Equilibrium Selection in Participation Games: A Unified Framework

David M. Frankel (Melbourne Business School)*

May 26, 2024

Abstract

In many applied settings, an activity or project requires a critical mass of participants to be worthwhile. This property can give rise to multiple equilibria. We study seven well-known equilibrium selection theories: two heuristic arguments, two models with rational players, and three from the evolutionary literature. With one exception, each relies on strategic complementarities. We relax this to a weak single crossing property and show that the theories' predictions have a simple common form: an agent plays a best response to some fictional distribution of the participation rate of her opponents.

JEL: C72, C73, C79.

Keywords: equilibrium selection, evolutionary games, replicator dynamic, mutations, trembles, shocks, stochastic stability, ergodic distribution, global games, dynamic games, single crossing, implementation theory, mechanism design, principal-agents problem.

---

*200 Leicester Street, Carlton, VIC 3053, Australia. Email: d.frankel@mbs.edu. I thank seminar participants at Berkeley, Hebrew U., Melbourne, Monash, Northwestern, Paris School of Economics, Penn, Princeton, Sorbonne, Stanford GSB, Tel Aviv, UCL, and Wisconsin. This paper has also benefited from the comments and suggestions of Sylvain Chassang, David Dillenberger, Federico Echenique, Faruk Gul, Christian Hellwig, Philippe Jehiel, Simon Loertscher, George Mailath, Benny Moldovanu, Stephen Morris, Ady Pauzner, Alessandro Pavan, Philip Reny, Marzena Rostek, Chris Shannon, Lones Smith, Jakub Steiner, Jeroen Swinkels, Michal Szkup, Steve Tadelis, Juuso Toika, Eyal Winter, Ming Yang, Leeat Yariv, and Peyton Young.

1 Introduction

Many activities offer benefits that depend on the number of others who participate. Examples include investing in a project; joining an online platform; buying an electric vehicle; and attacking a currency or political regime.

If participating is worthwhile only if enough others participate, then both all-participate and none-participate are self-fulfilling prophecies. This multiplicity is problematic for a principal who wishes to induce the agents to participate in the activity. To choose an optimal scheme, the principal must know which schemes will induce the agents to participate: she needs a theory of equilibrium selection.

The literature gives us a choice of such theories. These include heuristics and the predictions of evolutionary and global games. In applied work, most researchers choose a single theory, applying it throughout their analysis. A concern is that the resulting predictions may depend nontrivially on this choice.

A more robust approach would be to prove results that hold for some large class of selection theories. However, a researcher who pursued such a strategy would need to allocate scarce space in her paper to explaining the various theories. Little space would remain to present her own findings. Unsurprisingly, no one (to our knowledge) has chosen this route.

In addition, many selection theories rely on restrictive payoff assumptions that hamper their application. The most common is strategic complementarities: the payoff from participating is nondecreasing in the overall participation rate. This property is violated by many common schemes.[1]

In summary, in order for a multi-criterion approach to be practical, one must identify a class of selection criteria that have a simple common form that can be easily embedded in a larger model. And to permit the study of real-world schemes that violate strategic complementarities, one must rely on a substantially weaker

[1] Examples appear in §6 and in Frankel [22].

payoff assumption. This paper provides both such advances.

We study seven iconic selection theories: two heuristics and five noncooperative models. Of the latter, two assume rational players and three are from the evolutionary literature. We show that these seven theories all yield criteria of a simple common form: an agent plays a best response to some distribution of the proportion of others who participate.[2] This distribution is common across the agents and does not vary with the game's payoffs. We refer to it as the agents' *fictional beliefs* as it need not coincide with the true distribution or with the agents' actual beliefs.

In the case of discrete agents, we assume only a weak single crossing property that Athey [5] calls "weak SC1": if the payoff from participating is positive for one other-agent participation rate, then it is not negative for any higher rate.[3] We discuss two real-world examples that satisfy this property but not strategic complementarities.[4]

Prior work has shown that in two-agent participation games, the risk-dominant equilibrium is selected in each of the five noncooperative theories that we study. We provide a unified explanation for these results: it is because these five theories yield criteria with our simple common form and, moreover, treat the two actions symmetrically. Any such theory must reduce to risk dominance in the two-agent case. In contrast, the heuristic theories treat the two actions asymmetrically and do not reduce to risk dominance.

Our study is motivated by the problem of a principal who wishes to design a scheme to induce a set of agents to participate in a joint activity. In the online appendix, we develop an algorithm to solve this problem. Frankel [22] uses this method to study three versions of a security design problem. One of these versions satisfies weak SC1 but no stronger property, and thus exploits the full strength of our

[2]The literature typically refers to those who play the participation game as "players" as they are considered in isolation. In our setting, it makes more sense to refer to them as "agents" as most applications will also have a principal who, technically, is a player as well. Thus we will typically use "players" to refer to past results that have no principal and "agents" to refer to our own framework.

[3]With infinitesimal agents, somewhat stronger conditions are needed; see §5.

[4]See n. 1.

results.

In order to provide a uniform sufficient condition for our main result, we extend the various selection theories. For the evolutionary theories of Foster and Young [19] and Kandori, Mailath and Rob [35], as well as the dynamic rational-player theory of Matsui and Matsuyama [39], the $n \geq 2$ player case was previously studied by Kim [37] under the assumption of strategic complementarities. We relax this to weak SC1. The evolutionary theory of Fudenberg and Harris [25] has been studied only in the case of $n = 2$ players under strategic complementarities; we extend their result to $n \geq 2$ players relying only on weak SC1. Finally, we extend the literature on global games with single crossing properties.[5]

The rest of this paper is as follows. We set out the model in §2. We present the two heuristic theories in §3 and the five noncooperative theories in §4. A detailed summary of results appears in §5. Two examples appear in §6, while §7 concludes. The online appendix (Frankel [23]) includes an analysis of the principal-agents problem, the longer proofs from the paper, and a notation guide.

2 The Model

The game is played by a set I of *ex-ante* identical agents. The agents may be either discrete ($I = \{1, ..., n\}$) or infinitesimal ($I = [0, 1]$). They have aggregate measure one, so the measure of a single agent is $1/n$ in the discrete case.[6]

An agent chooses an action $a \in \{0, 1\}$, where action 1 (resp., 0) is interpreted as

[5] See §4.3, n. 46.

[6] We must remain vague for now about the context in which the game of I agents is played, as well as the agents' information sets and degrees of rationality. For instance, in evolutionary games, the game is repeatedly played by random sets I of symmetrically informed but boundedly rational agents who are drawn from a larger population. In global games, the game is played once with a fixed set I of imperfectly informed but fully rational agents.

(not) participating in a given activity. From the perspective of a given agent, let[7]

$$\ell \in \Lambda \overset{d}{=} \begin{cases} [0,1] & \text{if agents are infinitesimal} \\ \lambda \overset{d}{=} \left\{ \frac{i}{n-1} : i = 0, ..., n-1 \right\} & \text{if agents are discrete} \end{cases}$$

denote the proportion of the other agents who participate. We refer to ℓ as the *other-agent participation rate*. An agent's utility is given by a function $u(a, \ell)$: she cares about the proportion of others who participate but not their identities. An important role will be played by an agent's net payoff

$$\pi^u(\ell) \overset{d}{=} u(1, \ell) - u(0, \ell) \tag{1}$$

from participating for a given other-agent participation rate ℓ. We will refer to π^u as the *payoff function* and sometimes write it as π, suppressing the underlying utility function u.

2.1 Single Crossing Properties

In the case of discrete agents, our results will assume only that the payoff function π satisfy the following property.[8]

Definition 1. (Athey [5]) A function $h : S \to \Re$ on a set $S \subseteq \Re$ satisfies *weak single crossing in a single variable* ("weak SC1") on S if for all $s_H > s_L$ both in S, $h(s_L) > 0$ implies $h(s_H) \geq 0$.

That is, if π it is positive for some other-agent participation rate ℓ, then it is not negative for any higher rate. Some of our results for infinitesimal agents will rely on the following strengthening of weak SC1.

[7] The notation "$\overset{d}{=}$" denotes a definition. In the definition of λ, i is interpreted as the number of others who participate while $n-1$ is the number of other agents.

[8] Definitions 1 and 2 are from the working paper version (Athey [5, p. 6]) of Athey [6]. Athey [7] has verified that all of the results in her published paper remain valid under these earlier definitions, which we find more intuitive.

Definition 2. (Athey [5]) A function $h : S \to \Re$ satisfies *single crossing in a single variable* ("SC1") on S if for all $s_H > s_L$ both in S, $h(s_L) \geq 0$ implies $h(s_H) \geq 0$ and $h(s_L) > 0$ implies $h(s_H) > 0$.

That is, if π is nonnegative (positive) for some other-agent participation rate ℓ, then it is also nonnegative (resp., positive) for any higher rate. Finally, some results from the literature rely on a strengthening of SC1:

Definition 3. (Athey [5]) A function $h : S \to \Re$ satisfies *strict single crossing in a single variable* ("strict SC1") on S if for all $s_H > s_L$ both in S, $h(s_L) \geq 0$ implies $h(s_H) > 0$.

That is, if π is nonnegative for some other-agent participation rate ℓ, then it is positive for any higher rate.

These properties are illustrated in Figure 1. In the left panel, weak SC1 holds: negative values do not follow positive ones. However, SC1 fails since (a) negative values follow zeroes and (b) zeroes follow positive values. In the middle panel, SC1 holds but strict SC1 fails as zeroes follow zeroes. (Under strict SC1, only positive values can follow a zero.) In the third panel, all three properties hold.[9]

The strongest condition that we will assume is SC1. In contrast, the prior literature relies largely on strategic complementarities, which is stronger even than strict SC1.[10,11] Hence, our results constitute a substantial increase in flexibility.

[9] As the names suggest, strict SC1 implies SC1 which, in turn, implies weak SC1. Furthermore, while the curves in Figure 1 happen to be continuous, none of the three properties requires continuity.

[10] More precisely, most prior selection results have assumed *strict* strategic complementarities (i.e., that the payoff function $\pi(\cdot)$ is increasing). This property implies strict SC1. In the global games literature, it is more common to assume only strategic complementarities, which means "no *decreasing* segments" (Frankel, Morris, and Pauzner [24, p. 4]). Strategic complementarities implies SC1, but is orthogonal to strict SC1 as the latter allows decreasing segments but not segments in which the function is identically zero.

[11] The only exception is global games; see §1.

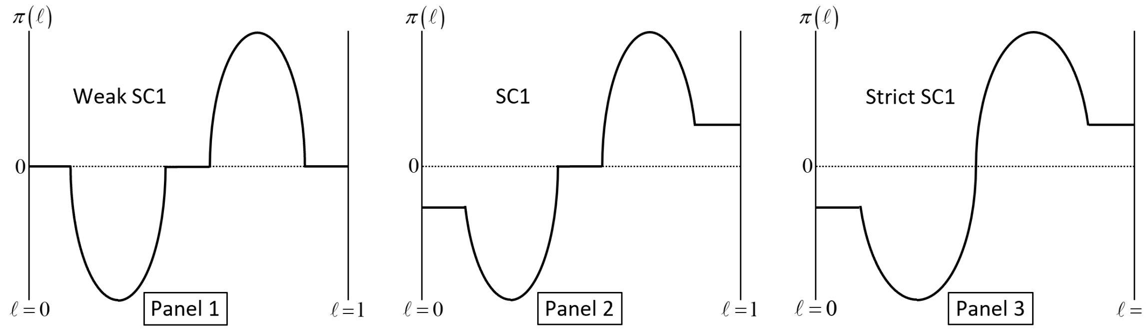

Figure 1: Sample payoff functions $\pi : [0, 1] \to \Re$ that satisfy the three types of single crossing properties defined by Athey [5] and used in this paper. While the functions shown are continous, the properties do not require this.

2.2 Selection Criteria and Theories

Let U be some set of utility functions $u : \{0, 1\} \times \Lambda \to \Re$. A *selection criterion* (or simply "criterion") on U is a partition of U into three sets. In one, all agents participate; in another, no agents participate; and in a third, the criterion makes no prediction.[12] More precisely:

Definition 4. A *selection criterion* Θ on U is a partition of U into three sets Θ_0, $\Theta_{1/2}$, and Θ_1 with the following interpretation.

- If u is in Θ_1 (resp., Θ_0), the criterion predicts that the agents will (not) participate.
- If u is in $\Theta_{1/2}$, no prediction is made.

