ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

THE EVOLUTION OF COLLECTIVE ACTION

David P. Myatt and Chris Wallace

Number 190

April 2005

Manor Road Building, Oxford OX1 3UQ

The Evolution of Collective Action

David P. Myatt
Department of Economics and St. Catherine's College, University of Oxford
david.myatt@economics.ox.ac.uk

Chris Wallace
Department of Economics and Trinity College, University of Oxford
christopher.wallace@economics.ox.ac.uk

Revised: April 2005.[1]

Abstract. A public good is produced if and only if a team of m or more volunteers contribute to it. An equilibrium-selection problem leads to the questions: will the collective action succeed? If so, who will participate in the team? The paper studies the evolution of collective action: as part of a strategy-revision process, updating players choose quantal responses to existing play. With symmetric players, success depends upon the cost of contribution, the benefit from provision, and the critical team-size m; the relative variability of costs and benefits, and their correlation, are also critical. When players differ, successful teams consist of either the most efficient contributors, or those with the most idiosyncratic preferences. The addition of a single "bad apple" (for instance, an individual whose costs are particularly variable) to a population in which a successful team operates may result in destabilisation: over time, the bad apple might supplant an existing contributor, prompting a collapse.

1. Collective Action and Coordination

This paper studies the evolving play of a collective-action game in which the successful provision of a public good requires the voluntary participation of a critical number of individuals. For such a game, there are equilibria in which a successful team of contributors is formed, and an equilibrium for which the collective action fails. The consequent equilibrium-selection problem leads to the study of a simple strategy-revision process that is subject to noise. Following this approach, a number of interesting questions arise. Will the collective action succeed? If so, then who participates? How do the answers depend upon the costs and benefits involved; the variability and correlation of payoffs; the team size required for success; the return of any excessive or unused contributions; and the set of potential contributors?

Formally, collective actions are interpreted as n-player binary-action games in which m or more voluntary contributions are required for the successful production of a public good; such games include the "threshold" provision game of Palfrey and Rosenthal (1984). The

[1]**Date Printed**. April 19, 2005 (EVOCAMar2005.tex). **JEL Classification**. C72, C73, and H41.
Keywords. Collective action, evolution, teams, equilibrium selection, exponential cost, rooted trees.
Acknowledgements. For their helpful comments, thanks are due to seminar participants at Keele University, Oxford University, and the "Teamwork" conference held at St. Catherine's College, Oxford in 2003. For more detailed suggestions, thanks are offered to David Levine and three anonymous reviewers.

possibility of teamwork suggests a resolution to the usual problem of public-good under-provision: by making an entire project depend in a pivotal way on each team member, incentives are dramatically enhanced, so long as others are expected to play their part.

There is, of course, a coordination problem. When $m \geq 2$, so that genuine teamwork is a prerequisite for success, there are $\binom{n}{m}$ (pure-strategy) Nash equilibria in which the public good is provided; there is a single equilibrium in which no player participates; and finally, as Palfrey and Rosenthal (1984) have shown, there is a number of mixed-strategy equilibria.[2]

One approach to equilibrium selection is an appeal to symmetry. Dixit and Olson (2000) invoked a "pragmatic" argument in favour of a symmetric mixed-strategy Nash equilibrium. (In their game, so long as at least m players participate, the cost of public-good provision is shared by the volunteers.) Their appeal stemmed from (p. 318) "the difficulty of the coordination that is required for choosing one among the asymmetric pure strategy equilibria when a subset of a larger group is designated to do one thing and the rest another" and, in their context, "the argument for symmetry is stronger because the coordination problem is harder: the potential participants are not even identified until they show up ..."

This appeal to symmetry faces a number of critiques. First, it is limited to games that are exactly symmetric. Second, it simply ducks the problem of equilibrium selection: the "difficulty of coordination" reflects the existence of the problem, and does not imply a symmetric solution. Third, the problem of coordination is not limited to asymmetric pure equilibria: for the play of a mixed-strategy Nash equilibrium, the players must randomise with knife-edge probabilities, or at least expect others to do so. Fourth, such mixed equilibria are inherently unstable: any slight deviation might prompt a sequence of reactions that lead away to an (asymmetric) pure equilibrium. Fifth, mixed-strategy equilibria exhibit unappealing comparative statics, and particularly so when applied to asymmetric games.[3] Sixth, an appeal to symmetry is not universally applicable: in a threshold public-good game, there are *no* symmetric mixed equilibria for a wide range of parameter configurations.[4]

Fortunately, recent theoretical advances afford new perspectives on the equilibrium-selection problem. One approach is to view the game in question as an instance of a "global" game (Carlsson and van Damme, 1993).[5] A global game will often have a unique equilibrium, even when there is a selection problem within the "local" instances of the game in question. A second approach, following Foster and Young (1990), Young (1993a), Kandori, Mailath, and Rob (1993), and others, is to study the long-run behaviour of a strategy-revision process that is subject to noise. As noise is driven from the model, certain strategy profiles (often a unique pure-strategy equilibrium) are almost always played in the long-run.

[2]Coordination problems arise when teamwork is unnecessary. If $m = 1$, the players face a volunteers' dilemma (Diekmann, 1985, for example); the volunteer's identity must still be ascertained (Myatt and Wallace, 2005a).
[3]For example, players with higher contribution costs may particiopate with higher probability.
[4]Mixed equilibria involve symmetry only among a subset who mix (Palfrey and Rosenthal, 1984).
[5]Morris and Shin (2003) and Myatt, Shin, and Wallace (2002) offered surveys of the global-games literature.

The latter approach is taken here. At each point in time, a player is chosen, at random, to select a revised strategy. This revising player chooses a quantal response (McKelvey and Palfrey, 1995); equivalently, this is a myopic best-response given payoffs that are subject to perturbations. Both normal and logistic payoff disturbances are considered; these yield probit and logit quantal responses, respectively.[6] The strategy-revision process differs from much of the existing literature in two ways. First, there is no artificial population of agents waiting to fill each player position; the state space is simply the set of strategy profiles.[7] Second, the ergodicity of the process is not due to mistakes, mutations, or experimentation; the necessary noise stems from payoff variability. Allowing the variance of payoffs to vanish yields, in the limit, a one-step-at-a-time best-response process. Furthermore, the corresponding ergodic distribution places all weight on a single pure-strategy Nash equilibrium. Which one?

Heuristically, the success of a collective action depends upon the relative ease of transitions between the state of play in which no player participates (so that the action fails) and the $\binom{n}{m}$ states in which a successful team operates. To leave failure behind, a team must be built. This takes $m-1$ difficult steps: at each step a non-participant must overcome a private cost (perhaps due to a transient idiosyncratic private benefit) and join in. (The mth team member is pivotal to the collective action and hence joins with high probability.) To escape from a successful state, a team member must quit. This is a single difficult step, perhaps due to a transient high provision cost. Will the collective action succeed? This will hinge upon whether the $m-1$ steps needed to build the team are easier than the single step needed to destroy it. Since the specification considered here permits consideration of asymmetric collective-action games, a secondary question arises: if the collective action succeeds, then who will contribute?[8] The answer, of course, is the players who find it easiest to overcome their private costs and join in. More subtly, they might not be the players with the lowest costs of provision, but might instead be those with the most variable payoffs.

Clearly, it is easiest to build a team when the average contribution cost is low. A successful team is relatively stable when the average value placed on the public good is high. The intuition given above suggests, and the formal results confirm, that the relative variability of costs and benefits is also important. Positive correlation between costs and benefits enhances the stability of a successful team: a high-cost draw is likely associated with a high-benefit draw, and so the contribution incentive (benefit minus cost) of an active team member is likely to remain positive. On the other hand, negative correlation undermines stability.

---

[6]Blume (1993, 1995, 1997, 2003) pioneered logit-driven strategy-revision processes, and employed an elegant Gibbs representation of the long-run play of potential games. Here, the probit permits consideration of heteroscedasticity and correlation between the costs of contribution and the benefits of team success.

[7]Kandori, Mailath, and Rob (1993) specified a population of potential occupants for each player position. For Young (1993a,b, 1996, 1998), the state-space corresponded to the history of play for each position.

[8]This differs from the work of Diermeier and Van Mieghem (2000). They considered the evolving play of a symmetric threshold game. Restricting to logit quantal-responses, they were able to determine when a team will succeed, but symmetry precluded the identification of the typical team members.

The success of a collective action will also depend upon the critical team-size m. An increase in m, other things being equal, will increase the number of steps needed to build a team. On the other hand, it might also enhance individual benefits from a consequently higher-value public good, hence reinforcing team stability. When this is so, the picture becomes slightly more complex; it is not always true that larger teams are less successful. For instance, large teams may actually have the advantage when value is "not too concave" in team size.

Success will also depend upon the precise specification of the game. In the benchmark threshold game, any unused (when a collective action fails) or excessive (when there are too many volunteers) contributions are discarded. Changing to a game in which such contributions are (partially) refunded can improve matters. On the hand, the introduction of a refundable subscription scheme (described in detail later in the paper) can harm the chances of success.

When players are asymmetric, their identities matter. Perhaps surprisingly, the addition of a single "bad apple" (for instance, an individual whose payoffs have relatively high variance) to a population in which a successful team operates may result in the team's destabilisation. The reason is that, over time, the bad apple might supplant an existing contributor, subsequently leave the team, and therefore prompt its collapse.

The paper proceeds by developing a notation for binary-action games, specifying the process by which play evolves, and introducing the equilibrium-selection device (Section 2). Section 3 applies the framework to the collective-action game of interest and contains the central result of the paper (Theorem 1). Sections 4, 5, and 6 respectively offer an analysis of the factors that affect team success: correlation, concavity, and project components; alternative game specifications; and the presence of "bad apples" amongst the set of players. Related strands of literature from sociology and political science are drawn together in the concluding remarks.

2. Binary-Action Games and Exponential Cost

This section introduces a framework for considering the evolving play of binary-action games. In Section 3, this framework will be applied to the problems of collective action.

2.1. **Binary-Action Games.**

The collective-action games of interest will be a subset of the class of simultaneous-move binary-action games. In this wider class, player $i \in \{1, \ldots, n\}$ takes an action $z_i \in \{0, 1\}$, where in later sections $z_i = 1$ will be thought of as contributing toward the production of a public good. A pure-strategy profile is then $z \in Z \equiv \{0, 1\}^n$. Looking forward to the strategy-revision process described below, a strategy profile z will be described as the state of play, and the set Z will be referred to as the state space. Finally, the payoff enjoyed by player i given the state of play z is simply $u_i(z)$.

The following notation proves useful. Beginning with a state of play z, write z^{i+} for the new state obtained by setting $z_i = 1$. Similarly, write z^{i-} for the new state obtained by setting $z_i = 0$. Notice that the states of play z^{i+}, and z^{i-} differ only by the action of player i, and

$z \in \{z^{i-}, z^{i+}\}$. Given this notation, and fixing a state z, player i's incentive to play $z_i = 1$ rather than $z_i = 0$ (what might be termed a "contribution incentive") is

$$\Delta u_i(z) \equiv u_i(z^{i+}) - u_i(z^{i-}).$$

There is little loss of generality in assuming that the contribution incentive is always non-zero; the set of pure-strategy Nash equilibria is $Z^* = \{z : \text{for all } i,\ z_i = 1 \Leftrightarrow \Delta u_i(z^*) > 0\}$.

2.2. **Evolving Play.** Play evolves via a one-step-at-a-time strategy-revision process. At time $t \in \{1, \ldots, \infty\}$ the state of play is $z_t \in Z$.[9] A single player i is randomly selected (with probability $1/n$) and selects a revised action based solely on the current state. The state is updated accordingly, and so $z_{t+1} \in \{z_t^{i+}, z_t^{i-}\}$. This process generates a Markov chain on the state space Z, with transition probabilities $\Pr[z \to z'] \equiv \Pr[z_{t+1} = z' \mid z_t = z]$.

It remains to specify how strategy revisions are made. One possibility would be for a revising player to choose a best response to the current state.[10] Given such a process, any pure-strategy Nash-equilibrium state of play will be absorbing. When there are multiple equilibria (as there are for a threshold game) the long-run behaviour of the process is path dependent.

To yield ergodicity, a revising player i chooses a "quantal response" to the state of play.[11] This carries a random-utility interpretation: the revised action is based upon a "noisy" contribution incentive, so that $z_{i(t+1)} = 1$ if and only if $\Delta\tilde{u}_i(z_t) > 0$ where $\Delta\tilde{u}_i(z) \sim f_i(\,\cdot\,; z, \varepsilon)$. The density f_i depends upon (i) the player, (ii) the state, and (iii) a parameter ε. Furthermore, it is assumed to have full support on the real line for $\varepsilon > 0$. Transitions are now easy to characterise. For example, to move from z^{i-} to z^{i+}, player i must be chosen (this occurs with probability $1/n$) and must then choose $z_i = 1$, which occurs when $\Delta\tilde{u}_i > 0$. In fact,

$$\Pr[z^{i-} \to z^{i+}] = \frac{\int_0^\infty f_i(x; z, \varepsilon)dx}{n} \quad \text{and} \quad \Pr[z^{i+} \to z^{i-}] = \frac{\int_{-\infty}^0 f_i(x; z, \varepsilon)dx}{n}.$$

A transition away to a neighbouring state (that is, a state in which one player's action differs) always has positive probability. There is also positive probability that the process remains in the same state. These features ensure that the process is ergodic for $\varepsilon > 0$. The results will focus on the associated ergodic distribution as $\varepsilon \to 0$. In the limit, noise in

[9] There is no artificial population of agents ready to fill each player position; the set of players *is* the population. As a result, revising players respond to the actual pure-strategy profile currently in use.

[10] Formally, when i is chosen to revise, $z_{t+1} = z_t^{i+}$ if $\Delta u_i(z_t) > 0$, and $z_{t+1} = z_t^{i-}$ if $\Delta u_i(z_t) < 0$.

[11] McKelvey and Palfrey (1995) defined the notion of a "quantal-response equilibrium" as a way of weakening the stark predictions of the Nash-equilibrium solution concept in normal-form games. Their main focus was a logit specification, leading to the notion of a "limiting logit equilibrium" as an equilibrium selection device. The use of quantal responses as part of a strategy-revision process builds upon earlier work by Blume (1993, 1995, 1997, 2003), Brock and Durlauf (2001), Blume and Durlauf (2001, 2003), Diermeier (2001), and Diermeier and Van Mieghem (2000). For instance, Blume (1993) departed from the earlier work of Young (1993a) and Kandori, Mailath, and Rob (1993) by supposing that the log likelihood ratio of a player's revised choice is a linear function of the expected payoff difference between the two available actions. Diermeier (2001, 2003), and Diermeier and Van Mieghem (2000) applied this idea to collective-action games.

the quantal response will be allowed to vanish so that $\Delta\tilde{u}_i(z_t)$ converges in probability to $\Delta u_i(z_t)$. Equivalently, $\Pr[z^{i-} \to z^{i+}] \to \frac{1}{n}$ if $\Delta u_i(z) > 0$ and $\Pr[z^{i-} \to z^{i+}] \to 0$ otherwise.

2.3. **Logit and Probit Specifications.** The results of the paper will be applied to two different specifications for f_i. The first, often applied to quantal responses, is the logit.[12] Specifically, the noisy payoff difference $\Delta\tilde{u}_i(z)$ is equal to $\Delta u_i(z)$ plus a logistic noise term with zero mean and variance $2(\pi\varepsilon)^2$.[13,14] This yields transition probabilities

$$\Pr[z^{i-} \to z^{i+}] = \frac{\exp(\Delta u_i(z)/2\varepsilon)}{n[1+\exp(\Delta u_i(z)/2\varepsilon)]} \text{ and } \Pr[z^{i+} \to z^{i-}] = \frac{1}{n[1+\exp(\Delta u_i(z)/2\varepsilon)]}. \quad (1)$$

A second random-utility specification is the probit, so that $\Delta\tilde{u}_i(z) \sim N\left(\Delta u_i(z), \varepsilon \operatorname{var}_i(z)\right)$ where $\operatorname{var}_i(z) > 0$.[15] The transition probabilities of interest take the form

$$\Pr[z^{i-} \to z^{i+}] = \frac{1}{n}\Phi\left(\frac{\Delta u_i(z)}{\sqrt{\varepsilon \operatorname{var}_i(z)}}\right) \text{ and } \Pr[z^{i+} \to z^{i-}] = \frac{1}{n}\left[1-\Phi\left(\frac{\Delta u_i(z)}{\sqrt{\varepsilon \operatorname{var}_i(z)}}\right)\right], \quad (2)$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. The probit specification is richer than the logit, since the variance of the noisy contribution incentive $\varepsilon \operatorname{var}_i(z)$ is allowed to depend on the state z.[16] As will be shown just below, this specification allows for correlation between different payoff elements, such as cost and benefits, the consideration of heteroscedastic payoff disturbances, and proves convenient when the collective action arises from the integration of several additive components.

