumptions (A1) and (A2) the events  $\{A_k\}$  defined by

$$A_{k+1} \equiv \left\{ \phi_{k+1}(\cdot, p^{t(k)}) \in B_\delta^i(p^{t(k)}) \right\}$$

are mutually independent. Furthermore, by (A3) there is some  $d > 0$  such that  $\pi(A_k | \alpha) \geq d$ . Since  $\sum_k \pi(A_k | \alpha) = \infty$ , we invoke the Second Borel Cantelli Lemma to assert that

$$\pi(A_k, k - \text{infinitely often} | \alpha) = 1.$$

Now define  $\bar{\alpha}(t) = \sum_i \alpha_i(t)$  denoting the latest individual to be born by time  $t$ . Define the subsequence  $\{k_m\}$  of  $k = 1, 2, \dots$  recursively by

$$\begin{aligned} & k_1 \\ \bar{\alpha}(t(k_1)) &= k_2 \\ \bar{\alpha}(t(k_2)) &= k_3 \\ & \vdots \end{aligned}$$

Each successive  $k_m$  denotes a complete turnover in the population of all agent types  $i$ . As before,  $\pi(A_{k_m} | \alpha) \geq d > 0$ , and the sequence  $\{A_{k_m}\}_{m=1}^{\infty}$  of sets of beliefs is independent. Consequently, the corresponding  $\{C_{k_m}\}$  are mutually independent. The Second Borel-Cantelli Lemma is invoked once again as

$$\sum_{m=1}^{\infty} \mu_{\pi}(C_{k_m}) \geq \sum_{m=1}^{\infty} \gamma d = \infty,$$

proving that

$$\mu_{\pi}(C_{k_m}, m - \text{infinitely often}) = 1.$$

Hence, there is corresponding infinite set of stopping times  $\{t_m(p)\}$  with  $t_1(p) < t_2(p) < t_3(p) < \dots$  on a set of  $\mu_{\pi}$ -probability one paths  $p$ , such that  $(x_{t_m(p)}, s_{t_m(p)}) \in E_M$  for each  $m$ , and for any  $t$  there is some  $m$  for which  $t_m(p) > t$ . This proves Claim 1.  $\square$

**Claim 2**  $\mu_{\pi}(\{p : \forall m, \exists \tau > t_m(p), s(\tau) \notin E_M\}) = 0$ .

**proof of Claim 2** Given the construction of stopping times above, fix a stopping time  $t_m$ . Let  $(x_{t_m}, s_{t_m}) \in E_M$  denote the profile at  $t_m$ . Let  $i^* \in M$  denote the first progeny currently in  $M$  who has a revision opportunity to depart from  $M$  to  $V$  or change his vote in  $M$ . There are only two scenarios in which  $i^*$  will switch his behavior.

- (a) The long run payoff that  $i^*$  anticipates in  $V$  exceeds his current payoff from continuing in  $E_M$ .
- (b) Agent  $i^*$  incurs a punishment for a failure to defect from  $E_M$ .

Consider first part (a). As  $\lambda \rightarrow 0$ ,  $i^*$ 's long run payoff in  $V$  will exceed that in  $M$  if the long run average time spent as a contributor in  $V$  is strictly less than  $\frac{\pi^*}{m}$ , that is, if he has a strategy  $f_{i^*}$  that satisfies

$$E_{\phi_{i^*}(f_{i^*}, p^t)} \left[ \int_t^{\infty} e^{-\tau(\tau-t)} u_{i^*}(p_t(\tau)) d\tau \right] \rightarrow (v_{i^*} - c_{i^*} \beta_{i^*}) \text{ as } \lambda \rightarrow 0 \text{ for some } \beta_{i^*} < \frac{\pi^*}{m}.$$

However, the time averaged lottery  $\beta_{i^*}$  achieved by  $i^*$  can be achieved by any other agent since the game is anonymous. To see this, observe that since  $\beta_{i^*} < \frac{x^*}{m}$  there must be some other progeny  $i$  with  $\beta_i > \frac{x^*}{m}$ . However, by anonymity,  $i$  can mimic the behavior of  $i^*$  without detection from anyone but  $i^*$ . Since the behavioral path is consistent with a SCE, other individuals' forecasts about the public path do not change. Therefore, the continuation after  $i$ 's defection must be an anonymous lottery  $\bar{\beta}$  which is preferable to  $i^*$  than any punishment phase induced by  $i^*$ 's public defection. Hence  $\beta_i = \bar{\beta}$  for all  $i$  where  $\bar{\beta}$  is public knowledge.