We further define:

Definition 5. A *selection theory* is a rationale for a selection criterion.

For brevity, we will often refer to a selection theory as a "theory" and to a selection criterion as a "criterion".

[12] The five noncooperative theories select only equilibria in which all agents choose the same action. To ensure that the two heuristic theories share this property, we impose an additional assumption labeled "A0" in §3 below.

Remark 6. Selection criteria select among equilibria. Thus, if a criterion Θ predicts that the agents will all choose action $a \in \{0, 1\}$, then "all play a" must be a Nash equilibrium.

We study two types of theories. A *heuristic* theory is an informal argument that a criterion is reasonable. A *noncooperative* theory supports a criterion Θ by embedding the participation game in a more elaborate model. The model may be dynamic and may include one or more sources of randomness, such as shocks, noise, trembles, or mutations. It is then shown that the agents behave according to the criterion Θ in some limit - e.g. in the infinite horizon, as the agents become infinitely patient, or as the randomness shrinks to zero.

2.3 Preview of Results

We will show that seven iconic theories each implies a criterion in which an agent chooses a best response to some distribution Γ of other-agent participation rates ℓ. This distribution depends on the criterion but is common across agents and utility functions u. That is, each theory yields a criterion of the form

$$u \text{ is in } \left\{ \begin{array}{c} \Theta_0 \\ \Theta_{1/2} \\ \Theta_1 \end{array} \right\} \text{ as } \int_{\ell=0}^{1} \pi^u(\ell)\, d\Gamma(\ell) \left\{ \begin{array}{c} < \\ = \\ > \end{array} \right\} 0 \tag{2}$$

for a fixed distribution Γ that depends on the criterion but not on u.

We refer to the distribution Γ as the agents' *fictional beliefs*.[13] In general, it does

[13] While Kim [37] proves precursors of some of our results, the simple form (2) is not immediate in his work. Rather, Kim writes (2) as

$$\text{the agents (do not) participate if } \sum_{i=1}^{n} w_i \left[\pi_i^H - \pi_i^L\right] > (<)\, 0 \tag{3}$$

where the w_is are weights that depend on the criterion and π_i^a is the payoff from playing $a = L, H$ if the total number (including oneself) who play a is i. The relation of Kim's π_i^a to our utility function u is thus $\pi_i^H = u\left(1, \frac{i-1}{n-1}\right)$ and $\pi_i^L = u\left(0, \frac{n-i}{n-1}\right)$. Accordingly, unlike our (2), the summand in (3) evaluates u at two *different* other-agent participation rates. Nevertheless, they are not coincide with the true distribution of other-agent participation rates or with the agents' actual beliefs (if they have any) over this distribution.[14] Rather, it merely measures the relative importance that the theory assigns to different segments of the payoff function in predicting which equilibrium the agents will select.

3 The Two Heuristic Theories

In most applications of our tools, there is a principal who wishes to induce the agents to participate in an activity.[15] Our first heuristic theory, the Partial Implementation (PI) theory, is the most optimistic for the principal. It holds that the agents will choose the equilibrium that she prefers: if "all participate" is a strict Nash equilibria, the agents will select it. More precisely:[16]

The Partial Implementation (PI) theory is the argument that the agents will choose "all participate" if it is a strict Nash equilibrium.

We define an associated selection criterion:

The PI criterion is the criterion given by (2) for the fictional beliefs

equivalent. The reason is that for the selection criteria in question, the weights are symmetric: for each i, w_i equals w_{n+1-i}. Accordingly, (3) can be written as $\sum_{i=1}^{n} w_i \left[\pi_i^H - \pi_{n+1-i}^L\right]$ or, equivalently, as $\sum_{i=1}^{n} w_i \left[u\left(1, \frac{i-1}{n-1}\right) - u\left(0, \frac{i-1}{n-1}\right)\right]$, which equals (2) if we define $\Gamma(\ell)$ to be the partial sum $\sum_{i=0}^{\lfloor (n-1)\ell \rfloor} w_{i+1}$ of weights w_i.

[14]For generic payoffs π, the integral in (2) will be either positive or negative: either all agents will participate, or none will. The true distribution thus puts all of its weight on one or zero. While there are fictional beliefs with this property, most put positive weight on intermediate other-agent participation rates.

[15]In some cases, such as exchange rate attacks and bank runs, a principal prefers that the agents *not* participate. To apply our model to such activities, simply rename "participate" to "not participate" and vice-versa.

[16]If "all participate" is only a weak Nash equilibrium ($\pi^u(1) = 0$), the PI theory makes no prediction. Intuitively, what is reasonable in this case depends not only on the utility function u, but also on the wider context in which the participation game is played. For a discussion, see §9 of our online appendix.

$$\Gamma^{\mathrm{PI}}(\ell) \stackrel{d}{=} \begin{cases} 0 & \text{if } \ell < 1 \\ 1 & \text{if } \ell = 1. \end{cases} \tag{4}$$

That is, agents who rely on the PI criterion will play a best response to the belief that all others will participate.

Our other heuristic theory - the Full Implementation (FI) theory - is the most pessimistic theory for the principal.[17] It holds that the agents will choose the equilibrium that she least prefers: if "none participate" is a strict Nash equilibria, the agents will select it. More precisely:

The Full Implementation (FI) theory is a heuristic argument that states that the agents will choose "none participate" if it is a strict Nash equilibrium.[18]

We define an associated selection criterion:

The FI criterion is the criterion given by (2) for the fictional beliefs

$$\Gamma^{\mathrm{FI}}(\ell) \stackrel{d}{=} 1 \text{ for all } \ell. \tag{5}$$

That is, agents who rely on the FI criterion will play a best response to the belief that *no* others will participate.

The predictions of the two heuristic theories depend not only on the game's payoff function π, but also on the principal's preferred outcome: PI predicts this outcome while FI predicts its alternative. In contrast, the predictions of the five noncooperative theories depend only on the game's payoffs: the principal's preferences play no role. This will imply that the noncooperative theories reduce to risk-dominance in the case of two agents (§5.1).

[17] The "Partial" and "Full" implementation terminology is due to Bergemann and Morris [10, p. 1773].

[18] The FI theory has been used in participation games by Bernstein and Winter [11], Halac, Kremer, and Winter [28], Halac, Lipnowski, and Rappoport [29], Segal [45], and Winter [50]. The usual argument is not that the agents will choose the bad equilibrium, but rather that the principal wishes to rule this out.

For payoffs that satisfy weak SC1, the noncooperative theories predict only pure equilibria: if they make a prediction, it is that either all or none of the agents will participate. It is less clear what the heuristic theories should predict for payoffs that merely satisfy weak SC1. For instance, for any $c \in (0,1)$, the payoff function $\pi_c(\ell) \stackrel{d}{=} -1_{\ell>c}$ satisfies weak SC1 and any participation rate in $[0,c)$ is a stable Nash equilibrium.[19] The five noncooperative theories put positive weight on other-agent participation rates over c and thus predict that no agents will participate. It is less clear what the two heuristic theories would predict for π_c: they are usually applied to coordination games, in which mixed equilibria are unstable and thus ruled out. To pin down their predictions under weak SC1, and to ensure consistency with the noncooperative theories, we thus strengthen them by assuming that the agents will coordinate if possible:

A0 If all-participate or none-participate is a Nash equilibrium, the agents will not mix.

With this addition, the PI theory is equivalent to the PI criterion under weak SC1:

Claim 7. *Assume A0. If π satisfies weak SC1, the PI theory implies the PI criterion and vice-versa.*

Proof. If $\int_{\ell=0}^{1} \pi(\ell)\, d\Gamma^{\mathrm{PI}}(\ell) = \pi(1) > 0$ then the PI criterion predicts that the agents will all participate. Since, moreover, "all participate" is a strict Nash equilibrium, the PI criterion agrees with the PI theory. If $\pi(1) < 0$, the PI criterion predicts that the agents will not participate. This also agrees with the PI theory under A0. Why? First, "all participate" is not a Nash equilibrium in this case so the PI theory cannot predict it by Remark 6. And $\pi(0) \leq 0$ by weak SC1, so "none participate" *is* a Nash equilibrium whence, by A0, the agents will select it. Finally, if $\pi(1) = 0$ then both theory and criterion make no prediction. □

[19] This example assumes infinitesimal agents. The notation "1_{P}" denotes the indicator function, which equals one (zero) if statement P is true (false).

An analogous result holds for the FI theory and criterion:

Claim 8. *Assume A0. If π satisfies weak SC1, the FI theory implies the FI criterion and vice-versa.*

Proof. The proof follows that of Claim 7 *mutatis mutandis.* If $\int_{\ell=0}^{1} \pi(\ell)\, d\Gamma^{\mathrm{FI}}(\ell) = \pi(0) < 0$ then the FI criterion predicts that no agents will participate. Since, moreover, "none participate" is a strict Nash equilibrium, the FI criterion agrees with the FI theory. If $\pi(0) > 0$, the FI criterion predicts that the agents will participate. This also agrees with the FI theory combined with A0. Why? First, "none participate" is not even a weak Nash equilibrium in this case: the FI theory cannot predict it by Remark 6. Moreover, $\pi(1) \geq 0$ by weak SC1 so "all participate" is a Nash equilibrium: by A0, the agents must select it. Finally, if $\pi(0) = 0$, neither theory nor criterion makes any prediction. $\square$

4 The Five Noncooperative Theories

4.1 Preliminaries

For discrete agents, we assume only that the payoff function $\pi(\ell)$ satisfies weak SC1. This property is too weak to establish much on its own. However, the five noncooperative theories each involve some sort of random matching. Hence an agent cares not about her realized payoff $\pi(\ell)$ *per se*, but rather about its expectation $\gamma(z) = E[\pi(\ell)|z]$ where z is the participation rate in the whole population. The key result, which extends a claim in Athey [6], is that if π satisfies weak SC1 in ℓ then γ satisfies SC1 and is Lipschitz-continuous in z. These properties of γ suffice to obtain selection results for the five theories.

The details are as follows. Assume an agent plays the participation game against $n-1$ opponents, each of whom has an independent probability $z \in [0,1]$ of participating. The probability that exactly $i \in \{0, 1, ..., n-1\}$ others will participate is simply the binomial density:

$$b(i; n-1, z) \overset{d}{=} \begin{pmatrix} n-1 \\ i \end{pmatrix} z^i (1-z)^{n-1-i}, \tag{6}$$

which is the probability of i successes in $n-1$ independent trials each with success probability z. The agent's expected payoff $\gamma(z)$ is the sum, over all i from zero to $n-1$, of her realized payoff $\pi\left(\frac{i}{n-1}\right)$ if i others participate weighted by this probability:

$$\gamma(z) \overset{d}{=} \sum_{i=0}^{n-1} b(i; n-1, z)\, \pi\left(\frac{i}{n-1}\right). \tag{7}$$

We can write (7) more compactly as an integral, suppressing n:

$$\gamma(z) = \int_{\ell \in \lambda} \pi(\ell)\, \kappa(\ell; z)\, d\mu(\ell) \tag{8}$$

where the measure

$$d\mu(\ell) \overset{d}{=} 1_{\ell \in \lambda} \tag{9}$$

assigns unit (resp., zero) weight to rates ℓ (not) in λ and

$$\kappa(\ell; z) \overset{d}{=} b((n-1)\ell; n-1, z) \tag{10}$$

is the probability of $(n-1)\ell$ successes in $n-1$ trials each with success probability z.

The following claim, which extends a result in Athey [6], states that if the realized payoff function π satisfies weak SC1 in ℓ, then the expected payoff function γ defined in (7) is Lipschitz-continuous and satisfies SC1 in z.[20] The proof appears at the end of this section.

Claim 9. *If π satisfies weak SC1 on λ, then:*

1. *γ is Lipschitz-continuous on $[0,1]$ and satisfies $\gamma(z) = \pi(z)$ for $z = 0, 1$;*

2. *γ satisfies SC1 on $(0,1)$, and weak SC1 on $[0,1]$.*

Figure 2 shows a calculated example. A payoff function $\pi(\ell)$ - the collection of blue dots - has been chosen that satisfies only weak SC1. The associated expected

[20] Omitted proofs are in our online appendix (Frankel [23]).

