The Economic Journal, **119** (*January*), 61–90. © The Author(s). Journal compilation © Royal Economic Society 2009. Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

EVOLUTION, TEAMWORK AND COLLECTIVE ACTION: PRODUCTION TARGETS IN THE PRIVATE PROVISION OF PUBLIC GOODS*

David P. Myatt and Chris Wallace

Collective-action problems arise when private actions generate common consequences; for example, the private provision of a public good. This article asks: what shapes of public-good production function work well when play evolves over time, and hence moves between equilibria? Welfare-maximising public-good production functions yield nothing when combined efforts fall below some threshold but otherwise maximally exploit the production-possibility frontier. They generate multiple equilibria: coordinated teamwork is integral to successful collective actions. Optimal thresholds correspond to the output that individuals who pay all private costs but enjoy only private benefits would be just willing to provide.

1. Public-good Production and Collective Action

A classic collective-action problem arises when 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 a private component, specific to the individual and depending only on that player's action, and a public component, common to all players and depending on all actions. 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.

The public component of payoffs corresponds to a public-good production function. Given a family of such functions, the member that leads to the highest welfare is sought. A 'long-run' welfare measure is employed, since play evolves according to a strategy-revision process.

Of particular interest is the shape of public-good production functions. For example, consider a 'Cournot contributions' game (Shibata, 1971; Warr, 1983; Bergstrom *et al.*,

* The authors thank editors, anonymous referees and colleagues, particularly Ken Binmore, Larry Samuelson and Peyton Young, as well as seminar participants in Bristol, Copenhagen, Oxford and Helsinki for their comments and advice. Somewhat confusingly, earlier versions of this article were circulated under the titles 'Production Targets and Free Disposal in the Private Provision of Public Goods' and 'An Evolutionary Justification for Thresholds in Collective Action Problems.' This article was submitted before David Myatt was invited to become an editor of this Journal.

1986, 1992) in which public-good output is a concave function of the contributions made simultaneously by players. Such a game typically has a unique Nash equilibrium, and provision falls short of the social optimum. Next, consider a production function for which output is zero unless players' contributions reach a critical threshold (Palfrey and Rosenthal, 1984). If individual players are unable or unwilling to provide the threshold level of contribution unilaterally, then there is a Nash equilibrium in which output is zero. On the other hand, there can also be an equilibrium in which a team of players jointly share the burden of provision. A contributing team member bears only a moderate cost but enjoys the full non-excludable benefit of provision; dropping out causes production to collapse, and so each player is pivotal to the success of the collective action. This 'good' equilibrium could involve a higher level of public-good provision than that achieved with a concave production function. However, there is the risk that the 'bad' equilibrium is played.

To illustrate, consider the academic committees that will be familiar to many readers. Voluntary attendance involves a private cost and (arguably) generates a public benefit; alas, free riding may lead to sub-optimal attendance.[1] Of course, many committees employ quora. A quorum is a production target: a collective action (the committee's decision) goes ahead only if a threshold (the quorum) is reached. The quorum generates a particular non-concave public-good production function, where output from a sub-quorate committee is zero. This is technologically inefficient, since output is effectively discarded (decisions are put on hold) when the collective inputs fall below the threshold. Nevertheless, the use of a quorum might be seen as a response to a classic moral-hazard-in-teams problem: by setting the threshold equal to the socially optimal level of participation, a social planner makes each attendee pivotal to the collective action. However, this may also create an equilibrium in which nobody attends the meeting. Unless society can choose which equilibrium is played, the outcome is indeterminate, and welfare may be higher or lower than in a world without a quorum.

As a second illustration, consider an environment without a social planner so that incentives cannot be imposed arbitrarily: the voluntary provision of open-source software. Certain open-source activities, such as bug-fixing, can be conducted individually, and their success does not depend upon the actions of others.[2] The creation of complex software, however, involves specialisation; the entire system works only if all team members pull their weight.[3] The need for teamwork generates a collective-action game with 'good' and 'bad' equilibria; this scenario bears a striking semblance to that of a quorum-constrained committee.

Moving beyond the examples considered above, an assessment of threshold effects is a special case of the general question addressed by the article: which public-good production technologies work well? Thresholds, production targets, or quora are not imposed; rather, these features appear naturally from a welfare ordering of public-good production technologies.

[1] Conversely, too many cooks might spoil the broth; over-attendance is also addressed within this article.

[2] Johnson (2002) modelled open-source software provision as a voluntary-action game, while Myatt and Wallace (2008*a*) used an evolutionary analysis of the volunteer's dilemma to identify the likely providers.

[3] Studies based on data from the sourceforge.net collaborative development environment (Giuri *et al.*, 2006) have observed that teamwork is fundamental to the open-source model.

Three observations emerge from these introductory remarks. First, to ascertain welfare performance it is necessary to determine which actions will be taken: in the examples above, the Nash solution concept does not always yield a unique prediction of play in a one-shot game. Second, inferior technologies which do not fully exploit the boundaries of the production-possibility frontier (in the case of a sub-quorate committee, some output is effectively thrown away) might well yield higher welfare by encouraging teamwork. Third, some of these production technologies involve a threshold rule: output is zero unless the inputs rise above a critical level. These three observations lead to three questions. Which strategy profiles (a voluntary contribution from each individual) will be played and how often? To what extent is the use of technologically inefficient production processes socially desirable? Finally, if a threshold (such as a quorum) could be imposed upon an otherwise standard production function, then at what level should this threshold be set?

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 randomly selected player enjoys a strategy-revision opportunity and chooses a quantal response (a 'smoothed' best reply) to the contemporary decisions of others.[4] The analysis of this process leads to a unique characterisation of play via a probability distribution over strategy profiles (inputs to production) which reflects the frequency with which each is played in the long run. This (ergodic) distribution takes a simple form since the collective-action game considered here is a potential game (Monderer and Shapley, 1996): a game in which players act as though they are jointly maximising a single real-valued function (a potential function) of their combined actions.[5] The potential is the private benefit enjoyed by a single individual minus the sum of all private costs. Under the strategy-revision process considered here, the long-run log likelihood of a strategy profile (and so the corresponding public-good provision) is proportional to its potential.

Turning to the second question, a family of feasible public-good production functions is considered. As an example, suppose that feasibility is determined by a production-possibility frontier: a function $\bar{G}(z)$ where z represents the players' combined inputs. (Equivalently, z is a strategy profile from the collective-action game.) A feasible production function $G(z)$ satisfies $0 \leq G(z) \leq \bar{G}(z)$. Given this family, the welfare-maximising member is sought, where aggregate welfare is the long-run average of the sum of the players' payoffs. The article's analysis reveals that the optimal production function attains the frontier ($G(z) = \bar{G}(z)$) for some input combinations, and generates nothing ($G(z) = 0$) for others; however, it never takes intermediate values, so that $0 < G(z) < \bar{G}(z)$ is ruled out.

To see why, notice that a reduction in $G(z)$ has two effects. Firstly, it directly lowers welfare whenever z is played. Secondly, it makes z less attractive (formally, it reduces its potential) and so other strategy profiles are played more often. If these other

[4] Specifically, players update according to a multinomial-logit choice rule (McKelvey and Palfrey, 1995). These choices may be generated by an underlying random-utility model, or may be viewed as 'smoothed' best-replies. 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.

[5] 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).

profiles yield higher welfare, then the aggregate impact (direct and indirect) of the two effects might be to enhance welfare. If this is so, then any further reduction of production also must be (in the aggregate) welfare enhancing: the negative direct effect is less severe since z is played less frequently. Continuing this argument, if it is welfare-increasing to reduce $G(z)$ then it is optimal to reduce it all the way to zero; similarly, if it is optimal to increase $G(z)$ then it is optimal to raise it all the way to the upper bound $\bar{G}(z)$.

A direct conclusion is that welfare-maximising production functions exhibit discontinuous steps and so lack the concavity of traditional textbook specifications. This further implies that the associated collective-action games tend to exhibit multiple equilibria; teamwork (coordination on 'good' equilibria) is a necessary component of a successful collective action.

The answer to the third and final question is that an optimally shaped public-good production function implements a welfare-based threshold rule. This rule is easiest to describe when the evolution of play approximates a best-reply process; that is, when the noise associated with quantal-response strategy revisions is small. Using the notation developed above, the optimal production function satisfies $G(z) = \bar{G}(z)$ if and only if the welfare generated by the play of z exceeds a critical target. This target threshold may be calculated via the following two-step procedure. First, consider the input combinations for which an individual's private benefit from public-good production $\bar{G}(z)$ exceeds the total private cost. (In the language of potential games, these are the strategy profiles for which the potential is non-negative.) These production plans are interpreted as privately feasible, in the sense that a private operation, by someone who bears all private costs but cannot capture non-excludable spillovers, would generate positive profits. Second, calculate the maximum welfare achieved by maximising across the set of privately feasible input combinations: this maximum is the critical welfare target. This welfare threshold typically falls short of the social optimum. Were a higher threshold to be chosen, however, the process would languish in a state of low production for much of the time. Setting a lower target, whilst less ambitious, helps the process to spend more time in 'team success' rather than 'team failure' states of play.

