ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

PRODUCTION TARGETS AND FREE DISPOSAL IN THE PRIVATE PROVISION OF PUBLIC GOODS

David P. Myatt and Chris Wallace

Number 231

March 2005

Manor Road Building, Oxford OX1 3UQ

Production Targets and Free Disposal in the Private Provision of Public Goods

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

University of Oxford Economics Discussion Paper 231. March 2005.[1]

Abstract. In a collective-action game a player's payoff is the sum of (i) a private component that depends only on that player's action, and (ii) a public component, common to all players and dependent upon all actions. A classic application is the private provision of a public good. Play evolves: strategy revisions are made according to a multinomial-logit choice rule. Long-run behaviour is determined by a potential function, which incorporates the private (not social) benefits of activity. Behaviour may be influenced only by reducing public-good output (an application of a free-disposal property). When welfare is the expected time average of aggregate payoffs, it is socially optimal to either leave production well alone, or damage it as much as possible. This often takes the form of a production target, where all output is discarded unless some threshold is reached, potentially generating an equilibrium-selection problem. When the evolution of play approximates a best-reply process, the optimal threshold corresponds to the output level that an individual who pays all private costs but enjoys only private benefits would be just willing to provide.

1. Public-Good Production and Collective Action

In this paper, a collective action is a situation in which private actions lead to common consequences. Examples include the private provision of a public good, or the private exploitation of a common resource. For the former case, the central issue addressed is this: what types of public-good production technology are conducive to the success of a collective action? More concretely, a general game is considered in which a player's payoff is the sum of (i) a private component, specific to the individual and depending only on that player's action, and (ii) a public component, common to all players and depending on all actions.[2]

---

[1]**Date Printed**. March 24, 2005 (`Prodv1.tex`). **JEL Classification**. C72, C73, and H41.
Keywords. Public goods, potential games, evolution, quantal response, equilibrium selection, thresholds.
Acknowledgements. The authors would like to thank colleagues, particularly Ken Binmore and Peyton Young, as well as seminar participants in Bristol and Copenhagen for their comments and advice.

[2]This distinction follows the tradition established by Olson (1968), for whom a collective action was defined by the separation of individual and common interests. It is also related to the notion of teamwork employed by Marschak (1955, p. 128), who defined a team as "[...] a group of persons each of whom takes decisions about something different but who receive a common reward as the joint result of all those decisions."

This latter dependence corresponds to a public-good production function. Given a family of such functions, the member that leads to the highest social welfare is sought.

To motivate, consider the academic committees that will be familiar to many readers. Often, committee members voluntarily attend such meetings; this involves a private cost and (arguably) generates a public benefit. Individuals will be tempted to free ride on the attendance of their colleagues, and hence participation may be below the social optimum. The quality of committee decisions could be raised if only more people were to turn up.[3] Many committees employ quora so that no decision can be taken without sufficient participation. A quorum may be viewed in two ways: first, as a production target, so that public-good production (the committee's decision) goes ahead only if a threshold (that is, the quorum) level of participation is reached; second, as a particular public-good production function, where output (any decision taken by a sub-quorate committee) is sometimes thrown away. However, a quorum generates a coordination problem: committee members attend if and only if sufficiently many of their colleagues do so. Whilst such issues motivate the paper, attention is not limited to quora, thresholds, and production targets. Rather, such targets emerge from a general analysis of public-good production functions.

A first intuition might be that, fixing inputs, more production is always better. To see why this might not be the case, consider individuals who voluntarily contribute toward the production of a non-excludable public good. Contributions will be made up to the point at which private marginal cost is equal to private marginal benefit. Of course, further contributions are socially preferable, since a private individual does not account for positive spillovers. Turn now to a different system in which any contributions below the social optimum are discarded.[4] The new production function is everywhere (weakly) below the original one. Nevertheless, there may now be an equilibrium in which socially optimal contributions are forthcoming, since each individual is pivotal to the collective action. If this equilibrium were to be played, then the change in production technology has certainly improved matters. However, if no individual is willing to unilaterally contribute the minimum required to generate positive production, then there will be another, zero-production, equilibrium.

Three observations emerge. First, to ascertain welfare performance it is necessary to determine which actions will be taken: in the example above, Nash equilibrium does not always yield a unique prediction of play. Second, social improvements may arise from output destruction: the use of a free-disposal opportunity. Third, the modified technology involved the elimination of output below a certain level of production: this is a threshold rule. These three observations lead to three questions. Which strategy profiles will be played and how often? To what extent is the exploitation of a free-disposal property socially desirable? For which levels of production (that is, for which threshold) should free disposal be employed?

[3] Of course, too many cooks might spoil the broth; over-attendance is also addressed within the paper.

[4] This is one solution to the problem of moral hazard in teams studied by Holmström (1982). This incentive mechanism requires a group of team members (agents) to dispose of output; Holmström (1982) argued that this is non-credible, and hence necessitates the introduction of a governing third-party (a principal).

An answer to the first question stems from the study of a strategy-revision process via which play evolves: at each point in time a player is selected at random and given a strategy-revision opportunity. The revising player updates according to a multinomial-logit choice rule.[5] This log-linear quantal-response might be generated by an underlying random-utility model, or it may be viewed as a "smoothed" best-reply. This leads to a well-defined long-run (that is, ergodic) distribution over the states of play. Aggregate welfare may then be defined as the expected sum of players' payoffs over time. Owing to the fact that the public-good provision game is a potential game, the ergodic distribution takes a simple form: the log likelihood of a strategy profile is proportional to its potential.[6] In turn, the potential is equal to the private benefit enjoyed by a single individual minus the sum of all private costs.

Turning to the second question, a production-function family is indexed by a set of parameters. An individual parameter has the same effect on production across a subset of the strategy profiles. Theorem 1 reveals that welfare is optimised by choosing each parameter to either maximise or minimise production in the corresponding subset. For instance, if free disposal is available, then it is socially optimal to eliminate production for certain strategy profiles, whilst retaining all output for the others. To see why, notice that reducing the production arising from a state of play has two effects: it lowers both the welfare and the potential (and hence relative likelihood) of that state. This second effect means that other strategy profiles are played more often. If these other profiles yield higher welfare, then the net effect might be to enhance aggregate welfare. If this is so, then any further destruction of production also must be (in the aggregate) welfare enhancing: the negative welfare impact of free disposal has been lessened, since the strategy profile in question is played less frequently.

Finally, it is socially optimal to operate a welfare-determined threshold rule. This rule is simple to interpret when the evolution of play approximates a best-reply process; that is, when the noise associated with quantal-response strategy revisions is small. For any given state of play, it is optimal to eliminate production when the welfare that would otherwise arise falls below a critical level. This threshold may be calculated in the following way. Consider the states of play for which an individual's private benefit from public-good production exceeds the total private cost. These states are interpreted as privately feasible, in the sense that a private operation (by someone who bears all private costs but cannot capture the non-excludable spillovers) would be profitable. The maximum welfare achieved over the privately feasible set is the threshold. Clearly, this threshold typically falls short of the social optimum. Were a higher threshold to be chosen, however, the strategy-revision process would languish in a state of low production for much of the time. Setting a lower target, whilst less ambitious, ensures that the process spends more time in higher-production states.

[5]Revising players choose quantal responses, in the sense of McKelvey and Palfrey (1995). Blume (1993, 1995, 1997, 2003), Blume and Durlauf (2001, 2003), Brock and Durlauf (2001), and Young (1998, 2001) have employed this choice rule in a variety of models.

[6]More precisely, the game considered here is an exact potential game (Monderer and Shapley, 1996). The elegant Gibbs representation of the ergodic distribution for such games was noted by Blume (1997).

The characterisation given above applies when the quantal-response noise is small. Even when it is not, the socially optimal choice of a public-good production function continues to involve thresholds. If the input to production is the simple sum of player contributions, output arises from a concave production function, and free disposal is available then, so long as the underlying production technology is sufficiently concave, it is socially optimal to set a contribution target (a minimum threshold) coupled with a contribution cap (a maximum). That is, public-good production goes ahead so long as total contributions fall within the consequent interval, and otherwise all production is thrown away.

Section 2 offers two examples that motivate and illustrate the paper. The class of games and the dynamic via which play evolves are described in Section 3. Section 4 contains the central results. Theorem 1 shows that it is optimal either to make full use of free disposal or to leave public-good production well alone. Theorem 2 calculates the threshold rule that determines when output should be thrown away. In Section 5 these theorems are brought to bear on the examples of Section 2. Section 6 relates the results to earlier literature.

2. Two Simple Worked Examples

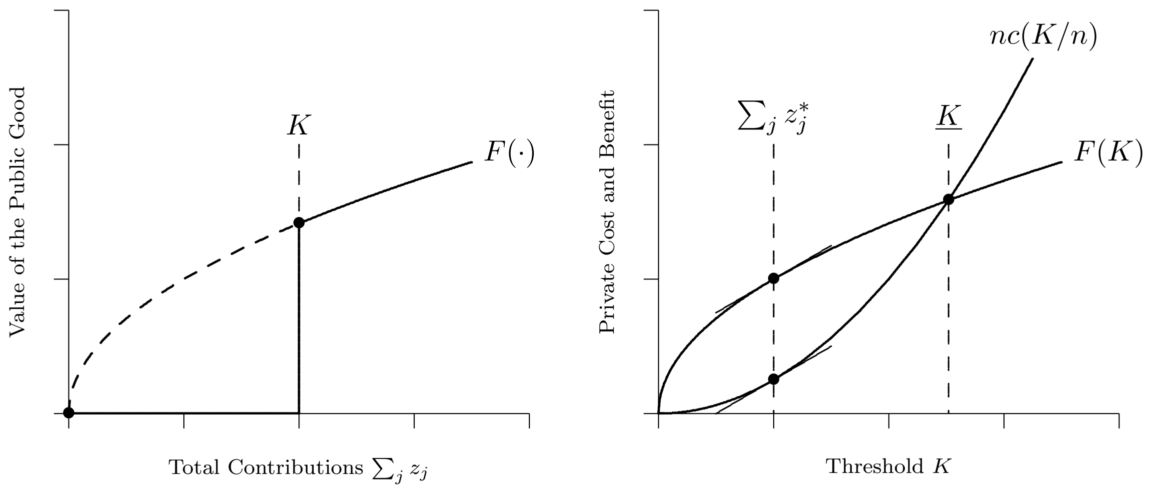
This section provides two simple motivating examples. The first illustrates the key questions of the paper. The second introduces the methodology used to answer these questions.

2.1. **A First Example.** Consider a simple Cournot-contributions game, as studied by Shibata (1971), Warr (1983), and Bergstrom, Blume, and Varian (1986, 1992), in which n players simultaneously, and voluntarily, contribute toward the private production of a public good.[7] Suppose that player i contributes $z_i \geq 0$. In doing so, the player incurs a private cost of $c_i(z_i)$. Players' contributions $z \in \mathcal{R}^n_+$ combine to yield a public good of value $F(\sum_j z_j)$. The non-excludable public good is enjoyed equally by all players. Costs and benefits are additively separable, and hence player i seeks to maximise $F(\sum_j z_j) - c_i(z_i)$.