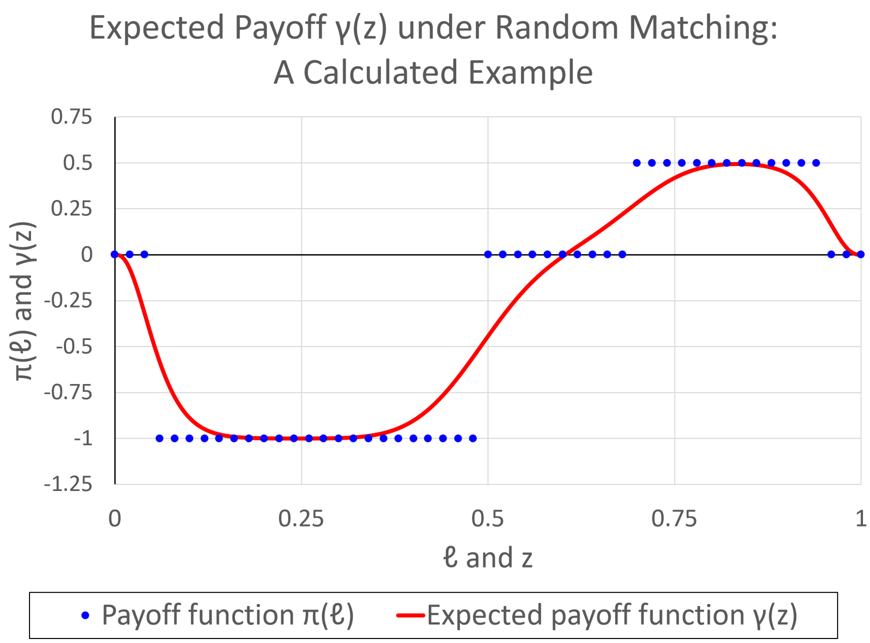

Figure 2: Illustration of Claim 9 for the case of $n-1=50$ opponents. The assumed payoff function π is defined on the discrete set λ of other-agent participation rates ℓ. The associated expected payoff function γ, the red curve, is defined on the full set $[0,1]$ of population participation rates z.

payoff function γ - the red curve - is continuous, satisfies SC1 on $(0,1)$ and weak SC1 on $[0,1]$, and coincides with π at the endpoints.[21]

When agents are infinitesimal, each agent plays against the entire population: we cannot rely on random matching to pass to a smoother payoff function. Thus, we will usually require π itself to satisfy the properties of γ that are derived in Claim 9.

Proof of Claim 9. Part 1. With a success chance of zero (one), none (all) of $n-1$ trials must succeed so $\gamma(z) = E[\pi(\ell)|z]$ equals $\pi(z)$ for $z = 0,1$. As for Lipschitz continuity, let $z_0 < z_1$ each lie in $[0,1]$. By (6), (7), and the triangle inequality,

$$|\gamma(z_1) - \gamma(z_0)| \leq c_0 \sum_{i=0}^{n-1} \left| \begin{array}{c} z_1^i (1-z_1)^{n-1-i} \\ -z_0^i (1-z_0)^{n-1-i} \end{array} \right| . \tag{11}$$

[21]In this example, γ satisfies strict SC1 on $(0,1)$. However, this is not so if π is identically zero. Thus, we show only SC1 in Claim 9 as it suffices for our purposes.

where c_0 is the finite constant $\max_{i=0,...,n-1} \binom{n-1}{i} \max_{\ell\in\lambda} |\pi(\ell)|$. By the triangle inequality,

$$\begin{aligned} \left| \begin{array}{c} z_1^i (1-z_1)^{n-1-i} \\ -z_0^i (1-z_0)^{n-1-i} \end{array} \right| &\leq \left| z_1^i - z_0^i \right| (1-z_1)^{n-1-i} + z_0^i \left| \begin{array}{c} (1-z_1)^{n-1-i} \\ -(1-z_0)^{n-1-i} \end{array} \right| \\ &\leq \left| z_1^i - z_0^i \right| + \left| (1-z_1)^{n-1-i} - (1-z_0)^{n-1-i} \right|. \end{aligned} \tag{12}$$

For $i = 0$, $\left| z_1^i - z_0^i \right|$ is zero .[22] For $i > 0$, the slope of z^i is iz^{i-1} which, since $z \in [0,1]$, lies in $[0,i]$. Thus, the slope of z^i over all $i = 0,...,n-1$ lies in $[0, n-1]$. Similarly, the slope of $(1-z)^{n-1-i}$ lies in $[-(n-1), 0]$ for $i = 0,...,n-1$. Thus, by (11) and (12), γ is Lipschitz-continuous with Lipschitz constant $2c_0 n(n-1)$.

Part 2. The following definition and lemma are from Athey [6].

Definition 10. (Athey [6, Def. 3 (p. 191) & n. 17 (p. 195)]) Let $Z, S \subseteq \Re$ and let μ be a measure on S. A nonnegative function $h : S \times Z \to \Re$ is *log-supermodular* a.e.-μ if, for every pair $z_H > z_L$ each in Z, $h(s_H, z_H) h(s_L, z_L) \geq h(s_H, z_L) h(s_L, z_H)$ for μ-a.e. pair $s_H > s_L$ each in S.[23]

Lemma 11. *(Athey [6, Extension (iii) to Lemma 5, p. 201]) Let $Z, S \subseteq \Re$ and let μ be a measure on S. Let $g : S \to \Re$ and $K : S \times Z \to \Re$. Assume that (A) $supp(K(\cdot, z))$ is constant in $z \in Z$. If (H1) g satisfies weak SC1 on S and (H2) K is log-supermodular a.e.-μ, then (C) $G(z) = \int_{s\in S} g(s) K(s,z) d\mu(s)$ satisfies SC1 on Z.*

Now interpret (Z, S, g, K) as $((0,1), \lambda, \pi, \kappa)$ and define μ as in (9), whence γ corresponds to G. We now verify the conditions of the Lemma. Condition A holds as $\kappa(\ell, z) > 0$ for all $\ell \in \lambda$ and $z \in (0,1)$ by (6) and (10). H1 holds by assumption:

[22] For all $z \in [0,1]$, $z^0 = 1$, *including* $z = 0, 1$. This ensures that $b(i; n-1, 0)$ and $b(i; n-1, 1)$ take the correct values.

[23] A measure μ on S is extended to a product measure on S^2 in the usual way: for any $\sigma \subseteq S^2$ we let $\mu(\sigma) = \int_{x\in S} \mu(\{x' \in S : (x, x') \in \sigma\}) d\mu(x)$ whenever the integral exists. (If it does not, the set σ is not measurable under μ.)

π satisfies weak SC1 on λ. As for H2, let $z_L, z_H \in (0,1)$ and $\ell_L, \ell_H \in \lambda$ satisfy

$$z_H > z_L \text{ and } \ell_H > \ell_L.$$

Since $\kappa(\ell, z_H) > 0$ for $\ell \in \{\ell_L, \ell_H\}$, we need only verify $\frac{\kappa(\ell_L, z_L)}{\kappa(\ell_L, z_H)} \geq \frac{\kappa(\ell_H, z_L)}{\kappa(\ell_H, z_H)}$ or, equivalently, by (6) and (10), $\left(\frac{z_H}{z_L}\frac{1-z_L}{1-z_H}\right)^{s_H - s_L} \geq 1$ which holds by hypothesis. Lemma 11 now implies that γ satisfies SC1 on $(0,1)$ as claimed. Finally, γ satisfies weak SC1 on $[0,1]$ by the following result, interpreting h as γ.

Lemma 12. *Let $h : [0,1] \to \Re$ be continuous on $[0,1]$ and satisfy SC1 on $(0,1)$. Then h satisfies weak SC1 on $[0,1]$.*

Proof. Let $z_H > z_L$ both lie in $[0,1]$. Let $\varepsilon = (z_H - z_L)/2$. If $h(z_L) > 0$ then there is an $\iota \in (0, \varepsilon)$ such that for all $z \in (z_L, z_L + \iota)$ and all $z' \in (z_H - \iota, z_H)$, $h(z) > 0$ by continuity of h and thus $h(z') > 0$ by SC1 (since $z < z'$ and both lie in $(0,1)$). Hence, by continuity of h, $h(z_H) \geq 0$ as required. □

Q.E.D.$_{\textbf{Claim 9}}$

4.2 Dynamic Theories

We now present four dynamic theories. In the discrete case, the participation game is played by groups of $n \geq 2$ agents who are drawn at random from a much larger population.[24] Our discussion in §4.1 then implies:

Claim 13. *In games in which groups of n agents are randomly matched from a much larger population to play the participation game, an agent's expected static payoff from participating is given by $\gamma(z)$ defined in (7) where $z \in [0,1]$ is the participation rate in the population.*

[24] This population is finite in KMR and a continuum in the other three theories.

4.2.1 The Kandori-Mailath-Rob (KMR) Theory

The KMR theory is an evolutionary, discrete-time model in which random groups of n boundedly rational agents are selected, in each period, to play the participation game.[25] The action with the higher static payoff in a given period is chosen by more agents in the next. There are trembles: each agent has a small chance of choosing the suboptimal action.

More precisely, we study the following generalization of the evolutionary model of Kandori, Mailath, and Rob [35].[26,27] The population consists of a large finite number N of agents. In each period $t = 0, 1, ...$, random groups of $n \geq 2$ agents are selected to play the participation game. By Claim 13, an agent's expected payoff from participating is thus $\gamma(z)$ where $z \in [0, 1]$ is the participation rate in the population. The model we study is in the following class.

Definition 14. (Ellison [18]; Young [52]) A Model of Evolution with Noise (MEN) is a triple $(Z, P, P(\varepsilon))$ consisting of a state space Z (a finite set); a Markov transition matrix P on Z that gives the transition probabilities in the absence of mutations; and a family of perturbed Markov transition matrices $P(\varepsilon)$ for each tremble probability $\varepsilon \in [0, \bar{\varepsilon}]$ such that (a) $P(\varepsilon)$ is ergodic[28] for all ε and (b) $P(\varepsilon)$ is continuous in ε with $P(0) = P$.

In KMR [35] without trembles, the agents play a static best response to the prior period's population participation rate z. This is a special case of the following class of deterministic dynamics, which we use instead.

[25] The KMR theory is defined only for discrete agents.

[26] KMR [35] prove their result for two-player, two-action coordination games. Kim [37] generalizes this result to coordination games with $n \geq 2$ players. We replace Kim's coordination game assumption with weak SC1 and relax KMR's [35] assumption on dynamics.

[27] While KMR [35] is often grouped with Young [52], the models are not equivalent (Jacobsen, Jensen, and Sloth [33]).

[28] A Markov chain is ergodic if every state is eventually reached from every other state with positive probability.

Definition 15. (KMR [35]) The deterministic dynamic P in Def. 14 is *Darwinian* with respect to the relative payoff function γ if the following conditions hold.

1. If one action has a higher payoff under γ given the current participation rate z, the proportion playing that action rises in the next period (unless everyone is already playing that action, in which case they keep doing so).

2. If the two actions give the same payoff under γ given the current participation rate z, then the proportion playing a given action either (i) stays the same in the next period or (ii) could rise or fall, but does not move solely in one direction.

As in KMR [35], we assume further that the perturbed matrix $P(\varepsilon)$ is generated by independent random trembles (IRT): each agent plays according to the deterministic dynamic with probability $1-\varepsilon$ and randomizes 50-50 over the two actions with probability ε. Let a MENI be a MEN whose noise is due to IRT trembles. For any MENI $(Z, P, P(\varepsilon))$ and any $\varepsilon > 0$, the perturbed model has a unique limiting distribution μ^{ε} over Z, which is invariant to initial conditions (Ellison [18, sec. 2.2, p. 21]). We focus on the limit of this distribution as the trembles go to zero:

Definition 16. (Foster and Young [19]; Young [52]) The *long-run stochastically stable set* is the set of states z for which $\mu^*(z) > 0$ where $\mu^* \stackrel{d}{=} \lim_{\varepsilon \downarrow 0} \mu^{\varepsilon}$.