In some circumstances the optimal public-good production function can be characterised more simply. If the input to production is the simple sum of players' contributions and the production possibility frontier is sufficiently concave then the socially optimal production function couples a contribution target (a minimum threshold) with a contribution cap (a maximum). That is, fully efficient public-good production goes ahead if and only if total contributions fall within a specified interval. The contribution target typically falls below the socially optimal level of provision. In the context of the academic committees mentioned above, this contribution target corresponds precisely to the use of a quorum.

The remainder of the article is organised as follows. Section 2 studies a pair of worked examples which serve to motivate the questions asked by the article and illustrate the methods used to answer them. The class of collection-action games and the process via which play evolves are both described in Section 3. Section 4 contains the central results: Proposition 1 shows that an optimal public-good production function either makes full use of the feasible technology or produces nothing, while Proposition 2 calculates the welfare-based threshold rule that determines when production opportunities should be thrown away. Section 5 applies the results to a class of Cournot-contribution games: Proposition 3 demonstrates the optimal combination of a contribution target and cap. Section 6 relates the results to earlier literature, drawing particular connections with sociological theories of critical mass.

2. Two Simple Worked Examples

This Section presents two simple worked examples. The first motivates the key questions of the article, while the second illustrates the methodology used to answer these questions.

2.1. *A Simple Cournot-contributions Game*

Consider a game in which n symmetric players simultaneously and voluntarily contribute toward the production of a public good. Player i contributes z_i and so incurs a private cost $c(z_i)$. The contributions yield a public good of value $\bar{F}(Z)$ where $Z \equiv \sum_{j=1}^{n} z_j$. The non-excludable public good is enjoyed equally by everyone and so player i seeks to maximise $\bar{F}(Z) - c(z_i)$. Given the convexity of $c(\cdot)$ and concavity of $\bar{F}(\cdot)$, together with a few technical assumptions, there is a unique pure-strategy Nash equilibrium. This equalises (private) marginal cost and (private) marginal benefit. Alas, production falls short of the social optimum: ideally players would contribute more.

Now consider a second production function $F_K(\cdot)$ which abandons the concavity associated with $\bar{F}(\cdot)$: suppose that $F_K(\cdot)$ yields the same output as $\bar{F}(\cdot)$ if and only if total contributions exceed some threshold K, and otherwise yields nothing (Figure 1):

$$F_K(Z) \equiv \begin{cases} \bar{F}(Z) & \text{if } Z \geq K, \text{ or} \\ 0 & \text{otherwise.} \end{cases}$$

$F_K(\cdot)$ is (trivially) non-concave and is less productive than $\bar{F}(\cdot)$. Nevertheless, it can enhance performance by encouraging teamwork. To see this, suppose that K exceeds the total Nash contribution under $\bar{F}(\cdot)$ and satisfies $c(K) > \bar{F}(K) > c(K/n)$. Under $F_K(\cdot)$ there is a Nash equilibrium in which each player contributes K/n. Successful

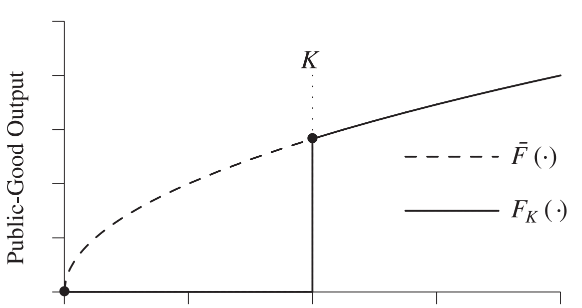

Fig. 1. *A Production Function with a Threshold*

production depends pivotally on each player; the inequality $\bar{F}(K) > c(K/n)$ ensures that team members do not deviate.[6] However, a player contributes if and only if others do so; the inequality $c(K) > \bar{F}(K)$ ensures that it is unprofitable to be the sole contributor, and so there is an equilibrium involving no contributions.[7] So, a coordination problem arises: while the threshold in $F_K(\cdot)$ could lead to greater contributions, it might also lead to the collapse of production. A proper welfare comparison of $\bar{F}(\cdot)$ and $F_K(\cdot)$ needs to address an equilibrium-selection problem.

The comparison of $\bar{F}(\cdot)$ and $F_K(\cdot)$ is closely related to a hypothetical social-planning problem. Given the production function $\bar{F}(\cdot)$, a benevolent planner might wish to impose the socially optimal contributions. However, suppose that players cannot be coerced but instead the planner is only able to exploit a free disposal opportunity by throwing away output. One possibility would be to discard contributions which fall below K: the idea is to force players to work together as a coordinated team. Indeed, the planner might consider equating K to the socially optimal aggregate contribution.[8] But, as the target K becomes more ambitious, it may become difficult for players to successfully coordinate on the 'good' equilibrium.

To explore the equilibrium-selection problem further, consider a simple two-player scenario. The 'team success' and 'team failure' equilibria correspond to $z_1 = z_2 = K/2$ and $z_1 = z_2 = 0$, respectively. This scenario can be thought of as a 2×2 coordination game (Figure 2).[9]

One equilibrium-selection criterion is risk dominance (Harsanyi and Selten, 1988). In a symmetric 2×2 game, the risk-dominant equilibrium involves pure strategies which are best replies given the belief that opposing actions are equally likely; heuristically, a risk-dominant equilibrium is relatively safe. So, if Player 1 believes that $z_2 = K/2$ and $z_2 = 0$ are equally likely, then an inspection of Figure 2 confirms that $z_1 = K/2$ is a strict best reply if and only if $F(K)/2 > c(K/2)$. Equivalently, the 'team success' equilibrium is risk dominant if and only if $F(K) \geq 2c(K/2)$. This says that the individual private (not social) benefit from attaining the threshold exceeds the total private cost. With n players, the analogous condition would be $F(K) \geq nc(K/n)$. In the context of a production target, this condition is not necessarily satisfied when K is set at the socially efficient level. Thus, for the hypothetical social-planning problem described earlier, a planner might temper ambition with realism by setting K such that this condition is met (Figure 3).

Amongst the justifications for the use of risk dominance as an equilibrium-selection criterion: two are described here. When rational players lack common knowledge of the payoffs, Carlsson and Van Damme (1993) showed that the risk-dominant strategy profile of a 2×2 game is almost always played. Kandori *et al.* (1993) and Young (1993)

[6] Since K already exceeds the original Nash aggregate contribution, no player will push beyond the threshold.

[7] In addition to these two symmetric equilibria, an asymmetric pure strategy profile is an equilibrium if it satisfies $Z = K$, $\bar{F}'(K) \leq c'(z_i)$ and $\bar{F}(K) \geq c(z_i)$ for all i. If $\bar{F}(K) < c(K/n)$ then the unique equilibrium involves no contributions, whereas if $\bar{F}(K) > c(K)$ then all equilibria satisfy $Z \geq K$.

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

[9] 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.

	$z_1 = \frac{K}{2}$	$z_2 = 0$
$z_1 = \frac{K}{2}$	$\bar{F}(K) - c\left(\frac{K}{2}\right)$ $\bar{F}(K) - c\left(\frac{K}{2}\right)$	0 $-c\left(\frac{K}{2}\right)$
$z_1 = 0$	$-c\left(\frac{K}{2}\right)$ 0	0 0

Fig. 2. *Multiple Equilibria in a 2 × 2 Collective-Action Game*

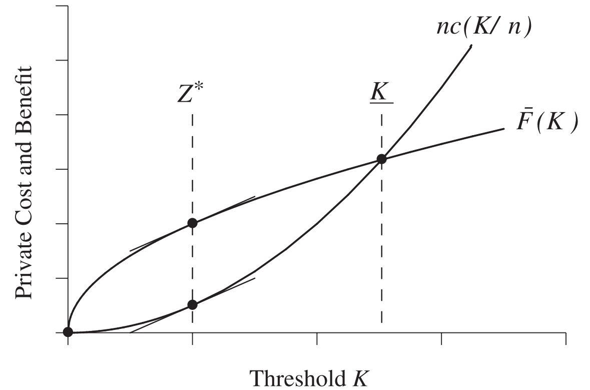

Fig. 3. *Feasible Thresholds*

Note: This Figure illustrates the conjectured equilibrium-selection constraint. Z^* is the aggregate Nash contribution under the production function $\bar{F}(\cdot)$. Turning to the production function $F_K(\cdot)$, for thresholds $K > Z^*$, there can be an equilibrium where the players jointly contribute K toward the public good. However, if $K > \underline{K}$ where $\bar{F}(\underline{K}) = nc(\underline{K}/n)$ the discussion of Section 2.1 suggests that play will switch to the 'team failure' equilibrium, and so overly ambitious thresholds are self-defeating.

took a very different approach, in which play evolves via a stochastic strategy-revision process. 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, while suggestive, is not the sole focus of this article. Rather than specifying a precise selection criterion, this article instead admits that behaviour may vary over time, and so play may move between equilibria. The appropriate solution concept, then, is a non-degenerate probability distribution over the different strategy profiles (equivalently, combinations of inputs to the production) which represents the long-run frequency of different modes of play. Such a probability distribution emerges from the analysis of a strategy-revision process. In the context of such a process the switch between production functions from $\bar{F}(\cdot)$ to $F_K(\cdot)$ has two effects. First, it reduces welfare whenever contributions amount to less than K. Second, it influences the evolution of play and so changes the frequencies with which strategy profiles are played. The next worked example explores this idea further.