Moreover, since  $(x_{t_m}, s_{t_m})$  is a Nash equilibrium of the stage game, the current set of residents of  $M$  satisfies  $M = \{i \mid v_i - c_i \frac{x^*}{m} > 0\}$  while the current set of residents in  $V$  satisfies  $V = \{i \mid v_i - c_i \frac{x^*}{m+1} < 0\}$ . Therefore,  $\bar{\beta} \geq \frac{x^*}{m}$  as none of the current residents of  $V$  would participate in a time averaged lottery in  $V$ . Since  $\lambda > 0$  creates a delay in reaching a profile in  $E_V$ , it follows that

$$\{f_{i^*} \mid s_{i^*t} = V\} \max_{s_{i^*t} = V} E_{\phi_{i^*}(j_{i^*}, p^t)} \left[ \int_t^\infty e^{-\tau(\tau-t)} u_{i^*}(p_t(\tau)) d\tau \right] < v_{i^*} - c_{i^*} \frac{x^*}{m}$$

and so  $i^*$  is never better off in  $V$ .

Now consider case (b). In this case,  $i^*$  might consider a deviation to  $V$  if he believes that failure to defect will lead to a punishment more severe than spending time in  $V$ . Let  $A_{t_m} = \{p(\tau) \notin E_M, \text{ for some } \tau \geq t_m\}$  denoting the event of a departure from  $E_M$  after  $t_m$ .

Define the belief  $\phi^* \in \Phi$  to satisfy

$$\phi_{i^*}^*(A_{t_m} \mid \hat{f}_i, p^{t_m}) = 0$$

for some strategy  $f_{i^*}^*$  consistent with this belief. This belief places probability zero on a departure from  $E_M$  after  $t_m$ . Clearly, there exists a best response  $\hat{f}$  to  $\phi^*$  which is consistent with  $\phi_{i^*}^*$ : a best response to this belief is to keep  $(x_{it_m}, s_{it_m}) = (1, M)$  at least until one observes a defection from  $E_M$ .

Suppose that the  $k$ th individual to enter society is of the  $i^*$ th progeny. Given  $\hat{f}_{i^*}$  and some real number  $\eta > 0$ , define the set

$$N_{\eta, \hat{f}_{i^*}}(\phi^*) \equiv \left\{ \phi \in \Phi : \left| E_{\phi^*(\hat{f}_{i^*}, p^{t(k-1)})}[U_{i^*}] - E_{\phi(\hat{f}_{i^*}, p^{t(k-1)})}[U_{i^*}] \right| < \eta, \text{ for bounded and continuous } U_{i^*} \right\}$$

This set,  $N_{\eta, \hat{f}_{i^*}}(\phi^*)$  is open in the product topology on  $\Phi$ .

Observe then by the argument in Claim 1 that, given an arrival realization  $\alpha$  there is a set of  $\phi$  with  $\pi(\cdot \mid \alpha)$ -probability one for which the set  $N_{\eta, \hat{f}_{i^*}}(\phi^*)$  must occur infinitely often.

Now observe that there exists some  $\eta > 0$  such that  $N_{\eta, \hat{f}_{i^*}}(\phi^*) \cap \{\phi_{i^*} : \phi_{i^*}(A_{t_m} | \hat{f}_{i^*}, p^{t_m}) > 0\} = \emptyset$ . That is, beliefs in a particular neighborhood  $N_{\eta, \hat{f}_{i^*}}(\phi^*)$  of  $\phi^*$  cannot place positive probability on the event that a departure from  $E_M$  occurs after any time  $\tau > t_m$ , since we can construct a particular  $U_i$  by varying discount rate  $r$  so that  $r \rightarrow 0$  (i.e., no discounting), where defection sometime after  $t_m$  under any  $\phi_i \in \{\phi_i : \phi_i(A_{t_m} | \hat{f}_i, p^{t_m}) > 0\}$  produces large loss in expected utility.