Given a few technical assumptions, such as the concavity of the production function $F(\cdot)$ and the convexity of each cost function $c_i(\cdot)$, there will be a unique pure-strategy Nash equilibrium z^* involving the equality of (private) marginal cost and (private) marginal benefit, so that $F'(\sum_j z^*_j) = c'_i(z^*_i)$.[8] Alas, production falls short of the social optimum. If social welfare is the simple sum of the players' utilities, then the socially optimal strategy profile z^s entails $nF'(\sum_j z^s_j) = c'_i(z^s_i)$, so that $z^s_i > z^*_i$. A hypothetical social planner might wish to impose such contributions. However, here it is assumed that players cannot be coerced. Instead, a planner is only able to manipulate the production function, and may do so only by exploiting the free-disposal property. How might the planner induce increased contributions?

[7]For diagrammatic analyses of such games, see Ley (1996) and Cornes and Sandler (1996).

[8]Sufficient conditions would be as follows: (i) $F(\cdot)$ is strictly increasing, strictly concave, continuously differentiable, $F'(0) > 0$, and $\lim_{x\to\infty} F'(x) = 0$; (ii) for each i, $c_i(\cdot)$ is strictly increasing for $z_i > 0$, strictly convex, continuously differentiable, and $c'_i(0) = 0$. These may be easily replaced with weaker conditions.



(a) Denting the Production Function

(b) Choosing the Threshold

Panel (a) illustrates the "denting" of a production function. The solid line represents the production function after the threshold K has been introduced; to the left of K the free-disposal property has been maximally used. Panel (b) illustrates the constraints faced when selecting a threshold. For thresholds $K > \sum_j z_j^*$, there can be an equilibrium where the players jointly contribute K toward the public good. However, if $K > \underline{K}$ where $F(\underline{K}) = nc(\underline{K}/n)$ then risk dominance will select the equilibrium in which nothing is produced, and so such thresholds are self-defeating.

Figure 1. Optimal Coordination-Problem Generation

To answer this, suppose that if total contributions fall short of some threshold $K > \sum_i z_i^*$ then they are thrown away. K serves as a production target: it "dents" the production function (Figure 1), and leads player i to maximise $\mathcal{I}[\sum_j z_j \geq K] \times F(\sum_j z_j) - c_i(z_i)$, where $\mathcal{I}[\cdot]$ is the indicator function. Players will realise that only by (jointly) contributing K will any production take place. However, this technique may also generate a coordination problem: a player may be willing to contribute to the project if and only if others do so. The threshold's presence may generate an equilibrium in which no contributions are forthcoming.

To explore this, assume, for simplicity, that players share the same cost function $c(\cdot)$, and that the threshold is chosen to satisfy $c(K) > F(K) \geq c(K/n)$. The inequality $c(K) > F(K)$ ensures that no individual would we be willing to be the sole contributor of the minimum contribution K, and hence there is an equilibrium involving no contributions. On the other hand, since $c(K/n) \leq F(K)$, there is an equilibrium in which each player sets $z_i = K/n$.[9] A player realises that a reduced contribution will result in the complete failure of production. Furthermore, since K was chosen to be larger than the original equilibrium level of production, so that $c'(K/n) > F'(K)$, no player will wish to increase their contribution.

[9]Of course, there are many other equilibria. For instance, take any strategy profile z satisfying $\sum_j z_j = K$, $F'(K) \leq c'(z_i)$ and $F(K) \geq c(z_i)$ for all i. These are not considered here; however, Section 3 will allow for general functional forms, asymmetric players and equilibria, and a more general notion of free disposal.

This discussion leads to the following observation: while a threshold might yield an equilibrium with greater contributions, it may also generate the possibility of zero production.[10] To assess the success of a production target it is necessary to determine which equilibrium (if any) is played. Two questions arise. First, is the introduction of a production target the best way forward? Second, how high should a threshold be?

Comprehensive answers to both of these questions emerge from the analysis of later sections. Here, however, the second question is briefly considered. An obvious answer would be to choose the threshold $K = \sum_j z_j^s$, so that the socially optimal profile becomes an equilibrium.[11] But, as the target becomes more ambitious, it may be more difficult for players to successfully coordinate on the "good" equilibrium. To explore this issue, consider the simplest two-player scenario; this is, effectively, a 2×2 coordination game,

	$z_2 = K/2$	$z_2 = 0$
$z_1 = K/2$	$F(K) - c(K/2)$ $F(K) - c(K/2)$	0 $-c(K/2)$
$z_1 = 0$	$-c(K/2)$ 0	0 0

which, by inspection, has two pure-strategy Nash equilibria.[12] One way of choosing between these equilibria is to employ the concept of risk dominance (Harsanyi and Selten, 1988). In a symmetric 2×2 game, the risk-dominant equilibrium involves strategies which are best responses to Laplacian (that is, uniform) priors over the opponents' actions. Heuristically, players might be expected to play such strategies because they are "safer," in that they make sense when players are uncertain of their opponents' likely decisions.

In this case, the equilibrium in which both players contribute $z_i = K/2$ is risk dominant if and only if $F(K) \geq 2c(K/2)$. This says that the private (not social) benefit from attaining the threshold exceeds the private cost. With n players, the analogous condition would be $F(K) \geq nc(K/n)$. Notice that if the production target is set at the socially efficient level, so that $K = \sum_j z_j^s$, then this is not necessarily satisfied. Thus, a hypothetical social planner might wish to temper ambition with realism by setting K such that this condition is met.

A number of recent approaches to equilibrium selection provide support for the use of the risk-dominance criterion. Prominent amongst these is the global-games methodology of Carlsson and van Damme (1993) and Morris and Shin (2003). When rational players lack common knowledge of the game being played, Carlsson and van Damme (1993) showed that the risk-dominant strategy profile of a 2×2 game is almost always played. Kandori, Mailath, and Rob (1993) and Young (1993) provided an alternative approach, in which play evolves

[10] This claim holds when $c(K) > F(K) \geq c(K/n)$. If $F(K) < c(K/n)$ then the unique equilibrium involves no contributions, whereas if $F(K) > c(K)$ then the unique equilibrium outcome satisfies $\sum_j z_j \geq K$.

[11] This corresponds to the principal-driven solution to the moral-hazard-in-teams problem (Holmström, 1982).

[12] Clearly, a player may choose a contribution other than the two highlighted here. There is no further insight to be gained by including these and, furthermore, no such restriction is made in the remainder of the paper.

via a stochastic-adjustment dynamic. Boundedly rational players adapt myopically to recent play, and their actions are subject to noise. When noise is small, and in the long run, play almost always corresponds to the risk-dominant equilibrium of a 2×2 coordination game.

Equilibrium selection via risk dominance, while suggestive, is not the focus of this paper. Instead, this paper admits the possibility that different strategy profiles might be played at different times. The introduction of a threshold, or some other manipulation of the public-good production function, might influence not only the social welfare of a particular strategy profile, but also the proportion of time for which that profile is played. The next example explores this idea. Furthermore, it sheds light on the desirability of production targets.[13]

2.2. **A Second Example.** This example illustrates the methodology used to tackle the questions arising from Section 2.1, and helps to build intuition for the answers. Two players must non-cooperatively choose whether or not to contribute toward public-good production, so that their pure-strategy sets take the binary form $z_i \in \{0, 1\}$. A contributor incurs a private cost of c. When both players contribute, the public good has value v_H to both players, whereas a single contribution reduces its value to $\theta \times v_L$. Crucially, the parameter $\theta \in [0, 1]$ will be varied in the subsequent analysis. Setting $\theta = 1$ corresponds to leaving the production function well alone, whereas $\theta = 0$ is equivalent to choosing a threshold (as discussed in the previous subsection) of $K = 2$.

In strategic form, this simple 2×2 game can be represented in the following matrix.

	$z_2 = 1$	$z_2 = 0$
$z_1 = 1$	$v_H - c$, $v_H - c$	$\theta v_L - c$, θv_L
$z_1 = 0$	θv_L, $\theta v_L - c$	0, 0

To make the game interesting, suppose that $v_L > c > v_H - v_L > c/2 > 0$, so that, for $\theta = 1$, (i) the private benefit of the first contribution exceeds its private cost, but (ii) the private benefit of the second contribution falls short of its private cost, and (iii) it is socially optimal for both players to contribute. For θ sufficiently large, there will be pure-strategy Nash equilibria in which only one player contributes. For θ sufficiently small, however, there will be one equilibrium in which both players contribute, and one in which neither do so.

Welfare is the simple sum of the players' payoffs. Hence, using an obvious notation, $w_{11} = 2(v_H - c)$, $w_{10} = w_{01} = 2\theta v_L - c$, and $w_{00} = 0$. Clearly, welfare depends upon the strategies played, and hence a notion of aggregate or expected welfare requires a probability distribution π_z over the set of strategy profiles $Z = \{0, 1\}^2$. One approach would be to simply

[13]Here, the threshold takes the form of a minimum level of public-good production. Since it is also inefficient for players to contribute too much, it might be worthwhile to impose a cap on contributions. In fact, the analysis of later sections demonstrates that this is so.

select a strategy profile, perhaps by using an equilbrium-selection criterion such as the risk-dominance concept described in Section 2.1. Instead, serious heed is paid to the possibility the game is played frequently, and the players adopt different strategies at different times.

Concretely, play of this public-good production game evolves via the following strategy-revision process. At (discrete) time t, the state of play $z^t \in Z$ is the strategy profile in use. One of the players is randomly chosen to revise their strategy, and does so via logit quantal-response (McKelvey and Palfrey, 1995). For example, suppose that player 2 is currently contributing ($z_2^t = 1$) whilst player 1 is not ($z_1^t = 0$). Player 1 is given the opportunity to revise. By choosing to contribute, player 1 will enjoy a payoff $v_H - c$. On the other hand not contributing will yield θv_L. A logit quantal-response means that, rather than choosing a simple best response, the log odds ratio of player 1 choosing to contribute versus not is linear in the payoff difference between the two actions. Hence, the probability that the process moves from the state $z^t = (0, 1)$ to $z^{t+1} = (1, 1)$ is

$$p_{01\to11} = \frac{1}{2} \times \frac{\exp[\lambda(v_H - c)]}{\exp[\lambda\theta v_L] + \exp[\lambda(v_H - c)]}. \tag{1}$$

The $\frac{1}{2}$ term is the probability that player 1 is chosen to revise. λ indexes the degree to which this is a model of "smoothed" best response. If $\lambda = 0$ then a revising player chooses at random, whereas when $\lambda \to \infty$ the revising player almost always chooses a best response.

As with econometric models of discrete choice, the logit quantal-response admits a random-utility interpretation. Suppose that a revising player's payoffs are subject to noise. Specifically, the payoff difference between the two available actions is logistically distributed with parameter λ. The player then chooses the pure strategy with the higher payoff realisation.

The strategy revision process described above is a Markov chain on the state space Z, and the probability in Equation 1 is one of the associated transition probabilities. The other transition probabilities may be calculated similarly. By inspection, this process is ergodic; that is, its long-run behaviour is independent of initial conditions.