2.2. *Evolving Play and the Shape of the Production Function*

This example (a special case of the collective-action game considered previously) illustrates the methodology used to tackle the questions arising from Section 2.1, and helps to build intuition for the answers. Two players each choose either to contribute to a public good ($z_i = 1$) or to free ride ($z_i = 0$). A contributor incurs a private cost of $c > 0$, and so $c(z_i) = cz_i$.

Turning to the production technology, when both players contribute ($Z = 2$) the public good has value $F_\theta(2) \equiv v_H$ to both players. A single contribution ($Z = 1$) reduces its value to $F_\theta(1) \equiv \theta v_L$ where $v_L < v_H$ and where $0 \le \theta \le 1$. If neither player contributes ($Z = 0$) then nothing is produced, so that $F_\theta(0) \equiv 0$. Summarising, the production function satisfies

$$F_\theta(Z) = \begin{cases} v_H & \text{if } Z = 2, \\ \theta v_L & \text{if } Z = 1, \text{ and} \\ 0 & \text{if } Z = 0. \end{cases}$$

Crucially, the parameter $\theta \in [0,1]$ can be varied, and so it indexes a family of feasible production functions (Figure 4). Setting $\theta = 1$ pushes $F_\theta(1)$ against the frontier $\bar{F}(1) = v_L$, whereas setting $\theta = 0$ is equivalent to specifying a threshold of $K = 2$. This game can be simply represented as a 2×2 strategic form (Figure 5). To make it interesting, assume that

$$v_L > c > v_H - v_L > c/2 > 0.$$

These inequalities ensure that, for $\theta = 1$: the private benefit of the first contribution exceeds its private cost; the private benefit of the second contribution falls short of its private cost; but it is socially optimal for both players to contribute. For θ sufficiently large (that is, when $\theta \ge (v_H - c)/v_L$) there are pure-strategy Nash equilibria in which only one player contributes. For smaller θ, however, there is one equilibrium in which both players contribute and another in which neither do so. For either case there are multiple equilibria, and hence the Nash solution concept cannot provide a unique prediction. To address this problem, serious heed is paid to the twin possibilities that

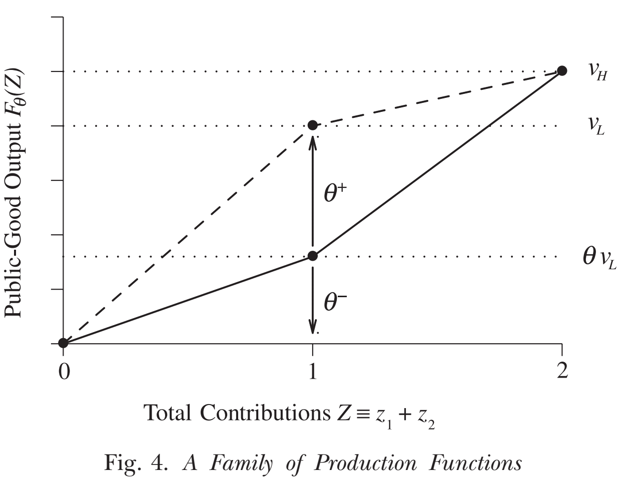

Fig. 4. *A Family of Production Functions*

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

Strategic Form

	$z_2 = 1$	$z_2 = 0$
$z_1 = 1$	$w_{11} = 2(v_H - c)$	$w_{10} = 2\theta v_L - c$
$z_1 = 0$	$w_{01} = 2\theta v_L - c$	$w_{00} = 0$

Welfare

Fig. 5. *Equilibria and Welfare in a Collective-Action Game*

the game is played frequently and that the players adopt different strategies at different times.

Concretely, play evolves via the following strategy-revision process. At (discrete) time t, the state of play z^t is the strategy profile in use, and so the state space is simply the collection of the possible profiles: (0,0), (1,0), (0,1) and (1,1). A randomly chosen player enjoys a strategy-revision opportunity, and plays a quantal response (McKelvey and Palfrey, 1995) to contemporary play. For example, suppose that Player 2 is currently contributing whilst Player 1 free rides, so that $z^t = (0,1)$. Player 1 is given the opportunity to revise. By choosing to contribute, Player 1 enjoys 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 best reply, the log odds ratio of player 1 choosing to contribute versus not is linear in the payoff difference between the two actions. That is,

$$\log \frac{\Pr(\text{Contribute})}{\Pr(\text{Free Ride})} = \lambda(v_H - \theta v_L - c). \tag{1}$$

Similarly, beginning from the state of play (0,0) in which neither player contributes, this log odds ratio is $\lambda(\theta v_L - c)$. The parameter λ indexes the degree to which this is a model of 'smoothed' best replies. If $\lambda = 0$ then a revising player chooses at random whereas when $\lambda \to \infty$ the quantal response is almost always a myopic best reply. As with econometric models of discrete choice, the logit quantal-response admits a random-utility interpretation.[10]

This strategy-revision process is a simple Markov chain. Using expressions such as (1) transition probabilities $p_{z \to z'} \equiv \Pr(z^{t+1} = z' \mid z^t = z)$ are easily calculated. This process is ergodic; that is, its long-run behaviour is independent of initial conditions. This means that the long-run frequency $\pi_z \equiv \lim_{t \to \infty} \Pr(z^t = z)$ with which each strategy profile is played (this is the ergodic distribution of the Markov chain) is uniquely defined.

Calculating the long-run strategy frequencies is relatively simple, as 'detailed-balance conditions' apply: the relative likelihood of two different strategy profiles (and hence input combinations) is determined by the relative likelihood of directly jumping back and forth between them. That is, $\pi_z p_{z \to z'} = \pi_{z'} p_{z' \to z}$ for all z and z'. For

[10] Suppose that a revising player's payoffs are subject to noise: the payoff difference between the two available actions is logistically distributed with parameter λ. This yields Equation (1).

example, taking the transition from (0,1) to (1,1) discussed previously, and its complement,

$$\frac{\pi_{11}}{\pi_{01}} = \frac{p_{01\to 11}}{p_{11\to 01}} = \mathrm{e}^{\lambda(v_H - \theta v_L - c)}.$$

Using these detailed-balance conditions, it is straightforward to find the ergodic distribution. The long-run probabilities associated with the various strategy profiles,

$$\pi_{11} \propto \mathrm{e}^{\lambda(v_H - 2c)}, \quad \pi_{10} = \pi_{01} \propto \mathrm{e}^{\lambda(\theta v_L - c)} \quad \text{and} \quad \pi_{00} \propto \mathrm{e}^0 = 1,$$

depend on the gap between the private benefit to a single player and the total private cost.

With the properties of long-run play established, attention can turn to social welfare. Defining welfare as the sum of players' payoffs and using an obvious notation,

$$w_{11} = 2(v_H - c), \quad w_{10} = w_{01} = 2\theta v_L - c \quad \text{and} \quad w_{00} = 0.$$

An appropriate measure of aggregate expected welfare requires a probability distribution over the set of strategy profiles. The distribution derived above yields the welfare measure

$$W_\lambda(\theta) \equiv \sum_z \pi_z w_z = \frac{2(2\theta v_L - c)\mathrm{e}^{\lambda(\theta v_L - c)} + 2(v_H - c)\mathrm{e}^{\lambda(v_H - 2c)}}{2\mathrm{e}^{\lambda(\theta v_L - c)} + \mathrm{e}^{\lambda(v_H - 2c)} + 1}. \tag{2}$$

Welfare depends upon the production function in use (determined by θ) and the level of noise in the quantal responses (determined by λ). The search for the welfare-maximising public-good production function reduces to an examination of how $W_\lambda(\theta)$ responds to changes in θ. Such changes feed through both the welfare terms w_z and the probabilities π_z.