2.4. **Ergodicity, Trees, and Exponential Cost.** For $\varepsilon > 0$, the strategy-revision process is ergodic. This ensures that there is a unique distribution ϖ_z across the state space Z satisfying $\varpi_z = \lim_{t\to\infty} \Pr[z_t = z]$, independent of any initial conditions. The objective here will be to characterise this ergodic distribution in the limit as $\varepsilon \to 0$, and hence the noise in the quantal response of a revising player vanishes.

Following earlier literature, a graph-theoretic technique will be used to characterise the ergodic distribution. A tree rooted at z is a directed graph (that is, a subset $h \subseteq Z \times Z$) such that each node $z' \neq z$ has a unique successor. All sequences of edges lead to z, which

[12]Anderson, Goeree, and Holt (1999) and Goeree and Holt (2000) used the logit to model noisy decision-making in binary-choice games. Their focus, however, was on symmetric mixed-strategy equilibria. Motivated by the counter-intuitive and counter-evidential properties of mixed equilibria, they responded by retaining a restriction to symmetry, whilst employing the quantal-response equilibrium solution concept.

[13]The multiplicative constant, 2, is for expositional convenience (see the definition of Section 2.4).

[14]This specification yields "state-dependent mutations" in the sense of Bergin and Lipman (1996). With positive probability, a revising player i will choose $z_i = 1$ even though $\Delta u_i(z) < 0$; this is a "mutation." The probability of such a mutation will vanish to zero as $\varepsilon \to 0$. The mutation rate is the rate at which this probability vanishes. Mutations are state dependent when this rate varies with z, as they do here.

[15]Blume (2003) found that the usual risk-dominance selection result is robust to distributions of payoff differences that are "skew-symmetric". A probit need not fall into this category (Myatt and Wallace, 2004).

[16]It would be more general to specify $\Delta\tilde{u}_i(z) \sim N\left(\Delta u_i(z), \operatorname{var}_i(z,\varepsilon)\right)$ where $\operatorname{var}_i(z,\varepsilon) \to 0$ as $\varepsilon \to 0$. This leads, however, to insufficient structure on the implicit mutation rates. By manipulation of the shape of the dependence of $\operatorname{var}_i(z,\varepsilon)$ on ε, any selection result may be obtained (Bergin and Lipman, 1996).

has no successor. The collection of trees rooted at z is H_z. The following lemma is obtained from Freidlin and Wentzell (1998).

Lemma 1. *ϖ satisfies $\varpi_z = q_z / \sum_{z' \in Z} q_{z'}$, where $q_z = \sum_{h \in H_z} \prod_{(s,s') \in h} \Pr[s \to s']$.*

The relative likelihood of two states z and z' may be assessed by considering the ratio $q_z/q_{z'}$. The problem is that the expression in Lemma 1 is very complicated. However, as the limit $\varepsilon \to 0$ is taken, only certain trees will matter, greatly simplifying the calculations.

The next task is to measure the importance of rooted trees using the concept of "exponential cost" (Myatt and Wallace, 2003). To motivate this idea, consider a state z where the contribution incentive of player i is strictly negative, so that $\Delta u_i(z) < 0$. A transition from z^{i-} up to z^{i+} (so that the previously inactive player i begins to contribute) has probability

$$\frac{1}{n}\Phi\left(\frac{\Delta u_i(z)}{\sqrt{\varepsilon \operatorname{var}_i(z)}}\right) = \frac{1}{n} \times \frac{1 - \Phi(|\Delta u_i(z)|/\sqrt{\varepsilon \operatorname{var}_i(z)})}{\phi(|\Delta u_i(z)|/\sqrt{\varepsilon \operatorname{var}_i(z)})} \times \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\Delta u_i(z)^2}{2\varepsilon \operatorname{var}_i(z)}\right), \quad (3)$$

where $\phi(\cdot)$ is the density of the standard normal. This transition probability converges to zero as $\varepsilon \to 0$. Interest here lies in the *rate* at which it vanishes to zero. The second term on the right-hand side of (3) is the reciprocal of the hazard rate of a normal distribution. As $\varepsilon \to 0$, the argument of the hazard rate approaches infinity ($|\Delta u_i(z)|/\sqrt{\varepsilon \operatorname{var}_i(z)} \to \infty$). However, the hazard rate of the normal distribution is asymptotically linear in its argument.[17] That is, as $\varepsilon \to 0$ the second term of (3) converges to $[|\Delta u_i(z)|/\sqrt{\varepsilon \operatorname{var}_i(z)}]^{-1}$. The final term, however, is exponential in ε, and so will dominate in the limit. This means that, for very small ε, the transition probability behaves like $\exp(-\mathcal{E}/2\varepsilon)$, where $\mathcal{E} = \Delta u_i(z)^2 / \operatorname{var}_i(z)$. The larger $\mathcal{E}$ is, the faster the rate at which the transition probability vanishes to zero. Calling $\mathcal{E}$ the "exponential cost" of this particular transition, the following is a formal definition, adapted from Myatt and Wallace (2003, Definition 2).

Definition. *The* exponential cost $\mathcal{E} \in \mathbb{R}_+ \cup \{\infty\}$ *of a continuous function $g(\varepsilon)$ satisfies: (i) if $g(0) > 0$ then $\mathcal{E} = 0$; (ii) if $g(\varepsilon) = 0$ for all ε sufficiently small then $\mathcal{E} = \infty$; and (iii) if $g(0) = 0$ but $g(\varepsilon) > 0$ for all $\varepsilon > 0$ but sufficiently small (so that $g(\varepsilon) \to 0$ as $\varepsilon \to 0$) then the function has exponential cost $\mathcal{E} > 0$ if, for arbitrarily small $\alpha > 0$,*

$$\lim_{\varepsilon \to 0} g(\varepsilon) \exp\left(\frac{\mathcal{E} + \alpha}{2\varepsilon}\right) = \infty \quad \textit{and} \quad \lim_{\varepsilon \to 0} g(\varepsilon) \exp\left(\frac{\mathcal{E} - \alpha}{2\varepsilon}\right) = 0.$$

This property is denoted $g(\varepsilon) = \tilde{o}(\mathcal{E})$. Alternatively, write $\mathcal{E}(g(\cdot)) = \mathcal{E}$.

Hence $g(\varepsilon)$ has exponential cost $\mathcal{E}$ if it behaves as $\exp(-\mathcal{E}/2\varepsilon)$ does when ε vanishes. The exponential-cost property has parallels with the standard $o(\cdot)$ and $O(\cdot)$ notation familiar from the asymptotic behaviour of functions and sequences. Notice that exponential cost is not defined for all functions g, and the behaviour of $g(\varepsilon)\exp(\mathcal{E}/2\varepsilon)$ is undefined.

[17] This is a critical feature of the distribution chosen for payoff noise. Any other distribution with hazard rates that behave less than exponentially in the limit will be amenable to this sort of analysis.

Lemma 2. *For a set of exponential costs* $\{\mathcal{E}_l\}$*, the following properties hold:*

$$\prod_l \tilde{o}(\mathcal{E}_l) = \tilde{o}\left(\sum_l \mathcal{E}_l\right), \quad \textit{and} \quad \sum_l \tilde{o}(\mathcal{E}_l) = \tilde{o}\left(\min_l \mathcal{E}_l\right), \quad \textit{and} \quad a \times \tilde{o}(\mathcal{E}) = \tilde{o}(\mathcal{E}).$$

Further, taking ratios of functions of ε*:* $\mathcal{E}_l > \mathcal{E}_{l'} \Rightarrow \lim_{\varepsilon\to 0}[\tilde{o}(\mathcal{E}_l)/\tilde{o}(\mathcal{E}_{l'})] = 0$.

By inspection of (1), exponential cost may be applied to the logit as well as the probit.

Lemma 3. *Consider a strategy-revision process with either logit or probit responses. If* $\Delta u_i(z) > 0$ *then* $\mathcal{E}(\Pr[z^{i-} \to z^{i+}]) = 0$*. If* $\Delta u_i(z) < 0$*, then* $\mathcal{E}(\Pr[z^{i-} \to z^{i+}]) = |\Delta u_i(z)|$ *for logit quantal-response and* $\mathcal{E}(\Pr[z^{i-} \to z^{i+}]) = (\Delta u_i(z))^2/\operatorname{var}_i(z)$ *for probit quantal response.*

The next section turns to the task of finding the exponential costs of rooted trees for the dynamic of interest. With the properties outlined in Lemma 2, it is possible to infer that only the least exponential-cost rooted trees matter as $\varepsilon \to 0$.

2.5. **Minimal Exponential-Cost Trees and Equilibrium Selection.** Define $\mathcal{E}_{zz'} = \mathcal{E}(\Pr[z \to z'])$, the exponential cost of a transition from z to z'. From Lemma 3, exponential cost is defined for all transitions under logit and probit specifications.[18] Recalling the definition of a tree $h \subseteq Z \times Z$ used in Lemma 1, and abusing notation somewhat, write the exponential cost of such a tree as $\mathcal{E}_h = \sum_{(z,z')\in h} \mathcal{E}_{zz'}$. This is simply the exponential cost of the product of the transition probabilities taken from the branches of the tree.

Lemma 1 may now be employed, and the ergodic distribution examined as $\varepsilon \to 0$. First, only the least-cost rooted trees for each state z are of importance in the limit. To see this, use the properties of exponential cost from Lemma 2:

$$\mathcal{E}(q_z) = \mathcal{E}\left(\sum\nolimits_{h\in H_z}\prod\nolimits_{(s,s')\in h} \Pr[s \to s']\right) = \min_{h\in H_z} \mathcal{E}\left(\prod\nolimits_{(s,s')\in h} \Pr[s \to s']\right) = \min_{h\in H_z} \mathcal{E}_h.$$

Second, if there is a tree with a root at state z that has a lower exponential cost than any tree rooted at state z', then state z has infinitely more weight in the limit. Once again from Lemma 2: $\mathcal{E}(q_z) < \mathcal{E}(q_{z'}) \Rightarrow \lim_{\varepsilon\to 0}[q_{z'}/q_z] = 0$. Attention is focused, therefore, on the set of states $Z^\ddagger$ with least exponential-cost rooted trees, where

$$Z^\ddagger = \left\{ z \in Z \,:\, \min_{h\in H_z}\{\mathcal{E}_h\} \leq \min_{z'\in Z}\min_{h'\in H_{z'}}\{\mathcal{E}_{h'}\} \right\}.$$

Allowing $\varepsilon \to 0$, these states are "selected" in the ergodic distribution.

Lemma 4. *States in* $Z^\ddagger$ *attract all probability in the limit:* $\lim_{\varepsilon\to 0} \sum_{z\in Z^\ddagger} \varpi_z = 1$.

Lemma 4 allows selection to take place. If $Z^\ddagger$ consists of a single state corresponding to a pure-strategy Nash equilibrium, then (for sufficiently small ε) that equilibrium will be played almost always in the long run. Finding such an equilibrium involves a search for the tree

[18] For completeness, for two states $z' \neq z$ that differ by more than one player's contribution, set $\mathcal{E}_{zz'} = \infty$.

with the lowest (exponential) cost, following Young (1993a) and Kandori, Mailath, and Rob (1993). The difference here is that the cost of each non-best-response is no longer a constant.

3. Threshold Games and Equilibrium Selection

For many specifications, collective-action success will depend on the number of contributors (Schelling, 1973, 1978). For this reason, it is helpful to partition the state space Z into a collection of "layers." To this end, the kth layer of the state space is $Z_k = \{z : |z| = k\}$, where $|z| \equiv \sum_{i=1}^{n} z_i$. These are the states in which exactly k players contribute.

3.1. **Threshold Games.** One specification in which the outcome depends upon the number of contributors is the threshold game of Palfrey and Rosenthal (1984): by contributing, player i pays a private cost $c_i > 0$; this results in public-good provision, and hence a benefit $v_i > c_i$, if and only if m or more players participate, where $m \geq 2$.[19] Formally,

$$u_i(z) = \mathcal{I}\left[|z| \geq m\right] v_i - z_i c_i \quad \Rightarrow \quad \Delta u_i(z) = \mathcal{I}\left[z^{i+} \in Z_m\right] v_i - c_i, \tag{4}$$

where $\mathcal{I}[\cdot]$ is the indicator function. Since $m \geq 2$, a single contribution involves a cost without a benefit. Hence it is a Nash equilibrium for no player to contribute. On the other hand, if exactly m players contribute then further contributions generate no extra benefit, and hence a non-contributing player would not wish to participate. Likewise, a participant would not wish to stop contributing, as to do so would reduce their payoff by $v_i - c_i > 0$. So, there are $\binom{n}{m} + 1$ pure-strategy equilibria, and hence an equilibrium-selection problem.

The specification of (4) implies that any unused (when $0 < |z| < m$) or excess (when $|z| > m$) contributions are lost. Other specifications are possible. Such contributions might be fully or partially returned. When $|z| > m$, the m contributors required might be picked at random from those who volunteer. These specifications are all considered in Section 5.

3.2. **Equilibrium Selection.** The following specification generates a random-utility model of the kind discussed in Section 2.3: $\tilde{c}_i \sim N(c_i, \varepsilon\xi_i^2)$ and $\tilde{v}_i - \tilde{c}_i \sim N(v_i - c_i, \varepsilon\kappa_i^2)$. The mean contribution incentive is $\Delta u_i(z) = v_i \mathcal{I}(z^{i+} \in Z_m) - c_i$. At z the variances are

$$\mathrm{var}_i(z) = \begin{cases} \kappa_i^2 & \text{if } \sum_{j \neq i} z_j = m - 1, \\ \xi_i^2 & \text{if } \sum_{j \neq i} z_j \neq m - 1. \end{cases}$$

Transition probabilities follow from (2). The relevant exponential costs are as follows.

Definition. *For the probit case, the* birth *and* death *costs of player i are, respectively,*

$$\beta_i = \frac{c_i^2}{\xi_i^2}, \quad \textit{and} \quad \delta_i = \frac{(v_i - c_i)^2}{\kappa_i^2}.$$

[19] The benefit v_i does not depend upon the providers' identities; any "warm glow" from provision (Bernheim, 1986; Bergstrom, Blume, and Varian, 1986) is incorporated into c_i. Setting $m \geq 2$ implies that genuine teamwork is required for success; the $m = 1$ case is considered by Myatt and Wallace (2005a).

For the logit case, these birth and death costs are $\beta_i = c_i$ *and* $\delta_i = v_i - c_i$*. Without loss of generality, it is assumed that* $\beta_1 < \beta_2 < \cdots < \beta_n$*. Call* $z^\ddagger = \{z \in Z_m : z_i = 1\,,\, i \leq m\}$*.*

A player's birth cost is the exponential cost of a transition which involves that player choosing to contribute when there *are not* exactly $m-1$ other players contributing. A player's death cost is the exponential cost of a transition which involves that player choosing not to contribute when there *are* exactly $m-1$ other players contributing. In other words, they are the exponential costs of choosing a strategy "against the flow of play".

Ordering the players so that player 1 finds it easiest to choose against the flow of play, then player 2, and so on, motivates the definition of state $z^\ddagger$. This is the state in which the "most efficient" m players successfully form a team (for example, and for the probit case, when ξ_i is constant across players, this would be the team that involves the m players with the lowest costs, c_i, participating). Lemma 5 formalises the relationship between the exponential cost of moving from state z to some state $z' \neq z$ and the notions of birth and death cost.

Lemma 5. *Suppose* $z' \neq z$*. If there is no* i *s.t.* $z' = z^{i+}$ *or* $z' = z^{i-}$ *then* $\mathcal{E}_{zz'} = \infty$*. Else,*

$$z' = z^{i+} \Rightarrow \mathcal{E}_{zz'} = \begin{cases} \beta_i & z \notin Z_{m-1}, \\ 0 & z \in Z_{m-1} \end{cases} \quad \textit{and} \quad z' = z^{i-} \Rightarrow \mathcal{E}_{zz'} = \begin{cases} \delta_i & z \in Z_m, \\ 0 & z \notin Z_m. \end{cases}$$

Interest lies in the relative likelihood of two important states. The first is $z^\ddagger$, where a team comprising the players with the m lowest birth costs successfully operates. The second is z^0, the unique element in Z_0, where no player contributes. A state z is said to be "selected" if

$$\lim_{\varepsilon \to 0} \left[\lim_{t \to \infty} \Pr[z_t = z] \right] = 1.$$

Theorem 1 is concerned with which of these two states ($z^\ddagger$ and z^0) is selected. From Lemma 4, if $\mathcal{E}(z') < \mathcal{E}(z)$ for all $z \neq z'$, then z' is selected.[20] With this in mind, define $\iota = \arg\min_{i \leq m} \delta_i$ (the player within the "most efficient" team who has the lowest death cost).