Interest lies in the long-run, or ergodic, distribution of the process. That is, the unique distribution π_z such that $\pi_z = \lim_{t\to\infty} \Pr[z^t = z]$. Hence π_z is the probability that strategy profile z is played in the long run. Calculating the ergodic distribution for the example at hand is relatively simple, since "detailed-balance conditions" apply. In other words, for a process such as this, $\pi_z/\pi_{z'} = p_{z'\to z}/p_{z\to z'}$. As a result, it is possible to write down the ratios of the ergodic probabilities as ratios of the transition probabilities. For example, taking the probability from Equation 1, and its complement,

$$\frac{\pi_{11}}{\pi_{01}} = \frac{p_{01\to11}}{p_{11\to01}} = \frac{\exp[\lambda(v_H - c)]}{\exp[\lambda\theta v_L]} = \frac{\exp[\lambda(v_H - 2c)]}{\exp[\lambda(\theta v_L - c)]}.$$

Working with these detailed-balance conditions, it is straightforward to write down the ergodic distribution of the logit quantal-response process. This satisfies

$$\pi_{11} \propto \exp[\lambda(v_H - 2c)], \quad \pi_{10} = \pi_{01} \propto \exp[\lambda(\theta v_L - c)], \quad \text{and} \quad \pi_{00} \propto \exp[0] = 1.$$

Observe that the log-likelihood of a particular state of play is determined by the difference between the *private* benefit to a single player of public-good provision and the total private cost. This contrasts with the social welfare of a state of play which is the *social* benefit minus the total private cost. Inspecting these probabilities, together with the welfare terms w_z described earlier, consider the implications of changing θ.

A reduction in θ, following the exploitation of the free-disposal property, has two effects. First, it reduces the social welfare of states $z = (0, 1)$ and $z = (1, 0)$. Second, it reduces the relative likelihood of these states and, therefore, increases the relative likelihood of state $z = (1, 1)$ in which welfare is maximised. An investigation of the combined effects requires an inspection of aggregate (or average) welfare. In the long run, expected welfare is simply

$$W_\lambda(\theta) = \frac{2(2\theta v_L - c)\exp[\lambda(\theta v_L - c)] + 2(v_H - c)\exp[\lambda(v_H - 2c)]}{2\exp[\lambda(\theta v_L - c)] + \exp[\lambda(v_H - 2c)] + 1}. \tag{2}$$

In fact, for any $\lambda \geq 0$, it can be shown that $W_\lambda(\cdot)$ is quasi-convex in θ, and thus welfare is maximised by setting either $\theta = 0$ or $\theta = 1$ (this follows from Theorem 1). These two options correspond to either leaving the production alone, or maximally employing the free-disposal property to introduce a threshold at $K = 2$.

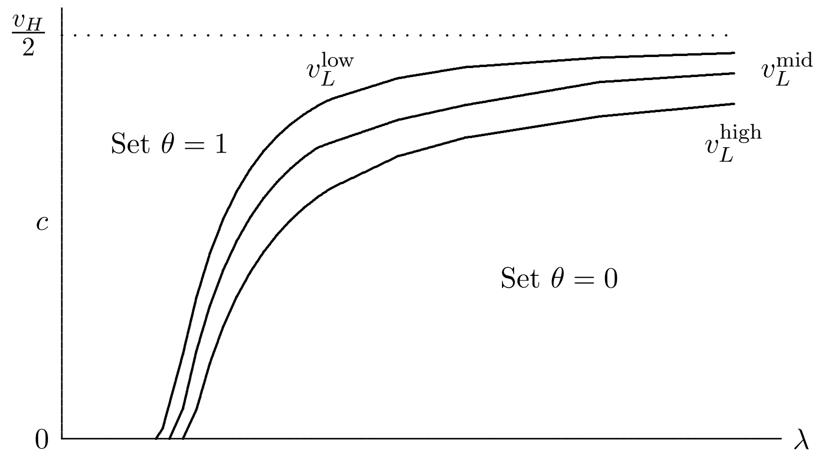
A standard approach would have been to let $\lambda \to \infty$, driving the noise out of the system, and examine the ergodic distribution in the limit. Here, however, there was no need to do this in order to obtain an interesting result. It is nonetheless instructive to look at the limiting behaviour of the expression in Equation 2 to give an intuition.

Consider the case when $v_H > 2c$. By setting $\theta = 0$, $\lim_{\lambda\to\infty} W_\lambda(0) = 2(v_H - c)$. Given the various payoff elements, this is clearly the highest that welfare can be. It is precisely in this case that society can gain by "throwing away production," and creating a coordination problem that, in the limit, involves the pure-strategy equilibrium $(1, 1)$ being played with probability one. However, when $v_H < 2c$, the "bad" equilibrium would be selected, and society has no interest in generating such a coordination problem. Indeed, when $v_H < 2c$,

$$\lim_{\lambda\to\infty} W_\lambda(\theta) = \begin{cases} 2\theta v_L - c & \text{if } \theta > c/v_L, \\ \frac{2}{3}(2\theta v_L - c) & \text{if } \theta = c/v_L, \\ 0 & \text{if } \theta < c/v_L; \end{cases}$$

which is simply maximised by setting $\theta = 1$.[14] In the limit, then, it is clear that there is some c^* such that for $c < c^*$, $\theta = 0$ is optimal, and otherwise $\theta = 1$ is optimal. In fact, this is true away from the limit as well. Consider $\lambda = 0$. Here, decisions are "all noise". Either

[14]There is also a final knife-edge case with $v_H = 2c$. Here, if $\theta > c/v_L$, $\lim_{\lambda\to\infty} W_\lambda(\theta) = 2\theta v_L - c$, and if $\theta \leq c/v_L$, $\lim_{\lambda\to\infty} W_\lambda(\theta) = c$. Again, the social planner would maximise expected welfare by setting $\theta = 1$.



The figure shows the $c^*(\lambda)$ function for three different values of v_L, where $v_L^{\text{high}} > v_L^{\text{mid}} > v_L^{\text{low}}$. All three lines asymptote at $v_H/2$. In the area underneath the curve, it is socially optimal to set $\theta = 0$. In the area above the graph, it is socially optimal to set $\theta = 1$. As $v_L \to v_H$, the former area disappears.

FIGURE 2. The Cut-Off Values $c^*(\lambda)$

action is taken with probability half, and $W_0(\theta) = \theta v_L - c + v_H/2$. This is maximised by setting $\theta = 1$. More generally, there will be a cut-off level of the contribution cost $c^*(\lambda)$, such that for all c lower than this it is optimal to leave the production function alone, and for all c higher than this it is optimal to introduce a production target. Moreover, from the earlier calculation, $c^*(0) = 0$ and $\lim_{\lambda\to\infty} c^*(\lambda) = v_H/2$. Figure 2 illustrates.

2.3. **Some Preliminary Conclusions.** The example of Section 2.1 illustrates the motives for generating a coordination problem (where previously none existed) when attempting to improve social welfare in the classic private-provision-of-a-public-good setting. By "denting" a production function, the social optimum may become an equilibrium of the game; however, another equilibrium might arise where no provision takes place at all. As a result, if generating such a coordination problem is indeed warranted, it must be the case that the "good" equilibrium is played: else the situation likely will be even worse than prior to intervention.

Should the production function be damaged in this fashion? The example of Section 2.2 provides some preliminary answers. A notion of expected welfare across different strategy profiles is introduced, based upon the a logit quantal-response process via which play evolves. This generates long-run probabilities for the play of each strategy profile, which can be used to construct aggregate social welfare. This function is quasi-convex in a policy parameter that reflects the free-disposal property. For a given strategy profile, it is socially optimal to either leave the production function alone or throw all of the production away. The subsequent analytical framework shows that this result turns out to be quite general.

3. The Evolution of Production

This section presents the general framework. First, a finite strategic-form game is introduced to model the private provision of a public good. This is shown to be a potential game, in the sense of Monderer and Shapley (1996). Second, a multinomial-logit quantal-response process is described via which play evolves. Finally, it is shown that long-run play is described by a Gibbs representation of the ergodic distribution, determined by the potential function.

3.1. **Private Provision of a Public Good.**

Consider an n-player simultaneous-move game in which player i chooses their pure strategy from a finite set: $z_i \in Z_i$. The set of pure-strategy profiles is therefore $Z = \times_{i \in N} Z_i$. Looking forward to the quantal-response process formulated later, $z^t \in Z$ will be referred to as "the state of play" at t, and Z as the state space. The payoff for player i is written $u_i(z) : Z \mapsto \mathcal{R}$. Of interest will be the set of strategy profiles $\Delta(z)$ that are accessible from a state z following the deviation of a single player. Defining $\Delta_i(z)$ as those profiles that involve the deviation of player i from z, $\Delta(z) = \bigcup_{i \in N} \Delta_i(z)$.[15]

The particular game under consideration involves the private provision of a public good. To this end, players' strategies might be interpreted as "contributions" toward the production of a public good. Such contributions cost the individual concerned $c_i(z_i) : Z_i \mapsto \mathcal{R}$. Notice that this cost function depends only upon the *individual* player's strategy. On the other hand, the value of the public good to each player depends upon the *aggregate* contribution profile, and is written $G(z) : Z \mapsto \mathcal{R}$. Thus, the payoff to player i may be written

$$u_i(z) = G(z) - c_i(z_i). \tag{3}$$

$G(\cdot)$ is taken to represent the production function for the public good. Unlike the examples considered in Section 2, there is no structure imposed on either the production technology or the cost functions.[16] The crucial assumption is that payoffs are additively separable, and players value the public good $G(z)$ identically. For instance, it may well be the case that particular individuals enjoy contributing to the public good. Similarly, the payoff component $G(z)$ may represent a public bad (such as pollution) rather than a good.

The next step is to show that this formulation admits an exact potential function $\psi(z)$. This is a real-valued function defined on the space of strategy profiles which captures the essential strategic properties of a game. Somewhat more formally, a game is an exact potential game if and only if a function $\psi(z)$ can be found such that for any pair of strategy profiles z and z', where $z' \in \Delta(z)$, and for any player i,

$$z' \in \Delta_i(z) \quad \Rightarrow \quad u_i(z') - u_i(z) = \psi(z') - \psi(z).$$

[15]Formally, $\Delta_i(z) = \left\{ z' \in Z \,|\, z'_j = z_j \ \forall j \neq i, \text{ and } z'_i \neq z_i \right\}$. Notice that $z' \in \Delta_i(z) \iff z \in \Delta_i(z')$.

[16]The action sets described in Section 2.1 are continuous. To obtain the results of that example within the present framework would simply require the construction of a sufficiently fine grid of feasible contributions.

Consider a player i contemplating a choice between two pure strategies z_i and z_i'. Fixing the behaviour of all others, player i will prefer z_i over z_i' if and only if the potential of the strategy profile with the former is larger than that with the latter. Thus players might be expected to act as if they are jointly attempting to maximise the potential function.