Consider a reduction in θ. This has two effects: it reduces the welfare of states (0,1) and (1,0); but it also reduces the likelihood of these states and so increases the relative likelihood of the state (1,1) in which welfare is maximised. Suppose the welfare gain from the latter effect is larger than the loss from the former. Then an additional reduction in θ increases aggregate expected welfare further: the states (0,1) and (1,0) in which welfare is lowered are played less frequently (as a result of the first decrease in θ) and so the negative impact is smaller than before and must again be outweighed by the gain from the increased frequency of (1,1) play. Thus, if it is beneficial to lower θ a little, it will be optimal to lower it maximally. Similarly, if it is beneficial to raise θ a little, it is optimal to raise it maximally.

This argument (or rather a formal version of it) establishes that welfare is maximised by setting either $\theta = 1$ or $\theta = 0$. These two options respectively correspond to the maximal exploitation of production opportunities, or the use of a threshold rule (at $K = 2$) in which all inputs falling below K produce no output. So which option is best?

Later Sections provide a complete answer to this question. Here, however, it is instructive to examine what happens when λ is large, so that quantal responses closely approximate myopic best replies. An inspection of (2) reveals that $W_\lambda(1) \to 2v_L - c$ as

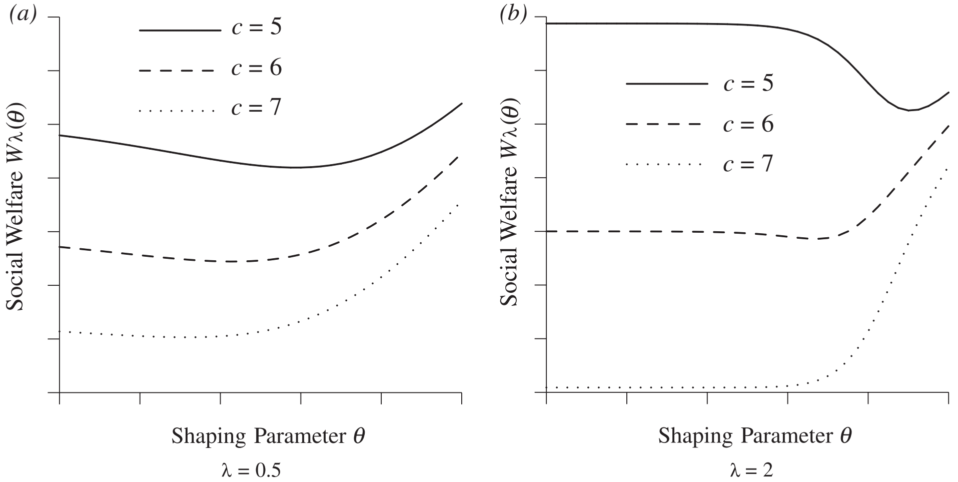

Fig. 6. *Quasi-Convexity of Welfare in θ*
Note: This Figure illustrates the quasi-convexity of social welfare $W_\lambda(\theta)$ in θ for various parameter values. The specification of Figures 4 and 5 is used with $v_H = 12$ and $v_L = 8$. When noise is high ($\lambda = 0.5$) then welfare is maximised by choosing $\theta = 1$ for each value of c displayed; production possibilities should be exploited maximally. However, when λ is larger, so that strategy revisions are closer approximations to myopic best replies, it is optimal to set $\theta = 0$ and move to a threshold-based production function, so long as c is small enough.

$\lambda \to \infty$, so that the process spends almost all time in the states (0,1) and (1,0). However,

$$W_\lambda(0) \to \begin{cases} 2v_H - c & \text{if } v_H > 2c, \text{ and} \\ 0 & \text{if } v_H < 2c. \end{cases}$$

If $v_H > 2c$ then society can gain by using an inefficient technology ($\theta = 0$) which throws away output by imposing the threshold $K = 2$. The inequality $v_H > 2c$ has the 'private feasibility' interpretation (Section 2.1): v_H is the benefit of the full ($Z = 2$) public good to a single player whereas $2c$ is the total private cost of provision. In contrast, when $v_H < 2c$ (so that $Z = 2$ is privately infeasible) welfare is optimised by fully exploiting the production-possibility frontier: if $\theta = 0$ the 'bad' equilibrium is played almost all of the time.

The conclusion here is that in the limit (as $\lambda \to \infty$) a threshold is socially optimal so long as c is small enough. This is true generally: there is a cut-off level of the contribution cost c^*_λ, such that for all c higher than this it is optimal to exploit production opportunities maximally, and for all c lower it is optimal to introduce a threshold. Figure 7 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 using a 'dented' production

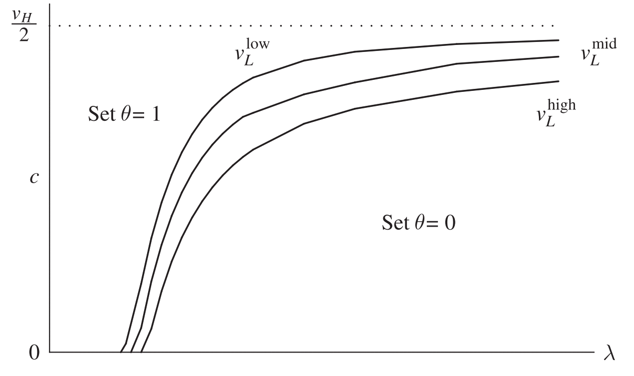

Fig. 7. *The Cut-Off Values* $c^*(\lambda)$

Note: The Figure plots c^*_λ for three different values of v_L, where $v_L^{\text{high}} > v_L^{\text{mid}} > v_L^{\text{low}}$. Note that $c^*_\lambda \to v_H/2$ as $\lambda \to \infty$. Beneath the curve it is socially optimal to set $\theta = 0$, but above it is socially optimal to set $\theta = 1$. As $v_L \to v_H$, the former region disappears.

function, the social optimum may become a 'team success' equilibrium of the game; however, another 'team failure' 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 it would be better to use a traditional concave production function which attains the production-possibility frontier.

Should an optimal production function exhibit the 'all or nothing' characteristics of a threshold rule? Section 2.2 provides a preliminary answer. A quantal-response process generates long-run probabilities for the play of each strategy profile, which generate an aggregate social welfare measure. For the relevant input combinations, it is socially optimal either to exploit fully the production-possibility frontier function, or to throw everything away.

3. The Evolution of Public-good Production

This Section presents an analytical framework within which the insights of Section 2 prove to be quite general. First, a general collective-action game is used to model the private provision of a public good. Second, a quantal-response strategy-revision process is described via which play evolves. Finally, long-run play is characterised.

3.1. *A General Collective-Action Game*

Consider an n-player simultaneous-move game in which player i chooses an action z_i from the finite set $\mathcal{Z}_i$. This action is interpreted as the player's contribution to the production of a public good, or participation in some other collective action. It generates a private cost $c_i(z_i) : \mathcal{Z}_i \mapsto \mathbb{R}$. This cost is independent of other players' actions, and hence captures the individual interests of player i.

The actions of the n players together form a pure strategy profile $z \in \mathcal{Z} \equiv \times_i \mathcal{Z}_i$ which represents the combination of inputs that feed into the public-good production process. These inputs generate a public-good output of $G(z) : \mathcal{Z} \mapsto \mathbb{R}$. This (pure) public good is enjoyed equally by all players, and hence $G(z)$ captures players' common interests.

Combining these elements, a collective-action game is obtained if the payoffs satisfy

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

for each player i. No structure is imposed on either the production technology or the cost functions. The key assumptions are first that payoffs are additively separable, and second that players value the public good identically. These assumptions are needed in order to generate a potential game. In a related paper, Myatt and Wallace (2008*b*), progress is made without these assumptions. Instead, that paper focuses on binary-action games. Here there are no restrictions placed upon the strategy space but payoffs must take the form in (3).

It may be that some individuals enjoy contributing to the public good. Similarly, $G(z)$ may represent a public bad (such as pollution) rather than a good. Thus the formulation of (3) goes beyond Cournot-contribution games and incorporates many other scenarios, including the private exploitation of a common resource.[11]

A collective-action game satisfying (3) proves easy to analyse since it is an exact potential game (Monderer and Shapley, 1996). Such a game admits an exact potential function $\psi(z) : \mathcal{Z} \mapsto \mathbb{R}$. This is a real-valued function defined on the space of strategy profiles which captures the essential strategic properties of a game: whenever a player deviates, the change in the potential function is precisely equal to the change in the player's payoff. More formally, if two strategy profiles z and z' differ only by the action taken by player i, then $u_i(z') - u_i(z) = \psi(z') - \psi(z)$. Roughly speaking, the players of a potential game act as if they are jointly attempting to maximise the potential function. Here, the separation of the players' payoffs into public and private components yields an exact potential function

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

$\psi(z)$ incorporates a single player's private benefit from the public good minus the total private cost of provision. This contrasts with a natural social-welfare function

$$w(z) \equiv nG(z) - \sum_{i=1}^{n} c_i(z_i) \tag{5}$$

which incorporates the social benefits of $nG(z)$ accruing to the entire set of players.