Theorem 1. *If* $\delta_\iota \geq \min\{\max_{i>m} \delta_i, \beta_\iota\}$*, then* $\min\{\mathcal{E}(z^0), \mathcal{E}(z^\ddagger)\} \leq \mathcal{E}(z)\ \forall z \in Z$*, and*

$$\mathcal{E}(z^\ddagger) < \mathcal{E}(z^0) \quad \Longleftrightarrow \quad \sum_{i=1}^{m-1} \beta_i < \min_{j>m} \{\beta_j + \delta_j - \min(\beta_m, \delta_j), \delta_\iota\}\,.$$

The first part of the theorem establishes a sufficient condition for one of the two states of interest ($z^\ddagger$ and z^0) to have positive weight in the ergodic distribution. Given this, the second part of the theorem offers a necessary and sufficient condition for which of these states is selected. It is clear that any state not in Z_0 or Z_m will not have positive weight in the limit. All such states have zero-cost exits (see Lemma 5) and are therefore very easy to escape from, whereas states in Z_0 or Z_m (corresponding to pure-strategy Nash equilibria) only have positive-cost exits. As a result, trees rooted at such states will have lower exponential costs.

[20] This further abuse of notation is natural: $\mathcal{E}(z)$ is the exponential cost of the least-cost tree rooted at z.

Now consider states in Z_0 and Z_m. Suppose, for the moment, that $\mathcal{E}(z^\ddagger) < \mathcal{E}(z^0)$. The sufficient condition requires that either (i) the sequence $(\delta_1, \ldots, \delta_n)$ satisfies a weak single-crossing property — the smallest of the death costs *within* the most efficient team must be larger than the largest death cost *without* the most efficient team, or that (ii) the smallest death cost in the most efficient team is larger than the birth cost of that player.

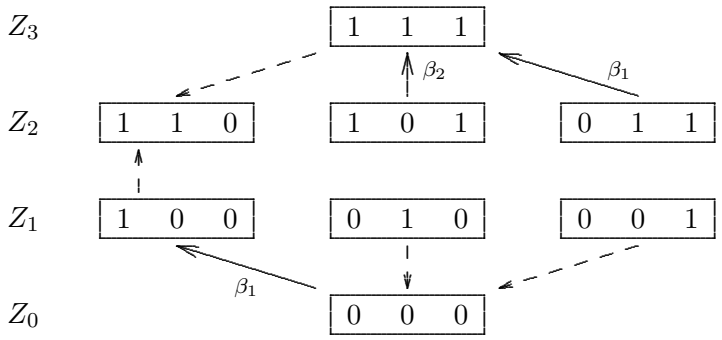
The former requirement (i) is easy to interpret: not only do the first m players have lower birth costs (by assumption), they also have higher death costs: the most efficient team is easiest to form and hardest to kill. The latter requirement (ii) states that it is cheaper to birth the "most-likely-to-die" player (within the most efficient team) than it is to kill them.

Suppose that one of these requirements is met, and consider the relative likelihood of states in Z_m. The cheapest way out of $z^\ddagger$ is either to kill player ι or birth player $m+1$ (who has the lowest birth cost of all the players not in the team). Now consider some state $z \neq z^\ddagger$ in Z_m. How easy is it to escape from this state relative to $z^\ddagger$? Suppose player ι is not in the team ($z_\iota = 0$). If so, birthing player ι costs β_ι. Even if this is not the cheapest way out, it is (weakly) cheaper than the cheapest way out of $z^\ddagger$ (since $\beta_\iota \leq \delta_\iota$ and $\beta_\iota < \beta_{m+1}$ by assumption). Now suppose that player ι is in the team (so that $z_\iota = 1$). Exiting this state as cheaply as possible will involve a death (at a cost of at most δ_ι) or a birth (at a cost of $\beta_j < \beta_{m+1}$, since there will be at least one player $j < m+1$ not in the team: $z \neq z^\ddagger$). Clearly, the smaller of these two numbers is less than the cost associated with the least-cost exit from $z^\ddagger$. Hence, even if this is not the least-cost exit from z, it is still (weakly) lower than the least-cost exit from $z^\ddagger$. It is always more difficult (higher cost) to exit the state containing the most efficient players than any other in Z_m.

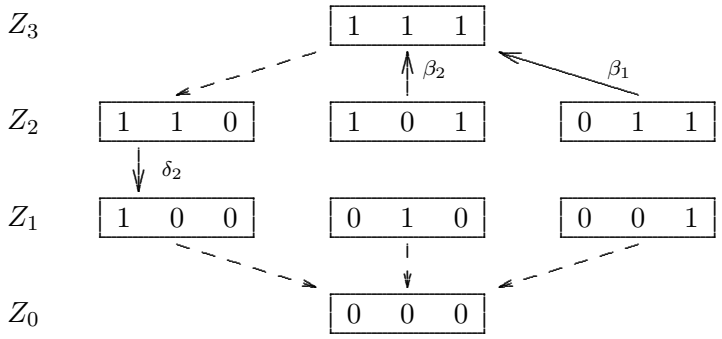
The above argument continues to hold if $\mathcal{E}(z^0) \leq \mathcal{E}(z^\ddagger)$. Whichever of these two states has the lower exponential-cost rooted tree, $\min\{\mathcal{E}(z^0), \mathcal{E}(z^\ddagger)\} \leq \mathcal{E}(z)$ for all other states $z \in Z$. Selection will then depend upon which of these two critical states has the lower associated exponential cost. The second part of Theorem 1 reveals the appropriate condition.

The logic behind this part of the theorem is as follows: first, in order to reach $z^\ddagger$ from z^0, $m-1$ costly steps must be taken. The cheapest way to do this is to birth the first $m-1$ players (who are the players with the lowest birth costs). From the state in which only these $m-1$ players contribute, there is then a zero exponential-cost step where player m joins in. Hence the total cost of transiting from z^0 to $z^\ddagger$ is $\sum_{i=1}^{m-1} \beta_i$. On the other hand, the cheapest way to get to z^0 from $z^\ddagger$ might involve a death. If so, it will surely be the team player with the lowest death cost, player ι, at a cost of δ_ι. If not, it must involve a birth.

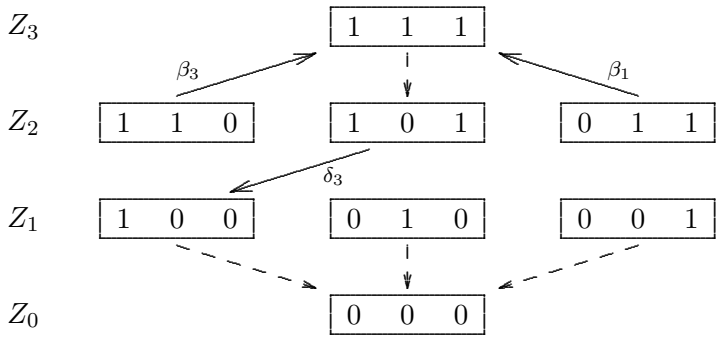
In this latter case, a birth leads the process into Z_{m+1}. The process must pass through a state in Z_m in order to return to the 0th layer. Thus there must be a death (at zero cost, since the process is moving from the $(m+1)$th layer to the mth layer). From this new state (which involves exactly m contributors), the process must again exit. Since only one costly death is required to move down a layer, there would be no sense in a path that required yet



(a) Minimum-Cost Tree Rooted at $z^{\ddagger}$



(b) Tree Rooted at z^0 via "Instant Death" Exit



(c) Tree Rooted at z^0 via "Bad Apple" Exit

The figures illustrate the rooted trees of interest for a threshold game with $n = 3$ and $m = 2$. The boxes represent states $z \in Z$, in which the binary entries are the players' actions. Dashed arrows are zero-cost transitions. The first figure displays a tree rooted at $z^{\ddagger}$: a birth cost of β_1 is needed to build the team. The second figure shows a tree rooted at z^0: the birth cost β_1 is removed, and a death cost of δ_2 is incurred for the team to collapse. The team would be successful if $\beta_1 < \delta_2$. The third figure shows an alternative exit from $z^{\ddagger}$: the "bad apple" player 3 supplants player 2, and subsequently exits. This is a less costly way to exit if $\delta_3 + \beta_3 - \beta_2 < \delta_1$. Hence, the team is successful if $\beta_1 < \min\{\delta_1, \beta_3 + \delta_3 - \beta_2\}$, as in Theorem 1.

FIGURE 1. Explaining the Proof to Theorem 1

another birth and death — it would have higher cost, and the player so acquired could have been acquired at the first birth. To recap, the least-costly exit must involve the birth and then death of some player j not in the most efficient team. The additional cost involved in this process justifies the expression $\{\beta_j + \delta_j - \min(\beta_m, \delta_j)\}$ in the theorem (see Figure 1).[21] Of course, only the smallest of these costs matters for the least-cost rooted tree, and so a minimum is taken across all players who potentially may be birthed at $z^\ddagger$: all players $j > m$.

3.3. **Sufficient Conditions for Success and Failure.** At this point, it might be noted that much of this argument does not depend upon the $\delta_\iota \geq \min\{\max_{i>m} \delta_i, \beta_\iota\}$ condition. Indeed, the minimum cost of a tree rooted at *any* state in Z_m has an upper bound provided by the exponential cost of the least-cost tree rooted at $z^\ddagger$ (an obvious point since $z^\ddagger \in Z_m$!) Thus if the least-cost tree rooted at z^0 has a higher cost than this upper bound, it is certain that z^0 will not be selected. A sufficient condition for collective-action success can be deduced.

Proposition 1. *For all $z \notin Z_m$, there is some $z^m \in Z_m$ such that*

$$\sum_{i=1}^{m-1} \beta_i < \min_{j>m} \{\beta_j + \delta_j - \min(\beta_m, \delta_j), \delta_\iota\} \quad \Rightarrow \quad \mathcal{E}(z^m) < \mathcal{E}(z).$$

Likewise, a sufficient condition for team failure is available. Here, the logic is clear. Take the state in Z_m that has the highest possible death cost. This will be the state where the m highest death-cost players contribute. The cheapest way out of such a state might be a death — in which case it has an exponential cost equal to the lowest death cost of the players in the team, which is the mth largest death cost in the population, denoted $\delta_{(m)}$. On the other hand, the cheapest way out could be a birth. The highest possible minimum birth cost from any state is β_{m+1} (attained at state $z^\ddagger$). This would require $\beta_{m+1} < \delta_{(m)}$, implying that $\beta_j < \delta_{(m)}$ for all $j \leq m$. The cheapest way to get from z^0 to Z_m still involves the birth of the $m-1$ lowest-cost contributors. Now, if the highest possible minimum-cost exit from Z_m is cheaper than the lowest possible minimum-cost exit from z^0, z^0 will certainly be selected.

Proposition 2. *Define $\delta_{(i)}$ to be the ith largest of $(\delta_1, \ldots, \delta_n)$. Then*

$$\delta_{(m)} < \sum_{i=1}^{m-1} \beta_i \quad \Rightarrow \quad \mathcal{E}(z^0) < \mathcal{E}(z) \quad \forall z \neq z^0.$$

A simple case of interest is when the entire population is required to contribute in order to successfully form a team. That is, when $m = n$. In this instance, the two propositions may be applied immediately. Note that $\delta_{(n)} = \delta_\iota$, and that $Z_m = Z_n$ has a single member where all players contribute. The only way out of a successfully formed team is via a death, the cheapest of which costs δ_ι. The cheapest way to get from state z^0 to Z_n is via the birth of the $n-1$ lowest cost players. Corollary 1 states the implied selection result.

[21] The reader is referred to the appendicised proof for a detailed explanation of the term $\{\min(\beta_m, \delta_j)\}$.

Corollary 1. *If $m = n$, $z^{\ddagger}$ is selected if $\sum_{i=1}^{n-1} \beta_i < \delta_\iota$, and otherwise z^0 is selected.*

Returning to the general case of $2 \leq m \leq n$, Section 6 looks at some particular specifications for the birth-cost and death-cost parameters. This provides insight into what the results presented above mean for equilibrium selection in collective-action problems.

4. Correlation, Concavity, and Project Components

An investigation into the factors that influence the equilibrium-selection criterion begins here. This section asks how selection depends upon the costs and benefits of public-good provision; the variability and correlation of payoffs; and the team size required for success.

4.1. **Symmetric Threshold Games.** Here, attention is confined to a situation in which players are symmetric, so that the birth costs and death costs β_i and δ_i take the same values β and δ for all players. The condition of Theorem 1 is always satisfied, and hence the collective action succeeds (that is, $\lim_{\varepsilon\to 0}\lim_{t\to\infty}\Pr[z_t \in Z_m] = 1$) if and only if $(m-1)\beta < \delta$: it takes $m-1$ births to form a successful team, and a single death to destroy it.

Proposition 3. *For the threshold game with symmetric players evolving via probit quantal response, the collective action succeeds* ($\lim_{\varepsilon\to 0}\lim_{t\to\infty}\Pr[z_t \in Z_m] = 1$) *if and only if*

$$\frac{(v-c)^2}{\kappa^2} - (m-1)\frac{c^2}{\xi^2} > 0. \tag{5}$$

Under logit quantal response, the collective action succeeds if and only if $v > mc$.

Unlike the logit, the probit specification permits an investigation of the impact of payoff idiosyncrasy, and hence the focus of this section will be on the determinants of success under probit quantal response. Success will be said to be increasing in a parameter if the left-hand side of (5) is increasing in that parameter. By inspection, success increases with v and ξ^2, and decreases in c, m, and κ^2. These observations may, however, be misleading. For instance, an increase in the average contribution cost c may well be accompanied by an increase in its variance. Similarly, an increase in the critical team size m might also reduce the average cost of each individual contribution c. Such issues are addressed below.

4.2. **Value, Cost, and Correlation.** The stability of a team (that is, when $z \in Z_m$) will be determined by the variance of a member's contribution incentive $\mathrm{var}[\tilde{v} - \tilde{c}] = \varepsilon \times \kappa^2$. Of course, this variance depends upon the joint distribution of $\tilde{v}$ and $\tilde{c}$. Suppose, for instance, that $\tilde{v}$ and $\tilde{c}$ are joint normal, that $\mathrm{var}[\tilde{v}] = \varepsilon \times \sigma^2$, and that their correlation coefficient is ρ. Then $\kappa^2 = \sigma^2 + \xi^2 - 2\rho\sigma\xi$, and (5) becomes, following re-arrangement,

$$m < 1 + \left[\frac{v}{c} - 1\right]^2 \frac{\xi^2}{\sigma^2 + \xi^2 - 2\rho\sigma\xi}. \tag{6}$$

By inspection, team success is increasing in ρ. Positive correlation between value and cost reduces the variability of $\tilde{v} - \tilde{c}$. The players become more reliable in the sense that high-cost shocks are likely to be offset by high-valuation shocks, preventing the collapse of a successfully formed team. Clearly, negative correlation has the opposite effect. Further inspection of (6) reveals the effects of the variance terms σ^2 and ξ^2.

Result 1. *Team success is increasing in the correlation ρ between value and cost. Fixing ρ, if $\rho < \sigma/\xi$, success is increasing in ξ^2 and decreasing in σ^2; if $\rho > \sigma/\xi$, success is decreasing and increasing, respectively. When the covariance $\rho\sigma\xi$ is fixed, rather than ρ, success is decreasing in σ^2. It is increasing in ξ^2 so long as $\sigma^2 > 2\rho\sigma\xi$, otherwise it is decreasing.*

Hence, so long as the correlation between the cost and benefit is not too large, then success is enhanced by private costs (and hence any "warm glow" from the act of provision) that are relatively variable, and benefits from the public good that are relatively stable.

4.3. **Concavity and Critical Team-Size.** The criterion (6) is simplest when $\sigma^2 = 0$, so that payoff variability stems entirely from $\tilde{c}$; it simplifies to $m < 1 + [(v/c) - 1]^2$. Success is, of course, decreasing in the critical team size m, so long as v and c are fixed. Here a setting is considered in which the value of the public good depends on m, so that $v = v(m)$. The idea is that m measures the size and importance of a collective action; for larger m the public good requires more contributions, while generating greater value.