Here, the separation of the players' payoffs into public and private components allows such a representation, and for the above specification yields an exact potential function

$$\psi(z) = G(z) - \sum_{i=1}^{n} c_i(z_i). \tag{4}$$

Notice that $\psi(z)$ incorporates a single player's *private* benefit from the public good minus the total private cost of provision. This stands in contrast to a social-welfare function which would incorporate the *social* benefits of $nG(z)$ accruing to the entire set of players. Thus, as is to be expected, an individual player will neglect the externalities associated with their actions when making a strategy choice.

Given the discussion above, the Nash equilibria of a potential game will be connected with the maximisation of the potential function. Two subsets of Z are of interest:

$$Z^{\ddagger} = \underbrace{\{z \in Z \,|\, \psi(z) \geq \psi(z'),\, \forall z' \in Z\}}_{\text{Potential-maximising strategy profiles}} \quad \subseteq \quad Z^{\dagger} = \underbrace{\{z \in Z \,|\, \psi(z) \geq \psi(z'),\, \forall z' \in \Delta(z)\}}_{\text{Nash-equilibrium strategy profiles}}.$$

Hence $Z^{\dagger}$ is the set of strategy profiles from which no player can unilaterally deviate and increase the potential. Since a change in a player's payoff following that player's deviation corresponds one-to-one with a change in potential, $Z^{\dagger}$ is also the set of pure-strategy Nash equilibria. Clearly, the strategy profiles $Z^{\ddagger}$ which globally (rather than unilaterally) maximise the potential form a subset of the (pure-strategy) Nash equilibria. Since $Z^{\ddagger}$ is non-empty, a potential game has at least one pure-strategy equilibrium.

For the public-good production game, $Z^{\ddagger} = \arg\max_{z \in Z}[G(z) - \sum_i c_i(z_i)]$. A member of this set has the following interpretation. Suppose that a private individual were to compensate other players for their costs. Since the public good is non-excludable, this individual could only extract the private benefit of provision, and hence would wish to choose $z \in Z^{\ddagger}$.

3.2. **Quantal Response.** Rather than restrict attention to equilibria, play of the game evolves according to a quantal-response process. In particular, at a given time t, a player i is drawn at random from N (with some fixed probability $\varrho_i > 0$), and given a strategy-revision opportunity. The strategy profile is updated accordingly, so that $z^{t+1} \in \Delta_i(z^t) \cup \{z^t\}$. Interest lies in the long-run properties of this "one-step-at-a-time" strategy-revision process.

One possibility would be for a revising player to choose a best response to the current strategy profile. The long-run behaviour described by such a process is potentially history-dependent. To see this, consider a strict Nash equilibrium (of which there may be many). Any player selected to revise their strategy whilst facing others playing their part in such an equilibrium will make no change. Once the process reaches such an equilibrium it will never leave.

To circumvent this problem, noise is required. This might be generated in a variety of ways, so long as it results in the possibility that play can move away from every equilibrium with some positive (but possibly small) probability. Here this is achieved through the introduction of noise to the payoff specification. When players are called upon to revise their strategies, their true payoffs are assumed to be given by

$$\tilde{u}_i(z) = u_i(z) + \varepsilon_i(z), \tag{5}$$

where the noise terms are identically and independently distributed according to the Gumbel with scale parameter λ. Player i, when required to make a strategy revision, chooses to maximise $\tilde{u}_i(z)$. As a result the transition probabilities between strategy profiles (states) follow a standard multinomial-logit distribution (McFadden, 1974). That is,

$$z^{t+1} \in \Delta_i(z^t) \quad \Rightarrow \quad \Pr[z^t \to z^{t+1}] = \varrho_i \times \frac{\exp[\lambda u_i(z^{t+1})]}{\sum_{z' \in \Delta_i(z^t) \cup \{z^t\}} \exp[\lambda u_i(z')]}.$$

Therefore, the ratio of the transition probability for moving from state z to z' where $z' \in \Delta_i(z)$ and the transition probability for moving from state z' to z, for some i, is

$$\frac{\Pr[z \to z']}{\Pr[z' \to z]} = \exp\Big[\lambda[u_i(z') - u_i(z)]\Big] = \exp\Big[\lambda[\psi(z') - \psi(z)]\Big]. \tag{6}$$

The last equality follows from the fact that $\psi(\cdot)$ is an exact potential function. This says that the relative probability of jumping back and forth between two neighbouring states is determined by the difference in potential. It does not depend on player i's payoff from taking an action other than z_i or z_i'; this reflects the independence-of-irrelevant-alternatives property of the multinomial logit.[17]

A variant of Equation 6 also holds for sequences of transitions. To see this, consider three states z, z', and z'' for which $z' \in \Delta_i(z)$ and $z'' \in \Delta_j(z')$. Furthermore, consider a two-step transition in which the process moves from z to z'' via the intermediate state z'. Clearly,

$$\frac{\Pr[z \to z'] \times \Pr[z' \to z'']}{\Pr[z'' \to z'] \times \Pr[z' \to z]} = \exp\Big[\lambda[u_i(z') - u_i(z)] + \lambda[u_j(z'') - u_j(z')]\Big] = \exp\Big[\lambda[\psi(z'') - \psi(z)]\Big].$$

This means that the relative probability of following a path between two states versus following the reverse path depends only upon the difference in potential at the start and end of the path: the exact route taken does not matter. This key property, which stems in turn from the use of an exact potential game, allows an easy characterisation of long-run behaviour.

[17] This property, while mathematically convenient, is also restrictive. In particular, the assumed independence of the noise terms entering Equation 5 rules out payoff correlation. This is a common critique of the multinomial logit as an econometric model (Hausman and McFadden, 1984). An econometric alternative that allows for such correlation is the multinomial probit (Hausman and Wise, 1978). Unfortunately, closed-form solutions for this alternative are not available. Nevertheless, it is possible to obtain results for a probit specification, but only for vanishingly small noise. This is the approach taken by Myatt and Wallace (2004), who study the evolving play of a range of Palfrey and Rosenthal (1984) threshold games.

3.3. **Long-Run Behaviour.** The (multinomial) logit quantal-response process described above is an ergodic Markov chain on the state space Z; that is, there is a unique probability distribution π_z over Z such that $\pi_z = \lim_{t\to\infty} \Pr[z^t = z]$, independent of initial conditions. Here, and following Blume (1997), the distribution takes a simple Gibbs representation.[18]

Lemma 1. *Consider an n-player strategic-form game that admits an exact potential $\psi(\cdot)$. If play evolves by multinomial-logit quantal-response then the ergodic distribution satisfies*

$$\pi_z \equiv \lim_{t\to\infty} \Pr[z^t = z] = \frac{\exp[\lambda\psi(z)]}{\sum_{z'\in Z}\exp[\lambda\psi(z')]}. \tag{7}$$

The long-run log likelihood-ratio of states is determined by the potential difference.

A probability distribution π_z is said to be in "detailed balance" with respect to two states of a Markov chain if $\pi_z \Pr[z \to z'] = \pi_{z'} \Pr[z' \to z]$. For logit quantal-response and the distribution of Equation 7, notice that if $z' \in \Delta(z)$ then this holds automatically (Equation 6). For $z = z'$, the equality holds trivially, and for any other pairs of states the transition probabilities are zero, and the equality holds once again. If a distribution is in detailed balance with respect to all pairs of states of an ergodic Markov chain, then it is the unique ergodic distribution (Grimmett and Stirzaker, 2001, p. 238, Theorem 4). This proves Lemma 1.

The section draws to a close with the observation that, as $\lambda \to \infty$, all weight in the ergodic distribution falls on the states with maximum potential: $\lim_{\lambda\to\infty}\left[\lim_{t\to\infty}\Pr[z_t \in Z^\ddagger]\right] = 1$. That is, when noise is arbitrarily small, the strategy profiles that maximise potential are "selected". These will, of course, correspond to the potential-maximising Nash equilibria. In the public-good production game of interest, this is the equilibrium that maximises the difference between the private value of public-good production to an individual, $G(\cdot)$, and the total private cost of provision. That is, the equilibrium selected corresponds to the one that would be chosen were a private individual to bear the total costs of production (for example, by compensating every player for their cost of contribution).

4. Social Welfare and Coordination

This section contains the central results of the paper. Social welfare is considered, and the reaction of the welfare criterion to parameter changes is ascertained: when a parameter influences public-good output for a particular state of play (or subset of states of play), it is optimal to either maximise or minimise such output; intermediate values should not be chosen. For a formal justification, it is first necessary to define a notion of social welfare.

[18]Specifically, Blume (1997) discussed the relationship between potential games and play that evolves via a log-linear choice rule. He observed that detailed-balance conditions for such processes are satisfies if and only if the game admits an exact potential function.

4.1. **Social Welfare.** Social welfare at a particular strategy profile is $w(z)$. For the public-good production game, the natural candidate for $w(z)$ is the sum of players' payoffs,

$$w(z) = nG(z) - \sum_{i=1}^{n} c_i(z_i). \tag{8}$$

If strategy profile z were always played (following, perhaps, from an equilibrium-selection argument), then $w(z)$ would be a sufficient welfare measure. Given the evolution described in Section 3, however, play will evolve.[19] In the long-run, profile z will be played with probability π_z (Lemma 1), and the time average of welfare will be $W = \sum_{z \in Z} \pi_z w(z)$:[20]

$$W = \sum_{z \in Z} \left[\frac{\exp(\lambda \psi(z))}{\sum_{z' \in Z} \exp(\lambda \psi(z'))} \times w(z) \right] \quad \text{where} \quad \psi(z) = G(z) - \sum_{i=1}^{n} c_i(z_i), \tag{9}$$

and where $w(z)$ is taken from Equation 8. Equation 9 reveals that $G(z)$ influences aggregate welfare in two ways. First, it directly affects welfare $w(z)$ when z is played. Second, it also determines the potential $\psi(z)$, and hence the relative likelihood that the process is in state z. Both effects must be considered when evaluating a public-good production function.

4.2. **Influencing Production.** Taking W as social welfare, suppose that a parameter $\theta \in [0, 1]$ affects public-good production $G(z)$, but not costs. For instance, θ might be a policy tool deployed by a social planner. Whereas a single parameter is considered here, Theorem 1 will apply to every member of a larger set of parameters. Furthermore, restricting the support of the parameter to $\theta \in [0, 1]$ is without loss of generality.

Given the parameter θ, expand the production function to $G(z, \theta)$ and, similarly, expand welfare and potential to $w(z, \theta)$ and $\psi(z, \theta)$ respectively. Assume that $G(z, \cdot)$ is monotonic (this is without loss of generality) and twice continuously differentiable in θ for all z. Then

$$\frac{\partial \psi(z, \theta)}{\partial \theta} = \frac{\partial G(z, \theta)}{\partial \theta} \quad \text{and} \quad \frac{\partial w(z, \theta)}{\partial \theta} = n \times \frac{\partial G(z, \theta)}{\partial \theta}.$$

This means that the effect of θ on welfare will be proportional to its effect on the potential function. Of course, changing θ might have an impact at all states $z \in Z$. Define the set of states such that a change in the parameter has an impact on production as $Z(\theta)$. For instance, it might well be the case that θ only influences production $G(z, \theta)$ for a single state of play z, in which case the set $Z(\theta)$ will be a singleton. By inspection of the example

[19] When there is very little noise ($\lambda \to \infty$) then (generically) a single pure-strategy Nash equilibrium will be played almost all of the time, in the *very* long run. Expected transition times between equilibria, however, may be question-beggingly large (Ellison, 2000) and hence initial conditions may be the critical factor. The long run is not so long when noise is larger, and the analysis of this paper applies for all levels of noise.