To the KMR theory we define an associated criterion:

The KMR criterion is the criterion given in (2) for the fictional beliefs

$$\Gamma_n^{\mathrm{KMR}}(\ell) \stackrel{d}{=} \sum_{i=0}^{\lfloor (n-1)\ell \rfloor} W_{i,n}^{\mathrm{KMR}} \quad \text{where} \quad W_{i,n}^{\mathrm{KMR}} \stackrel{d}{=} \begin{pmatrix} n-1 \\ i \end{pmatrix} \left(\frac{1}{2}\right)^{n-1}. \tag{13}$$

In the limit as the number n of agents grows, Γ_n^{KMR} puts all of its weight on 50% of the other agents participating:

Claim 17. *In the limit as $n \to \infty$, $\Gamma_n^{KMR}(\ell)$ converges to*

$$\Gamma_\infty^{KMR}(\ell) \stackrel{d}{=} \begin{cases} 0 & \text{if} \quad \ell < 1/2 \\ 1 & \text{if} \quad \ell \geq 1/2. \end{cases} \tag{14}$$

Proof. Online appendix (Frankel [23]). □

In this limit, the KMR criterion selects the equilibrium with the larger basin of attraction - or, equivalently, the action that is a best response to an other-agent participation rate of one half.

If π satisfies weak SC1 then, in any MENI with Darwinian dynamics, the agents' participation decision is given by (2) for the fictional beliefs Γ_n^{KMR} if the population size N is large enough:[29]

Theorem 18. *Let π satisfy weak SC1 on λ and let γ be the expected relative payoff function defined in (7). For each aggregate population size N, let $\left(Z^N, P^N, P^N(\varepsilon)\right)$ be a MENI whose deterministic dynamic P^N is Darwinian with respect to γ. Then if $\int_{\ell=0}^{1} \pi(\ell)\, d\Gamma_n^{KMR}(\ell)$ is positive (resp., negative), there is an $N^* < \infty$ such that for all $N > N^*$, the only element of the long-run stochastically stable set of $\left(Z^N, P^N, P^N(\varepsilon)\right)$ is $z = 1$ (resp., $z = 0$).*

Proof. Online appendix. □

An intuition is as follows. Let $c \approx 0.6$ be the point in Figure 2 at which γ crosses zero from below. The basin of attraction of the endpoint $z = 0$ is $[0, c)$ and that of $z = 1$ is $(c, 1]$. If the state is in an endpoint's basin of attraction, then the Darwinian dynamics will push it towards that endpoint in the next period. Thus, as the trembles vanish, the state will spend nearly all its time *at* one or the other of the two endpoints.

Yet transitions between the two endpoints will sometimes occur as long as the tremble probability ε is positive. This is because trembles are independent, so there exists a positive possibility that enough agents will *simultaneously* tremble in one period to move the state to the other endpoint's basin of attraction. Darwinian dynamics will then push the state towards the other endpoint.

[29] Kim [37, Prop. 2, p. 211] shows this result for KMR when payoffs satisfy strategic complementarities. We generalize Kim's result to all MENIs with Darwinian dynamics in which π satisfies weak SC1. The proof shows that the size of an action's basin of attraction is determined by the sign of γ at the midpoint $z = 1/2$ and then applies Ellison's [18] radius-coradius theorem.

As the trembles vanish, the endpoint with the larger basin of attraction becomes infinitely harder to leave relative to the other endpoint. This is because (a) more trembles are needed to leave it and (b) the chance of i simultaneous trembles is proportional to ε^i, so the ratio of probabilities of i' vs. i trembles is proportional $\varepsilon^{i'-i}$ which goes to zero if $i' > i$. Thus, the system spends nearly all of its time in the endpoint with the larger basin of attraction.

Moreover, one can easily see from Figure 2 that zero (one) has a larger basin if $\gamma(1/2)$ is negative (positive). And since, by (8), $\gamma(1/2)$ equals $\int_{\ell=0}^{1} \pi(\ell)\, d\Gamma_n^{\mathrm{KMR}}(\ell)$, the agents behave according to (2) with fictional beliefs Γ_n^{KMR}.[30] Finally, as $n \to \infty$, the realized proportion ℓ of an agent's opponents who participate converges in probability to the population participation rate z by the law of large numbers. Thus, roughly speaking, $\gamma(1/2)$ converges to $\pi(1/2)$: the fictional beliefs converge to a step function at $\ell = 1/2$. This completes the intuition.

4.2.2 The FY Theory

Foster and Young [19] (FY) study an evolutionary, continuous-time model with infinitesimal agents.[31] At each time $t \in \Re_+$, random groups of n boundedly rational agents are selected to play the participation game. Let η_t (resp., ν_t) denote the measure of agents who (do not) participate at time t and let $z_t = \frac{\eta_t}{\eta_t + \nu_t}$ denote the proportion who participate. The the growth rate of each population is assumed to equal its expected payoff in the participation game:

$$d\eta = \eta\gamma_1(z)\, dt \quad \text{and} \quad d\nu = \nu\gamma_0(z)\, dt \tag{15}$$

[30] Intuitively, if $n-1$ agents are chosen at random from an infinite population that is evenly split between the two actions ($z = 1/2$) then the probability that exactly i of them participate is just the binomial coefficient $W_{i,n}^{\mathrm{KMR}}$ defined in (13). Hence, "participate" has a larger basin of attraction if an agent would choose it under the belief that, for each $i = 0, ..., n-1$, the probability is $W_{i,n}^{\mathrm{KMR}}$ that exactly i of her $n-1$ opponents participate.

[31] FY [19] focus on two-player, two-action coordination games. Kim [37] generalizes their result to coordination games with $n \geq 2$ players, while we replace Kim's coordination game assumption with weak SC1.

where

$$\gamma_a(z) \stackrel{d}{=} \int_{\ell=0}^{1} u(a,\ell)\,\kappa(\ell;z)\,d\mu(\ell) \tag{16}$$

is an agent's gross expected utility from playing action $a = 0, 1$.[32] Solving (15) for dz, we obtain the deterministic replicator dynamic: $dz = z(1-z)\gamma(z)\,dt$.[33] To this, FY add Brownian shocks:

$$dz = z(1-z)\gamma(z)\,dt + \sigma dw, \tag{17}$$

where $\sigma > 0$ is a constant and w is a Brownian motion.[34] FY study the long-run distribution of this model in the limit as the shocks dw shrink to zero. They show that the model has an ergodic distribution, in which the agents select an equilibrium according to the following criterion:

The FY criterion is the criterion given by (2) for the fictional beliefs

$$\Gamma_n^{\mathrm{FY}}(\ell) \stackrel{d}{=} \sum_{i=0}^{\lfloor (n-1)\ell \rfloor} W_{i,n}^{\mathrm{FY}} \quad \text{where} \quad W_{i,n}^{\mathrm{FY}} \stackrel{d}{=} \frac{6(i+1)(n-i)}{n(n+1)(n+2)} \tag{18}$$

in the discrete case and $\Gamma_\infty^{\mathrm{FY}}(\ell) \stackrel{d}{=} 3\ell^2 - 2\ell^3$ in the infinitesimal case.

Theorem 19. *Assume π satisfies weak SC1 on λ. Then for $\sigma > 0$, equation (17) has an ergodic distribution. In the long run, in the limit as σ shrinks to zero, the state z spends nearly all of its time in a small neighborhood of one (resp., zero) if $\int_{\ell=0}^{1} \pi(\ell)\,d\Gamma_n^{FY}(\ell)$ is positive (resp., negative).*

Proof. Online appendix. □

[32] The formula for $\gamma_a(z)$ is obtained by replacing $\pi(\ell)$ in (8) by $u(a,\ell)$.

[33] To see this, let $\phi(\eta,\nu) = \frac{\eta}{\eta+\nu}$ whence $dz = d\phi(\eta,\nu) = \phi_1 d\eta + \phi_2 d\nu$ where subscripts indicate partial derivatives. Thus, since $\phi_1 = \frac{\nu}{(\eta+\nu)^2}$ and $\phi_2 = -\frac{\eta}{(\eta+\nu)^2}$, dz equals $\frac{\nu}{(\eta+\nu)^2}d\eta - \frac{\eta}{(\eta+\nu)^2}d\nu$ which can be rearranged to $z(1-z)\left(\frac{d\eta}{\eta} - \frac{d\nu}{\nu}\right) = z(1-z)\gamma(z)\,dt$ as $\gamma(z) = \gamma_1(z) - \gamma_0(z)$ by (1) and (16).

[34] FY assume that z reflects at the boundary of $[\Delta, 1-\Delta]$ for small $\Delta > 0$. Relying on a different mathematical result, we let z reflect at the boundaries of $[0,1]$. This does not affect the results: Kim's [37] generalization of FY to n-player coordination games, which also restricts $[\Delta, 1-\Delta]$, finds the same selection criterion as our Theorem 19.

Theorem 19 assumes that each agent interacts with a random set of $n-1$ other agents. It thus corresponds to the discrete case. To address the infinitesimal case, we assume now that an agent interacts with the whole population: her flow payoff is $\pi(z_t)$ rather than $\gamma(z_t)$. Equation (17) is thus replaced with

$$dz = z(1-z)\,\pi(z)\,dz + \sigma dw. \tag{19}$$

As this is the only change, a result analogous to Theorem 19 holds if we assume that π has the properties of γ that we use in the proof of that result:

Corollary 20. *Let $\pi : [0,1] \to \Re$ be Lipschitz-continuous on $[0,1]$ and satisfy SC1 on $(0,1)$. Then for $\sigma > 0$, equation (19) has an ergodic distribution. In the long run, in the limit as σ shrinks to zero, the state z spends nearly all of its time in a small neighborhood of one (resp., zero) if $\int_{\ell=0}^{1} \pi(\ell)\,d\Gamma_{\infty}^{FY}(\ell)$ is positive (resp., negative).*

Proof. Online appendix. □

Moreover, the discrete criterion converges to the continuous one as n grows:

Claim 21. *For all $\ell \in [0,1]$, $\lim_{n\to\infty} \Gamma_n^{FY}(\ell) = \Gamma_\infty^{FY}(\ell)$.*

Proof. Online appendix. □

We defer an intuition for these results to §4.2.4.

4.2.3 The FH Theory

While FY add shocks directly to the state z, Fudenberg and Harris [25] (FH) instead add shocks to the sizes of the populations playing the two actions. They study a sequence of variants of this two-player random matching model until they obtain an ergodic distribution. We extend this final variant to the case of $n \geq 2$ players.[35]

[35]The FH model has been solved only in the case of two players. This contrasts with the other noncooperative theories, which Kim [37] solves for $n \geq 2$ players under the assumption of strategic complementarities. Hence, our technical contribution is largest in the case of FH.

In their final variant, FH assume that

$$d\eta = \eta\left[\gamma_1(z)\,dt + \sigma_\eta dw_\eta\right] + \left(\lambda_\nu \nu - \lambda_\eta \eta\right) dt \tag{20}$$

and

$$d\nu = \nu\left[\gamma_0(z)\,dt + \sigma_\nu dw_\nu\right] + \left(\lambda_\eta \eta - \lambda_\nu \nu\right) dt \tag{21}$$

where w_η and w_ν are independent Brownian motions.[36] This model differs from FY in two ways. First, there are shocks (dw_η and dw_ν) to the masses (η and ν, resp.) playing each action rather than to the proportion z playing action one. FH [25, Prop. 2] show that such shocks alone do not give rise to an ergodic distribution. Intuitively, the implied shocks to z are of order $z(1-z)$ which vanishes as the state z approaches zero or one. Hence, the two endpoints of the state space $[0,1]$ are absorbing. To avoid this, FH also add individual random trembles: each agent playing action one (zero) exogenously switches to action zero (one) according to a Poisson process with arrival rate λ_η (λ_ν). These trembles keep the participation rate z away from the endpoints, thus yielding an ergodic distribution.

Define the constant

$$\sigma \overset{d}{=} \sqrt{\sigma_\eta^2 + \sigma_\nu^2}. \tag{22}$$

Let w denote $\frac{\sigma_\eta}{\sigma} w_\eta - \frac{\sigma_\nu}{\sigma} w_\nu$, which is a standard Brownian motion.[37] By (20) and (21) and Ito's Lemma, the state $z = \frac{\eta}{\eta+\nu}$ changes according to[38]

$$dz = \alpha(z)\,dt + \beta(z)\,dw \tag{25}$$

[36] The functions γ_i for $i = 0, 1$ are defined in (16).

[37] As w_η and w_ν have continuous paths and independent increments, so does w. And since w_η and w_ν are independent Brownian motions, $\sigma dw = \sigma_\eta dw_\eta - \sigma_\nu dw_\nu$ is normal with mean zero and variance σ^2. Thus, w is a standard Brownian motion.