Given the discussion above, the Nash equilibria of a potential game are associated with the maximisation of the potential function. Two subsets of $\mathcal{Z}$ are of interest. First, $\mathcal{Z}^\dagger$ 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 a deviation corresponds one-to-one with a change in potential, $\mathcal{Z}^\dagger$ is also the set of pure-strategy Nash equilibria. Second, $\mathcal{Z}^\ddagger$ is the set of strategy profiles from which players cannot jointly deviate and increase the potential. These form a subset of the (pure-strategy) Nash

[11] Cournot-contributions games often involve continuous action sets and hence to place them within the present framework simply requires the construction of a sufficiently fine grid of feasible contributions.

equilibria. Since $\mathcal{Z}^{\ddagger}$ is non-empty, a potential game has at least one pure-strategy Nash equilibrium.[12]

For a collective-action game, a member of $\mathcal{Z}^{\ddagger} = \arg\max_{z \in Z}[G(z) - \sum_{i=1}^{n} c_i(z_i)]$ has the following interpretation. Suppose that a private entrepreneur were to hire the players (contributors to public-good production) and pay for their costs. Clearly, the entrepreneur's costs would be $\sum_{i=1}^{n} c_i(z_i)$. The entrepreneur would then sell the public good. Since it is non-excludable, this sale would extract only a single private benefit $G(z)$ of provision as revenue, yielding profits of $\psi(z) = G(z) - \sum_{i=1}^{n} c_i(z_i)$. Maximising profits, the entrepreneur would choose $z \in \mathcal{Z}^{\ddagger}$. That is, $\mathcal{Z}^{\ddagger}$ is the set of privately profit maximising input combinations. This set also has relevance for the evolving play of the game.

3.2. *Evolving Play*

A standard approach to the analysis would be to consider the Nash equilibria in $\mathcal{Z}^{\dagger}$. However, this approach has weaknesses. Collective-action games often have multiple equilibria and so equilibrium-selection problems arise. Furthermore, a consideration of long-run behaviour must admit the possibility that play may change over time.

Here, play evolves according to a strategy-revision process. At each discrete time a player i is drawn at random (with probability $\varrho_i > 0$) and given a strategy-revision opportunity. One possibility would be for a revising player to choose a myopic best reply to the current strategy profile. The long-run behaviour described by such a process is potentially history-dependent, since it will 'lock in' to a strict Nash equilibrium (of which there may be many).

To circumvent this problem, noise is required. This might be generated in a variety of ways, so long as play can move away from every strategy profile with some positive (but possibly small) probability. Here this is achieved through a logit quantal-response specification: strategy revisions follow a familiar multinomial-logit distribution (McFadden, 1974).[13] Somewhat more formally, write z^t for the state of play (a strategy profile, or combination of inputs) at time t, and $p_{z \to z'} \equiv \Pr(z^{t+1} = z' | z^t = z)$ for the transition probability of moving from z to z'. If states of play z' and z differ only by the action of player i then

$$p_{z \to z'} = \varrho_i \times \frac{e^{\lambda u_i(z')}}{\sum_{z'' \in \Delta_i(z)} e^{\lambda u_i(z'')}} \tag{6}$$

where $\Delta_i(z)$ are strategy profiles that differ from z by at most the action of player i. The parameter λ indexes the degree to which the quantal response is a smoothed best reply to the play of others: if $\lambda = 0$ then a revising player chooses at random (that is, equiprobably), whereas in the limit as $\lambda \to \infty$ a revising player almost always chooses a best reply.

[12] More formally, $\mathcal{Z}^{\ddagger} = \{z \in \mathcal{Z} \mid \psi(z) \geq \psi(z'), \forall z' \in \mathcal{Z}\} \supseteq \mathcal{Z}^{\dagger} = \{z \in \mathcal{Z} \mid \psi(z) \geq \psi(z'), \forall z' \in \Delta(z)\}$, where $\Delta(z)$ is the set of 'neighbouring' strategy profiles obtained via a single-player deviation from z.

[13] This has the usual random-utility interpretation: if payoffs are subject to identically and independently distributed Gumbel shocks with scale parameter λ, then the multinomial logit is obtained.

Logit quantal-response strategy revisions are easy to work with when applied to an exact potential game. Taking two states of play z and z' that differ only by the action of player i,

$$\log\left(\frac{p_{z\to z'}}{p_{z'\to z}}\right) = \lambda[u_i(z') - u_i(z)] = \lambda[\psi(z') - \psi(z)]. \tag{7}$$

The second equality follows since $\psi(\cdot)$ is an exact potential function. Equation (7) says that the relative probability of jumping back and forth between neighbours is determined by the difference in potential and not by player i's payoff from actions other than z_i or z_i'; this reflects the independence-of-irrelevant-alternatives property of the multinomial logit.[14]

Alternative noise specifications are not amenable to the analysis presented here. However, progress can be made with a variety of distributional forms (including the probit) for vanishingly small noise. Such limiting results are exactly the focus of Myatt and Wallace (2008*b*), which studies the evolving play of various Palfrey and Rosenthal (1984) threshold games. Analytical results are available when attention is restricted to either binary-action games under general noise specifications with vanishingly small noise, or potential games with completely general strategy spaces under the logit with any level of noise. This paper studies the latter, whilst the former is analysed in Myatt and Wallace (2008*b*): the results are complementary.

A variant of (7) holds for sequences of transitions. For example, a two-step transition in which the process moves from z to z'' via an intermediate state z' satisfies

$$\log\left(\frac{p_{z\to z'} \times p_{z'\to z''}}{p_{z''\to z'} \times p_{z'\to z}}\right) = \lambda[\psi(z'') - \psi(z)].$$

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.

3.3. *Long-run Behaviour*

The strategy-revision process described above is an ergodic Markov chain on the state space $\mathcal{Z}$ of strategy profiles. What this means is that the long run frequencies $\pi_z \equiv \lim_{t\to\infty} \Pr(z^t{=}z)$ with which strategy profiles are played (these form the ergodic distribution) are uniquely defined and independent of initial conditions. Following Blume (1997), the ergodic distribution takes a simple Gibbs representation.[15]

[14] This property is somewhat restrictive: assuming independence of payoff shocks rules out payoff correlation. This is a common critique of the multinomial logit as an econometric model (Hausman and McFadden, 1984). An alternative allowing for correlation is the multinomial probit (Hausman and Wise, 1978). Unfortunately, closed-form solutions for the general class of public-good provision games examined here are not available.

[15] Blume (1997) discussed the relationship between potential games and log-linear choice rules. He observed that detailed-balance conditions are satisfied if and only if the game admits an exact potential function.

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*

$$\pi_z \equiv \lim_{t\to\infty} \Pr(z^t = z) = \frac{e^{\lambda\psi(z)}}{\sum_{z'\in\mathcal{Z}} e^{\lambda\psi(z')}}. \tag{8}$$

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

Heuristically, at least, this says that the long-run frequency of each strategy profile is determined by the private profitability of the corresponding input combination.

Lemma 1 is straightforward to prove. A probability distribution π_z is said to be in 'detailed balance' with respect to two states of a Markov chain if $\pi_z p_{z\to z'} = \pi_{z'} p_{z'\to z}$. This says that the average flow out from z to z' is equal to the average flow in from z' to z. Combining (7) and (8) it is easy to see that this detailed-balance condition holds whenever two states differ by a single action. For $z = z'$, the condition holds trivially and for any other pairs the transition probabilities are zero, so the condition 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).

The Section draws to a close with the observation that, as $\lambda \to \infty$, all weight falls on the states with maximum potential: $\lim_{\lambda\to\infty}[\lim_{t\to\infty} \Pr(z_t \in Z^{\ddagger})] = 1$. That is, when noise is small, these strategy profiles (the potential-maximising Nash equilibria) are 'selected'. In the collective-action game of interest, these equilibria maximise 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 equilibria selected correspond to the 'entrepreneurial' ones that would be chosen were a private individual to bear the total costs of production by compensating every player for their cost of contribution.

4. Social Welfare and Coordination

This Section contains the central results of the article. Social welfare is considered and the reaction of the welfare criterion to parameter changes is ascertained: when a parameter shifts public-good output across a set of states of play it is optimal either to maximise or to minimise output; intermediate values should not be chosen.