Given this approach, the range of critical team-sizes for which (6) holds depends on the exact shape of $v(m)$. Consider the specification $v(m) = m^\alpha$, so that (6) becomes

$$m^\alpha > c\left[1 + \sqrt{m-1}\right]. \tag{7}$$

Inspection of this inequality suggests that the effect of m depends upon whether or not $\alpha < \frac{1}{2}$. If so, as m increases the right-hand side increases more quickly than the left-hand side; the production technology is "too concave" and large teams fail. Although a large team results in a high-value public good, and thus a successful team is relatively stable, the initial formation of this team is far less likely. When $\alpha > \frac{1}{2}$, the opposite is true.

Result 2. *Suppose that $\sigma^2 = 0$ and $v = m^\alpha$. If $\alpha > \frac{1}{2}$ then the collective action succeeds for all m sufficiently large. If $\alpha < \frac{1}{2}$, then the collective action fails for all m sufficiently large.*

4.4. **Project Components and Correlation.** The assumption $\sigma^2 = 0$ used in Result 2 ensures that an increase in m can only enhance the stability of a successfully formed team via an increase in v. However, an increase in m plausibly might also increase σ^2. To explore this issue, the analysis is simplified by setting $\rho = 0$, so that $\kappa^2 = \sigma^2 + \xi^2$. Allowing σ^2 to depend on m yields the selection criterion

$$\frac{(v(m) - c)^2}{\xi^2 + \sigma^2(m)} > (m-1) \times \frac{c^2}{\xi^2}. \tag{8}$$

An increase in m now has three effects: it increases the number of steps needed to build a team; it increases $v(m)$ and hence the stability of a successful team; it increases the variability $\sigma^2(m)$ of team-members' payoffs, and hence reduces the stability of a successful team.

The criterion (8) is further explored via the following specification. Suppose that the public good is a combination of m components. An individual's valuation for the jth component is $\tilde{w}_j \sim N(\bar{v}, \varepsilon \times \bar{\sigma}^2)$, and $\text{cov}[\tilde{w}_j, \tilde{w}_k] = \varepsilon \times \gamma\bar{\sigma}^2$ for $k \neq j$. An individual's valuation for the public good is the sum $\tilde{v} = \sum_{j=1}^{m} \tilde{w}_j$, so that an increase in m corresponds to the addition of a new component to the bundle. This yields $v(m) = \bar{v}m$ and $\sigma^2(m) = \bar{\sigma}^2[\gamma m^2 + (1-\gamma)m]$. Observe that $\sigma^2(m)$ is increasing in γ and hence team success is decreasing in γ. As valuations for the components become more correlated, an individual's valuation for the integrated project becomes more variable.

It remains to see how an increase in m affects team success. This will depend upon whether γ is positive or zero; that is, whether or not the project components are correlated.

Result 3. *Suppose that $\rho = 0$, $v = \bar{v}m$, and $\sigma^2 = \bar{\sigma}^2[\gamma m^2 + (1-\gamma)m]$. If $\gamma = 0$, there exists m^* such that the team succeeds for all $m > m^*$ if $\bar{v}/\bar{\sigma} > c/\xi$, and fails for all $m > m^*$ if $\bar{v}/\bar{\sigma} < c/\xi$. If $\gamma > 0$ then there exist three thresholds $m_1^* \leq m_2^* \leq m_3^*$ such that the team is successful if $m < m_1^*$ or if $m_2^* < m < m_3^*$, and the team fails if $m_1^* < m < m_2^*$ or if $m > m_3^*$.*

To understand this result, inspect (8). Notice that the right-hand side (which represents the difficulty of building a successful team) increases linearly with m. Team success for large m will therefore depend upon how the left-hand side changes with m.

When $\gamma = 0$, valuations for the m components of the integrated public good are uncorrelated, and hence $\sigma^2(m) = m\bar{\sigma}^2$, which is linear in m. Thus, for large m, the left-hand side of (8) is $m(\bar{v}/\bar{\sigma})^2$. So, the difficulty of building a successful team (the $m-1$ steps required) and the stability of a team once so built both increase linearly with m. The success of larger teams depends upon whether $\bar{v}/\bar{\sigma}$ is bigger than c/ξ. The first of these terms indexes the probability $1 - \Phi(\bar{v}/\bar{\sigma}\sqrt{\varepsilon})$ that an individual dislikes a project component. Similarly, the second term indexes the probability $1 - \Phi(c/\xi\sqrt{\varepsilon})$ that an individual has a negative cost of contributing to the collective action (equivalently, any private benefit exceeds any private cost). If it is more likely that a player enjoys contributing than dislikes a component of the public good, then larger teams (of size $m > m^*$) are successful. If not, then there is an upper bound to the size of a successful team.

When $\gamma > 0$, both the mean and the standard deviation of an individual's valuation for the public good rise linearly with m as m grows large. This means that, for larger m, the left-hand side of (8) converges to a constant. However, increases in m always increase the number steps required to build a team and thus, for m large, the team must fail.

Results 2 and 3 can be related to the classic work of Olson (1968). He argued that "unless the number of individuals in a group is quite small ... rational, self-interested individuals

will not act to achieve their common or group interests." Result 2 shows that, so long as the production function is not too concave, an increased group size may assist collective action. On the other hand, Result 3 suggests that, when components of a larger public good are correlated, there will be an upper bound to the size of a successful team.

5. Refunds, Team Formation, and Subscriptions

In the Palfrey and Rosenthal (1984) threshold game, public-good provision depends upon the number of contributors, but the private cost borne by an individual contributor does not. This can be restrictive; for instance, a contributor might not be required to be pay the entire private cost if the collective action fails. Here, three alternative specifications are considered in which excessive or unused contributions are (partially) returned.

5.1. **Refundable Contributions and Team Formation.** Palfrey and Rosenthal (1984) introduced a second public-good provision game in which the private cost of provision is refunded if the collective action fails.[22] This may be combined with the regular threshold game to yield an encompassing "refundable contributions" game with specification

$$u_i(z) = \mathcal{I}\left[|z| \geq m\right](v_i - r_i z_i) - z_i c_i. \tag{9}$$

Here a contributor pays a (perhaps small) "transaction" cost of c_i. If more than m players volunteer, they each pay an additional "production" cost of r_i; however, if the collective action fails (when $|z| < m$) this production cost is not paid. Thus, the production cost r_i is a refundable element of the contribution. So long as $c_i > 0$ and $v_i > r_i + c_i > 0$ for all i, there are $\binom{n}{m} + 1$ pure-strategy Nash equilibria; the selection problem remains.[23]

As an alternative, consider a "team formation" game. The decision to contribute is interpreted as volunteering to participate in a team, and carries a transaction cost of $c_i > 0$. If m or more players volunteer then m of the volunteers are randomly chosen to form the team, and the remainder are sent home. A chosen team member then pays a cost r_i. (It may well be the case that $r_i < 0$, if team membership carries a private benefit.) This yields

$$u_i(z) = \mathcal{I}\left[|z| \geq m\right]\left(v_i - z_i \frac{m r_i}{|z|}\right) - z_i c_i, \tag{10}$$

[22]This followed van de Kragt, Orbell, and Dawes (1983). Similarly, Dawes, Orbell, Simmons, and van de Kragt (1986) studied the "relative effectiveness of ... modifying incentives in step-level dilemmas." One modification consisted of "ensuring players that they will not lose a contribution if the group effort fails."

[23]Setting $c_i = 0$ for each i ensures that states of play in layers below Z_m become weak Nash equilibria. This yields a game similar to that studied by Bagnoli and Lipman (1989), who allowed players to contribute any non-negative amount toward the provision of a public good. If combined contributions are insufficient for provision, then they are refunded. For undominated perfect equilibria, the public good is efficiently provided. A related framework was adopted by Admati and Perry (1991). They analysed a game in which two players alternate in making (sunk) contributions to a public project until completion. The unique subgame-perfect outcome is inefficient; however, when contributions are refunded in the event of failure, efficiency is restored.

where once again, if $c_i > 0$ and $v_i > r_i + c_i > 0$, there are $\binom{n}{m} + 1$ pure-strategy equilibria. This game involves the (partial) refund of contributions when the team fails ($|z| < m$) and the refund of excessive contributions when the team succeeds ($|z| \geq m$).

Whilst Theorem 1 applies only to the threshold game, its proof (Appendix A) is stated broadly enough to encompass the refundable-contributions and team-formation games. For each of these, and for ease of exposition, the potentially refundable element r_i will be fixed for each player. For $\tilde{v}_i$ and $\tilde{c}_i$, a probit specification is first considered.

Inspecting (9) and (10), and for $|z^{i+}| < m$, the contribution incentive satisfies $\Delta u_i(z) = -c_i$. Thus, below the mth layer of the state space a birth cost satisfies $\beta_i = c_i^2/\xi_i^2$, as before. Turning to death costs, for $z^{i+} \in Z_m$, the contribution incentive for (9) and (10) satisfies $\Delta u_i(z) = v_i - r_i - c_i$. It follows that δ_i must be modified to $\delta_i = (v_i - r_i - c_i)^2/\kappa_i^2$.

It remains to consider births that move the process up beyond the mth layer. Recall that the easiest exit from $z^\ddagger$ might involve a transition up from Z_m and into Z_{m+1}. For the refund game, and for $|z^{i+}| = m+1$, the relevant contribution incentive satisfies $\Delta u_i(z) = -(c_i + r_i)$. For the team-formation game it is $\Delta u_i(z) = -(c_i + mr_i/(m+1))$. These birth costs differ from those that apply below the mth layer. This leads to the following definition.

Definition. *For the refundable-contribution and team-formation games, define the* exit-birth cost $\hat{\beta}_i$ *as the exponential cost of a transition from $z^{i-} \in Z_m$ up to $z^{i+} \in Z_{m+1}$ in the layer above via the addition of player i. This satisfies $\hat{\beta}_i = (c_i + r_i)^2/\xi_i^2$ for the refund case, and $\hat{\beta}_i = (c_i + mr_i/(m+1))^2/\xi_i^2$ for team formation, when play evolves by probit quantal response.*

Theorem 1 may now be modified: the term $\sum_{i=1}^{m-1} \beta_i$ continues to use the "ordinary" birth costs β_i, while remaining birth-cost terms are replaced by the exit-birth costs $\hat{\beta}_i$ above.

To understand the difference between these games and the threshold game, it is convenient to restrict attention to the symmetric case. Appropriately modified, Theorem 1 demonstrates that a team will succeed if and only if $(m-1)\beta < \delta$.

Proposition 4. *For the refundable-contributions and team-formation games with symmetric players evolving via probit quantal response, the collective action succeeds if and only if*

$$\frac{(v-c-r)^2}{\kappa^2} - (m-1)\frac{c^2}{\xi^2} > 0. \tag{11}$$

Under logit quantal response, the collective action succeeds if and only if $v > mc + r$.

These specifications include two distinct cost elements: a transaction cost of c, and a potentially refundable production cost r. In the threshold game, both costs are captured by c; to obtain an appropriate comparison, replace c with $c + r$ so that the selection criterion of Proposition 3 becomes $(v-c-r)^2/\kappa^2 > (m-1)(c+r)^2/\xi^2$. Clearly, this is more difficult

to satisfy than (11) above. Hence, a move from a threshold game to either a refundable-contributions or team-formation game has an identical effect: they both make building a team easier, but have no impact whatsoever on the stability of a successfully formed team.[24]

Similar logic applies for a logit quantal-response process. Replacing c with $c+r$ once again, the selection criterion for a threshold game would be $v > m(c+r)$, whilst for the alternative specifications (Propostion 4) it is $v > mc + r$.

5.2. **Refundable Subscriptions.** As a final specification, consider a situation in which an individual may contribute a financial subscription r towards the production of the public good.[25] Funds of mr are required for successful provision, and any excessive contributions are refunded. The (implicit) planner in charge of such a mechanism cannot ascertain the exact identity of the contributors, and hence any excess funds are divided equally between the entire group of players. This leads to a "refundable subscription" game, where

$$u_i(z) = \mathcal{I}\left[|z| \geq m\right]\left(v_i - \frac{mr}{n}\right) + \frac{|z|r}{n} - z_i(r + c_i). \tag{12}$$

A contributor pays a transaction cost c_i, plus a subscription of r. A player receives a $1/n$ share of the subscription fund $|z|r$, via the refund mechanism. If, however, there are m or more contributions, a player receives the benefit v_i from public-good provision, but loses a $1/n$ share of the funds mr spent. So long as $v_i > c_i + (1 + (m-1)/n)\, r$, the set of pure-strategy Nash equilibria is unchanged.

Theorem 1 may be applied directly to this refundable-subscription game. Inspecting (12), observe that the contribution incentive satisfies

$$\Delta u_i(z) = \left(v_i - \frac{mr}{n}\right)\mathcal{I}\left[z^{i+} \in Z_m\right] - \left(c_i + \frac{(n-1)r}{n}\right),$$

and so, following the logic used earlier, birth and death costs are simply

$$\beta_i = \frac{1}{\xi_i^2}\left(c_i + \frac{(n-1)r}{n}\right)^2 \quad \text{and} \quad \delta_i = \frac{1}{\kappa_i^2}\left(v_i - c_i - \frac{(n+m-1)r}{n}\right)^2.$$

In fact for the symmetric case, and setting $\kappa^2 = \xi^2$, the relevant selection criterion for a probit quantal response reduces to $\left(v - c - (n+m-1)r/n\right)^2 > (m-1) \times \left(c + (n-1)r/n\right)^2$. This inequality can be rearranged to yield

$$(v - c - r) - (c + r)\sqrt{m-1} > \frac{r}{n}\left[(m-1) - \sqrt{m-1}\,\right]. \tag{13}$$

For the threshold game, the relevant criterion would be for the left-hand side of (13) to be greater than zero. For $m > 2$, the right-hand hand of (13) is strictly positive. Therefore,

[24]When part of the contribution cost is moved from the transaction-cost element $\tilde{c}_i$ to the potentially refundable element $\tilde{r}_i$, then the variance terms ξ^2 and κ^2 might be expected to change. Thus, for this argument to work, it is implicitly assumed that all noise comes from either the transaction cost or from the benefit term; equivalently, there is no noise applied to the potentially refundable production element.

[25]The "subscription" terminology was used by Menezes, Monteiro, and Temimi (2001), and others.

the introduction of refundable subscriptions makes team success less likely! The reason is that although refundable subscriptions make a team slightly easier to build, they also make it much easier to destroy a successfully operating team. A departing team member will lose the benefit v of public-good provision, but will gain mr/n in refunded contributions.

Under a logit quantal-response specification, the introduction of refundable subscriptions yields a selection criterion $v > m(c + r)$. This is identical to the criterion for the threshold game. Thus, for the logit, refundable subscriptions make no difference whatsoever.[26]

6. Team Disruption and Bad Apples

Attention now turns to a final issue: how is the success of a collective action influenced by the set of potential contributors? This section shows that an otherwise-successful (symmetric) team can be disrupted by three different types of asymmetry. In each case, this is done by the introduction of a single "bad apple."

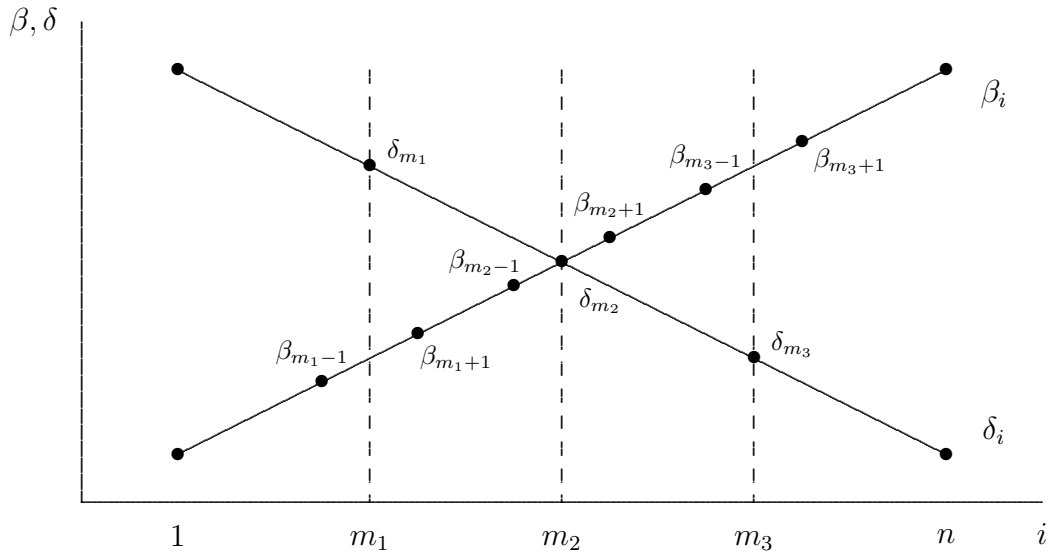
6.1. **Two Illustrative Examples.** Recalling the definition of birth and death costs in Section 3.2, two natural specifications for δ_i come to mind. Suppose, for example, that players differ only in their cost-of-contribution parameters c_i, so that all other parameters are constant across the population ($v_i = v$, and, for the probit case, $\xi_i = \xi$ and $\kappa_i = \kappa$ for all i). In this case, since $\beta_1 < \beta_2 < \cdots < \beta_n$ by construction, $\delta_1 > \delta_2 > \cdots > \delta_n$. An alternative specification (applying only to probit) would be to allow the population to differ according to their variances — more variable players have relatively high values of ξ_i and κ_i. This leads to the opposite situation where $\delta_1 < \delta_2 < \cdots < \delta_n$. Section 6.1.1 deals with the former, Section 6.1.2 deals with the latter.