[38] To see this, define $\phi(\eta, \nu) = \frac{\eta}{\eta+\nu}$ which equals z. By Ito's Lemma in two dimensions,

$$dz = \phi_1 d\eta + \phi_2 d\nu + \frac{1}{2}\left[(d\eta)^2 \phi_{11} + 2(d\eta)(d\nu)\phi_{12} + (d\nu)^2 \phi_{22}\right], \tag{23}$$

where numbered subscripts denote partial derivatives. In the present setting, $(d\eta)^2 = (\sigma_\eta \eta)^2 dt$, $(d\nu)^2 = (\sigma_\nu \nu)^2 dt$, $(d\eta)(d\nu) = 0$, $\phi_1 = \frac{\nu}{(\eta+\nu)^2}$, $\phi_2 = -\frac{\eta}{(\eta+\nu)^2}$, $\phi_{11} = -\frac{2\nu}{(\eta+\nu)^3}$, and $\phi_{22} = \frac{2\eta}{(\eta+\nu)^3}$ and with coefficients

$$\alpha(z) = \; z(1-z)\left[\gamma(z) + (1-z)\sigma_\nu^2 - z\sigma_\eta^2\right] + \lambda_\nu(1-z) - \lambda_\eta z \tag{26}$$

and

$$\beta(z) = z(1-z)\sigma. \tag{27}$$

We assume that σ_η^2, σ_ν^2, λ_η, and λ_ν are all constant multiples of a common parameter, which we take to be σ^2 to save notation:

$$\sigma_\eta^2 = s_\eta\sigma^2, \;\; \sigma_\nu^2 = s_\nu\sigma^2, \;\; \lambda_\eta = l_\eta\sigma^2, \text{ and } \lambda_\nu = l_\nu\sigma^2 \tag{28}$$

where $s_\eta, s_\nu, l_\eta, l_\nu \in \Re_{++}$ are fixed.[39]

To the FH theory we define a corresponding criterion:

The FH criterion is the criterion given by (2) for the fictional beliefs

$$\Gamma^{\mathrm{FH}}(\ell) \overset{d}{=} \begin{cases} 1/2 & \text{if } \ell < 1 \\ 1 & \text{if } \ell = 1. \end{cases}. \tag{29}$$

That is, agents who rely on the FH criterion will play a best response to the belief that either no others or all others will participate, with equal probabilities.

In the FH model, in the long run in the limit as the noise parameter σ shrinks to zero, the agents play according to the FH criterion. More precisely, let Q be the long-run distribution of the state: for any $y \in [0,1]$, $Q(y) \overset{d}{=} \lim_{t\to\infty} \Pr(z_t \le y)$. For any $y \in (0,1)$, define the two quantities

$$\chi_1^y \overset{d}{=} \frac{Q(y)}{1-Q(y)} \text{ and } \chi_2^y \overset{d}{=} \frac{1-Q(1-y)}{Q(1-y)}. \tag{30}$$

---

thus, using $\gamma(z) = \gamma_1(\eta) - \gamma_0(z)$,

$$dz = z(1-z)\left(\frac{d\eta}{\eta} - \frac{d\nu}{\nu}\right) + z(1-z)\left[(1-z)\sigma_\nu^2 - z\sigma_\eta^2\right]dt. \tag{24}$$

We then use (20) and (21) to substitute for $d\eta$ and $d\nu$ in equation (24) to obtain (25). One can easily see that (25) is equivalent to FH's equation (8).

[39]The parameters s_η and s_ν in (28) must satisfy $s_\eta + s_\nu = 1$ as $\sigma_\eta^2 + \sigma_\nu^2 = (s_\eta + s_\nu)\sigma^2$ equals σ^2 by (22).

Intuitively, χ_1^y (resp., χ_2^y) gives the *odds* that the state lies in $[0, y]$ (resp., in $(1-y, 1]$). The formal result is as follows.

Theorem 22. *Assume either that (a) agents are discrete and π satisfies weak SC1 or (b) agents are infinitesimal and π is Lipschitz-continuous on $[0,1]$ and satisfies SC1 on $(0,1)$. Then for any $y \in (0,1)$, if $\int_{\ell=0}^{1} \pi(\ell)\, d\Gamma^{FH}(\ell)$ is negative (resp., positive) then $\lim_{\sigma \downarrow 0} \chi_1^y = \infty$ (resp., $\lim_{\sigma \downarrow 0} \chi_2^y = \infty$).*

Proof. Online appendix. □

4.2.4 FY vs. FH: Discussion

The FY and FH theories both build on the replicator dynamic but lead to starkly different criteria. The limiting fictional density $d\Gamma_\infty^{\mathrm{FY}}(\ell)/d\ell$ in the FY criterion is continuous and proportional to $\ell(1-\ell)$. In contrast, the fictional density in the FH criterion is discrete with weight 1/2 on each endpoint. Why?

In both FY and FH, the evolution of the state z is given by $dz = \alpha(z)\, dt + \beta(z)\, dw$ where w is a Brownian motion, $\beta(z)$ is the diffusion term, and $\alpha(z)$ is the drift which equals $\gamma(z)\, z(1-z)$. To transition from zero to one, the state z must pass through every value in $(0,1)$. Positive drift raises the chance of such a transition and negative drift lowers it. However, the shocks act as a foil: the larger is the diffusion term $\beta(z)$ at a state z, the less the drift matters at that state.[40] In particular, the drift's contribution to the transition probability is proportional to the ratio $\alpha(z)/\beta^2(z)$ of the drift to the squared diffusion term.[41]

In FY, the diffusion term $\beta(z)$ is a constant so the drift at each state has equal weight. Thus, the probability of a transition from zero to one is proportional to the mean of the drift $\gamma(z)\, z(1-z)$ over all states z. In the limit as $n \to \infty$, ℓ converges

[40] Similarly, Beggs [8] finds that an equilibrium is more likely to be selected if the shocks are smaller in its basin of attraction. More generally, Bergin and Lipman [9] and Binmore and Samuelson [12] show that equilibrium selection depends on the sizes of the shocks at different states.

[41] This is seen in Lemmas 45 and 46 in §11 of the online appendix.

in probability to z so the mean drift converges to the integral of $\pi(\ell)\,\ell(1-\ell)$ over all $\ell \in [0,1]$. This is why the fictional density is proportional to $\ell(1-\ell)$.

In FH, in contrast, $\beta(z)$ goes to zero as z approaches either endpoint at the same rate. This leads to the weights on the drift $\alpha(z)$ at the two endpoints to grow without bound, while remaining equal to one another. As a result, the agents play a best response to the mean of the payoffs at the two endpoints.[42]

4.2.5 The Matsui-Matsuyama (MM) Theory

In the MM theory of Matsui and Matsuyama [39], the agents are rational and forward looking.[43] At each time $t \in \Re_+$, random groups of n agents are selected from the unit interval $[0,1]$ to play the participation game. Hence, as in FY and FH, an agent's flow payoff from participating is $\gamma(z)$ where z is the population participation rate.

In MM, each agent gets chances to switch actions according to an independent Poisson process with arrival rate $p > 0$. The rate of change of the state z thus satisfies $\frac{dz_t}{dt} \in [-pz_t, p(1-z_t)]$ so for any given initial state $z_0 = \zeta \in [0,1]$, the path $z_\cdot = (z_t)_{t\geq 0}$ satisfies $z_t \in \left[\underline{z}_t^{\zeta}, \overline{z}_t^{\zeta}\right]$ where the lower and upper bounds are given by

$$\underline{z}_t^{\zeta} = \zeta e^{-pt} \quad \text{and} \quad \overline{z}_t^{\zeta} = 1 - (1-\zeta)\, e^{-pt}, \tag{31}$$

respectively. Let $r > 0$ be the rate of time preference. As agents have perfect foresight, the benefit of switching from "not participate" to "participate" at time t, if an agent expects the path $(z_s)_{s\geq 0}$ to be played, is proportional to

$$V_t^r(z_\cdot) \stackrel{d}{=} (p+r) \int_{s=t}^{\infty} e^{-(p+r)(s-t)} \gamma(z_s)\, ds \tag{32}$$

---

[42] Blume [13] studies a discrete-time model in which groups of $n \geq 2$ players are randomly matched to play a two-action coordination game. Relative to our version of KMR, Blume makes stronger assumptions on payoffs but imposes weaker conditions on noise. He shows that either the FH criterion [13, Thms. 3 & 4] or the KMR criterion [13, Thm. 5] may emerge, depending on the form of the noise.

[43] MM [39] focus on two-player games. Kim [37, Proposition 1] generalizes their result to coordination games with $n \geq 2$ players and two actions. We replace Kim's coordination game assumption with weak SC1.

since, for each time $s \geq t$, the expected flow payoff from participating is $\gamma(z_s)$ and the chance that the agent will not receive another revision opportunity *before* time s is $e^{-p(s-t)}$. An agent with an action revision opportunity at time t will thus choose (not to) participate if (32) is positive (resp., negative). One can easily calculate that

$$\text{if } z_t = z \text{ for all } t \geq 0, \text{ then } V_t^r(z_\cdot) = \gamma(z). \tag{33}$$

MM define the following concepts.

Definition 23. A state $z \in [0,1]$ is *accessible from* $z' \in [0,1]$ if there exists an equilibrium path from z' that reaches or converges to z. The state z is *globally accessible* if it is accessible from any $z' \in [0,1]$.

Definition 24. A state $z \in [0,1]$ is *absorbing* if there is a neighborhood of z such that any equilibrium path originating in this neighborhood must converge to z. It is *fragile* if it is not absorbing.

To the MM theory we define an associated selection criterion:[44],[45]

The Laplace criterion is the criterion given in (2) for the fictional beliefs

$$\Gamma_n^{\text{Laplace}}(\ell) \stackrel{d}{=} \sum_{i=0}^{\lfloor (n-1)\ell \rfloor} \frac{1}{n} \tag{34}$$

in the discrete case and

$$\Gamma_\infty^{\text{Laplace}}(\ell) \stackrel{d}{=} \ell \tag{35}$$

in the infinitesimal case.

Agents who rely on the Laplace criterion play a best response to the belief that all other-agent participation rates are equally likely.

[44] Kim [37] first identified the Laplace criterion and showed that it is implied by the GG and MM theories in coordination games. An intuition for the former result appears in MS [41, pp. 61-63].

[45] The role of the floor function in (34) is to ensure that the beliefs are defined for all $\ell \in [0,1]$ and thus that the integral in (2) is well-defined.

Theorem 25. *Let* $\pi : \lambda \to \Re$ *satisfy weak SC1. In the above version of MM with random groups of* n *agents who play the stage game, if* $\int_{\ell=0}^{1} \pi(\ell)\, d\Gamma_n^{Laplace}(\ell)$ *is positive (resp., negative), there is an* $r^* > 0$ *such that for all* $r \in (0, r^*)$, $z = 1$ *(resp.,* $z = 0$*) is absorbing and globally accessible while* $z = 0$ *(resp.,* $z = 1$*) is fragile and not globally accessible.*

Proof. Online appendix. □

With infinitesimal agents, an agent plays against the whole population $[0, 1]$. Her flow payoff is thus $\pi(z)$ rather than $\gamma(z)$. Accordingly, the same results hold if π has the properties of γ on which we rely in the proof of Theorem 25:

Corollary 26. *Let* $\pi : [0, 1] \to \Re$ *be continuous on* $[0, 1]$ *and satisfy SC1 on* $(0, 1)$*. In the modified version of MM with flow payoff* $\pi(z)$ *rather than* $\gamma(z)$*, if*

$$\int_{\ell=0}^{1} \pi(\ell)\, d\Gamma_\infty^{Laplace}(\ell)$$

is positive (resp., negative), then there is an $r^* > 0$ *such that for all* $r \in (0, r^*)$, $z = 1$ *(resp.,* $z = 0$*) is absorbing and globally accessible while* $z = 0$ *(resp.,* $z = 1$*) is fragile and not globally accessible.*

Proof. Online appendix. □

4.3 A Static Theory: Global Games

The global games (GG) theory is a static model in which the payoff function π depends not only on ℓ but also on an agent's private signal of an unobserved "fundamental" θ, such that (not) participating is strictly dominant for sufficiently high (low) signals. A contagion argument then pins down the agents' behavior for almost any fundamental.[46,47]

Following MS [41, Lemma 2.3, p. 70], we assume there is a random, unobserved state θ that is uniformly distributed on the whole real line.[48,49] Agents may be either discrete or infinitesimal. Each agent i sees a noisy signal $x_i = \theta + \sigma\varepsilon_i$ of the state θ, where $\sigma > 0$ is a scalar. An agent's net payoff from participating is a function $\pi(\ell, x_i)$ of the other-agent participation rate ℓ and the agent's private signal x_i.[50] The

[46]Global games were first studied for two-player, two-action games by Carlsson and van Damme [14] and for n-player, two-action coordination games by Kim [37]. The coordination-game assumption was first relaxed by Goldstein and Pauzner [26] who showed that there is a unique Nash equilibrium under strict SC1 if signal errors are uniformly distributed. Morris and Shin [41, p. 70] then showed the existence of a unique *threshold* equilibrium for general signal errors under strict SC1. Next Szkup [46] extend this uniqueness result to uniform signal errors and a specific payoff function that satisfies only weak SC1. Finally, we extend the result to general signal errors and a general payoff function that satisfies weak SC1. An example in which weak SC1 fails and there is no threshold equilibrium appears in Karp, Lee, and Mason [36].