4.1. *Social Welfare*

Given the play of a strategy profile z a natural welfare measure is the sum $w(z)$ of payoffs defined in (5). If z were always played then $w(z)$ could be a sufficient measure. However, play evolves; in the long run, z is played with probability π_z (Lemma 1), and so the long-run time-average of welfare is

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

and where

$$w(z) = nG(z) - \sum_{i=1}^{n} c_i(z_i). \tag{10}$$

This is the welfare criterion that is used in the remainder of the article. Note, however, that there are other welfare measures that could be employed. One candidate is the present-discounted value of $w(z)$. Given a discount factor δ, in per-period terms this satisfies

$$W_\delta \equiv \frac{1}{1-\delta} \sum_{t=0}^{\infty} \delta^t \left[\sum\nolimits_{z \in \mathcal{Z}} \Pr(z^t = z) \times w(z) \right].$$

Of course, the probabilities in this expression depend upon the initial conditions of the process. Two ways of resolving this problem are these: first, placing the distribution π_z across the initial states of play yields $\mathrm{E}(W_\delta) = W$. Second, for any initial conditions $W_\delta \to W$ as $\delta \to 1$. Hence W can be interpreted as the objective of either

(*i*) a social planner who is uncertain of initial conditions and uses the ergodic distribution to capture this uncertainty; or
(*ii*) a patient social planner.

More generally, all of the results hold for the welfare measure W_δ, so long as δ is large enough. W is less appropriate as a welfare measure when the hypothetical social planner is impatient and λ is very large. When there is very little noise then the short-run properties of the strategy-revision process are determined by initial conditions, and so welfare depends heavily upon the initial state of play.

4.2. *Influencing Production*

Attention now turns to the construction of a welfare ordering over a family of public-good production functions. Notice that the production function $G(z)$ influences aggregate welfare in two ways. First, $G(z)$ directly affects welfare $w(z)$ whenever z is played. Secondly, $G(z)$ determines the potential $\psi(z)$, and hence the relative likelihood that the process visits state-of-play z. These effects must be considered jointly.

To make progress, suppose that a parameter $\theta \in [0,1]$ affects production. (Restricting the support to $\theta \in [0,1]$ is without loss of generality.) The common component to payoffs becomes $G_\theta(z)$; similarly, welfare and potential become $w_\theta(z)$ and $\psi_\theta(z)$ respectively. Assume that $G_\theta(z)$ is strictly increasing (this is without loss of generality) and twice continuously differentiable in θ for all z. Define the states of play for which θ has an effect as $\mathcal{Z}_\theta$.

Definition. *θ is an output shifter if it affects $G_\theta(z)$ in the same way for all states in $\mathcal{Z}_\theta$.*

More formally, this definition is satisfied if $G_\theta(z) - G_\theta(z')$ is constant with respect to θ for all pairs z and z' in $\mathcal{Z}_\theta$; this is automatically true when $\mathcal{Z}_\theta$ is a singleton. If θ is an output shifter then reducing it is equivalent to eliminating output for all input combinations in the set $\mathcal{Z}_\theta$. This means that an investigation of θ can address the question:

should a hypothetical social planner ever choose to exploit a free disposal opportunity and throw output away? Proposition 1 provides an answer. (Formal proofs are contained in the Appendix.)

PROPOSITION 1. *If the parameter θ is an output shifter then social welfare W is a quasi-convex function of θ, and hence is maximised by choosing either $\theta = 0$ or $\theta = 1$.*

The conclusion 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.

To see why, suppose that $\partial W/\partial\theta < 0$. A reduction in θ directly harms $w_\theta(z)$ for $z \in \mathcal{Z}_\theta$. However, the potential $\psi_\theta(z)$ also falls, whence the probabilities with which states in $\mathcal{Z}_\theta$ are played diminish, pushing play elsewhere. Since $\partial W/\partial\theta < 0$, this second effect must help welfare and so states outside $\mathcal{Z}_\theta$ must yield higher welfare. Crucially, further reductions in θ must also be beneficial: the cost of doing so (the reduction in $w_\theta(z)$ for $z \in \mathcal{Z}_\theta$) falls since Z_θ is visited less often. Thus, if it is optimal to reduce θ, then it is optimal to do so maximally.

4.3. *Influencing Production Based on Inputs*

It is instructive to consider a specific application that also expands the set of parameters. Suppose that $\theta \in [0,1]^{|\mathcal{Z}|}$ is now a vector of parameters, one for each state of play, and that $G_\theta(z)$ satisfies

$$G_\theta(z) = \theta_z \bar{G}(z)$$

for all $z \in \mathcal{Z}$. This application has two features. First, the parameter θ_z only affects the output of the public good emerging from the collection of inputs z. This means that the production function may be manipulated very finely. Second, the specification implements a free-disposal property. Specifically, $\bar{G}(z)$ represents a production-possibility frontier, and any production function lying below this frontier is feasible.

Proposition 1 applies to each component of the vector θ: for each z it is optimal either to leave production unhindered, or to discard it completely. This generates a corollary.

COROLLARY. *Suppose that the family of production functions comprises all production functions G(z) satisfying $0 \leq G(z) \leq \bar{G}(z)$. Then the socially optimal member $G^*(z)$ satisfies*

$$G^*(z) = \begin{cases} \bar{G}(z) & \text{if } z \in \mathcal{Z}^*, \text{ and} \\ 0 & \text{if } z \notin \mathcal{Z}^*, \end{cases}$$

for some subset $\mathcal{Z}^*$ of feasible input combinations. The corresponding socially optimal parameter choice θ^* satisfies $\theta^*_z = 1$ if $z \in z^*$ and $\theta^*_z = 0$ otherwise.

A similar claim also holds if the lower bound to the family of public-good production functions is something other than zero. For instance, if $G_\theta(z) = \theta_z\bar{G}(z) + (1 - \theta_z)\underline{G}(z)$ where $\bar{G}(z) \geq \underline{G}(z)$, then a similar corollary obtains where $G^*(z) = \underline{G}(z)$ for $z \notin \mathcal{Z}^*$.

Note that a direct implication of this corollary is that high performing public-good production functions tend to exhibit discontinuous steps. Thus, a successful collective action entails teamwork: players need to coordinate on play within the set $\mathcal{Z}^*$ in order to succeed.

4.4. *Influencing Production Based on Feasible 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 capture subtle variations in the production technology. For instance, in the committee example (Section 1) production can depend on the precise identities of those who show up to the meeting. In contrast, a second application is considered now in which the influence of parameter changes is rather blunt.

Suppose that the private actions $z \in \mathcal{Z}$ combine to yield an intermediate public good $y \in \mathcal{Y}$ via a production function $y = H(z)$. The value of the final public good y to an individual is $F_\theta(y) = \theta_y \bar{F}(y)$ for some parameter vector $\theta \in [0,1]^{|\mathcal{Y}|}$.[16] This yields a final public-good production function of $G_\theta(z) = F_\theta[H(z)]$. 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, $\mathcal{Z}_{\theta_y} = \{z \in \mathcal{Z}\colon y = H(z)\}$. By inspection, θ_y affects the value of the public good in the same way for all $z \in \mathcal{Z}_{\theta_y}$, and so each element of the parameter vector is an output shifter. Hence, Proposition 1 applies.

Corollary. *Suppose that the family of production functions comprises all production functions $G(z)$ satisfying $G(z) = F[H(z)]$, where the value of the public good $F(z)$ satisfies $0 \leq F(y) \leq \bar{F}(y)$ for all $y \in \mathcal{Y}$. Then the socially optimal member $G^*(z)$ satisfies*

$$G^*(z) = \begin{cases} \bar{F}[H(z)] & \text{if } H(z) \in \mathcal{Y}^*, \text{ and} \\ 0 & \text{if } H(z) \notin \mathcal{Y}^*, \end{cases}$$

for some subset $\mathcal{Y}^*$ of feasible public-good intermediate output. The corresponding socially optimal parameter choice θ^* satisfies $\theta^*_y = 1$ if $y \in \mathcal{Y}^*$ and $\theta^*_y = 0$ otherwise.

A concrete illustration is a contributions game in which $H(\cdot)$ takes the form $H(z) = \sum_i z_i$. In the committee example, this would hold when production depends solely on the number of attendees at the meeting, and not on their precise identities.

4.5. *Welfare-Based Thresholds*

Proposition 1 and its corollaries reveal that welfare is optimised when $G(z)$ either fully exploits the production possibility frontier or produces nothing. It does not, however, specify which input combinations should be subject to the complete disposal of public-good output. This issue is addressed here.

For the purposes of this Section and, in particular, for Proposition 2 below, the focus is on the specification from Section 4.3 used in the first corollary to Proposition 1. That

[16] This specification extends straightforwardly to $F_\theta(y) = \theta_y \bar{F}(y) + (1 - \theta_y)\underline{F}(y)$ where $\bar{F}(y) \geq \underline{F}(y)$.

is, the family of feasible public-good production functions consists of all function $G(z)$ satisfying $0 \leq G(z) \leq \bar{G}(z)$, where $\bar{G}(z)$ is a production-possibility frontier. Equivalently, the production function is influenced by a parameter $\theta \in [0,1]^{|\mathcal{Z}|}$ and satisfies $G_\theta(z) = \theta_z \bar{G}(z)$.[17]

The statement of the results is assisted by some simple notation. Abusing earlier notation a little, write $\psi_1(z)$ and $\psi_0(z)$ for the maximum and minimum potential and $w_1(z)$ and $w_0(z)$ 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 choice of $G(z)$, or equivalently the socially optimal choice of the set $\mathcal{Z}^*$. Finally, and without loss of generality, suppose that $\psi_1(z) \neq \psi_0(z')$ for all $z \neq z'$. This last assumption is made solely to ease the statement of Proposition 2.