6.1.1. *Decreasing Death Costs.* The first thing to notice is that, since the death costs are decreasing, it is always the case that $\delta_\iota = \delta_m > \delta_{m+1} = \max_{i>m} \delta_i$. As a result, the condition of Theorem 1 is met. Thus either $z^\ddagger$ is selected, or z^0 is selected, but never any other state in Z. The equation of the theorem is necessary and sufficient to examine selection.

[26] When $v_i = v$ (and $r_i = r$) and the payoffs are unperturbed, all of the collective-action games described here are (exact) potential games, in the sense of Monderer and Shapley (1996); Diermeier (2003), in a theory of consumer boycotts, made this observation for the symmetric case of Palfrey and Rosenthal's (1984) threshold game. Even when costs c_i (but not the benefits v_i) are asymmetric, the threshold game admits a potential function $\psi(z) = v\mathcal{I}[|z| \geq m] - \sum_{i=1}^{n} z_i c_i$. Applying the notation $\Delta_i\psi(z) = \psi(z^{i+}) - \psi(z^{i-})$, observe that $\Delta_i\psi(z) = \Delta u_i(z)$, so that the real-valued potential function $\psi(z)$ captures the strategic incentives of the game. Similar potential functions may be constructed for the other collective-action games of interest. The potentials for the refund and team-formation games are

$$\psi(z) = v\mathcal{I}[|z| \geq m] - \sum_{i=1}^{n} z_i c_i - r\mathcal{I}[|z| \geq m] \times \begin{cases} |z| - m + 1 & \text{refund game,} \\ m \sum_{k=m}^{|z|} \frac{1}{k} & \text{team-formation game.} \end{cases}$$

Finally, the potential function for the refundable-subscription game mirrors that of the threshold game: simply replace the value v with the term $v - mr/n$, and the cost c_i with $c_i + (n-1)r/n$. Myatt and Wallace (2005b) use the exact-potential feature of a general public-good provision game (focusing on a logit specification) to generate more general results away from the limit.



The players are ordered on the horizontal axis (from $i = 1$ to $i = n$). Birth and death costs appear on the vertical axis. By assumption, the birth costs are increasing in i, and the death costs are decreasing. There are three cases. For $m = m_1$, the death cost of interest (that is, δ_m) is greater than both the birth costs of interest. For $m = m_2$, the death cost lies between the infra-marginal and extra-marginal birth costs. Finally, for $m = m_3$, both birth costs exceed the relevant death cost.

FIGURE 2. Team Success when $\delta_1 > \delta_2 > \cdots > \delta_n$.

One possibility is that all the birth costs exceed the death costs. That is $\beta_1 > \delta_1$. In which case, $\delta_m < \beta_{m+1}$ and therefore $\delta_\iota < \beta_j + \delta_j - \min(\beta_m, \delta_j)$ for all j. The condition becomes

$$\sum_{i=1}^{m-1} \beta_i < \delta_m \quad \iff \quad \mathcal{E}(z^\ddagger) < \mathcal{E}(z^0), \tag{14}$$

so z^0 is selected (as $\beta_1 > \delta_m$). When all the birth costs exceed all the death costs, the team fails. This is quite straightforward — individual costs of contribution (normalised by their variability) are far too high relative to the benefits that would accrue (again normalised by variability). The more interesting possibility then, is where (at least) some of the death costs exceed the birth costs. Figure 2 illustrates the possible cases.

Consider first $m = m_3$. In this instance, the same logic applies as above (note that $\delta_m < \beta_{m+1}$, both indicated on the diagram), and the critical condition remains (14). The relevant comparison is between the cost of building the team and the lowest-cost exit from the team (which is the death of player m). As long as the latter exceeds the former, a successful team will form. This intuition will follow for any m such that $\delta_m < \beta_{m+1}$. Of course, for a team to succeed it is necessary that $\beta_{m-1} < \delta_m$, so in fact, at m_3, the team fails. If m lies in the neighbourhood of the crossing point (for example, $m = m_2$) then this necessary condition will be met, and the team may succeed.

A final case is illustrated for $m = m_1$. Here $\beta_{m-1} < \delta_m$ for sure. However, $\beta_{m+1} < \delta_m$ as well. It is now cheaper to exit the most efficient state via a birth, and not a death. The expression in (14) may no longer provide a necessary and sufficient condition. Notice that it can only become (weakly) *harder* to ensure a team's stability; the left-hand side of the statement (that is, $\sum_{i=1}^{m-1} \beta_i < \delta_m$) must still hold for the formation of a successful team. However, if there exists a player outside the team ($j > m$) such that $\beta_j + \delta_j - \min(\beta_m, \delta_j) < \delta_m$, the minimum of this former expression over j replaces δ_m in (14), see Theorem 1.

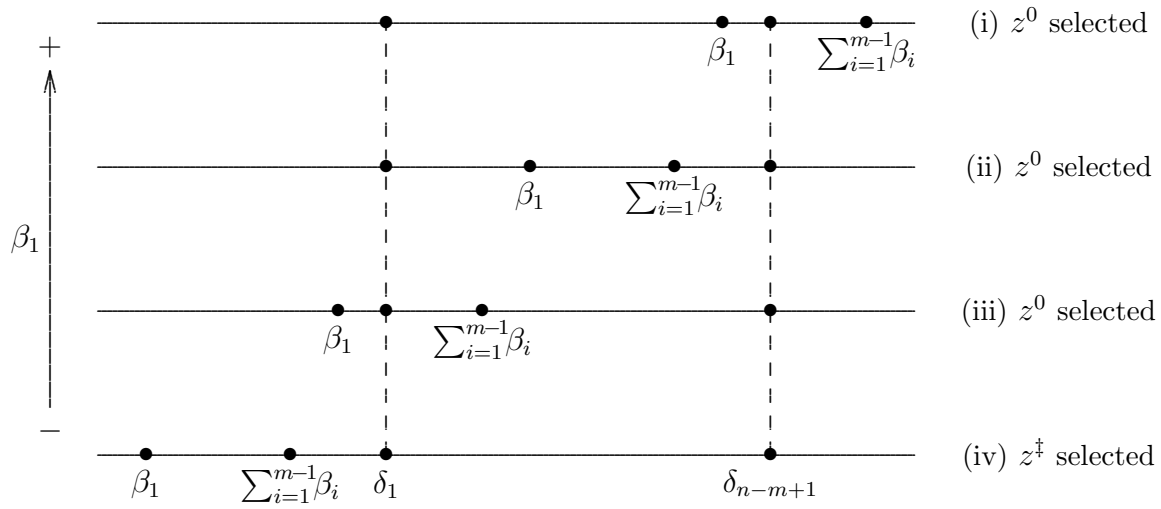
There need not exist such a player. In this case, the left-hand side of (14) remains the selection criterion and it is now relatively easy for a successful team to form. The relevant death cost is high, and each of the birth costs required for team formation is low. If there does exist such a player, however, there are several possibilities. First, notice that the relevant expression is bounded below by β_{m+1}. If there is a big enough jump between β_{m-1} and β_{m+1}, the team will succeed. This is when there are m "natural team players" (with low costs of participation) and the rest of the population is really of no consequence: they would never join a team and (potentially) disrupt it; their contribution costs are simply too high. If this jump is not sufficient, however, to guarantee successful team formation, the rest of the population do play a critical role.

In particular, the player $j > m$ with the lowest value of $\beta_j + \delta_j - \min(\beta_m, \delta_j)$ might be termed a "bad apple". Sections 6.2.1, 6.2.2, and 6.2.3 examine this idea in more detail. Meanwhile, it is easy to note here that the smallest possible value of this expression is β_{m+1}. The worst bad apple (perhaps the "rotten" apple) is therefore a player i who has a death cost of $\delta_i \leq \beta_m$ and birth cost β_m (the smallest possible value β_{m+1} can take, by construction). Although this player would never participate in a successful team, their birth cost is as low as it could be (given they are not in the most efficient team), and hence they find it rather easy to join. However, (following the zero-cost exit of another player $i \leq m$) the rotten apple finds it even easier to leave — hence disrupting the team.

6.1.2. *Increasing Death Costs.* The first thing to note in this case is that the term $\beta_j + \delta_j - \min(\beta_m, \delta_j)$ (with $j > m$) can be bounded below by δ_{m+1}. To see this, notice that the smallest values that β_j and δ_j can take are β_{m+1} and δ_{m+1} respectively. Next, notice that if β_m is smaller than δ_j for all $j > m$ then the term is bounded by $\beta_{m+1} + \delta_{m+1} - \beta_m \geq \delta_{m+1}$. If, on the other hand, there is some $j > m$ such that $\delta_j \leq \beta_m$ then the expression will only become larger. As a result $\beta_j + \delta_j - \min(\beta_m, \delta_j) \geq \delta_{m+1}$ for all $j > m$. Given that $\delta_\iota = \delta_1$, the expression on the left-hand side of the statement in Theorem 1 always takes the form

$$\sum_{i=1}^{m-1} \beta_i < \delta_1. \tag{15}$$

Suppose first that this inequality holds. This will certainly imply that $\beta_1 < \delta_1 < \delta_n$. Thus, the condition for Theorem 1 holds, and $z^\ddagger$ is selected. The inequality in (15) is sufficient



The figure illustrates four real lines, which are ordered by β_1. Two death costs are illustrated: $\delta_1 < \delta_{n-m+1}$. State z^0 is selected for cases (i)–(iii), and state $z^\ddagger$ is selected for case (iv). These cases are discussed in the main text.

FIGURE 3. Team Success when $\delta_1 < \delta_2 < \cdots < \delta_n$.

for team success. As long as it is cheaper to build the team than to destroy it, players successfully coordinate on the most efficient (in the sense of birth cost) equilibrium.

Now suppose that the inequality in (15) is reversed. Either it is still the case that $\beta_1 \leq \delta_1 < \delta_n$, or not. If so, the theorem continues to apply, z^0 is selected, and the team fails. If not, the theorem no longer applies. In this case, Proposition 2 is of some assistance. The mth largest value of δ_i is $\delta_{(m)} = \delta_{n-m+1}$. If this value is less than the sum of the lowest $m-1$ birth costs, the team will fail and z^0 is selected. However, this leaves a range of values for which additional analysis is required. Figure 3 illustrates the following corollary.

Corollary 2. *When $\delta_1 < \delta_2 < \cdots < \delta_n$, $z^\ddagger$ is selected if $\sum_{i=1}^{m-1} \beta_i < \delta_1$, and z^0 otherwise.*

The figure shows δ_1, δ_{n-m+1}, and the sum of the lowest $m-1$ birth costs for four different values of β_1. The bottom line, (iv), has the lowest value, so that $\beta_1 < \delta_1$. It is here that a successful team operates. The top line, (i), has the highest value of β_1, but it is not really player 1's fault that the team fails. The problem here is that the sum of costs (including β_1, of course) is simply far too large to allow *any* team to succeed, regardless of whether or not it contains player 1.

The second line from the bottom, (iii), illustrates the case where there is team failure precisely because of player 1's involvement in the team. Theorem 1 applies, and so the most efficient team would be the one selected were a team to successfully operate. However, the presence of player 1 in this team is fatal. The low death cost makes exit too easy — even though the player finds it easiest to join. Critically, it is easier to join than to exit for player 1; ensuring

participation in the (failing) team. Player 1 is, therefore, a bad apple who disrupts the team from within. The impact of such bad apples does not end here, however: it continues to have disastrous consequences *even when they would not be part of a successful team.*

Consider case (ii). Here player 1 causes the disruption of the team from without. Since Theorem 1 no longer applies, it is no longer certain that the $\mathcal{E}(z^{\ddagger}) \leq \mathcal{E}(z)$ for other $z \in Z_m$. Indeed, it is now perfectly possible that some other state has a lower exponential-cost rooted tree than the most efficient state. States in Z_m that exclude player 1 will certainly have *more* costly minimum-cost exits than states that do not.[27] This is because birth is more costly than death for player 1, and the minimum death cost from any state excluding player 1 is higher by definition. Such a state, in which m other players contribute, might stand a chance of success. By excluding the (perhaps over-excitable) bad apple, the rather more reliable (but perhaps higher-cost) team members may be able to coordinate successfully. However, the presence of the bad apple outside the team is enough to disrupt it. With such low birth costs, player 1 finds it easy to join in an already functioning team. Following the (zero-cost) exit of a reliable team member, player 1 (with such low death costs) then finds it very easy to exit. By persistently joining teams and then leaving them, the bad apple prevents all states with successfully operating teams from playing any role in the ergodic distribution, and z^0 is selected.

6.2. **Bad Apples.** In each of the following three sections, a single bad apple is introduced to a population in which previously, a successful team operated. Regardless of whether this additional player would play any role in the most stable team within the new population, the player's presence alone can be shown to destabilise the team and result in collective-action failure. Throughout this section, attention will be restricted to the probit quantal-response dynamic. For illustrative purposes, it is simplest to consider initially a symmetric situation with $\beta_i = c^2$ and $\delta_i = (v-c)^2$, and $\xi_i = \kappa_i = 1$, for all i. Further, suppose that, in the initially symmetric population, the team succeeds:

$$(v-c)^2 > (m-1)c^2 \quad \Longleftrightarrow \quad (v-c)^2 + c^2 > mc^2 \quad \Longleftrightarrow \quad \frac{v}{c} > 1 + \sqrt{m-1}, \qquad (16)$$

from Proposition 3. Three different asymmetries are considered. First, a high-cost bad apple with a cost parameter $c_i = \hat{c} > c$ is introduced, then a low-value bad apple with a value parameter $v_i = \hat{v} < v$, and finally a high-variance bad apple with $\xi_i = \kappa_i = \hat{\sigma} > 1$.

6.2.1. *The High-Cost Bad Apple.* Since $\hat{c} > c$, the bad apple will have the highest birth cost. Therefore $i = n$ in the notation of the paper. Since the death cost of this player is lower

[27]To tie down exactly which team this might be would require the construction of least-cost rooted trees for general $z \in Z_m$. This is a complex task that would not add much in the way of insight (beyond that which has already been offered). Instead, this issue is returned to in more detail following the example presented in Section 6.2.3. Myatt and Wallace (2005a) provides further discussion of this point when $m = 1$. For such games it *is* possible to find the least-cost rooted trees for any state $z \in Z_m$.

than the other players' death costs, $(v-\hat{c})^2 < (v-c)^2$, Theorem 1 applies. Indeed, this is an example of the sort considered in Section 6.1.1. The condition of the theorem becomes

$$(m-1)c^2 > \min\left\{\hat{c}^2 + (v-\hat{c})^2 - \min\{c^2, (v-\hat{c})^2\}, (v-c)^2\right\}$$

for the team to fail. The other terms $n > j > m$ in the first part of the right-hand side expression of Theorem 1 are all equal and larger (else the bad apple would have no impact). It is equally clear that if $\hat{c} \geq v - c$, the minimum on the left-hand side would reduce to $(v-c)^2$ and the bad apple would have no impact. Immediately then, it is apparent that the addition of a very high-cost bad apple would not have any effect upon the successful operation of the team. Such a player would play no role in team formation, and finds it too costly to join in the first place. As a result, the fact that it is very cheap for a bad apple to leave a functioning team is neither here nor there.

Suppose therefore that $\hat{c} \in (c, v-c)$. The team will fail if and only if $mc^2 > \hat{c}^2 + (v-\hat{c})^2$. If the player has such a cost parameter, $\hat{c}$, then upon entering a symmetric population (where $(v-c)^2 + c^2 > mc^2 > \hat{c}^2 + (v-\hat{c})^2$) the previously successful team will now fail *even though the player would never take part in a stable and successful team.* The bad apple is bad enough to be sufficiently easy to kill, but good enough to enter the team in the first place with sufficiently high probability. Minimising the sum of the bad apple's birth and death costs yields what might be termed the "rotten apple" with $\hat{c} = v/2$. If a rotten apple can't disrupt the team, then nobody can.[28] Figure 4 illustrates. If a team is very stable there are no disruptive high-cost bad apples. This will not be the case for low-value bad apples.