[47]The problem of eliciting the participation of a group of agents in a global games setting has been studied in the context of bailouts (Frankel [21]), bank runs (Goldstein and Pauzner [26]), foreign direct investment (Dasgupta [16]), investment subsidies (Sákovics and Steiner [44]), platform competition (Argenziano [4]; Jullien and Pavan [34]), and monopoly pricing (Frankel [20]). Other applications of global games include international contagion (Goldstein and Pauzner [27]), currency crises and market crashes (Morris and Shin [40, 42]), information acquisition (Yang [51]), merger waves (Toxvaerd [48]), and regime change (Angeletos, Hellwig, and Pavan [1, 2]; Szkup and Trevino [47]). Experimental support appears in Heinemann, Nagel, and Ockenfels [30, 31], while theoretical limitations are studied in Angeletos, Hellwig, and Pavan [1], Angeletos and Werning [3], Chassang [15], Hellwig, Mukherji, and Tsyvinski [32], and Weinstein and Yildiz [49].

[48]Such an "improper prior" can be seen as the limit, e.g., of a normal distribution as the variance grows without bound. The assumption of an improper prior greatly simplifies the analysis but is not essential to the results.

[49]In some applications, a principal designs a scheme to induce the agents to participate. The principal knows the agents' payoff function and thus can predict their response. We can obtain this feature in a global game by assuming that a public signal $\theta + \sigma'\varepsilon$ of the state is first observed, where $\sigma' > 0$ is a scalar and ε is noise. The principal chooses her scheme, after which each agent i sees a private signal $x_i = \theta + \sigma\varepsilon_i$ of the state. Taking the private signal error to zero ($\sigma \to 0$), one obtains a unique prediction for the agents' behavior. If we then take the public signal error to zero ($\sigma' \to 0$), the principal can estimate the state with arbitrarily high precision; she can thus predict how the agents will respond to any scheme.

[50]Frankel, Morris, and Pauzner [24, pp. 23 ff.] (FMP) give conditions under which, if π satisfies *strategic complementarities*, then the asymptotic (as $\sigma \to 0$) behavior of the model studied here coincides with that of a variant in which an agent's payoff is a function $\pi(\ell, \theta)$ not of her signal x_i but rather of the state θ. While we expect that their result holds also under weak SC1, the verification needed would be beyond the scope of this paper for two reasons. First, the Laplace criterion is just one of six criteria that we study and finds support also in the MM theory (§4.2.5). Second, real-world payoffs are plausibly heterogeneous, which is better captured by a model such as

idiosyncratic terms ε_i are independent of each other and of θ and have a continuous density f with full support on $\Re$ and corresponding distribution function F. We assume moreover that f satisfies

MLRP For all $x_H > x_L$ and $y_H > y_L$ in $\Re$,

$$f(s_H - z_H) f(s_L - z_L) \geq f(s_H - z_L) f(s_L - z_H).$$

An example with these properties is a normal distribution (with any mean and variance). MLRP appears as assumption A7 in MS [41, p. 69].

Fix an agent i. Suppose each agent $j \neq i$ participates (resp., does not participate) if her signal $x_j = \theta + \sigma\varepsilon_j$ exceeds (resp., is less than) some fixed threshold $k \in \Re \cup \{\pm\infty\}$. Given θ, the probability that agent j participates is then

$$\Pr(\theta + \sigma\varepsilon_j > k|\theta) = \Pr\left(\varepsilon_j > \frac{k-\theta}{\sigma}\middle|\theta\right) = 1 - F\left(\frac{k-\theta}{\sigma}\right) \tag{36}$$

We now turn separately to the cases of discrete and infinitesimal agents.

1. Discrete agents: $\Lambda = \lambda$. By (36), the probability (given θ) that a proportion $\ell \in \lambda$ of agent i's opponents participate is $\kappa\left(\ell, 1 - F\left(\frac{k-\theta}{\sigma}\right)\right)$ where κ is defined in (10). The density of the state θ at any realization θ_0 given the signal x_i is $\frac{1}{\sigma} f\left(\frac{x_i - \theta_0}{\sigma}\right)$. Thus, agent i's payoff from participating when her realized signal is x (and all others play according to the threshold k) is

$$\pi_\sigma^n(x,k) \stackrel{d}{=} \int_{\theta=-\infty}^{\infty} \frac{1}{\sigma} f\left(\frac{x-\theta}{\sigma}\right) \sum_{\ell\in\lambda} \kappa\left(\ell; 1 - F\left(\frac{k-\theta}{\sigma}\right)\right) \pi(\ell, x)\, d\theta. \tag{37}$$

2. Infinitesimal agents: agent i's payoff from participating when her signal realization is x is

$$\pi_\sigma^\infty(x,k) \stackrel{d}{=} \int_{\theta=-\infty}^{\infty} \frac{1}{\sigma} f\left(\frac{x-\theta}{\sigma}\right) \pi\left(1 - F\left(\frac{k-\theta}{\sigma}\right), x\right) d\theta \tag{38}$$

   as the other-agent participation rate ℓ equals the probability $F\left(\frac{k-\theta}{\sigma}\right)$ that a given opponent participates by the law of large numbers.

---

ours in which π depends partly on an idiosyncratic shock ε_i.

A threshold equilibrium consists of a finite threshold k^* such that if an agent i believes that each opponent j will (not) participate if her signal x_j exceeds (is less than) k^*, then it is a best response for i (not) to participate if her signal exceeds (resp., is less than) k^*. More precisely, let $\pi^*_\sigma(x,k)$ denote $\pi^\infty_\sigma(x,k)$ in the infinitesimal case and $\pi^n_\sigma(x,k)$ in the discrete case. We define:[51]

Definition 27. A *Threshold Equilibrium* consists of a threshold $k^* \in \Re \cup \{\pm\infty\}$ such that $\pi^*_\sigma(x,k^*) \gtreqless 0$ for all $x \gtreqless k^*$ in $\Re$.

Theorem 28 below gives sufficient conditions for the game to have a unique threshold equilibrium, with a finite threshold k^*.[52] The conditions run roughly as follows. A1 states that π is bounded and satisfies weak SC1. A4 states that there are dominance regions: for a sufficiently high (low) signal, (not) participating is strictly dominant. A2 and A3 are technical continuity and boundedness properties, while A5 restates our above assumptions on the noise distribution.

Theorem 28. *Let* Λ *and* $\pi(\ell,x)$ *denote* λ *and* $\pi^n(\ell,x)$*, resp., in the discrete case, and* $[0,1]$ *and* $\pi^\infty(\ell,x)$*, resp., in the infinitesimal case. Assume:*

A1. *Single Crossing: for each* $x \in \Re$*,* $\pi(\ell,x)$ *satisfies weak SC1 in* $\ell \in \Lambda$ *and* $\sup_{\ell\in\Lambda}|\pi(\ell,x)|$ *is finite;*

A2. *Continuity:* $\pi(\ell,x)$ *is continuous in* $x \in \Re$ *uniformly in* $\ell \in \Lambda$*;*[53]

A3. *State Monotonicity:* $\pi(\ell,x)$ *is increasing in* x *and bounded in* $\ell \in \Lambda$ *for each* $x \in \Re$*;*

[51] The interpretation of $k^* = +\infty$ (resp., $k^* = -\infty$) is that the players never (resp., always) participate, regardless of their signals.

[52] By focusing on threshold equilibria, we enable applications that satisfy only weak SC1. GP [26, pp. 1325-6] give some rationales for this restriction. It can be omitted under strategic complementarities, as shown in FMP [24] and Kim [37]. Other global games papers that restrict to threshold equilibria include Angeletos, Hellwig, and Pavan [2], Mathevet and Steiner [38], and Morris and Shin [42].

[53] More precisely: for any $\varepsilon > 0$ there is a $\delta > 0$ such that, for any $\ell \in \Lambda$ and $x, x' \in \Re$ satisfying $|x - x'| < \delta$, $|\pi(\ell,x) - \pi(\ell,x')| < \varepsilon$.

A4. *Dominance Regions: there are* $\underline{x} < \overline{x}$ *both in* $\Re$ *such that, for all* $\ell \in \Lambda$, $\pi(\ell, \underline{x}) < 0 < \pi(\ell, \overline{x})$*; and*

A5. *MLRP and Full Support: the noise density* f *is positive and continuous on* $\Re$ *and satisfies MLRP.*

Then the game has a unique threshold equilibrium, whose threshold k^* *is finite and is the unique* $k \in \Re$ *that satisfies* $\pi^*_\sigma(k,k) = 0$. *In this equilibrium, an agent with signal* x *(does not) participate if* $\int_{\ell=0}^{1} \pi(\ell, x)\, d\Gamma^{Laplace}(\ell)$ *is positive (negative).*

Proof. Online appendix. □

To remain faithful to the global games literature we have assumed that each agent plays against all of the other agents rather than against a random sample as in KMR, FY, FH, and MM. Yet the expected payoff function in this setting is still a Bernoulli mixture over the realized payoff function π, so that weak SC1 suffices (assumption A1). Intuitively, an agent knows her opponents' identities but not their signals. Hence, each opponent's action is again the outcome of a random trial, where the success probability is the chance that the opponent's signal exceeds the participation threshold.

5 Detailed Summary of Results

Our results are summarized in Table 1.[54] With discrete agents, weak SC1 is a sufficient condition for each theory to yield its associated criterion. This condition is typically easy to verify.[55]

For each theory, existing results in the literature were extended to obtain the sufficient conditions in columns 3 and 4. The final column ("Prior Art") indicates

[54]Formal statements of these results appear in §3 above and in §4 below.

[55]See, e.g., Frankel [22].

		Sufficient Conditions for Theory to Imply Criterion		
Theory	Criterion	Discrete Agents	Infinitesimal Agents	Prior Art
A. Heuristic Arguments				
PI	PI	Weak SC1	Weak SC1	SC
FI	FI	Weak SC1	Weak SC1	SC
B. Rational-Player Models				
MM	Laplace	Weak SC1	Continuous on $[0, 1]$ & SC1 on $(0, 1)$	SC
GG	Laplace	Weak SC1	Weak SC1	See §1
C. Evolutionary Models				
KMR	KMR	Weak SC1	Theory requires discrete agents	SC
FY	FY	Weak SC1	Lipschitz on $[0, 1]$ & SC1 on $(0, 1)$	SC
FH	FH	Weak SC1	Lipschitz on $[0, 1]$ & SC1 on $(0, 1)$	SC, $n = 2$

Table 1: Summary of Results. "Prior Art": sufficient conditions identified by the prior literature. "SC": strategic complementarities. "Lipschitz": Lipschitz-continuous.

the state of the literature before we began this project, where "SC" means "strategic complementarities" (π increasing or at least nondecreasing in ℓ).

The most significant extension is for the FH theory. FH [25] show only that their theory implies the risk dominance criterion in two-player games. However, *all* of the five noncooperative theories that we study reduce to risk dominance in such games.[56] The result that FH generalizes differently is thus novel.

The six fictional beliefs appear in Figure 3 for $n = 500$ agents. FI (resp., PI) beliefs put all of their weight on no one else (resp., everyone else) participating: Γ is a point mass on zero (resp., one). Laplacian beliefs put uniform weight on every other-agent participation rate and is thus the red diagonal line. The magenta dashed curve depicts KMR beliefs, which converge to a point mass at $\ell = 1/2$ as $n \to \infty$. FY beliefs are the purple dotted line; while attaining a maximum weight at $\ell = 1/2$, they are more dispersed than the KMR criterion and do not converge to a step function as $n \to \infty$. Finally, FH beliefs put half weight on no one else participating and half weight on everyone else participating; they are the double blue curve.