The optimal production function depends upon λ. For instance, when $\lambda = 0$ each strategy profile is equiprobable (revising players choose randomly) and so discarding production does not influence long-run play. The only impact, therefore, of setting $\theta_z = 0$ is to reduce welfare when z is played. As a result, it is optimal to set $\theta_z = 1$ for all z; equivalently, $\mathcal{Z}^* = \mathcal{Z}$.

For smaller levels of noise it is optimal (as is confirmed formally in Proposition 2) to neglect production opportunites in some states of play, based on a welfare-based threshold rule. The threshold is closely related to a set of so-called 'privately feasible' input combinations.

Definition. *The set $\tilde{\mathcal{Z}}$ of privately feasible input combinations are those states z satisfying $\psi_1(z) > \psi_0(z')$ for all $z' \neq z$. The maximum privately feasible welfare is $\tilde{w} \equiv \max_{z \in \tilde{\mathcal{Z}}} w_1(z)$.*

To understand this definition, 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 $\psi_\theta(z)$. Furthermore, suppose that this player were to be asked to choose an input combination $z \in \mathcal{Z}$. When $z \in \tilde{\mathcal{Z}}$ the player can be induced to choose z by setting $\theta_z = 1$ and $\theta_{z'} = 0$ for all $z' \neq z$. ; if not, there is always another state with higher potential, that is therefore privately preferable.

Proposition 2. *Suppose that the family of production functions comprises all production functions G(z) satisfying $0 \leq G(z) \leq \bar{G}(z)$. Then the socially optimal member satisfies*

$$G^*(z) = \bar{G}(z) \Leftrightarrow \gamma(z) \geq W^*_\lambda \quad \textit{where} \quad \gamma(z) \equiv \frac{w_1(z)e^{\lambda\psi_1(z)} - w_0(z)e^{\lambda\psi_0(z)}}{e^{\lambda\psi_1(z)} - e^{\lambda\psi_0(z)}},$$

and $G^*(z) = 0$ otherwise. Allowing $\lambda \to \infty$, $\gamma(z) \to w_1(z)$ and $W^*_\lambda \to \tilde{w}$. For λ large enough

$$G^*(z) = \begin{cases} \bar{G}(z) & \textit{if } w_1(z) \geq \tilde{w}, \textit{ and} \\ 0 & \textit{if } w_1(z) < \tilde{w}. \end{cases}$$

Hence when quantal-response strategy revisions approximate best replies, the optimal public production implements a threshold rule: production possibilities are fully exploited (so that $z \in \mathcal{Z}^*$) if and only if welfare meets or exceeds the maximum privately feasible welfare.

[17] The arguments presented here, including Proposition 2, also hold when $G_\theta(z) = \theta_z \bar{G}(z) + (1 - \theta_z)\underline{G}(z)$.

To understand the role of the expression $\gamma(z)$ in this Proposition, 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 ratio $\gamma(z)$ combines these 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 Proposition, it is useful to inspect the effect of a local change in θ_z. The proof of Proposition 1 in the Appendix reveals that

$$\frac{\mathrm{d}W}{\mathrm{d}\theta_z} \propto \frac{\partial w_\theta(z)}{\partial \theta_z} + \lambda \frac{\partial \psi_\theta(z)}{\partial \theta_z}[w_\theta(z) - W].$$

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 $w(z) \geq W$. This is, of course, a local effect only; the proof of Proposition 2 reveals that the same is true when considering the choice between $\theta_z = 1$ and $\theta_z = 0$.

Finally, note 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 often based on a simple pairing of a production target (a minimum threshold) together with a cap on contributions (a maximum threshold).

5.1. *Choosing Welfare-based Thresholds*

When the family of public-good production functions consists of all those functions satisfying $0 \leq G(z) \leq \bar{G}(z)$, the socially optimal public-good production function is easy to interpret. Recall that this family forms the basis for Proposition 2 and that the set $\mathcal{Z}^*$ is determined (at least for large enough λ) by the maximum privately feasible welfare $\tilde{w}$. A first task is to calculate this maximum.

To do this, suppose (without loss of generality) that action sets are finite subsets of the positive real line ($\mathcal{Z}_i \subset \mathcal{R}_+$) with $0 \in \mathcal{Z}_i$ and that $c_i(z_i)$ is strictly increasing with $c_i(0) = 0$.[18] This means that players always have the option of doing nothing and so incurring no costs.

It is straightforward to observe that the minimum potential for a state z, achieved by discarding all production, satisfies $\psi_0(z) = -\sum_{i=1}^{n} c_i(z_i)$. Maximising this over all possible states leads to $\max_{z \in \mathcal{Z}} \psi_0(z) = 0$. What this means is that there is always a state

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

(namely, $z_i = 0$ for all i) with at least weakly positive potential. In turn, this means that if a state of play is to be visited with positive probability in the limit (that is, for large λ) then it must have a positive potential. Generically, therefore, the set of privately feasible states is

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

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. The welfare threshold $\tilde{w}$ is obtained by finding the maximum over $\tilde{\mathcal{Z}}$. Thus, it is optimal to maximise production alone whenever welfare attains or exceeds its privately feasible maximum.

Taking this one step further, suppose that $c_i(z_i) = cz_i$ and $\bar{G}(z) = \bar{F}(Z)$ for $Z \equiv \sum_{j=1}^{n} z_j$, where $F(\cdot)$ is a strictly increasing and concave production function. The set $\tilde{\mathcal{Z}}$ is then the set for which the (private) average benefit of contributions exceeds the private average cost. Given the concavity of $\bar{F}(\cdot)$, the maximum privately feasible level of contributions is above that which would be forthcoming in the absence of any production-function manipulation. This is illustrated in Figure 8(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 $\bar{F}(\cdot)$ means that welfare $w_1(z)$ is a concave function of Z. Thus, in general, the optimal choice of public-good function will be

$$G^*(z) = \begin{cases} \bar{F}(Z) & \text{if } \underline{K} \le Z \le \bar{K} \text{ and} \\ 0 & \text{otherwise,} \end{cases} \tag{11}$$

for some pair of thresholds $\underline{K}$ and $\bar{K}$; there may be a maximum threshold as well as a minimum. The key observation is that the welfare-based threshold rule translates into a production-based threshold rule in the limit. In fact, this is true away from the limit when a further requirement is met. The next Section investigates.

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 welfare-based threshold rule of Proposition 2 translates into a pair of production thresholds, as in (11).

When $\bar{F}(\cdot)$ is concave, this is true for sufficiently large λ. For general λ, however, it is true so long as the production function is sufficiently concave. The measure of concavity required here is the ρ-concavity notion exploited by Caplin and Nalebuff (1991).[19] 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 is log-concavity. Higher values of ρ correspond to more stringent concavity.

[19] Caplin and Nalebuff (1991) attributed the concept to Avriel (1972). It has found economic application in recent work by Anderson and Renault (2003), Cowan (2007), and Cowan and Vickers (2007).

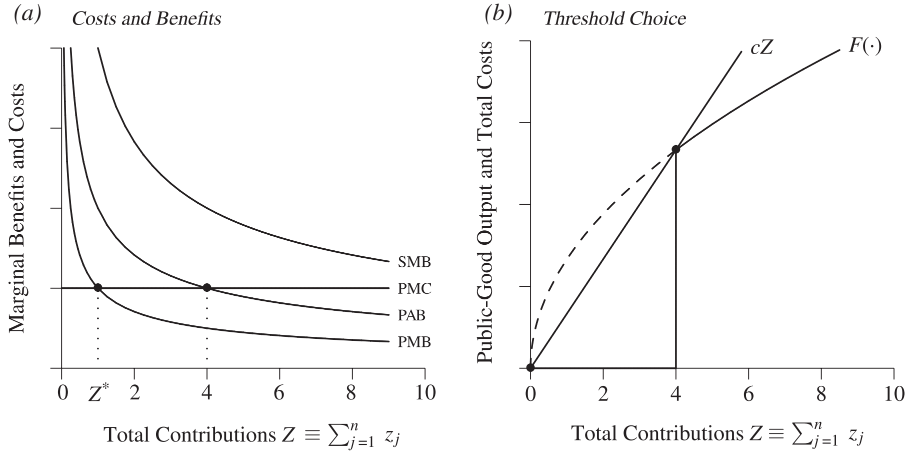

Fig. 8. *Choosing a Threshold*

Note: These Figures illustrate the socially optimal threshold for the specification $\bar{F}(Z), = \sqrt{Z}$ for $Z \equiv, \sum_{i=1}^{n} z_i$, linear costs with $c = \frac{1}{2}$, $n = 4$ players, and $\mathcal{Z}_i \subseteq [0,2.5]$. When the production frontier is attained, the unique equilibrium outcome (extending the action sets as necessary) is such that $Z^* = 1$; this involves the equality of marginal cost (PMC $= c$) and private marginal benefit (PMB $= \bar{F}'(\cdot)$). Since the social marginal benefit (SMB $= n\bar{F}'(\cdot)$) is everywhere above marginal cost, the social optimum involves maximum contributions from all players so that $Z = 10$. Finally, privately feasible states are those involving contributions where private average benefit (PAB) is greater than average cost (here equal to PMC). Thus $\tilde{\mathcal{Z}} = \{z : Z \leq 4\}$; for large enough λ it is optimal to introduce a production threshold at $K = 4$.