[20] Another candidate for W is the expected present-discounted value of $w(z)$. With a discount factor δ,

$$\frac{1}{1-\delta} \sum_{t=0}^{\infty} \delta^t \left[\sum_{z \in Z} \Pr[z^t = z] \times w(z) \right] \longrightarrow \sum_{z \in Z} \pi_z w(z) \quad \text{as } \delta \to 1.$$

presented in Section 2.2, there are also cases where θ affects multiple states. The following is the central theorem of the paper. A formal proof is contained in Appendix A.

Theorem 1. *If $z, z' \in Z(\theta) \Rightarrow G(z,\theta) - G(z',\theta) = \text{constant}$, then social welfare W is a quasi-convex function of θ, and hence is maximised by choosing either $\theta = 0$ or $\theta = 1$.*

The condition of the theorem says that for all states in $Z(\theta)$, changes in θ have the same effect on public-good production. This is automatically true when $Z(\theta)$ is a singleton. The conclusion of the theorem is that θ should be chosen as large or as small as possible, and never at some intermediate value; maximising expected social welfare always involves either damaging the production function as much as possible, or not damaging it at all.

4.3. **Free Disposal Based on Inputs.** Before explaining the intuition behind Theorem 1, it is instructive to consider a specific application that also expands the set of parameters. Suppose that θ is a vector of parameters, one for each state of play, so that $\theta \in [0,1]^{|Z|}$, and that public-good output satisfies

$$G(z,\theta) = \theta_z \overline{G}(z) + (1-\theta_z)\underline{G}(z) \quad \text{where} \quad \overline{G}(z) \geq \underline{G}(z).$$

This application has two features. First, the parameter θ_z only affects the output of the public good in state z. This means that the production function may be manipulated very finely; the effect is conditional on a particular set of inputs (the strategy profile z). Second, the specification implements the free-disposal property. Specifically, $\overline{G}(z)$ represents the public-good production function in the absence of any intervention, whereas $\underline{G}(z)$ is output when the production process has been intentionally damaged to the maximum extent. For instance, if production could be simply thrown away, then a specification $\underline{G}(z) = 0$ and hence $G(z,\theta) = \theta_z\overline{G}(z)$ would arise. Given this interpretation, setting $\theta_z = 1$ corresponds to leaving production well alone, whereas setting $\theta_z = 0$ makes full use of free-disposal.

Theorem 1 may be applied to each component of the parameter vector θ: for each individual state z it is socially optimal to either leave production unhindered, or to maximally damage output. To see why, suppose that it is socially optimal to dent output arising from the state of play z, so that $\partial W/\partial\theta_z < 0$. Doing so is directly costly, since $w(z)$ falls. However, the potential also falls, whence the probability with which z is played diminishes. This increases the probability placed upon the other states $z' \neq z$. Put more crudely, damaging production in state z also pushes probability away from z, and hence may increase welfare. Given that such denting is optimal, it must be optimal to continue: the cost of doing so (that is, the reduction in $w(z)$) has fallen, since strategy profile z is played less often. Thus, if it is optimal to reduce θ_z, then it is optimal to reduce it to the maximum extent. In other words, aggregate welfare is quasi-convex in θ_z, for all states z.

4.4. **Control Based on Output.** In Section 4.3, changes in a parameter are specific to output arising from a single strategy profile (interpreted as the privately-provided inputs) and hence represent subtle control of the production process. Here, a second application is considered in which such influence is rather blunt.

Suppose that private actions $z \in Z$ combine to yield a public good $y \in Y$ via a production function $y = H(z)$. The value of a public good y to an individual is $F(y, \theta)$:

$$F(y, \theta) = \theta_y \overline{F}(y) + (1 - \theta_y)\underline{F}(y), \quad \text{where} \quad \overline{F}(y) \geq \underline{F}(y),$$

and where $\theta \in [0, 1]^{|Y|}$. This combines to yield a final public-good production function of $G(z, \theta) = F(H(z), \theta)$. The key difference here is that θ_y affects the value of the public good when output y occurs. This is not specific to a strategy profile; there may be many states z that induce this outcome. In fact, $Z(\theta_y) = \{z \in Z : y = H(z)\}$. A concrete illustration is the example of Section 2.1. For that example, $H(\cdot)$ takes the simple form $H(z) = \sum_i z_i$. By inspection, θ_y affects the value of the public good in the same way for all $z \in Z(\theta_y)$. Hence, Theorem 1 applies: once again, it is socially optimal to either leave production alone, or damage it maximally. This is exactly the claim made for the example of Section 2.1.

4.5. **Free Disposal and Welfare Thresholds.** The specification of Section 4.3 allows control over the public-good production function for each strategy profile z. Theorem 1 says that welfare is optimised via the maximal use of free disposal for some states of play, and no intervention for others. It does not, however, specify which states of play should be subject to the disposal of public-good output. This issue is addressed here.

For the purposes of this section and, in particular, for Theorem 2 below, suppose that $\theta \in [0, 1]^{|Z'|}$ where $Z' \subseteq Z$ and where $Z(\theta_z) = \{z\}$ for each $z \in Z'$. This says that each individual parameter θ_z affects only a single state of play. A special case of this specification is that of Section 4.3, for which $Z' = Z$. The statement of the results is assisted by some simple notation. Write $\overline{\psi}(z)$ and $\underline{\psi}(z)$ for the maximum and minimum potential in state z, and $\overline{w}(z)$ and $\underline{w}(z)$ for the maximum and minimum welfare in state z. Such maxima and minima are achieved by setting $\theta_z = 1$ and $\theta_z = 0$ respectively. Fixing λ, write W^*_λ for the maximum welfare achievable via the socially optimal parameter choice θ^*. Finally, and without loss of generality, suppose that $\overline{\psi}(z) \neq \underline{\psi}(z')$ for all $z \neq z'$.

The optimal choice of θ will be dependent upon λ. For instance, suppose that $\lambda = 0$, so that revising players select a strategy at random. Each strategy profile is equiprobable in the ergodic distribution, and throwing away production does not influence this distribution. The only impact, therefore, of setting $\theta_z = 0$ is to reduce welfare when strategy profile z is played. As a result, it is optimal to set $\theta_z = 1$ for all z.

For smaller levels of noise it becomes optimal to throw away production in some states of play. An illustration of this is precisely the example of Section 2.2. Figure 2 suggests that the optimal choice of θ will be a threshold rule. This is confirmed formally in Theorem 2 below. The threshold is closely related to the subset of states of play $\tilde{Z}$ defined as

$$\tilde{Z} = \left\{ z \,:\, \overline{\psi}(z) > \max_{z' \neq z} \underline{\psi}(z') \right\}.$$

This may be interpreted as the set of privately feasible states. More explicitly, consider a player who is required to pay the private costs of all the other players. Such an individual would receive a net payoff of $\psi(z, \theta)$. Furthermore, suppose that this player were to be asked to choose an entire strategy profile $z \in Z$. When $z \in \tilde{Z}$ the player can be induced to choose z; if not, there is always another state with higher potential, that is therefore privately preferable. Finally, write $\tilde{w} = \max_{z \in \tilde{Z}} \overline{w}(z)$. This is the maximum privately feasible welfare.

Theorem 2. *The socially optimal parameter choice θ^* satisfies*

$$\theta^*_z = \mathcal{I}[\gamma(z) \geq W^*_\lambda] \quad \textit{where} \quad \gamma(z) = \frac{\overline{w}(z) e^{\lambda \overline{\psi}(z)} - \underline{w}(z) e^{\lambda \underline{\psi}(z)}}{e^{\lambda \overline{\psi}(z)} - e^{\lambda \underline{\psi}(z)}},$$

and where $\mathcal{I}[\cdot]$ is the indicator function. Furthermore, for this choice, and as $\lambda \to \infty$,

$$\gamma(z) \to \overline{w}(z) \quad \textit{and} \quad W^*_\lambda \to \tilde{w}.$$

Hence, for all λ sufficiently large, the optimal choice satisfies $\theta^*_z = \mathcal{I}[\overline{w}(z) \geq \tilde{w}]$.

For a state z, the production function is damaged as much as possible unless $\gamma(z)$ attains a welfare threshold. Recall that a switch from $\theta_z = 0$ to $\theta_z = 1$ has two effects: it enhances the welfare arising from the play of z, and also moves probability mass away from other states of play. The threshold rule described above incorporates these two effects. On the one hand, the numerator of $\gamma(z)$ reflects the increased welfare in state z (adjusted by its probability). On the other hand, the denominator represents the probability shift alone.

To understand the second part of the theorem, it is useful to inspect the effect of a local change in θ_z. From the proof of Theorem 1 in Appendix A, Equation 11 reveals that

$$\frac{dW}{d\theta_z} \propto \frac{\partial w(z)}{\partial \theta_z} + \lambda \frac{\partial \psi(z)}{\partial \theta_z} [w(z) - W].$$

Observe that the direct welfare effect (the first term) is invariant to λ. The probability shift effect, however, increases with λ. Thus, for large enough λ, aggregate welfare increases with θ_z if and only if $w(z) > W$. This is, of course, a local effect only; the proof of Theorem 2 reveals that the same is true when considering the choice between $\theta_z = 1$ and $\theta_z = 0$. Finally, the theorem shows that $W^*_\lambda \to \tilde{w}$ as $\lambda \to \infty$. For large λ (for which the quantal responses approximate best replies), the process spends most time in states that maximise the potential. If a state z maximises the potential then it must continue to do so when $\theta_z = 1$ and $\theta_{z'} = 0$ for all other z'. Thus, a state is only potential maximising if it is a member of $\tilde{Z}$; that is, only privately feasible states are played in the limit. As a result, the highest privately feasible welfare (that is, $\tilde{w}$) is the highest achievable aggregate welfare.

5. On the Private Provision of Public Goods

Further intuition is provided in this section via a return to the example of Section 2.1. In that case the threshold rule (for general λ) is based on a simple contributions target.