[56]This commonality is studied below in §5.1.

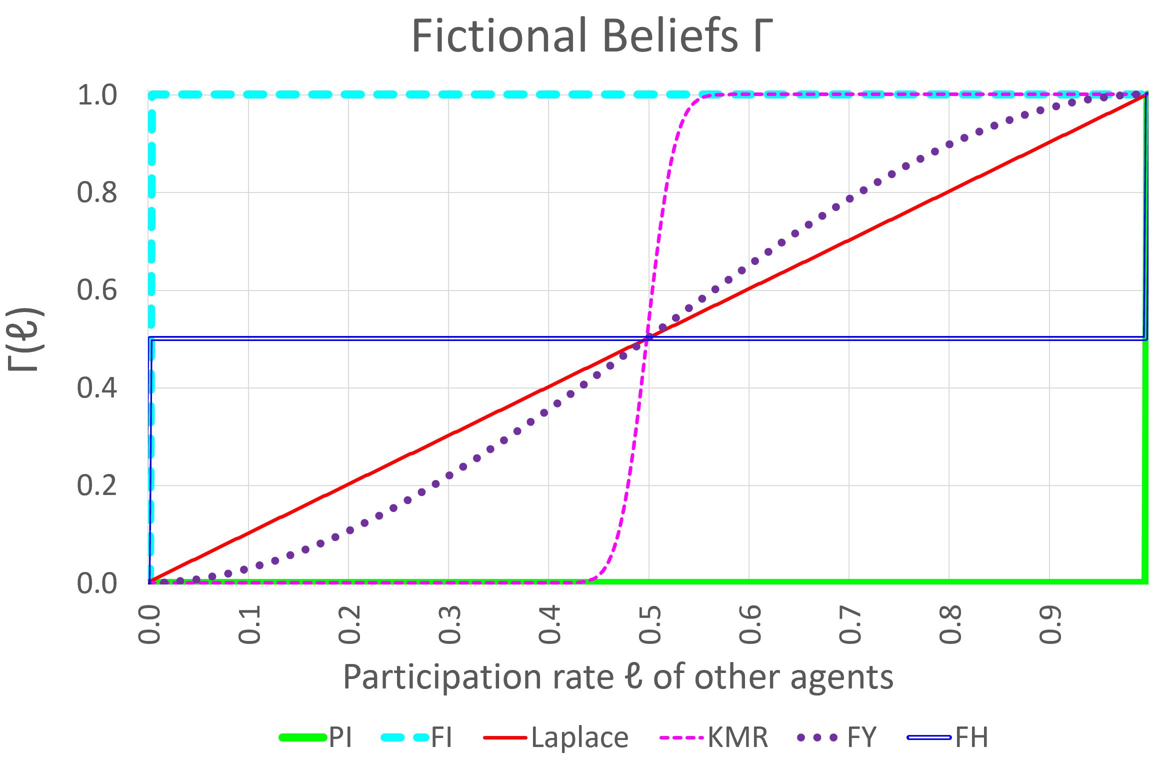

Figure 3: Fictional beliefs for each of the six criteria; $n = 500$ agents. The graph depicts the probability $\Gamma(\ell)$ that the proportion of others who participate is at most $\ell \in [0,1]$ according to the fictional beliefs Γ associated with the FI, PI, Laplace, KMR, FY, and FH criteria.

5.1 Relation to Risk Dominance

The first paper to study each of the five noncooperative theories shows that in two-player coordination games with two actions, the risk-dominant equilibrium is selected.[57] No common features of the five theories have been identified that lead to risk dominance. We now fill this gap.

The key property of the five theories that drives this result is that they satisfy (2) together with the following property.

Label Irrelevance (LI) For any $u, v \in U$, suppose the net payoff $\pi^u(\ell)$ under u from choosing action 1 if $\ell \in \lambda$ others choose action 1 equals the net payoff

[57] See Kandori, Mailath, and Rob [35], Foster and Young [19], Fudenberg and Harris [25], Matsui and Matsuyama [39], and Carlsson and van Damme [14].

$-\pi^v(1-\ell)$ under v from choosing action 0 if ℓ others choose action 0. Then the predictions of the criterion are reversed: for each $a \in \left\{0, \frac{1}{2}, 1\right\}$, $u \in \Theta_a$ if and only if $v \in \Theta_{1-a}$.

LI means there is no bias in favor of one action or the other - e.g., due to the preferences of an unmodeled principal. Only payoffs matter. The noncooperative theories satisfy LI: the underlying models treat the two actions symmetrically, with no bias in favor of one or the other. But the heuristic theories violate LI. To see this, suppose $u(1,\ell) = \ell$ and $u(0,\ell) = 1/2$. Then $\pi^u(\ell)$ equals $-\pi^u(1-\ell)$ as both equal $\ell - 1/2$. Applying LI with $v = u$ implies that no prediction can be made: if u is in Θ_a then it is also in Θ_{1-a} so $a = 1/2$. However, both "all participate" and "none participate" are strict Nash equilibria. Thus, the PI theory selects "all participate" while the FI theory selects "none participate": they violate LI.

The connection between LI and risk dominance is as follows. First, in the case of two agents, the only other-agent participation rates are 0 and 1, so $d\Gamma(0) + d\Gamma(1) = 1$. Second, LI implies that the fictional beliefs must assign the same weight to other-agent participation rates that are equidistance from 1/2.[58] This implies that $d\Gamma(0) = d\Gamma(1)$ so each weight equals 1/2. Thus, by (2), the agents play a best response to a 50-50 mixture: they select the risk-dominant equilibrium.

The formal results are as follows. We focus throughout on the case of discrete agents. Let U_0 be the set of utility functions u that satisfy weak SC1. Consider any criterion Θ on U_0 that satisfies (2) for some fictional beliefs Γ. We define two properties that are each equivalent to LI. The first is that if one rotates the fictional beliefs Γ by 180° in Fig. 3, one obtains the original beliefs.[59]

Rotation Invariance (RI) For each $\ell \in \lambda$, $\Gamma(\ell)$ equals $1 - \Gamma(1-\ell)$.

The second is that the weights at two points equidistant from 1/2 are equal:

[58]This is an implication of Claim 29 below.

[59]Casual observation of Fig. 3 suggests that the four noncooperative criteria satisfy RI while the two heuristic criteria do not. The results of this section prove this for any n.

Symmetry Around a Half (SAH) for each other-agent participation rate $\ell \in \lambda$, the weights $d\Gamma(\ell)$ and $d\Gamma(1-\ell)$ are equal.

RI implies that the slope of Γ at each pair of points that are equidistant from 1/2 is equal, as these equidistance points swap places under a 180° rotation. In fact, the three properties coincide:

Claim 29. *LI, RI, and SAH are all equivalent.*

Proof. Assume RI: for each $\ell \in \lambda$, $\Gamma(\ell)$ equals $1-\Gamma(1-\ell)$ or, equivalently, $\int_{\ell'=0}^{\ell} d\Gamma(\ell') = 1-\int_{\ell'=0}^{1-\ell} d\Gamma(\ell')$. Differentiating with respect to ℓ yields $d\Gamma(\ell) = d\Gamma(1-\ell)$: SAH holds. Conversely, assume SAH: for any $\ell' \in \lambda$, $d\Gamma(\ell')$ equals $d\Gamma(1-\ell')$. Integrating from zero to any given $\ell \in \lambda$ and letting $\ell'' = 1-\ell'$ yields

$$\Gamma(\ell) = \int_{\ell'=0}^{\ell} d\Gamma(\ell') = \int_{\ell'=0}^{\ell} d\Gamma(1-\ell') = \int_{\ell''=1-\ell}^{1} d\Gamma(\ell'') = 1-\Gamma(1-\ell),$$

which is just RI. Now let $\Xi(u) = \int_{\ell=0}^{1} \pi^u(\ell)\, d\Gamma(\ell)$ and assume SAH. Then for any u, v in U such that $\pi^u(\ell) = -\pi^v(1-\ell)$ for all $\ell \in \lambda$, one can easily verify that $\Xi(u) = -\Xi(v)$: LI holds. Conversely, suppose SAH fails: for some ℓ', $d\Gamma(\ell') \neq d\Gamma(1-\ell')$. Assume w.l.o.g. that $\ell' < 1/2$. The utility functions

$$u(a,\ell) = -d\Gamma(1-\ell')\, 1_{(a,\ell)=(1,\ell')} + d\Gamma(\ell')\, 1_{(a,\ell)=(1,1-\ell')}$$

and $v(a,\ell) = -d\Gamma(\ell')\, 1_{(a,\ell)=(1,\ell')} + d\Gamma(1-\ell')\, 1_{(a,\ell)=(1,1-\ell')}$ satisfy weak SC1. Moreover, $\pi^u(\ell) = -\pi^v(1-\ell)$ for all $\ell \in \lambda$. However, $\Xi(u) = 0$ while $\Xi(v) = [d\Gamma(1-\ell')]^2 - [d\Gamma(\ell')]^2 \neq 0$, so u is in $\Theta_{1/2}$ while v is not: LI fails as well. □

Moreover, only the noncooperative criteria satisfy the three properties:

Claim 30. *The noncooperative criteria in Table 1 satisfy LI, RI, and SAH while the heuristic criteria violate these properties.*

Proof. By Claim 29, it suffices to check any one of the properties. The example in the text shows that PI and FI violate LI. Further, Laplace and FH trivially satisfy SAH. Finally, let $\ell_i = \frac{i}{n-1}$. Substituting "KMR" and "FY" respectively for "crit",

(13) and (18) imply that $d\Gamma_n^{\text{crit}}(\ell_i) = W_{i,n}^{\text{crit}}$ and $d\Gamma_n^{\text{crit}}(1-\ell_i) = W_{n-1-i,n}^{\text{crit}}$. One can also verify using the same equations that $W_{i,n}^{\text{crit}} = W_{n-1-i,n}^{\text{crit}}$, so KMR and FY satisfy SAH as well. $\square$

Finally, SAH implies the risk-dominance criterion in the two-agent case:

Claim 31. *If Γ satisfies SAH and there are two agents ($n=2$), then $d\Gamma(0) = d\Gamma(1) = 1/2$: each agent plays a best response to a 50-50 mixture of the two actions.*

Proof. With two agents, the set of other-agent participation rates λ is $\{0,1\}$. SAH then implies $d\Gamma(0) = d\Gamma(1) = 1/2$. $\square$

Taken together, the claims in this section show that the noncooperative criteria reduce to risk-dominance in the two-agent case because they all satisfy (2) and are label-invariant.

6 Examples

To illustrate our results, we now review two devices that arise commonly in participation games that lead to violations of strategic complementarities, making prior results inapplicable. But as weak SC1 still holds, the devices can be analyzed using our tools.

The first device is rationing. For instance, Frankel [22] uses equation (2) to study the effects of share rationing in oversubscribed initial public offerings. Another example is found in the bank run model of Goldstein and Pauzner [26] ("GP"), which is based on Diamond and Dybvig [17]. That game is played among a set of patient agents who must decide whether to run (withdraw early) or wait (withdraw after the bank's projects have borne fruit). We identify participating with waiting, which is the desired action of the bank (the principal). Panel 1 in Figure 4 depicts a patient agent's net payoff $\pi(\ell)$ from waiting if a proportion ℓ of others wait.[60] If the propor-

[60]This payoff function is derived in our online appendix (§10).

tion of patient agents who wait is less than ℓ_0, the bank runs out of funds in period 1. These funds must therefore be rationed. This is accomplished via a lottery in which some agents are paid in full and others are not. A further fall in ℓ lowers the chance of being paid in full, which raising the relative payoff π from waiting: in the rationing region $\ell \in [0, \ell_0]$, there are strategic *substitutes*. Yet strict SC1 still holds as rationing occurs only in the region when the net payoff π is negative. Thus, equation (2) can be used to predict whether a bank run occurs.