Finally, we argue that for this  $\eta$

$$\pi(\{\phi_{k'}(\cdot, p^{t(k'-1)}) \in \{\phi_{i^*} : \phi_{i^*}(A_{t_m} | \hat{f}_{i^*}, p^{t_m}) > 0\} \mid \exists k < k', \phi_k \in N_{\eta, \hat{f}_{i^*}}(\phi^*)\}) = 0. \quad (9)$$

That is, no possible belief in  $\{\phi_{i^*} : \phi_{i^*}(A_{t_m} | \hat{f}_{i^*}, p^{t_m}) > 0\}$  will be adopted by any individual whose predecessors in the progeny  $i^*$  adopted belief in the particular neighborhood  $N_{\eta, \hat{f}_{i^*}}(\phi^*)$  of  $\phi^*$ .

To show this last step, suppose that (9) is violated. If  $\phi_k \in N_{\eta, \hat{f}_i}(\phi^*)$  for some  $k < k'$ , then for  $\eta$  is sufficiently small, it follows that  $k$  believes that the probability of defection from location  $M$  after  $t_m$  is zero. Now suppose that  $\phi_{k'}(A_{t_m} | \hat{f}_{k'}, p^{t_m}) > 0$  for some  $k' > k$ . Since there are no behavioral changes from  $t_m$  to the first time  $\tau$  that  $s_\tau \notin M$ , let  $k'$  denote the individual who departs from  $M$ . If  $\tau$  is a revision time for  $k'$ ,  $f_{k'}(p^\tau) \neq (M, 1)$  is a best response to belief  $\phi_{k'}$ . Since only calendar dates rather than revision and replacement times are publically observable, he must believe that some punishment ensues with high enough probability in the absence of a switch by any fixed calendar time  $\tau$ . However, since  $\lambda \neq 0$ , there is positive probability that  $k'$  has no such revision opportunity before any fixed calendar date  $\tau$ . Hence, such a punishment lies in the support of  $\mu_f$ . That is,

$$\mu_f(\{s_{it} \notin M, i \neq i^*, \text{ some } t > \tau\} \mid s_\tau = M) = 1 - \varepsilon \quad (10)$$

for some sufficiently small  $\varepsilon > 0$ .

However, if individual  $k$  is still alive and has beliefs  $\phi_k \in N_{\eta, \hat{f}_i}(\phi^*)$ , then by SCE the true distribution  $\mu_f$  must assign the probability of a departure from location  $M$  by time  $t$  to be zero. Since the revision and replacement times of any individual from progeny  $i^*$  are not observed by the other progenies, other individuals cannot condition on the first time  $t$  that  $k$  "dies." Recall the notation  $I(k)$  which denotes the set of individuals alive when  $k$  enters. Then  $I^{-1}(k)$  denotes the set of individuals who enter when  $k$  is alive. Hence, any contemporary of  $k$ , i.e. any individual  $k'$  where  $k' \in I^{-1}(k)$  cannot be the first to depart from  $M$  without violating the SCE assumption. In order for (10) to be satisfied, it must be the case that  $k'$  enters the world after all contemporaries of  $k$  have died. This means that  $k' \notin I^{-1}(j)$  for all  $j \in I^{-1}(k)$ . However, since  $j \in I^{-1}(k)$  never initiates a switch, no contemporary of  $j$  believes that  $j$  ever initiates a switch, and so SCE requires



that  $k' \notin I^{-1}(I^{-1}(k))$  where  $I^{-1}(I^{-1}(k))$  denotes the contemporaries of the contemporaries of  $k$ . Inducting on this definition, it follows that  $k' \notin I^{-1}(I^{-1}(\dots I^{-1}(k) \dots))$ . That is, since no contemporary of no contemporary of, ... individual  $k$  ever initiates a switch, individual  $k'$  cannot overlap with such an individual. This implies that  $k' = \infty$ .