5.1. **Choosing Welfare-Based Thresholds.** For large enough λ, the socially optimal choice of public-good production function satisfies $\theta^*_z = \mathcal{I}[\overline{w}(z) \geq \tilde{w}]$. This is easy to interpret when θ_z corresponds to free disposal, so that $G(z, \theta) = \theta_z \overline{G}(z)$. As in Section 2.1, suppose that strategy sets are finite subsets of the positive real line ($Z_i \subset \mathcal{R}_+$) with $0 \in Z_i$. Finally, suppose that $c_i(z_i)$ is strictly increasing with $c_i(0) = 0$.[21] Then,

$$\underline{\psi}(z) = -\sum_{i=1}^{n} c_i(z_i) \quad \Rightarrow \quad \max_{z \in Z} \underline{\psi}(z) = 0.$$

This means that if a state is to be played in the limit (that is, for large λ) then it must have a positive potential. Generically, therefore, the set of privately feasible states is

$$\tilde{Z} = \left\{ z \,:\, \overline{G}(z) > \sum_{i=1}^{n} c_i(z_i) \right\},$$

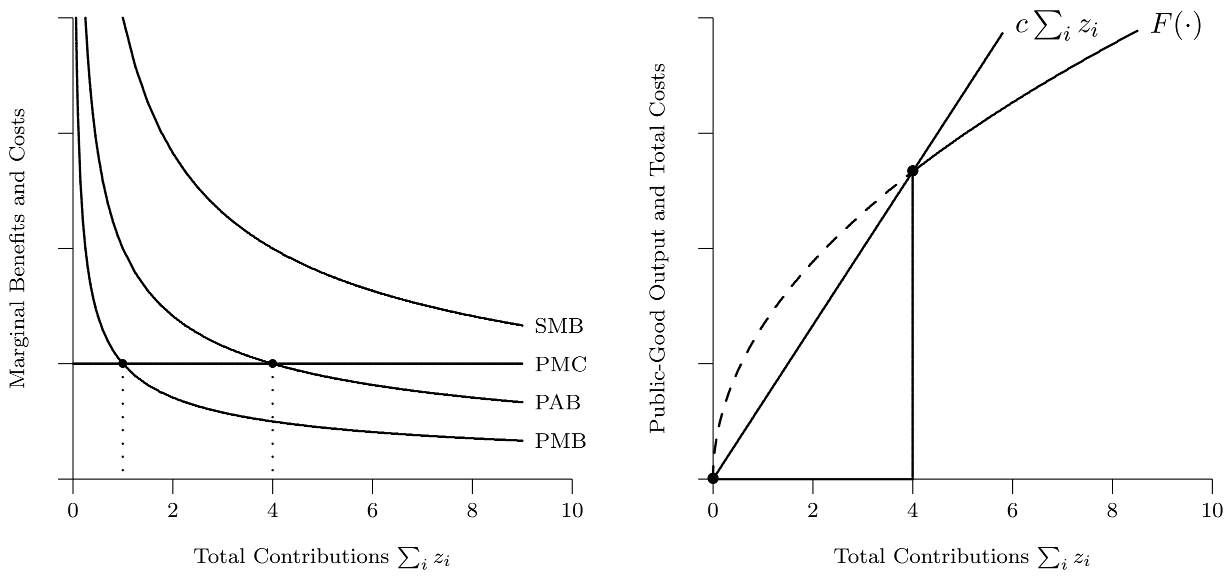
where the genericity requirement is simply that there is no positive-contribution state for which the public good's private value is exactly equal to the private cost of provision. Thus, under free disposal, it is optimal to leave production alone whenever welfare attains its privately feasible maximum.

Taking this one step further, suppose that $c_i(z_i) = cz_i$ and $\overline{G}(z) = F(\sum_i z_i)$, where $F(\cdot)$ is a strictly increasing and concave production function. The set $\tilde{Z}$ is then the set for which the (private) average benefit of contributions exceeds the private average cost. Given the concavity of $F(\cdot)$, the maximum privately feasible level of contributions is above that which would be forthcoming in the absence of any free disposal. This is illustrated in Figure 3(a). The associated production level corresponds to (so long as n is sufficiently large) the threshold below which it is socially optimal for production to be thrown away. In addition, the concavity of $F(\cdot)$ means that welfare $\overline{w}(z)$ is a concave function of total contributions. Thus, the optimal choice of θ^*_z will, in general, satisfy

$$\theta^*_z = \mathcal{I}\left[\underline{K} \leq \sum\nolimits_i z_i \leq \overline{K}\right]. \tag{10}$$

Thus there may also be a maximum threshold. The key observation is that the welfare-based threshold translates into a production-based threshold in the limit. In fact, this is true away from the limit when a further requirement is met. The next section investigates.

[21] This specification for strategy sets and costs is without loss of generality. A player's strategies may be relabelled with their associated costs, and the lowest cost may be set to zero.



(a) Costs and Benefits (b) Threshold Choice

These figures illustrate the choice of a socially optimal threshold for the specification $F(\sum_i z_i) = \sqrt{\sum_i z_i}$, linear costs with $c = \frac{1}{2}$, $n = 4$ players, and $Z_i \subset [0, 2.5]$. When production is left alone, the unique equilibrium outcome is such that $\sum_i z_i^* = 1$; panel 3(a) illustrates that this involves the equality of marginal cost (PMC $= c$) and private marginal benefit (PMB $= F'(\cdot)$). Since the social marginal benefit (SMB $= nF'(\cdot)$) is everywhere above marginal cost, the social optimum involves maximum contributions from all players so that $\sum_i z_i^s = 10$. Finally, the privately feasible states are those involving contributions where private average benefit (PAB) is greater than average cost (here equal to PMC). Thus $\tilde{Z} = \{z : \sum_i z_i \leq 4\}$. Turning to panel 3(b), it is optimal to introduce a production threshold at $K = 4$.

Figure 3. Choosing a Threshold

5.2. **Choosing Production-Based Thresholds.** Here, the restriction to a contributions game with symmetric linear costs and concave production continues to be made. The objective is to find conditions under which the threshold rule $\theta_z^* = \mathcal{I}[\gamma(z) \geq W_\lambda^*]$ of Theorem 2 translates into a pair of production thresholds (Equation 10). When $F(\cdot)$ is concave, this is always true for sufficiently large λ. For general λ, however, it is true so long as the production function is sufficiently concave. To be precise, the measure of concavity required here is the ρ-concavity notion exploited by Caplin and Nalebuff (1991).[22] A positive-valued function $F(\cdot)$ is ρ-concave if $F(\cdot)^\rho$ is concave. For $\rho = 1$, this is simply concavity. For $\rho \to 0$, it corresponds to log-concavity. Higher values of ρ correspond to more stringent concavity.

[22]Caplin and Nalebuff (1991) attributed the concept to Avriel (1972). This concept has found economic applications in recent work by Anderson and Renault (2003) and Cowan (2004), amongst others.

Theorem 3. *Suppose that the production function $F(\cdot)$ is ρ-concave, where*

$$\rho = \max_{r \geq 0} \left[1 + \frac{r^2(e^r + 1) - 2r(e^r - 1)}{(e^r - 1)(e^r - 1 - r)} \right] \approx 1.285.$$

Then, for all λ, n, and c, $\gamma(\cdot)$ is concave in total contributions $\sum_i z_i$. The optimal public-good production function employs production thresholds: $\theta^*_z = \mathcal{I}\left[\underline{K} \leq \sum_i z_i \leq \overline{K}\right]$ for $\underline{K} \leq \overline{K}$.

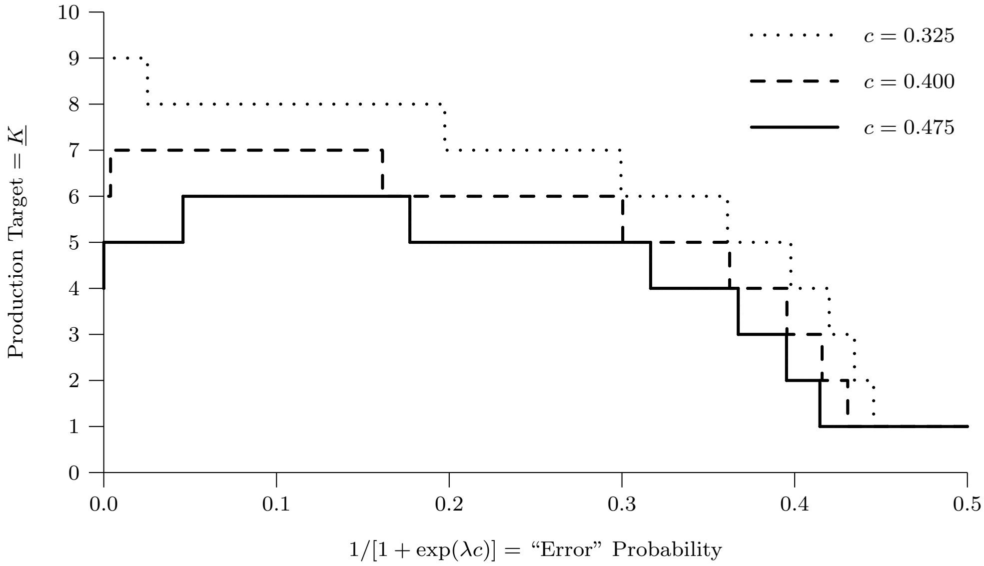
The ρ-concavity requirement of Theorem 3 has an alternative interpretation. Suppose that a public good of value $y > 0$ is to be produced. The private cost of the required contributions is $C(y) = cF^{-1}(y)$. $F(\cdot)$ is ρ-concave if and only if $-d \log C(y)/d \log y \geq \rho - 1$. This is a restriction on the elasticity of the slope of $C(y)$.[23] As an example, consider the production function $F(\sum_i z_i) = (\sum_i z_i)^\alpha$. This is ρ-concave if and only $\alpha \leq 1/\rho$.

The general conclusion emerging from both this section as well as the previous one is that the kind of public-good production technologies that work well are those that incorporate simple production targets. This agrees with the suggestions made in Section 2. A corollary is that aggregate-welfare-enhancing production functions will often lead to multiple equilibria in the underlying collective-action game. One issue that remains open is the way in which optimal threshold rules react to changes in the quantal-response parameter λ. This is the focus of the next section.

5.3. **Participation and Group Size.** It has already been argued that, when noise is very large, it is socially optimal to leave the production arising in each state of play undamaged. On the other hand, when noise is very small, only the states that result in production higher than that a private individual would be prepared to undertake should retain their full output. These observations suggest that as noise is reduced (that is, as λ grows larger) it is socially optimal to impose increasingly stringent production thresholds. An inspection of Theorem 2 explains why this might be expected. Note that $\gamma(z)$ is decreasing in λ. Hence if W^*_λ (optimal aggregate welfare) is locally increasing in λ, then the set of states for which production is left undamaged will shrink. However, were W^*_λ to be locally decreasing, this might not be the case. In fact, it is straightforward to construct examples for which the production threshold is non-monotonic in λ.

Consider a variant of the example introduced in Section 2.2, which takes the form of a simple participation game. That is, $Z_i = \{0, 1\}$ for all i, so that each player decides whether or not to contribute towards the collective action. Once again assume that the cost of participation is symmetric across the players, so that $c_i(z_i) = cz_i$. Public-good output arises from a concave production function F whose argument is the total number of contributions. For the purposes of the analysis and figure presented below, and as in the previous subsection, the specification $F(\sum_i z_i) = \sqrt{\sum_i z_i}$ is employed.

[23]This is also equivalent to the curvature measure (Robinson, 1933, pp. 40–41) known as "adjusted concavity."



The figure illustrates the socially optimal thresholds for three different values of the parameter c. The lower c, the higher will be aggregate welfare, and the higher will be the optimal threshold $\underline{K}$. Asymptotically (as $\lambda \to \infty$) the threshold becomes the highest integer x such that $x \leq 1/c^2$. For the three choices of c above, this is 9, 6, and 4 respectively. For $c = 0.325$, the threshold declines monotonically in noise. However, for $c = 0.400$ and $c = 0.475$, there is non-monotonicity in λ.