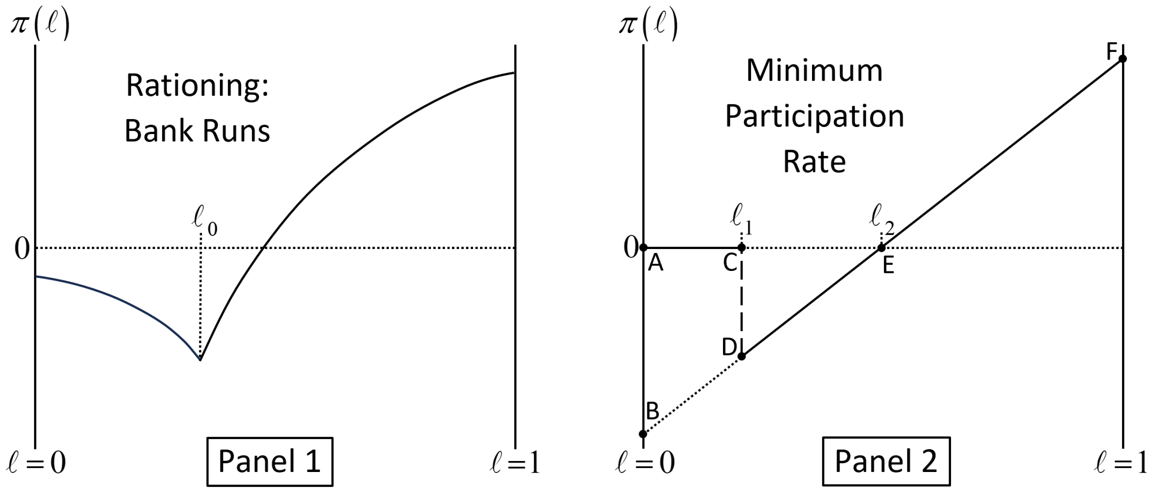

Figure 4: Panel 1: rationing in a bank run model. When fewer than ℓ_0 patient agents wait, the bank runs out of funds in period 1 and must ration them. A rise in the proportion ℓ who wait leads to less rationing so π is declining in $\ell \in [0, \ell_0]$. Panel 2: a minimum participation rate (MPR). The initial relative payoff π is given by BDEF. In a modified game with MPR ℓ_1, the activity is cancelled if fewer than ℓ_1 others participate so the relative payoff π from participating is zero in this range. This yields the payoff function ACDF. An optimal MPR is ℓ_2, which gives rise to the payoff function AEF.

The second device is a minimum participation rate ("MPR") below which the activity is canceled. An example is a minimum subscription rate in an initial public offering (Frankel [22]). A generic setting with a MPR appears in panel 2 of Figure 4. Without any MPR, the net payoff π from participating is given by the curve BDEF. If the principal specifies an MPR of ℓ_1, she commits to cancel the activity if fewer than ℓ_1 others participate. This sets the net payoff from participating to zero for other-agent participation rates below ℓ_1. The resulting payoff function π is thus the curve ACDF, which satisfies weak SC1 but no stronger property.

It is easy to see that the alternative MPR ℓ_2, which leads to the payoff function ACEF, maximizes the integral $\int_{\ell=0}^{1} \pi^u(\ell)\, d\Gamma(\ell)$ in (2) - and thus an agent's likelihood of participating - as it sets the negative values of $\pi(\ell)$ to zero. It is thus optimal for the principal for any fictional beliefs Γ. While the resulting payoff function ACEF satisfies strategic complementarities, one cannot know that ℓ_2 is optimal without knowing how the agents would respond to alternative MPRs (such as ℓ_1) that satisfy only weak SC1.

7 Concluding Remarks

In many applied settings, a group of agents choose whether or not to participate in a joint activity that is worthwhile only if enough participate. Examples include joining a platform, investing in a project, rebelling against a regime, attacking a currency, and leaving funds in a bank. Such games have multiple equilibria: both all-participate and none-participate are self-fulfilling prophecies. To make predictions, one thus needs a theory of equilibrium selection.

We study seven well-known selection theories from the literature: two based on heuristics, two with rational players, and three evolutionary models. We show that they all give rise to selection criteria with a simple common form: an agent plays a best response to some distribution of the proportion of others who participate. We refer to this distribution, which is common across agents and payoffs, as the "fictional beliefs" arising from the given criterion.

In showing this result, we assume only a weak single crossing property.[61] In contrast, most prior work on the seven selection theories has assumed strategic complementarities - a much stronger condition. By predicting the outcomes of a wider

[61] This holds for discrete agents. For the infinitesimal case, see Table 1.

set of participation games, our results let a researcher study more mechanisms that are used in practice and design better mechanisms for a principal or society.

The five noncooperative theories all imply that agents will select the risk-dominant equilibrium in two-player games. While this fact is well known, it has been regarded as a coincidence: no common reason has been given. We fill this gap. We show that a criterion with our simple common form that is also *label-invariant* must reduce to risk-dominance in the two-agent case. Label-invariance means that the criterion treats the two actions symmetrically: only payoffs matter.

In many settings, a principal devises a scheme to induce a set of agents to participate in some activity. Our online appendix develops an algorithm to solve this *principal-agents* game. Frankel [22] applies this algorithm to three security design settings, two of which violate strategic complementarities but satisfy our weaker single crossing condition.

References

[1] Angeletos, G.-M., C. Hellwig, and A. Pavan. 2006. "Signaling in a Global Game: Coordination and Policy Traps." Journal of Political Economy 114:452-84.

[2] —. 2007. "Dynamic Global Games of Regime Change: Learning, Multiplicity, and the Timing of Attacks." Econometrica 75:711-756.

[3] Angeletos, G.-M., and I. Werning. 2006. "Crises and Prices: Information Aggregation, Multiplicity, and Volatility." American Economic Review 96:1720-36.

[4] Argenziano, R. 2008. “Differentiated Networks: Equilibrium and Efficiency.” Rand Journal of Economics 39:747–769.

[5] Athey, S. 1996. “Comparative Statics Under Uncertainty: Single Crossing Properties and Log-Supermodularity.” M.I.T. Working Paper no. 96-22.

[6] ——. 2002. “Monotone Comparative Statics Under Uncertainty.” Quarterly Journal of Economics 117:187-223.

[7] ——. 4/25/2023. Private communication.

[8] Beggs, A.W. 2007. “Large Deviations and Equilibrium Selection in Large Populations.” Journal of Economic Theory 132:383-410.

[9] Bergin, J., and B.L. Lipman. 1996. “Evolution with State-Dependent Mutations.” Econometrica 64:943-956.

[10] Bergemann, D., and S. Morris. 2005. “Robust Mechanism Design.” Econometrica 73:1771-1813.

[11] Bernstein, S., and E. Winter. 2012. “Contracting with Heterogeneous Externalities.” American Economic Journal: Microeconomics 4:50-76.

[12] Binmore, K., and L. Samuelson. 1997. “Muddling Through: Noisy Equilibrium Selection.” Journal of Economic Theory 74:235-265.

[13] Blume, L.E. 2003. “How Noise Matters.” Games and Economic Behavior 44:251-271.

[14] Carlsson, H., and E. van Damme. 1993. “Global Games and Equilibrium Selection.” Econometrica 61:989-1018.

[15] Chassang, S. 2010. "Fear of Miscoordination and the Robustness of Cooperation in Dynamic Global Games With Exit." Econometrica 78:973-1006.

[16] Dasgupta, A. 2007. “Coordination and Delay in Global Games.” Journal of Economic Theory 134:195-225.

[17] Diamond, D.W., and P.H. Dybvig. 1983. “Bank Runs, Deposit Insurance, and Liquidity.” Journal of Political Economy 91:401-419.

[18] Ellison, G. 2000. “Basins of Attraction, Long-Run Stochastic Stability, and the Speed of Step-by-Step Evolution.” Review of Economic Studies 67:17-45.

[19] Foster, D., and P.H. Young. 1990. “Stochastic Evolutionary Game Dynamics.” Theoretical Population Biology 38:219–232.

[20] Frankel, D.M. 2012. “Recurrent Crises in Global Games.” Journal of Mathematical Economics 48:309-321.

[21] —–. 2017. “Efficient Ex-Ante Stabilization of Firms.” Journal of Economic Theory 170:112-144.

[22] —–. “How to Sell Equity if You Must.” Mimeo (www.dmfrankel.com/papers.html).

[23] —–. 2023. “Online Appendix to ‘Equilibrium Selection in Participation Games: A Unified Framework’.” Mimeo (www.dmfrankel.com/papers.html).

[24] Frankel, D.M., S. Morris, and A. Pauzner. 2003. “Equilibrium Selection in Global Games with Strategic Complementarities.” Journal of Economic Theory 108:1-44.

[25] Fudenberg, D., and C. Harris. 1992. “Evolutionary Dynamics with Aggregate Shocks.” Journal of Economic Theory 57:420-441.

[26] Goldstein, I., and A. Pauzner. 2005. “Demand-Deposit Contracts and the Probability of Bank Runs.” Journal of Finance 60:1293-1327.

[27] —–. 2004. "Contagion of Self-Fulfilling Financial Crises due to Diversification of Investment Portfolios." Journal of Economic Theory 119:151-83.

[28] Halac, M., I. Kremer, and E. Winter. 2020. “Raising Capital from Heterogeneous Investors.” American Economic Review 110:889-921.

[29] Halac, M., E. Lipnowski, and D. Rappoport. 2021. "Rank Uncertainty in Organizations." American Economic Review 111:757-786.

[30] Heinemann, F., R. Nagel, and P. Ockenfels. 2004. "The Theory of Global Games on Test: Experimental Analysis of Coordination Games with Public and Private Information." Econometrica 72:1583-1599.

[31] —–. 2009. "Measuring Strategic Uncertainty in Coordination Games." Review of Economic Studies 76:181-221.

[32] Hellwig, C., A. Mukherji, and A. Tsyvinski. 2006. "Self-Fulfilling Currency Crises: The Role of Interest Rates." American Economic Review 96:1769-1787.

[33] Jacobsen, H.J., M. Jensen, and B. Sloth. 1999. "On the Structural Difference Between the Evolutionary Approach of Young and that of Kandori, Mailath, and Rob." Mimeo, U. Copenhagen.

[34] Jullien, B., and A. Pavan. 2019. Information Management and Pricing in Platform Markets. Review of Economic Studies 86:1666-1703.

[35] Kandori, M., G. Mailath, and R. Rob. 1993. "Learning, Mutation and Long Equilibria in Games." Econometrica 61:27–56.

[36] Karp, L., I.H. Lee, and R. Mason. 2007. "A Global Game with Strategic Substitutes and Complements." Games and Economic Behavior 60:155-175.

[37] Kim, Y. 1996. "Equilibrium Selection in n-Person Coordination Games." Games and Economic Behavior 15:203-227.

[38] Mathevet, L., and J. Steiner. 2013. "Tractable Dynamic Global Games and Applications." Journal of Economic Theory 148:2583-2619.

[39] Matsui, A., and K. Matsuyama. 1995. "An Approach to Equilibrium Selection." Journal of Economic Theory 65:415-434.

[40] Morris, S., and H.S. Shin. 1998. "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks." American Economic Review 88:587-597.

[41] —–. 2003. “Global Games: Theory and Applications.” Pp. 56-114 in Dewatripont, M., L.P. Hansen, and S.J. Turnovsky. “Advances in Economics and Econometrics: Theory and Applications, [Econometric Society] Eighth World Congress, Volume 1.” Cambridge University Press.

[42] —–. 2004. "Liquidity Black Holes." Review of Finance 8:1-18.

[43] —–. 2006. “Catalytic Finance: When Does it Work?” Journal of International Economics 70:161-177.

[44] Sákovics, J., and J. Steiner. 2012. “Who Matters in Coordination Problems?” American Economic Review 102:3439-3461.

[45] Segal, I. 1999. “Contracting with Externalities.” Quarterly Journal of Economics 114:337-388.

[46] Szkup, M. 2022. “Preventing Self-Fulfilling Debt Crises: A Global Games Approach.” Review of Economic Dynamics 43:22-55.

[47] Szkup, M., and I. Trevino. 2015. “Information Acquisition in Global Games of Regime Change.” Journal of Economic Theory 160:387-428.

[48] Toxvaerd, F. 2008. "Strategic Merger Waves: A Theory of Musical Chairs." Journal of Economic Theory 140:1-26.

[49] Weinstein, J., and M. Yildiz. 2007. "A Structure Theorem for Rationalizability n. with Application to Robust Prediction of Refinements." Econometrica 75:365-400.

[50] Winter, E. 2004. “Incentives and Discrimination.” American Economic Review 94:764–73.

[51] Yang, M. 2015. "Coordination with Flexible Information Acquisition." Journal of Economic Theory 158:721-738.

[52] Young, H.P. 1993. "The Evolution of Conventions." Econometrica 61:57-84.