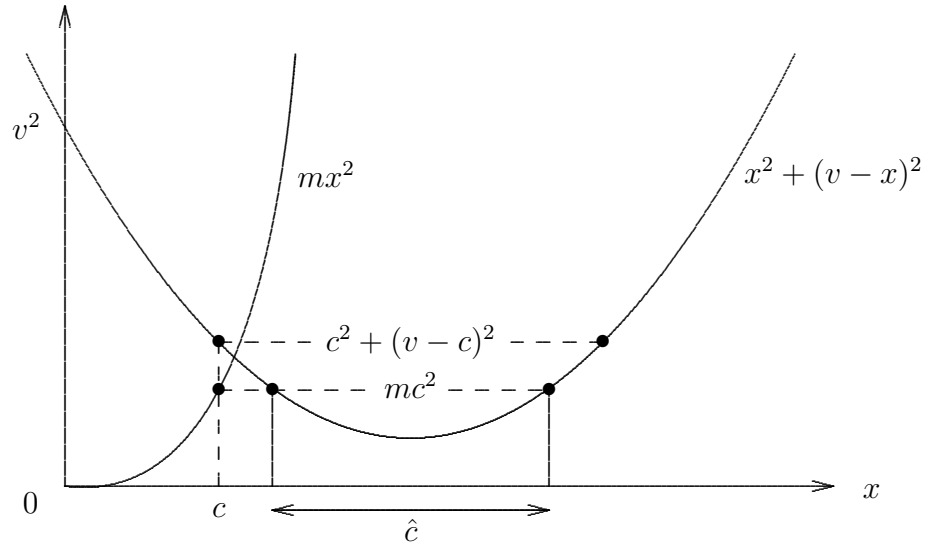
6.2.2. *The Low-Value Bad Apple.* Returning to a situation in which all players have the same cost parameter, consider introducing a player with a lower value than the original population, so that $v_i = \hat{v} < v$. The notation in use does not determine the ordering of the players — in fact $\beta_i = c^2$ for all i (including the new member of the population). It is possible, for example, to set $i = n$ and use Theorem 1. It is also possible to set $i = 1$ and use Corollary 2. Either way, the result will be the same (as might be hoped) and the low-value bad apple can disrupt a previously successful team from within or without. Applying the theorem, the condition for team failure post bad-apple introduction is

$$(m-1)c^2 > \min\left\{c^2 + (\hat{v}-c)^2 - \min\{c^2, (\hat{v}-c)^2\}, (v-c)^2\right\}. \tag{17}$$

[28]This will be trivially true when $m = 2$. Indeed, the smaller m, the more stable the team will be. When m reaches its minimum value, 2, it is no longer possible to find a bad apple that destabilises the team. In particular, the mx^2 curve in Figure 4 crosses the $x^2 + (v-x)^2$ curve at the latter's minimum. There is "no room" for a bad apple. It is straightforward to find bounds on c in terms of v and m for which it *is* possible to destabilise a successful team with at least a rotten apple. Some simple algebra confirms that if

$$c \in \left(\frac{v}{\sqrt{2m}}\,,\ \frac{v}{m-2}\left\{\sqrt{m-1}-1\right\}\right),$$

then there is an apple bad enough to destabilise the previously successful team.



Any value of $\hat{c}$ in the indicated region will destabilise the team. As c falls (or equivalently, v rises) the region indicated by the arrow shrinks until it is just the minimising value of the function $x^2 + (v-x)^2$, $v/2$, the so-called "rotten apple". Beyond this there is no bad-apple effect; a player added to the population with a higher cost parameter will have no impact whatsoever upon team formation.

FIGURE 4. The High-Cost Bad Apple with $c_i = \hat{c}$.

There are two ways in which a bad apple might have an impact. First, let $c < \hat{v} \leq 2c$, bearing in mind (16), so that $v > 2c$. In this case, $\min\{c^2, (\hat{v}-c)^2\} = (\hat{v}-c)^2$. The right-hand side of the above inequality becomes $\min\{c^2, (v-c)^2\} = c^2$. The inequality then holds and the team is disrupted. Note that c^2 is the cost of getting the bad apple *into* an already functioning team. Even if not in the team, the bad apples costs no more than any other player to birth, at which point such a player is extremely cheap to kill. The team is destabilised — from without. This is the equivalent of pushing δ_1 down to very low levels in Section 6.1.2 and applying Corollary 2; an example of case (ii) in Figure 3.

The second way in which the bad apple might have an impact is when $\hat{v} > 2c$. In this case, $\min\{c^2, (\hat{v}-c)^2\} = c^2$ and therefore the right-hand side of (17) becomes $\min\{(\hat{v}-c)^2, (v-c)^2\} = (\hat{v}-c)^2$. This is simply the cost of killing the bad apple. The team is disrupted from within, therefore, whenever $\hat{v} \in (2c, c\{1+\sqrt{m-1}\})$. Equivalently, in the language of Section 6.1.2; this is an example of case (iii) in Figure 3. Of course, if $\hat{v}$ is higher than this, the team continues to succeed despite the bad apple's presence. Summarising:

$$\hat{v} \in \begin{cases} (c, 2c] & \Rightarrow \quad \text{bad apple disrupts team from without,} \\ (2c, c\{1+\sqrt{m-1}\}) & \Rightarrow \quad \text{bad apple disrupts team from within,} \\ \big[c\{1+\sqrt{m-1}\}, v\big) & \Rightarrow \quad \text{no team disruption.} \end{cases}$$

In this case, no matter how stable was the previously successful team (in the sense of high v or low c), there will always exist an apple bad enough to overturn their efforts.[29]

6.2.3. *The High-Variance Bad Apple.* The final example presented here returns once again to the case where all players have the same cost and value parameters. This time, the bad apple will have higher variance than the initial population. Since it has been assumed $\xi_i = \kappa_i = 1$ for all i, consider the addition of a player with $\xi_j = \kappa_j = \hat{\sigma} > 1$. Ordering the players according to birth costs means that $\beta_i = c^2$ for $i > 1$, and $\delta_i = (v-c)^2$ for $i > 1$. Meanwhile $\beta_1 = c^2/\hat{\sigma}^2 < c^2$ and $\delta_1 = (v-c)^2/\hat{\sigma}^2 < (v-c)^2$. Therefore, the bad apple is player 1, and has a lower birth cost *and* a lower death cost than the other players. This is precisely the situation analysed in Section 6.1.2. Corollary 2 applies, and hence

$$\frac{c^2}{\hat{\sigma}^2} + (m-2)c^2 > \frac{(v-c)^2}{\hat{\sigma}^2} \quad \Rightarrow \quad \text{bad apple disrupts team.}$$

This can be rearranged so that the condition for the bad apple to disrupt the team becomes

$$\hat{\sigma} > \sqrt{\frac{v}{c} \times \frac{v/c - 2}{m-2}} > 1,$$

where the second inequality follows from the fact that the initial teams succeeds from (16). Note that $\delta_1 > \beta_1$ by assumption. So this bad apple always causes the disruption of the team from within, case (iii) in Figure 3. It would not be difficult to construct examples where high-variance bad apples can disrupt the team from without. The conditions required are explained in Section 6.1.2.

Notice again, that it will always be possible to find a $\hat{\sigma}$ large enough to destabilise the team (for given v and c).[30] No matter how stable the situation prior to the introduction of a bad apple, there will always exist a sufficiently variable player whose appearance in the population results in team disruption. This is precisely the case where an enthusiastic, over-excitable, and low-cost player enters the population; the collective action would stand a better chance if the population of potential contributors were restricted to high-cost, low-variance, and consistently reliable players — who will get the job done.

[29]It is worth noticing here that, even in the case $m = 2$, it is possible to destabilise any successful symmetric team via the addition of a player with a lower value than the other members of the population. However, when $m = 2$, although it initially seems that the team is disrupted from without, it does not really make sense to use this language here. The problem is that $\sum_{i=1}^{m-1} \beta_i = \beta_1$, and so cases (ii) and (iii) cannot be separated in the way suggested by Figure 3.

[30]Except when $m = 2$. In this instance, it is not really the case that the team is so much more stable for the smallest possible team-size. Rather, note that it is player 1 who now dominates the team-building and team-destroying exercise. Only the birth cost and death cost of player 1 matter for selection. Since $\kappa_1 = \xi_1 = \hat{\sigma}$ the critical inequality is unchanged. However, were this player to have *different* variance parameters on c and $v - c$, they might influence selection either way. In particular, it might be easier to satisfy the critical inequality of Corollary 2 — not so much a bad apple, as a "good egg". The same observation can be made when $m > 2$, although the good egg's impact is tempered by the need to get other players to contribute.

7. Some Concluding Remarks

Olson (1968) viewed a collective action as the private provision of a public good.[31] He argued that group members would not necessarily act in their collective interest; the "impossibility of exclusion" (Hardin, 1982) prevents capture of the positive externalities from provision. Sociologists, often motivated by experimental evidence (for instance, Marwell and Ames 1979, 1980), suggested that groups might develop mechanisms to induce contributions.[32]

One such mechanism involves the interdependence of contributions: when public-good provision requires the participation of a critical number of team members, an individual may be pivotal to a team's success. This is the "minimal contributing set" of van de Kragt, Orbell, and Dawes (1983) and Rapoport (1985).[33] The idea is neatly captured by the threshold game of Palfrey and Rosenthal (1984). Whilst some equilibria offer a resolution to the collective-action problem, there is another that induces failure.

The equilibrium-selection problem was (implicitly) recognised in sociological studies of "critical mass" models (Oliver and Marwell, 2001). Granovetter (1978), Oliver and Marwell (1988), Oliver, Marwell, and Teixeira (1985), and others, used numerical simulations, experiments, and informal theory to investigate the dynamics of groups involved in collective actions. One interpretation of the critical-mass idea is that a small number of initial contributions may raise the effectiveness of further contributions and induce mass participation. Heckathorn (1993, 1996), Macy (1990, 1991a,b), and others developed these ideas.

The approach taken here is related to this sociological work. Not only do average cost and value play a critical role in selection; their relative variability and the correlation between them are also of key importance. Teams succeed when there is positive correlation between the cost of contribution and the value players receive when the public good is provided. When the value received increases with the team size, it is not always true that larger teams are less stable. Large groups may have the advantage when value is "not too concave" in team size, or when the values generated by different contributions are uncorrelated. The format of the game matters; refundable contributions can sometimes help, sometimes hinder.

In asymmetric settings, a "bad apple" may disrupt a successful team. Such a player may have a higher average cost of contribution, a lower average valuation, or higher associated variances. Disruption may occur from within; the bad apple is a player who would have participated in a successful team, but whose presence alone is enough to tip the selection criterion the other way. More interestingly however, disruption may occur from *outside*

[31] In a comprehensive survey, Ostrom (2000) traced the development of the associated literature.

[32] For experimental evidence, see Ledyard (1995), Andreoni (1988), Cadsby and Maynes (1999), and others.

[33] van de Kragt, Orbell, and Dawes (1983, p. 113, original emphasis) explained: "Its essentials are very simple: *when the cost of the group's public good is assigned to a subset of members each of whose contribution is necessary to the production of the good, it is reasonable for each to contribute* ... [this logic] does not appear to have been specified in the literature on public-good provision." It may be optimal to rest the fate of a team upon the shoulders of each member; however, this works only if others participate as expected.

the team; even when the player would play no role in a successful team, the bad apple's destabilising presence forces the collective action to fail. On the other hand, it is equally possible to add a single "good egg" to a population (a player with a particularly low average cost of contribution, a high average value, or, most critically, low variance parameters) whose presence on the team results in its success and whose absence would be fatal.

Appendix A. Omitted Proofs

A.1. **Lemmas 1–5.** Lemma 1 follows Freidlin and Wentzell (1998, Chapter 6, Lemma 3.1). Lemma 2 follows Myatt and Wallace (2003, Lemma 2). Lemmas 3 and 4 follow from arguments in the text.

Proof of Lemma 5. Since the strategies are revised one step at a time, it must be the case that $z_{t+1} \in \{z_t^{i-}, z_t^{i+}\}$ for some i. Hence, if $z' \notin \{z^{i-}, z^{i+}\}$ for every i then $\Pr[z \to z'] = 0$ for all values of ε. From the definition of exponential cost, this implies that $\mathcal{E}_{zz'} = \infty$. Next, suppose that $z = z^{i-}$ and $z' = z^{i+}$ for some i. Focusing on probit quantal responses, if $z \in Z_{m-1}$ then

$$\Pr[z \to z'] = \frac{1}{n}\Phi\left(\frac{\Delta u_i(z)}{\sqrt{\varepsilon \operatorname{var}_i(z)}}\right) = \frac{1}{n} \times \begin{cases} \Phi\left((v_i - c_i)/\sqrt{\varepsilon \times \kappa_i^2}\right) & \text{if } z \in Z_{m-1}, \text{ and} \\ 1 - \Phi\left(c_i/\sqrt{\varepsilon \times \xi_i^2}\right) & \text{if } z \notin Z_{m-1}. \end{cases}$$

By inspection $\Pr[z \to z'] \to 1$ as $\varepsilon \to 0$ for $z \in Z_{m-1}$. This implies that $\mathcal{E}_{zz'} = 0$, as claimed. For $z \notin Z_{m-1}$, decompose the transition into a hazard rate and density (3) to obtain

$$\Pr[z \to z'] = \frac{1}{n} \times \frac{1 - \Phi\left(c_i/\sqrt{\varepsilon \times \xi_i^2}\right)}{\phi\left(c_i/\sqrt{\varepsilon \times \xi_i^2}\right)} \times \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{c_i^2}{2\varepsilon\xi_i^2}\right).$$

This tends to zero as $\varepsilon \to \infty$. However, as noted in the text following (3), the exponential term dominates in the limit. By inspection, this has exponential cost $\beta_i = c_i^2/\xi_i^2$.

It remains to consider the case where $z = z^{i+}$ and $z' = z^{i-}$ for some i. The claims of the lemma may be verified by a procedure analogous to that used above. For the logit specification, use (1) for the transition probabilities, and the lemma follows immediately. This lemma can be straightforwardly extended to incorporate the extensions to birth and death costs discussed in Section 5. □

A.2. **Theorem 1.** The proof of Theorem 1 will compare the least exponential-cost rooted trees at different states. Although the statement of the theorem given in the text restricts, for simplicity, to the Palfrey and Rosenthal (1984) threshold game, the proof given below applies to all four of the games considered in this paper. In particular, the exit-birth cost $\hat{\beta}_i$ of Section 5 is used throughout. For the theorem as stated, simply set $\hat{\beta}_i = \beta_i$ for all i.

Before proceeding to the main proof, a sequence of lemmas is proved that calculate the least-cost rooted trees for the states of interest. Each lemma will include the term A, where

$$A = \sum_{z \in Z_m} \min_{i \in N} \left\{ z_i \delta_i + (1 - z_i)\hat{\beta}_i \right\}.$$

This is the sum of the least exponential-cost exits from each state in Z_m. To see this, begin with a state $z \in Z_m$. An exit from such a state may involve a move up to a state $z' \in Z_{m+1}$ where $z' = z^{i+}$

for some i where $z_i = 0$. This will have a cost of $\hat{\beta}_i$. Alternatively, such an exit may involve a move down to a state $z' \in Z_{m-1}$ where $z' = z^{i-}$ for some i where $z_i = 1$. This will have a cost of δ_i. Hence $\min_{i\in N}\{z_i\delta_i + (1-z_i)\hat{\beta}_i\}$ is the minimum-cost transition away from any state $z \in Z_m$. The term A is now obtained by summing across all states in Z_m. Recall that $\iota = \arg\min_{i\leq m} \delta_i$.