FIGURE 4. Production Thresholds and "Error" Probabilities

This specification meets the curvature restriction of Theorem 3, and hence there will an upper and lower production threshold. With n agents, consider a cost parameter that satisfies $c^2 < n/4$. This ensures that, conditional on production being left undamaged in all states, welfare is increasing in the number of contributors (those players setting $z_i = 1$). Hence it is socially optimal for all n agents to join the collective action. This condition also implies that it is never optimal to impose a meaningful upper threshold. However, it is optimal to impose a lower threshold. In fact, in the limit as $\lambda \to \infty$, the socially optimal threshold will be the highest integer $x \in \{1, \ldots, n\}$ such that $x \leq 1/c^2$. More generally, the value of the optimal threshold will depend upon λ. Figure 4 illustrates.

The noise term is on the bottom axis: rather than λ itself, a monotonic transformation $1/[1 + \exp(\lambda c)]$ is used. This represents the probability that a revising player makes an "error". To see what is meant by this, consider a world in which nobody participates, and in which there are insufficient contributions to reach the threshold. A best reply for any revising player i would be $z_i = 0$. As λ increases, however, there is positive probability that such a player will choose to participate, and set $z_i = 1$ (that is, to act "against the flow of play"). The probability such a choice is made is given by $1/[1 + \exp(\lambda c)]$.

Figure 4 shows that the socially optimal production threshold is non-monotonic in noise. As noise increases (that is, as λ decreases) the optimal threshold rises, before falling to one.

Once there is sufficient noise it is no longer optimal to use a threshold at all; this is equivalent to setting a threshold at $\underline{K} = 1$ for the purposes of this example. For intermediate levels of noise, however, the threshold should be set more stringently than it would be for lower levels. To see why, consider the states of play above the threshold. When noise is reasonably large, but not overwhelming, a good deal of time is spent in such states. Thus by increasing the threshold beyond the level appropriate for vanishingly small noise, aggregate welfare may in fact be increased. When noise is extremely small, however, these states will never be played.

6. Discussion and Related Literature

In this paper a collective action is a game in which payoffs combine private and common interests. The common element arises from a public-good production technology, and the paper asks what types of production function are conducive to success. When free disposal is possible, it is socially optimal to either leave production at its highest possible level or damage it maximally. Following this procedure, thresholds (which often take the form of production targets) emerge as a robust feature of successful collective actions.

The results could be interpreted in the context of a hypothetical social planner's problem. In contrast to classical mechanism design, policies must be robust to evolving play. This is related to the evolutionary-implementation approach taken by Sandholm (2002, 2003, 2004), in which the desired outcome must be eventually learnt by players who follow reasonable myopic adjustment processes. It is not enough simply to create a socially optimal equilibrium; rather, the mechanism must ensure that such an equilibrium actually is played.[24]

The alternative interpretation favoured here does not rely upon the presence of a mechanism designer. Rather, the results help to explain why the shape of a public-good production function matters. This has been a long-standing issue for sociology. A leading example is "critical mass" theory (Marwell and Oliver, 1993).[25] Oliver, Marwell, and Teixeira (1985), for instance, considered a variety of production-function shapes, labelled as "accelerating" (convex), "decelerating" (concave), "general third-order" (S-shaped), and "step" functions. They relied upon a variety of numerical simulations, and hence it is difficult to provide a concise summary of their conclusions.[26] Nevertheless, they observed that "accelerating" and "step" production functions can lead to success by making individuals pivotal. However, a "critical mass" of participants is needed in order to reach this point. As Macy (1991) noted, this leads to "[...] a bistable system with cooperative and noncooperative equilibria." Put simply, the introduction of a threshold creates an equilibrium-selection problem.

[24] The evolutionary-implementation approach follows the lead of Cabrales (1999), who noted (p. 160) that "[...] very little attention has been paid to the issue of how equilibrium is reached, and whether it is stable."
[25] A retrospective assessment of this literature was provided by Oliver and Marwell (2001).
[26] This is also true of subsequent strands of literature, including Macy (1990) and Heckathorn (1993, 1996). The current paper is similar in spirit. Here, however, there is no need to resort to simulations: by leaning upon the theoretical advances of Blume (1997) and others, closed-form results are available.

This paper offers a single framework which both addresses the equilibrium-selection problem and provides a theoretical justification for its existence. The results show that the thresholds of Granovetter (1978) and the "minimal contributing sets" (that is, a minimum number of participants in a binary-action game) of van de Kragt, Orbell, and Dawes (1983) are precisely the features that are associated with the long-run robust success of collective actions. The paper also develops a criterion for setting such production targets. These targets should be chosen so that, in effect, voluntary participants are asked jointly to achieve a level of social welfare greater than that which could be profitably delivered by an individual facing all private costs but receiving only a single private benefit. Somewhat mischievously, they are asked to "beat the private sector."

Appendix A. Omitted Proofs

Proof of Theorem 1. From Equations 4, 8, and 9, and suppressing arguments z and θ;

$$W \sum_{z \in Z} \exp[\lambda \psi] = \sum_{z \in Z} w \exp[\lambda \psi].$$

Differentiating both sides with respect to θ yields

$$\frac{dW}{d\theta} \sum_{z \in Z} \exp[\lambda \psi] + \lambda W \sum_{z \in Z} \frac{\partial \psi}{\partial \theta} \exp[\lambda \psi] = \sum_{z \in Z} \exp[\lambda \psi] \left[\frac{\partial w}{\partial \theta} + \lambda w \frac{\partial \psi}{\partial \theta} \right].$$

Collecting potential terms on the right-hand side and rewriting;

$$\frac{dW}{d\theta} \sum_{z \in Z} \exp[\lambda \psi] = \sum_{z \in Z} \exp[\lambda \psi] \left[\frac{\partial w}{\partial \theta} + \lambda \frac{\partial \psi}{\partial \theta} [w - W] \right]. \tag{11}$$

Suppose for a moment that the policy parameter θ only affects production (and hence welfare and potential) in a single state z (so that $Z(\theta)$ is a singleton set). Equation 11 simplifies to

$$\frac{dW}{d\theta} \sum_{z' \in Z} \exp[\lambda \psi] = \exp[\lambda \psi] \left[\frac{\partial w}{\partial \theta} + \lambda \frac{\partial \psi}{\partial \theta} [w - W] \right],$$

which means that the first differential of expected welfare with respect to the policy parameter can be rewritten in a simple differential-equation format:

$$\frac{dW}{d\theta} = \pi_z \left[\frac{\partial w(z)}{\partial \theta} + \lambda \frac{\partial \psi(z)}{\partial \theta} [w(z) - W] \right]. \tag{12}$$

Differentiate Equation 11 again with respect to θ to obtain

$$\begin{aligned} &\frac{d^2 W}{d\theta^2} \sum_{z \in Z} \exp[\lambda \psi] + \lambda \frac{dW}{d\theta} \sum_{z \in Z} \frac{\partial \psi}{\partial \theta} \exp[\lambda \psi] \\ &\qquad = \lambda \sum_{z \in Z} \frac{\partial \psi}{\partial \theta} \exp[\lambda \psi] \left[\frac{\partial w}{\partial \theta} + \lambda \frac{\partial \psi}{\partial \theta} [w - W] \right] \\ &\qquad\qquad + \sum_{z \in Z} \exp[\lambda \psi] \left[\frac{\partial^2 w}{\partial \theta^2} + \lambda \frac{\partial \psi}{\partial \theta} \left[\frac{\partial w}{\partial \theta} - \frac{dW}{d\theta} \right] + \lambda \frac{\partial^2 \psi}{\partial \theta^2} [w - W] \right]. \end{aligned}$$

Next, evaluate this expression at a stationary point, so that $dW/d\theta = 0$. At such a θ,

$$\frac{d^2W}{d\theta^2}\sum_{z\in Z}\exp[\lambda\psi] = \lambda\sum_{z\in Z}\frac{\partial\psi}{\partial\theta}\exp[\lambda\psi]\left[\frac{\partial w}{\partial\theta}+\lambda\frac{\partial\psi}{\partial\theta}[w-W]\right] + \sum_{z\in Z}\exp[\lambda\psi]\left[\frac{\partial^2 w}{\partial\theta^2}+\lambda\frac{\partial\psi}{\partial\theta}\frac{\partial w}{\partial\theta}+\lambda\frac{\partial^2\psi}{\partial\theta^2}[w-W]\right]. \quad (13)$$

Once again, notice that this expression can be simplified greatly in the case when $Z(\theta)$ is a singleton set (when local changes in θ only have an effect upon a single state z). In fact, from Equation 12,

$$\frac{d^2W}{d\theta^2} = \frac{\partial\pi_z}{\partial\theta}\left[\frac{\partial w(z)}{\partial\theta}+\lambda\frac{\partial\psi(z)}{\partial\theta}[w(z)-W]\right] + \pi_z\left[\frac{\partial^2 w(z)}{\partial\theta^2}+\lambda\frac{\partial\psi(z)}{\partial\theta}\left[\frac{\partial w(z)}{\partial\theta}-\frac{dW}{d\theta}\right]+\lambda\frac{\partial^2\psi(z)}{\partial\theta^2}[w(z)-W]\right].$$

The first term on the right-hand side becomes zero when evaluated at a stationary point (see Equation 12) and thus, when $dW/d\theta = 0$, the second differential is

$$\frac{d^2W}{d\theta^2} = \pi_z\left[\frac{\partial^2 w(z)}{\partial\theta^2}+\lambda\frac{\partial\psi(z)}{\partial\theta}\frac{\partial w(z)}{\partial\theta}+\lambda\frac{\partial^2\psi(z)}{\partial\theta^2}[w(z)-W]\right].$$

For this case, now substitute $w(z) = nG(z) - \sum_{i\in N} c_i(z_i)$ and $\psi(z) = G(z) - \sum_{i\in N} c_i(z_i)$. Then

$$\begin{aligned}\frac{d^2W}{d\theta^2} &= \pi_z\left[n\frac{\partial^2 G(z)}{\partial\theta^2}+\lambda n\left[\frac{\partial G(z)}{\partial\theta}\right]^2+\lambda\frac{\partial^2 G(z)}{\partial\theta^2}[w(z)-W]\right]\\ &= \pi_z\left[\lambda n\left[\frac{\partial G(z)}{\partial\theta}\right]^2+\frac{\partial^2 G(z)}{\partial\theta^2}\underbrace{\Big[n+\lambda[w(z)-W]\Big]}_{\text{Zero from Equation 12}}\right] = \pi_z\lambda n\left[\frac{\partial G(z)}{\partial\theta}\right]^2 > 0,\end{aligned}$$

and hence, when $Z(\theta)$ is a singleton set, the result is proven. The remainder of the proof tackles the more general case by returning to the expression in Equation 13. First notice that, by definition, $z \notin Z(\theta) \Rightarrow \partial\psi(z,\theta)/\partial\theta = \partial w(z,\theta)/\partial\theta = 0$. Equation 11 explicitly becomes

$$\frac{dW}{d\theta}\sum_{z'\in Z}\exp[\lambda\psi(z')] = \sum_{z\in Z(\theta)}\exp[\lambda\psi(z)]\left[\frac{\partial w(z)}{\partial\theta}+\lambda\frac{\partial\psi(z)}{\partial\theta}[w(z)-W]\right].$$