We have therefore shown that there is some stopping time that hits beliefs  $\phi_{ij} \in N_{\eta, f_i}(\phi^*)$  for all  $ij$  alive when societal behavior lies in  $E_M$ , and given these beliefs, no future individual will ever adopt a belief that entails that he defect from  $M$ . This proves the second Claim.  $\square$

## References

### References

- [1] Battigalli, P. and P. Guatoli (1988), "Conjectural Equilibrium," mimeo.
- [2] Blume, L. (1993), "The Statistical Mechanics of Strategic Interaction," *Games and Economic Behavior*, 5: 387-424.
- [3] Blume, L. and D. Easley (1992) "Rational Expectations and Rational Learning," mimeo, Cornell University, July.
- [4] Buchanan, J. (1965), "An Economic Theory of Clubs," *Economica*, 33: 1-14.
- [5] Caplin, A. and B. Nalebuff (1992), "Individuals and Institutions," *American Economic Review*, 82: 317-22.
- [6] Caplin, A. and B. Nalebuff (1993), "Competition Among Institutions," mimeo, July.
- [7] Crawford, V. (1985), "Efficient and Durable Decision Rules: A Reformulation," *Econometrica*, 53: 817-35.
- [8] Ellison, G. (1993), "Learning, Local Interaction and Coordination," *Econometrica*, 61: 1047-71.
- [9] Fudenberg, D. and D. K. Levine (1993a), "Self-Confirming Equilibrium," *Econometrica*, 61:523-45.
- [10] Fudenberg, D. and D. K. Levine (1993b), "Steady State Learning and Nash Equilibrium," *Econometrica*, 61:547-73.
- [11] Gilboa, I. and A. Matsui (1991), "Social Stability and Equilibrium," *Econometrica*, 59: 859-67.
- [12] Greenberg, J. (1983), "Local Public Goods with Mobility: Existence and Optimality of a General Equilibrium," *Journal of Economic Theory*, 30, 17-33.
- [13] Groves, T. and J. Ledyard (1985), "Incentive Compatibility Ten Years After," Northwestern University Discussion Paper #648. April.
- [14] Holmström, B. and R. Myerson (1983), "Efficient and Durable Decision Rules with Incomplete Information," *Econometrica*, 51: 1799-1819.
- [15] Kalai, E. and E. Lehrer (1993a), "Rational Learning Leads to Nash Equilibrium," *Econometrica*, 61: 1019-46.

- [16] Kalai, E. and E. Lehrer (1993b), "Subjective Equilibrium in Repeated Games," *Econometrica*, 61: 1231-40.
- [17] Lagunoff, R. (1992), "Fully Endogenous Mechanism Selection on Finite Outcome Sets," *Economic Theory*, 2: 465-80.
- [18] Lagunoff, R. and A. Matsui (1995) "Evolution in Mechanisms for Public Projects," *Economic Theory*, forthcoming.
- [19] Ledyard, J. and T. Palfrey (1994), "Voting and Lottery Drafts as Efficient Public Goods Mechanisms," *Review of Economic Studies* 61: 327-55.
- [20] Mailath, G. (1992), "Introduction: Symposium on Evolutionary Game Theory," *Journal of Economic Theory*, 57: 259-77.
- [21] Matsui, A. (1992), "Best Response Dynamics and Socially Stable Strategies," *Journal of Economic Theory*, 57: 343-62.
- [22] Matsui, A. and R. Rob (1992), "Evolution and Rationality in Societal Games," CARESS Working Paper #91-31, University of Pennsylvania.
- [23] Nöldeke, G. and L. Samuelson (1993) "An Evolutionary Analysis of Backward and Forward Induction," *Games and Economic Behavior*, 5: 425-54.
- [24] Rubinstein, A. and A. Wolinsky (1990), "Rationalizable Conjectural Equilibrium: Between Nash and Rationalizability," *Games and Economic Behavior*, 2:
- [25] Scotchmer, S. (1985), "Profit Maximizing Clubs," *Journal of Public Economics*, 27: 25-45.
- [26] Scotchmer, S. and M. Wooders (1987), "Competitive Equilibrium and the Core in Club Economies with Anonymous Crowding," *Journal of Public Economics*, 34: 159-73.
- [27] Tiebout A. (1956), "A Pure Theory of Local Expenditures," *Journal of Political Economy*, 64: 416-24.
- [28] Vassilakis, S. (1990), "Rules for Changing Rules," University of Pittsburgh Working Paper #265, November.
- [29] Wooders M. (1978), "Equilibria, the Core, and Jurisdiction Structures in Economies with a Local Public Good," *Journal of Economic Theory*, 18, 328-48.
- [30] Wooders M. (1988), "Stability of Jurisdiction Structures in Economies with Local Public Goods," *Mathematical Social Sciences*, 15: 29-49.