Lemma 6. *The least exponential-cost tree rooted at state $z^\ddagger$ satisfies*

$$\mathcal{E}(z^\ddagger) = A - \min\{\delta_\iota, \hat{\beta}_{m+1}\} + \sum_{i=1}^{m-1} \beta_i.$$

Proof. Consider a tree rooted at $z^\ddagger$. There is no exit from $z^\ddagger$, but exits are required for all $z \in Z_m$ where $z \neq z^\ddagger$. Total costs of exits from Z_m are therefore at least

$$A - \min_{i\in N} \left\{ z_i^\ddagger \delta_i + (1 - z_i^\ddagger)\hat{\beta}_i \right\}.$$

The tree must involve a sequence of transitions from Z_0 into Z_m. Such a path must involve a transition from Z_0 to Z_1, from Z_1 to Z_2, and so on. Each such transition includes an additional birth cost. The least costly way to reach Z_{m-1} is to acquire the $m-1$ players with the least birth costs, and hence such a path has exponential cost of at least $\sum_{i=1}^{m-1} \beta_i$. A final step into Z_m may be taken at zero cost. For the last step, add player $i = m$. This gives a path that moves as quickly (and cheaply) as possible from Z_0 to $z^\ddagger \in Z_m$. Hence

$$\mathcal{E}(z^\ddagger) \geq A - \min_{i\in N} \left\{ z_i^\ddagger \delta_i + (1 - z_i^\ddagger)\hat{\beta}_i \right\} + \sum_{i=1}^{m-1} \beta_i. \tag{18}$$

This expression provides a lower bound to the exponential cost of any tree rooted at $z^\ddagger$. It incorporates transitions out of all states in Z_m, as well as the path from Z_0. Notice that the cheapest birth out of $z^\ddagger$ comes at a cost of $\hat{\beta}_{m+1}$, since $z_i^\ddagger = 1$ for all $i \leq m$ and $z^\ddagger = 0$ for all $i > m$. Similarly, the cheapest death out of $z^\ddagger$ comes at a cost of δ_ι. $\min_{i\in N}\{z_i^\ddagger\delta_i + (1-z_i^\ddagger)\beta_i\}$ in (18) can therefore be replaced by $\min\{\delta_\iota, \hat{\beta}_{m+1}\}$, yielding $\mathcal{E}(z^\ddagger) \geq A - \min\{\delta_\iota, \hat{\beta}_{m+1}\} + \sum_{i=1}^{m-1} \beta_i$.

It will now be shown that transitions from all remaining states may be constructed at zero additional exponential cost, so that a tree rooted at $z^\ddagger$ is obtained that attains the lower bound in (18). To do this, take a state $z \in Z_m$ where $z \neq z^\ddagger$ such that

$$\min_{i\in N} \left\{ z_i\delta_i + (1 - z_i)\hat{\beta}_i \right\} = \delta_j, \quad \text{for some} \quad z_j = 1.$$

Hence the least-cost exit from z is to move to $z' = z^{j-} \in Z_{m-1}$. One possibility is that z' is encountered on the upward path from Z_0 to $z^\ddagger$. In this case, there is path from z into $z^\ddagger$. A second possibility is that a zero-cost step upward might be taken directly into $z^\ddagger$. If neither of these opportunities is available, take a zero-cost step down into $z'' \in Z_{m-2}$. Now, z'' may be on the upward path from Z_0 to $z^\ddagger$. If not, then another zero exponential-cost step may be taken down into Z_{m-3}. This sequence of zero exponential-cost transitions continues until the graph hits either the upward path or Z_0. In either case, following the initial death of player j from state $z \in Z_m$, there is a path with zero additional exponential cost that leads into $z^\ddagger$.

Next, take a state $z \in Z_m$ where $z \neq z^\ddagger$ such that

$$\min_{i\in N} \left\{ z_i\delta_i + (1 - z_i)\hat{\beta}_i \right\} = \hat{\beta}_j \neq \delta_k, \quad \text{for some} \quad z_j = 0, \quad \text{and} \quad \forall z_k = 1.$$

Hence the least-cost exit from z is to move up to $z' = z^{j+} \in Z$. From state z', construct a zero exponential-cost transition back down into Z_m. To do this, pick $k = \arg\max_{i \in N}\{\hat{\beta}_i z_i\}$, and remove player k. Now $z \neq z^{\ddagger}$, so $\hat{\beta}_k > \hat{\beta}_j$.[34] Do this for all states $z \in Z_m$ where the least-cost exit is to move upward. Define $B(z) = \sum_{i \in N} \hat{\beta}_i z_i$. Notice that, as the process moves up from state $z \in Z_m$, and back down again, $B(z)$ is decreasing. Following this path, eventually state $z^{\ddagger} = \arg\min_{z \in Z_m} B(z)$ must be reached. Consequently there are no cycles, and there is a path from every state in Z_m to $z^{\ddagger}$. Finally, consider any states above layer Z_m that are unconnected. At zero cost, remove a contributor to transit to the layer below. Similarly, for states below Z_m that are unconnected, do the same. This is a rooted tree with an exponential cost that attains the lower bound. □

Lemma 7. *For all $z \in Z_m$, the exponential cost of a tree rooted at z is bounded below by*

$$A - \min_{i \in N} \left\{ z_i \delta_i + (1 - z_i)\hat{\beta}_i \right\} + \sum_{i=1}^{m-1} \beta_i.$$

Proof. Consider a tree rooted at z. There is no exit from z, but exits are required for all $z' \in Z_m$ where $z' \neq z$. Total costs of exits from Z_m are therefore at least

$$A - \min_{i \in N} \left\{ z_i \delta_i + (1 - z_i)\hat{\beta}_i \right\}.$$

The tree must involve a sequence of transitions from Z_0 into Z_m. The least-costly way to reach Z_{m-1} is still to acquire the $m-1$ players with the lowest birth costs, yielding an exponential cost of at least $\sum_{i=1}^{m-1} \beta_i$. The final step into Z_m has zero cost. This gives a path that moves as cheaply as possible from Z_0 to Z_m. Adding this cost to the above expression yields the result. □

Lemma 8. *Define $z^0 = z \in Z_0$. If $\delta_\iota \leq \hat{\beta}_{m+1}$, $\mathcal{E}(z^0) = A$. If $\delta_\iota > \hat{\beta}_{m+1}$,*

$$\mathcal{E}(z^0) = A - \hat{\beta}_{m+1} + \min_{j > m} \left\{ \hat{\beta}_j + \delta_j - \min(\hat{\beta}_m, \delta_j), \delta_\iota \right\}.$$

Proof. Consider a tree rooted at z^0. There must be a transition out of every state in Z_m, and hence such a tree has an exponential cost of at least A. For states $z \in Z_m$ where $z \neq z^{\ddagger}$ follow the procedure described in the proof of Lemma 6 to construct a sequence of zero-cost transitions that connect each such state to $z^{\ddagger}$. For any unconnected states above or below Z_m remove a player at zero cost from the active team to yield a transition to the layer below. Finally, consider state $z^{\ddagger}$. If $\delta_\iota \leq \hat{\beta}_{m+1}$ then the least exponential-cost exit is to move down to layer Z_{m-1}. This yields a tree rooted at z^0 with exponential cost A. The remainder of the proof is for the case $\delta_\iota > \hat{\beta}_{m+1}$.

Step (i). Construct a benchmark graph in the following manner. For each state in Z_m, take the least-cost exit. For all other states except z^0, take any zero-cost exit — there will always be at least one such transition available. The graph so connected has an exponential cost of A, by construction.

Step (ii). Consider a least exponential-cost tree rooted at z^0. (This will be compared to the benchmark graph described in step (i).) Without loss of generality, build transitions from all states in Z_{m-1} that point downwards, at zero exponential cost, into Z_{m-2}. Similarly, specify transitions downward from Z_{m-2} and so on.

[34] This is where the assumption that $\hat{\beta}_i \neq \hat{\beta}_j$ for all i and j comes in, but if there were to be ties then those tied players could be reordered by shoe size and the proof would continue unchanged.

Step (iii). With the tree described in step (ii) in mind, begin with state $z^\ddagger$. Follow the sequence of directed edges in the tree until the first node in Z_{m-1} is reached. The final edge in this sequence, which terminates at a state in Z_{m-1}, must originate from some state in Z_m. Call this originating state $z^\diamond$. From state $z^\diamond$, some player j satisfying $z_j^\diamond = 1$ must die at a cost of δ_j to move into layer Z_{m-1}. (In preparation for step (v) below, note that it might well be that $z^\diamond = z^\ddagger$.)

Step (iv). Take the set of states $z \in Z_m$ that are *not* encountered on the path from $z^\ddagger$ to $z^\diamond$. Take the least-cost exit from each of these. Add zero-cost links from other states not in Z_m to construct transitions that lead into the aforementioned path. (This follows the procedure described in the proof of Lemma 6, and can be completed at zero exponential cost.) Any difference in the exponential cost of the tree, relative to the benchmark graph, must arise on the path from $z^\ddagger$ to $z^\diamond$.

Step (v). Suppose that $z^\ddagger = z^\diamond$. If this is the least exponential-cost rooted tree, then the dying player j must be $j = \iota$, since this player faces the lowest death cost δ_ι among the team members at $z^\ddagger$. Relative to the benchmark, there is no exit upward. The exponential cost of this tree is

$$A + \delta_\iota - \hat{\beta}_{m+1}. \tag{19}$$

Step (vi). Suppose instead that $z^\ddagger \neq z^\diamond$. Write

$$\tilde{N} = \{i : z_i^\diamond = 1 \text{ and } z_i^\ddagger = 0\}.$$

$\tilde{N}$ is the set of team members acquired along the path from $z^\ddagger$ to $z^\diamond$. If $i \in \tilde{N}$ then player i must have been acquired at a cost of $\hat{\beta}_i \geq \hat{\beta}_{m+1}$. If this player was acquired following an exit from a state $z' \in Z_m$, then relative to the baseline directed graph the net exponential cost increase must be at least $\hat{\beta}_i - \hat{\beta}_{m+1}$. In particular, the acquisition of the player of interest (player j; this is the player who will die to enable the final escape from state $z^\diamond$) came at a positive net exponential cost. For $i \neq j$, these net costs could be saved by eliminating such acquisitions along the path. Hence, if the tree in question is the minimum-cost rooted tree, it must be the case that $\tilde{N} = \{j\}$. Hence the path from $z^\ddagger$ consists of two steps: The acquisition of player $j \geq m+1$, followed by the loss of some other player $i \leq m$ from the original team. In fact, player $i = m$ should be removed to ensure that the tree has minimum exponential cost — m is the player who ensures that the cheapest possible exit-birth cost (to be removed from $z^\diamond$) is as high as possible. Relative to the baseline graph (with exponential cost A) there are exponential-cost additions of $\hat{\beta}_j$ and δ_j, and exponential-cost deductions of $\hat{\beta}_{m+1}$ and $\min\{\delta_j, \hat{\beta}_m\}$. (This last point is because, from the viewpoint of state $z^\diamond$, player m is the player outside the team with the least exit-birth cost.) Hence, the exponential cost of this rooted tree is

$$A + \hat{\beta}_j + \delta_j - \hat{\beta}_{m+1} - \min\{\hat{\beta}_m, \delta_j\}.$$

Such a tree can be constructed by choosing *any* $j > m$ to be added to the original team $z^\ddagger$. Hence, if the tree in question really is the minimum exponential-cost tree, j must be chosen so that

$$j = \arg\min_{j>m} \left\{ \hat{\beta}_j + \delta_j - \min\{\hat{\beta}_m, \delta_j\} \right\}.$$

Step (vii). Add back in the term $A - \hat{\beta}_{m+1}$, to give

$$A - \hat{\beta}_{m+1} + \min_{j>m} \left\{ \hat{\beta}_j + \delta_j - \min\{\hat{\beta}_m, \delta_j\} \right\}.$$

Now $z^\diamond \neq z^\ddagger$ if and only if this term is smaller than the one given in (19) since only the least exponential-cost tree rooted at z^0 is of interest. Hence that cost is

$$\mathcal{E}(z^0) = A - \hat{\beta}_{m+1} + \min_{j>m} \left\{ \hat{\beta}_j + \delta_j - \min(\hat{\beta}_m, \delta_j), \delta_\iota \right\},$$

which is the required expression in the statement of the lemma. □

Proof of Theorem 1. Suppose $\delta_\iota \geq \min\{\max_{i>m} \delta_i, \hat{\beta}_\iota\}$. A first step is to show $\min\{\mathcal{E}(z^0), \mathcal{E}(z^\ddagger)\} \leq \mathcal{E}(z)$ for all $z \in Z$. Consider a state $z \in Z_k$ where $0 < k < m$. Clearly $\mathcal{E}(z^0) \leq \mathcal{E}(z)$. Consider a state $z \in Z_k$ where $m < k \leq n$, again it is straightforward that $\mathcal{E}(z^m) \leq \mathcal{E}(z)$ where $z^m \in Z_m$. It only remains to show that $\mathcal{E}(z^\ddagger) \leq \mathcal{E}(z)$ for $z \in Z_m$. Concentrating on such states, therefore, and using Lemmas 6 and 7, it will be shown that

$$A - \min\{\delta_\iota, \hat{\beta}_{m+1}\} + \sum_{i=1}^{m-1} \beta_i \leq A - \min_{i \in N} \left\{ z_i \delta_i + (1 - z_i)\hat{\beta}_i \right\} + \sum_{i=1}^{m-1} \beta_i$$

(for all $z \neq z^\ddagger$), since the latter term is a lower bound on $\mathcal{E}(z)$ for $z \in Z_m$. To proceed, assume the above inequality does not hold. Then

$$\min_{i \in N} \left\{ z_i \delta_i + (1 - z_i)\hat{\beta}_i \right\} > \min\{\delta_\iota, \hat{\beta}_{m+1}\}$$

Suppose first that $\hat{\beta}_{m+1} \leq \delta_\iota$. Now, in state z, there is some Player $j \leq m$ such that $z_j = 0$ (since $z \neq z^\ddagger$). The former minimum is then at most $\hat{\beta}_j$, but by assumption $\hat{\beta}_j < \hat{\beta}_{m+1}$, yielding a contradiction. Now consider $\hat{\beta}_{m+1} > \delta_\iota$. Again, looking to bound the former minimum above, suppose that $\delta_\iota \geq \max_{i>m} \delta_i$,

$$\min_{i \in N} \left\{ z_i \delta_i + (1 - z_i)\hat{\beta}_i \right\} \leq \min_{i \in N} \left\{ \delta_\iota, (1 - z_i)\hat{\beta}_i \right\} < \min\{\delta_\iota, \hat{\beta}_{m+1}\} = \delta_\iota$$

The first inequality follows from the fact that $z_j = 1$ for some $j > m$ (since $z \neq z^\ddagger$), and this player must have $\delta_j \leq \delta_\iota$. The second (strict) inequality follows from the fact that $z_j = 0$ for some $j \leq m$ (since $z \neq z^\ddagger$), and this player must have $\hat{\beta}_j < \hat{\beta}_{m+1}$. Finally, the equality follows by assumption, yielding another contradiction. Now (and lastly) suppose that $\hat{\beta}_\iota \leq \delta_\iota < \hat{\beta}_{m+1}$. Then if $z_\iota = 0$,

$$\min_{i \in N} \left\{ z_i \delta_i + (1 - z_i)\hat{\beta}_i \right\} \leq \hat{\beta}_\iota \leq \delta_\iota$$

and $\min\{\delta_\iota, \hat{\beta}_{m+1}\} = \delta_\iota$, yielding a contradiction. If, on the other hand, $z_\iota = 1$, then

$$\min_{i \in N} \left\{ z_i \delta_i + (1 - z_i)\hat{\beta}_i \right\} \leq \delta_\iota,$$

and the contradiction is immediate. Therefore $\mathcal{E}(z^\ddagger) \leq \mathcal{E}(z)$ for all $z \in Z_m$.

The second step (using Lemmas 6 and 8) is to show that

$$\mathcal{E}(z^\ddagger) < \mathcal{E}(z^0) \quad \iff \quad \sum_{i=1}^{m-1} \beta_i < \min_{j>m} \left\{ \hat{\beta}_j + \delta_j - \min(\hat{\beta}_m, \delta_j), \delta_\iota \right\}. \tag{20}$$

First suppose that $\delta_\iota \leq \hat{\beta}_{m+1}$. By Lemma 8, $\mathcal{E}(z^0) = A$. By Lemma 6,

$$\mathcal{E}(z^\ddagger) = A - \min\{\delta_\iota, \hat{\beta}_{m+1}\} + \sum_{i=1}^{m-1} \beta_i = A - \delta_\iota + \sum_{i=1}^{m-1} \beta_i.$$

Hence $\mathcal{E}(z^\ddagger) < \mathcal{E}(z^0)$ if and only if $\sum_{i=1}^{m-1} \beta_i < \delta_\iota$. Now notice that $\min_{j>m}\{\hat{\beta}_j + \delta_j - \min(\hat{\beta}_m, \delta_j)\} \geq \hat{\beta}_{m+1}$ since the sum of the second two terms is weakly positive, and the first is (weakly) bigger than $\hat{\beta}_{m+1}$ for all $j > m$. Since $\delta_\iota \leq \hat{\beta}_{m+1}$, (20) is verified. Now suppose $\delta_\iota > \hat{\beta}_{m+1}$. The exponential cost of a tree rooted at $z^\ddagger$ reduces to

$$\mathcal{E}(z^\ddagger) = A - \hat{\beta}_{m+1} + \sum_{i=1}^{m-1} \beta_i,$$

and a direct application of Lemma 8 verifies (20). This completes the proof. □

A.3. **Propositions 1 and 2.** Lemmas 6, 7, and 8 continue to hold, as does (of course) a slightly recast version of the first part of Theorem 1 (that is, $\min\{\mathcal{E}(z^0), \mathcal{E}(z')\} \leq \mathcal{E}(z)\ \forall z \notin Z_m$ and $z' \in Z_m$). It is sufficient to compare least exponential-cost trees rooted at states in Z_0 and Z_m.