Now impose again $\psi(z,\theta) = G(z,\theta) - \sum_{i\in N} c_i(z_i)$ and $w(z,\theta) = nG(z,\theta) - \sum_{i\in N} c_i(z_i)$, so that $\partial w/\partial\theta = n\partial\psi/\partial\theta = n\partial G/\partial\theta$. Then

$$\frac{dW}{d\theta}\sum_{z'\in Z}\exp[\lambda\psi(z')] = \sum_{z\in Z(\theta)}\exp[\lambda\psi(z)]\frac{\partial G(z)}{\partial\theta}\Big[n+\lambda[w(z)-W]\Big],$$

and Equation 13 becomes

$$\frac{d^2W}{d\theta^2}\sum_{z'\in Z}\exp[\lambda\psi] = \sum_{z\in Z(\theta)}\exp[\lambda\psi]\left[n\left[\frac{\partial G}{\partial\theta}\right]^2+\Big[n+\lambda[w-W]\Big]\left[\frac{\partial^2 G}{\partial\theta^2}+\lambda n\left[\frac{\partial G}{\partial\theta}\right]^2\right]\right].$$

Now, making use of the assumption that $z, z' \in Z(\theta) \Rightarrow G(z,\theta) - G(z',\theta) = \text{const}$, it follows that for any $z, z' \in Z(\theta)$, $\partial G(z,\theta)/\partial\theta = \partial G(z',\theta)/\partial\theta$ and $\partial^2 G(z,\theta)/\partial\theta^2 = \partial^2 G(z',\theta)/\partial\theta^2$. Hence

$$\frac{dW}{d\theta}\sum_{z'\in Z}\exp[\lambda\psi(z')] = \frac{\partial G(z)}{\partial\theta}\sum_{z\in Z(\theta)}\exp[\lambda\psi(z)]\Big[n+\lambda[w(z)-W]\Big],$$

where $\partial G(z)/\partial\theta \neq 0$, since $z \in Z(\theta)$. So at a stationary point $dW/d\theta = 0$, the second term on the right-hand side must be zero. As a result, the second derivative is

$$\frac{d^2W}{d\theta^2} = n\left[\frac{\partial G(z)}{\partial\theta}\right]^2 \sum_{z\in Z(\theta)}\exp[\lambda\psi(z)]\Big/\sum_{z'\in Z}\exp[\lambda\psi(z')] > 0,$$

and W is therefore quasi-convex in θ as required. □

Proof of Theorem 2. From Equation 9, aggregate welfare satisfies

$$W = \sum_{z\in Z}\frac{w(z)e^{\lambda\psi(z)}}{\sum_{z'\in Z}e^{\lambda\psi(z')}} = \frac{A}{B}, \quad \text{where} \quad A = \sum_{z\in Z} w(z)e^{\lambda\psi(z)} \quad \text{and} \quad B = \sum_{z\in Z} e^{\lambda\psi(z)}.$$

Suppose that θ^* is a collection of optimal parameters, and write W^*_λ for the associated aggregate welfare. Pick a state of play z where $\theta^*_z = 0$, if one exists. Since this is an optimal choice, a switch to $\theta_z = 1$ must result in (weakly) lower aggregate welfare. Thus,

$$\frac{A + \overline{w}(z)e^{\lambda\overline{\psi}(z)} - \underline{w}(z)e^{\lambda\underline{\psi}(z)}}{B + e^{\lambda\overline{\psi}(z)} - e^{\lambda\underline{\psi}(z)}} \leq \frac{A}{B} \quad \Leftrightarrow \quad \gamma(z) \leq \frac{A}{B},$$

which follows from simple algebraic manipulation. For a state z' where $\theta^*_{z'} = 1$, a switch to $\theta_{z'} = 0$ must result in (weakly) lower aggregate welfare, and hence $\gamma(z') \geq A/B$.

For the second part of the theorem, since $\overline{\psi}(z) > \underline{\psi}(z)$ and $\overline{w}(z) > \underline{w}(z)$, by inspection $\gamma(z) \to \overline{w}(z)$ as $\lambda \to \infty$. To show that $W^*_\lambda \to \tilde{w}$, as noted in the text, observe that when $\lambda \to \infty$ the process spends almost all of its time in the set $\tilde{Z}$. Hence $\lim_{\lambda\to\infty} W^*_\lambda \leq \max_{z\in\tilde{Z}} w(z) \equiv \tilde{w}$. Moreover, this bound can be attained by setting $\theta_z = 1$ for the maximiser of $w(z)$ over $\tilde{Z}$, and $\theta_{z'} = 0$ for all other states. For the final statement of the theorem, in states where $\overline{w}(z) > W^*_\lambda$ it must be the case that $\theta^*_z = 1$ for all λ large enough, and similarly $\theta_{z'} = 0$ when $\overline{w}(z') < W^*_\lambda$. It remains to be shown that $\theta^*_z = 1$ for the state z that achieves $w(z) = \tilde{w}$ within $\tilde{Z}$. The process stays almost always within $\tilde{Z}$ for large enough λ. Moreover, it must spend almost all of its time in the state of play z; if not, then by setting $\theta_z = 1$ and $\theta_{z'} = 0$ for all other z' welfare could be strictly increased. Given that the process is almost always at z for large λ, it makes no sense to set $\theta_z = 0$. □

Proof of Theorem 3. For a Cournot-contributions game with symmetric linear costs, the potential and welfare of a state z depend only upon the sum of contributions defined as $x \equiv \sum_i z_i$. In fact,

$$\gamma(z) = \hat{\gamma}(x) \quad \text{where} \quad \hat{\gamma}(x) = \frac{L(nF(x) - cx) + cx}{L-1} \quad \text{and} \quad L = e^{\lambda F(x)}.$$

To check concavity, differentiate twice to obtain

$$\hat{\gamma}'(x) = \frac{nLF'(x)[L-1-\lambda F(x)]}{(L-1)^2} - c, \quad \text{and}$$

$$\hat{\gamma}''(x) = \frac{\lambda nL[F'(x)]^2}{L-1} + \frac{nL[L-1-\lambda F(x)][\lambda[F'(x)]^2 + F''(x)]}{(L-1)^2} - \frac{2\lambda nL^2[F'(x)]^2[L-1-\lambda F(x)]}{(L-1)^3}.$$

For concavity, $\hat{\gamma}(x)$ must satisfy $\hat{\gamma}''(x) < 0$, which occurs if and only if

$$\frac{F''(x)}{F'(x)}(L-1)[L-1-\lambda F(x)] < \lambda F'(x)[2(L-1)-\lambda-\lambda LF(x)].$$

This will require $F(\cdot)$ to be sufficiently concave. In fact, writing $r = \lambda F(x)$, this requirement is

$$\frac{F(x)F''(x)}{[F'(x)]^2} \leq \frac{2r(e^r-1)-r^2(e^r+1)}{(e^r-1)(e^r-1-r)}.$$

A sufficient condition for concavity of $\hat{\gamma}(x)$ is for the above inequality to hold for all $r \geq 0$. Noting that $F(\cdot)$ is ρ-concave if and only if $-F(x)F''(x)/[F'(x)]^2 \geq \rho - 1$, this completes the proof. □

References

Anderson, S. P., and R. Renault (2003): "Efficiency and Surplus Bounds in Cournot Competition," *Journal of Economic Theory*, 113(2), 253–64.

Avriel, M. (1972): "r-Convex Functions," *Mathematical Programming*, 2, 309–23.

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.

——— (1992): "Uniqueness of Nash Equilibrium in Private Provision of Public Goods: An Improved Proof," *Journal of Public Economics*, 49(3), 391–92.

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.

Cabrales, A. (1999): "Adaptive Dynamics and the Implementation Problem with Complete Information," *Journal of Economic Theory*, 86(2), 159–84.

Caplin, A., and B. Nalebuff (1991): "Aggregation and Social Choice: A Mean Voter Theorem," *Econometrica*, 59(1), 1–23.

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

Cornes, R., and T. Sandler (1996): *The Theory of Externalities, Public Goods and Club Goods*. Cambridge University Press, London, 2nd edn.

Cowan, S. (2004): "The Welfare Effects of Third-Degree Price Discrimination," *Oxford University Department of Economics Discussion Paper*, 205.

Ellison, G. (2000): "Basins of Attraction, Long Run Stochastic Stability and the Speed of Step-by-Step Evolution," *Review of Economic Studies*, 67(1), 17–45.

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

Grimmett, G. R., and D. R. Stirzaker (2001): *Probability and Random Processes*. Oxford University Press, Oxford, 3rd edn.

Harsanyi, J. C., and R. Selten (1988): *A General Theory of Equilibrium Selection in Games*. MIT Press Classics, Cambridge, MA.

Hausman, J. A., and D. McFadden (1984): "Specification Tests for the Multinomial Logit Model," *Econometrica*, 52(5), 1219–40.

Hausman, J. A., and D. A. Wise (1978): "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences," *Econometrica*, 46(2), 403–26.

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.

Holmström, B. R. (1982): "Moral Hazard in Teams," *Bell Journal of Economics*, 13(2), 324–40.

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

Ley, E. (1996): "On the Private Provision of Public Goods: A Diagrammatic Exposition," *Investigaciones Económicas*, 20(1), 105–23.

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

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

Marschak, J. (1955): "A Theory of Teams," *Management Science*, 1(2), 127–37.

Marwell, G., and P. E. Oliver (1993): *The Critical Mass in Collective Action: A Micro-Social Theory.* Cambridge University Press, Cambridge and New York.

McFadden, D. (1974): "Conditional Logit Analysis of Quantitative Choice Behavior," in *Frontiers in Econometrics*, ed. by P. Zarambka. Academic Press, New York and London.

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

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., and C. Wallace (2004): "The Evolution of Collective Action," *Oxford University Department of Economics Discussion Paper*, 190.

Oliver, P. E., and G. Marwell (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.

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.

Robinson, J. (1933): *The Economics of Imperfect Competition.* MacMillan, London.

Sandholm, W. H. (2002): "Evolutionary Implementation and Congestion Pricing," *Review of Economic Studies*, 69(3), 667–89.

——— (2003): "Pigouvian Pricing and Stochastic Evolutionary Implementation," *University of Wisconsin.*

——— (2004): "Negative Externalities and Evolutionary Implementation," *University of Wisconsin*, forthcoming in the *Review of Economic Studies.*

Shibata, H. (1971): "A Bargaining Model of the Pure Theory of Public Expenditure," *Journal of Political Economy*, 79(1), 1–29.

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.

Warr, P. (1983): "The Private Provision of a Public Good is Independent of the Distribution of Income," *Economics Letters*, 13(2-3), 207–11.

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

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

——— (2001): "The Dynamics of Conformity," in *Social Dynamics*, ed. by S. N. Durlauf, and H. P. Young, chap. 5, pp. 133–53. MIT Press, Cambridge, MA.