Proof of Proposition 1. Suppose that the least-cost rooted tree in Z_m is at z^m. Then,

$$\mathcal{E}(z^m) \leq A - \min\{\delta_\iota, \hat{\beta}_{m+1}\} + \sum_{i=1}^{m-1} \beta_i.$$

This bound is attained at $z^\ddagger$ (see Lemma 6). Either $z^\ddagger$ has the lowest cost tree (and $z^m = z^\ddagger$), or there is another state with a lower cost tree — so this is indeed a strict upper bound. But then, using identical arguments to those presented in the proof of Theorem 1, it can be concluded that $\mathcal{E}(z^m) < \mathcal{E}(z^0)$ whenever $\sum_{i=1}^{m-1} \beta_i < \min_{j>m}\{\hat{\beta}_j + \delta_j - \min(\hat{\beta}_m, \delta_j), \delta_\iota\}$, simply by comparing the above bound with $\mathcal{E}(z^0)$ from Lemma 8. □

Proof of Proposition 2. By Lemma 7, for any $z \in Z_m$,

$$\mathcal{E}(z) \geq A - \min_{i \in N}\left\{z_i\delta_i + (1 - z_i)\hat{\beta}_i\right\} + \sum_{i=1}^{m-1} \beta_i.$$

Now, $\hat{\beta}_{m+1}$ is the highest minimum exit-birth cost possible from any state in Z_m (in particular, from $z^\ddagger$). Further, defining $\delta_{(i)}$ as the ith-highest death cost in the population, the highest minimum death cost possible from any state in Z_m is $\delta_{(m)}$ (from the state with $z_i = 1$ for each of the m highest death costs $\delta_{(j)}$, $j = 1, \ldots, m$). The above inequality becomes

$$\mathcal{E}(z) \geq A - \min\{\delta_{(m)}, \hat{\beta}_{m+1}\} + \sum_{i=1}^{m-1} \beta_i.$$

Notice that $\delta_{(m)} \geq \delta_\iota$. Now, assume that $\sum_{i=1}^{m-1} \beta_i > \delta_{(m)}$. First, consider $\hat{\beta}_{m+1} \geq \delta_{(m)}$. Hence $\hat{\beta}_{m+1} \geq \delta_\iota$ and by Lemma 8, $\mathcal{E}(z^0) = A$. Suppose that $\mathcal{E}(z^0) \geq \mathcal{E}(z)$. It follows that $A \geq A - \delta_{(m)} + \sum_{i=1}^{m-1} \beta_i$, a contradiction. Hence $\mathcal{E}(z^0) < \mathcal{E}(z)$. Second, consider $\delta_\iota \leq \hat{\beta}_{m+1} < \delta_{(m)}$. Again, by Lemma 8, $\mathcal{E}(z^0) = A$. Suppose that $\mathcal{E}(z^0) \geq \mathcal{E}(z)$. It follows that $A \geq A - \hat{\beta}_{m+1} + \sum_{i=1}^{m-1} \beta_i$, a contradiction (since $\hat{\beta}_{m+1} < \delta_{(m)}$). Hence $\mathcal{E}(z^0) < \mathcal{E}(z)$. Finally, consider $\hat{\beta}_{m+1} < \delta_\iota \leq \delta_{(m)}$. Now Lemma 8 yields $\mathcal{E}(z^0) = A - \hat{\beta}_{m+1} + \min_{j>m}\{\hat{\beta}_j + \delta_j - \min(\hat{\beta}_m, \delta_j), \delta_\iota\}$. Again, if $\mathcal{E}(z^0) \geq \mathcal{E}(z)$,

$$\sum_{i=1}^{m-1} \beta_i \leq \min_{j>m}\left\{\hat{\beta}_j + \delta_j - \min(\hat{\beta}_m, \delta_j), \delta_\iota\right\} \leq \delta_\iota.$$

But $\sum_{i=1}^{m-1} \beta_i > \delta_{(m)} \geq \delta_\iota$. This yields a final contradiction, and concludes the proof. □

A.4. **Corollaries 1 and 2.** Corollary 1 follows directly from earlier propositions.

Proof of Corollary 2. When $\beta_1 < \delta_1$, Theorem 1 applies, yielding the result. If $\sum_{i=1}^{m-1} \beta_i > \delta_{(m)} = \delta_{n-m+1}$, Proposition 2 applies. Finally, in the intermediate case (ii), apply Lemma 8, noticing that $\delta_1 \leq \beta_1 < \beta_{m+1}$ and hence $\mathcal{E}(z^0) = A$. From Lemma 7, note that if for some $z \in Z_m$, $z_1 = 0$, then

$$\mathcal{E}(z) \geq A - \min\{\underline{\delta}, \beta_1\} + \sum_{i=1}^{m-1} \beta_i,$$

where $\underline{\delta} = \{\min_{i \in N} \delta_i \,:\, z_i = 1\}$. If, on the other hand $z_1 = 1$, then

$$\mathcal{E}(z) \geq A - \delta_1 + \sum_{i=1}^{m-1} \beta_i.$$

Now suppose, contrary to the corollary's claim, that $\mathcal{E}(z^0) \geq \mathcal{E}(z)$ for some $z \in Z_m$. Then it must be the case that (if $z_1 = 0$) $A \geq A - \min\{\underline{\delta}, \beta_1\} + \sum_{i=1}^{m-1} \beta_i$, which implies that $\min\{\underline{\delta}, \beta_1\} \geq \sum_{i=1}^{m-1} \beta_i$, a contradiction. If, on the other hand, $z_1 = 1$, the contradiction is again immediate, since $\delta_1 < \sum_{i=1}^{m-1} \beta_i$. □

A.5. **Propositions 3 and 4 and Results 1, 2, and 3.**

Proof of Proposition 3. The proposition follows from setting $\beta_i = \beta$ and $\delta_i = \delta$ for all i, and the consequent substitution in the criterion of Theorem 1. □

Proof of Results 1 and 2. The results follow from an inspection of (6) and (7). □

Proof of Result 3. Set $c = \xi = 1$ without loss of generality. Setting $\gamma = 0$, (8) satisfies

$$1 < \frac{(\bar{v}m - 1)^2}{(1 + m\bar{\sigma}^2)(m-1)} \longrightarrow \frac{\bar{v}^2}{\bar{\sigma}^2} \quad \text{as} \quad m \to \infty.$$

For large m, this criterion is satisfied if $\bar{v} > \bar{\sigma}$ and is not if $\bar{v} < \bar{\sigma}$. More generally, team success can be ascertained by inspecting the quadratic $(\bar{v}^2 - \bar{\sigma}^2)m^2 + (\bar{\sigma}^2 - 2\bar{v} - 1)m + 2 > 0$. If $8(\bar{v}^2 - \bar{\sigma}^2) > (2\bar{v} + 1 - \bar{\sigma}^2)^2$, then there are no real roots. This implies that the team succeeds for all m if $\bar{v} > \bar{\sigma}$, and fails for all m if $\bar{v} < \bar{\sigma}$. If the quadratic has real roots, then such roots will determine the ranges of critical team-sizes for which there is success. For instance, writing the roots as $m_1^* < m_2^*$, when $\bar{v} > \bar{\sigma}$, the team fails when $m_1^* < m < m_2^*$. For $\gamma > 0$, replace the quadratic with the appropriate cubic, and the proof follows. □

Proof of Proposition 4. The proposition follows from the argument in the main text. □

References

Admati, A. R., and M. Perry (1991): "Joint Projects without Commitment," *Review of Economic Studies*, 58(2), 259–76.

Anderson, S. P., J. K. Goeree, and C. A. Holt (1999): "Stochastic Game Theory: Adjustment to Equilibrium Under Noisy Directional Learning," Department of Economics, University of Virginia.

Andreoni, J. (1988): "Why Free Ride? Strategies and Learning in Public Goods Experiments," *Journal of Public Economics*, 37(3), 291–304.

Bagnoli, M., and B. L. Lipman (1989): "Provision of Public Goods: Fully Implementing the Core through Private Contributions," *Review of Economic Studies*, 56(4), 583–601.

Bergin, J., and B. Lipman (1996): "Evolution with State-Dependent Mutations," *Econometrica*, 64(4), 943–56.

Bergstrom, T. C., L. Blume, and H. R. Varian (1986): "On the Private Provision of Public Goods," *Journal of Public Economics*, 29(1), 25–49.

Bernheim, B. D. (1986): "On the Voluntary and Involuntary Provision of Public Goods," *American Economic Review*, 76(4), 789–93.

Blume, L. E. (1993): "The Statistical Mechanics of Strategic Interaction," *Games and Economic Behavior*, 5(3), 387–424.

——— (1995): "The Statistical Mechanics of Best-Response Strategy Revision," *Games and Economic Behavior*, 11(2), 111–45.

——— (1997): "Population Games," in *The Economy as an Evolving Complex System II*, ed. by W. B. Arthur, S. N. Durlauf, and D. A. Lane. Westview Press, Boulder, CO.

——— (2003): "How Noise Matters," *Games and Economic Behavior*, 44(2), 251–71.

Blume, L. E., and S. N. Durlauf (2001): "The Interactions-Based Approach to Socioeconomic Behaviour," in *Social Dynamics*, ed. by S. N. Durlauf, and H. P. Young, chap. 2, pp. 15–44. MIT Press, Cambridge, MA.

——— (2003): "Equilibrium Concepts for Social Interaction Models," *International Game Theory Review*, 5(3), 193–209.

Brock, W. A., and S. N. Durlauf (2001): "Discrete Choice with Social Interactions," *Review of Economic Studies*, 68(2), 235–60.

Cadsby, C. B., and E. Maynes (1999): "Voluntary Provision of Threshold Public Goods with Continuous Contributions: Experimental Evidence," *Journal of Public Economics*, 71(1), 53–73.

Carlsson, H., and E. van Damme (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61(5), 989–1018.

Dawes, R. M., J. M. Orbell, R. T. Simmons, and A. J. C. van de Kragt (1986): "Organizing Groups for Collective Action," *American Political Science Review*, 80(4), 1171–85.

Diekmann, A. (1985): "Volunteer's Dilemma," *Journal of Conflict Resolution*, 29(4), 605–10.

Diermeier, D. (2001): "Spontaneous Coordination in Large Population Games with Persistent Randomness," Kellogg Graduate School of Management, Northwestern University.

——— (2003): "Consumer Boycotts," Kellogg Graduate School of Management, Northwestern University.

Diermeier, D., and J. A. Van Mieghem (2000): "Spontaneous Collective Action," *CMSEMS Discussion Paper*, 1302, Northwestern University.

Dixit, A. K., and M. Olson (2000): "Does Voluntary Participation Undermine the Coase Theorem?," *Journal of Public Economics*, 76(3), 309–35.

Foster, D., and H. P. Young (1990): "Stochastic Evolutionary Game Dynamics," *Theoretical Population Biology*, 38, 219–32.

Freidlin, M. I., and A. D. Wentzell (1998): *Random Perturbations of Dynamical Systems*. Springer-Verlag, Berlin/New York, 2nd edn.

Goeree, J. K., and C. A. Holt (2000): "An Explanation of Anomalous Behaivor in Binary-Choice Games: Entry, Voting, Public Goods, and the Volunteers' Dilemma," Department of Economics, University of Virginia.

Granovetter, M. (1978): "Threshold Models of Collective Behavior," *American Journal of Sociology*, 83(6), 1420–43.

Hardin, R. (1982): *Collective Action*. Johns Hopkins University Press, Baltimore.

Heckathorn, D. D. (1993): "Collective Action and Group Heterogeneity: Voluntary Provision Versus Selective Incentives," *American Sociological Review*, 58(3), 329–50.

——— (1996): "The Dynamics and Dilemmas of Collective Action," *American Sociological Review*, 61(2), 250–77.

Kandori, M., G. J. Mailath, and R. Rob (1993): "Learning, Mutation and Long-Run Equilibria in Games," *Econometrica*, 61(1), 29–56.

Ledyard, J. O. (1995): "Public Goods: A Survey of Experimental Research," in *The Handbook of Experimental Economics*, ed. by J. H. Kagel, and A. E. Roth, chap. 2, pp. 111–94. Princeton University Press, Princeton, NJ.

Macy, M. W. (1990): "Learning Theory and the Logic of Critical Mass," *American Sociological Review*, 55(6), 809–26.

——— (1991a): "Chains of Cooperation: Threshold Effects in Collective Action," *American Sociological Review*, 56(6), 730–47.

——— (1991b): "Learning to Cooperate: Stochastic and Tacit Collusion in Social Exchange," *American Journal of Sociology*, 97(3), 808–43.

Marwell, G., and R. E. Ames (1979): "Experiments on the Provision of Public Goods. I. Resources, Interest, Group Size, and the Free-Rider Problem," *American Journal of Sociology*, 84(6), 1335–60.

——— (1980): "Experiments on the Provision of Public Goods. II. Provision Points, Stakes, Experience, and the Free-Rider Problem," *American Journal of Sociology*, 85(4), 926–37.

McKelvey, R. D., and T. R. Palfrey (1995): "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, 10(1), 6–38.

Menezes, F. M., P. K. Monteiro, and A. Temimi (2001): "Private Provision of Discrete Public Goods with Incomplete Information," *Journal of Mathematical Economics*, 35(4), 493–514.

Monderer, D., and L. S. Shapley (1996): "Potential Games," *Games and Economic Behavior*, 14(1), 124–43.

Morris, S., and H. S. Shin (2003): "Global Games: Theory and Application," in *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, ed. by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, vol. 1, chap. 3, pp. 56–114. Cambridge University Press, London.

Myatt, D. P., H. S. Shin, and C. Wallace (2002): "The Assessment: Games and Coordination," *Oxford Review of Economic Policy*, 18(4), 397–417.

Myatt, D. P., and C. Wallace (2003): "A Multinomial Probit Model of Stochastic Evolution," *Journal of Economic Theory*, 113(2), 286–301.

——— (2004): "Adaptive Play by Idiosyncratic Agents," *Games and Economic Behavior*, 48(1), 124–38.

——— (2005a): "The Evolution of Voluntary Action," Oxford University, Department of Economics.

——— (2005b): "Production Targets and Free Disposal in the Private Provision of Public Goods," *Department of Economics Discussion Paper*, 231, Oxford University.

Oliver, P. E., and G. Marwell (1988): "The Paradox of Group Size in Collective Action: A Theory of Critical Mass. II.," *American Sociological Review*, 53(1), 1–8.

——— (2001): "Whatever Happened to Critical Mass Theory? A Retrospective and Assessment," *Sociological Theory*, 19(3), 293–311.

Oliver, P. E., G. Marwell, and R. Teixeira (1985): "A Theory of the Critical Mass. I. Interdependence, Group Heterogeneity, and the Production of Collective Action," *American Journal of Sociology*, 91(3), 522–56.

Olson, M. (1968): *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press, Cambridge, MA.

Ostrom, E. (2000): "Collective Action and the Evolution of Social Norms," *Journal of Economic Perspectives*, 14(3), 137–58.

Palfrey, T. R., and H. Rosenthal (1984): "Participation and the Provision of Discrete Public Goods: A Strategic Analysis," *Journal of Public Economics*, 24(2), 171–93.

Rapoport, A. (1985): "Provision of Public Goods and the MCS Experimental Paradigm," *American Political Science Review*, 79(1), 148–55.

Schelling, T. C. (1973): "Hockey Helmets, Concealed Weapons, and Daylight Saving: A Study of Binary Choices with Externalities," *Journal of Conflict Resolution*, 17(3), 381–428.

——— (1978): *Micromotives and Macrobehavior*. Norton, New York.

van de Kragt, A. J. C., J. M. Orbell, and R. M. Dawes (1983): "The Minimal Contributing Set as a Solution to Public Goods Problems," *American Political Science Review*, 77(1), 112–22.

Young, H. P. (1993a): "The Evolution of Conventions," *Econometrica*, 61(1), 57–84.

——— (1993b): "An Evolutionary Model of Bargaining," *Journal of Economic Theory*, 59(1), 145–68.

——— (1996): "The Economics of Convention," *Journal of Economic Perspectives*, 10(2), 105–22.

——— (1998): *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton University Press, New Jersey.