Proposition 3. *If $\bar{F}(\cdot)$ is sufficiently concave then, for all λ, n, and c, $\gamma(\cdot)$ is concave in total contributions, so that the optimal public-good production function takes the form:*

$$G^*(z) = \begin{cases} \bar{F}(Z) & \text{if } \underline{K} \leq Z \leq \bar{K}, \text{ and} \\ 0 & \text{otherwise} \end{cases}$$

for thresholds $\underline{K} \leq \bar{K}$. $\bar{F}(\cdot)$ is 'sufficiently concave' if it is ρ-concave for $\rho \approx 1.285$.

The ρ-concavity requirement 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 $-\mathrm{d}\log C(y)/\mathrm{d}\log y \geq \rho - 1$; a restriction on the elasticity of the slope of $C(y)$.[20] As an example, $F(Z) \propto Z^{\alpha}$ is ρ-concave if and only $\alpha \leq 1/\rho$.

A conclusion emerging from this Section is that public-good production technologies that work well are those that incorporate simple production targets. 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 noise parameter λ. This is the focus of the next Section.

[20] This is also equivalent to the curvature measure (Robinson, 1933, pp. 40–1) known as 'adjusted concavity'.

5.3. *Participation and Group Size*

It has already been argued that, when noise is very large, it is socially optimal to exploit any production possibilities fully. 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 (as λ grows larger) it is socially optimal to impose increasingly stringent thresholds. An inspection of Proposition 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 maximised shrinks. However, were W_λ^* to be locally decreasing, this might not be the case. In fact, it is possible to construct examples for which the production threshold is non-monotonic in λ.

Consider a game where $\mathcal{Z}_i = \{0,1\}$ for all i, so that each player decides whether or not to participate in the collective action. Once again assume costs are symmetric, 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 $F(Z) = \sqrt{Z}$ is employed.

This specification meets the curvature restriction of Proposition 3 and hence there will be an upper and lower production threshold. With n agents, suppose that $c^2 < n/4$. This ensures that, conditional on the full exploitation of production, welfare is increasing in the number of participants and so it is socially optimal for all n agents to join the collective action. This 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 weakly below $1/c^2$. More generally, the value of the optimal threshold will depend upon λ. Figure 9 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 9 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.

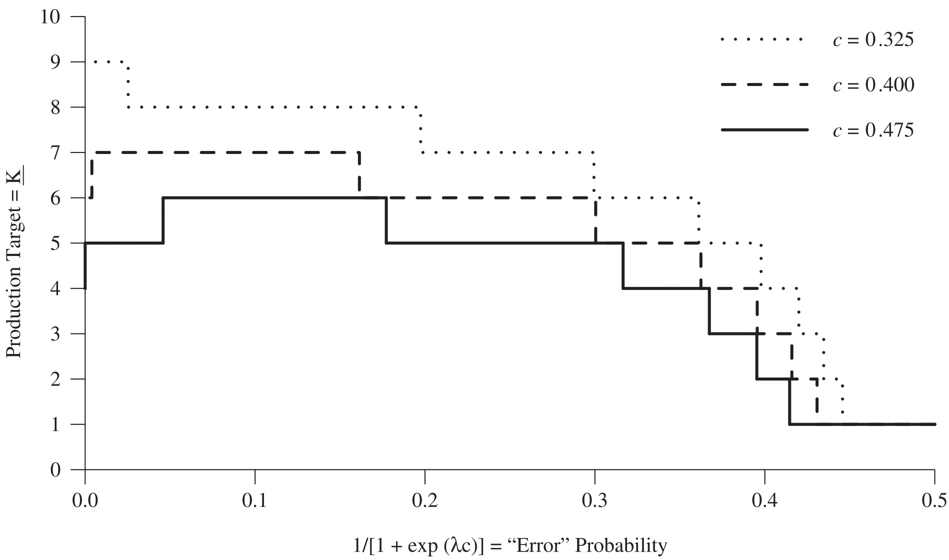

Fig. 9. *Production Thresholds and 'Error' Probabilities*

Note: 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.

6. Discussion and Related Literature

In this article 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 article asks what types of production function are conducive to success. The answer here envisages a family of production functions bounded above by a technological constraint. Inferior members of this family might be interpreted as functions that are obtained via the exploitation of a free-disposal property. It is socially optimal to either leave production at its highest possible level or damage it maximally. Following this procedure, thresholds (often taking the form of production targets) are a robust feature of successful collective actions.

The 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).[21] Oliver *et al.* (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

[21] A retrospective assessment of this literature was provided by Oliver and Marwell (2001).

concise summary of their conclusions.[22] 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.

This article 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 *et al.* (1983) are precisely the features that are associated with the long-run robust success of collective actions.[23] The paper also develops a criterion for setting targets. They should be chosen so that, in effect, voluntary participants are asked jointly to achieve a level of 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'.

An alternative perspective is that of a hypothetical social planner. The type-revelation problem of mechanism design is absent, and hence the planner simply wishes to impose the socially efficient outcome. Without coercion, the planner can create a socially efficient equilibrium but cannot enforce its play. Even if current players obey a suggestion to coordinate on the desired equilibrium, there is no guarantee that future players will do so. Therefore, policies must be robust to evolving play. This is the evolutionary-implementation approach taken by Sandholm (2002, 2005, 2007), in which the desired outcome must be learnt eventually by players who follow reasonable myopic adjustment processes.

Bagnoli and Lipman (1989), Admati and Perry (1991) and Marx and Matthews (2000) considered related models of collective action which feature threshold devices. They demonstrated that the shape of a production technology may influence incentives. These papers provide mechanisms that generate efficient equilibria. However, these authors, who focused on existence issues, did not consider whether an efficient equilibrium will in fact be played. In response, 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'. It is not enough to create a socially optimal equilibrium; rather, the mechanism must ensure that such an equilibrium actually is played.

Appendix. Omitted Proofs

Proof of Proposition 1. From (9) and (10), and suppressing arguments z and θ;

[22] This is also true of subsequent strands of literature, including Macy (1990) and Heckathorn (1993, 1996). The current article 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.

[23] van de Kragt *et al.* (1983) found that pre-play communication was important for the success of a collective action. The focus of their experiments is the relationship between communication and coordination and is related to a large literature on this topic, for example Cooper *et al.* (1992). Nonetheless, their conclusions stress the importance of the threshold structure discussed here.

$$W \sum_{z \in \mathcal{Z}} \exp(\lambda\psi) = \sum_{z \in \mathcal{Z}} w \exp(\lambda\psi).$$

Differentiating both sides with respect to θ yields

$$\frac{\partial W}{\partial \theta} \sum_{z \in \mathcal{Z}} \exp(\lambda\psi) + \lambda W \sum_{z \in \mathcal{Z}} \frac{\partial \psi}{\partial \theta} \exp(\lambda\psi) = \sum_{z \in \mathcal{Z}} \exp(\lambda\psi) \left(\frac{\partial w}{\partial \theta} + \lambda w \frac{\partial \psi}{\partial \theta} \right).$$

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

$$\frac{\partial W}{\partial \theta} \sum_{z \in \mathcal{Z}} \exp(\lambda\psi) = \sum_{z \in \mathcal{Z}} \exp(\lambda\psi) \left[\frac{\partial w}{\partial \theta} + \lambda \frac{\partial \psi}{\partial \theta} (w - W) \right] \tag{12}$$

$$= \sum_{z \in \mathcal{Z}_\theta} \exp(\lambda\psi) \frac{\partial G}{\partial \theta} [n + \lambda(w - W)], \tag{13}$$

where the second equality follows because $\partial w/\partial\theta = n\partial\psi/\partial\theta = n\partial G/\partial\theta$ and because $\partial\psi/\partial\theta = \partial w/\partial\theta = 0$ for $z \notin \mathcal{Z}_\theta$. *Differentiating again and evaluating at* $dW/d\theta = 0$,

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

θ is an output shifter and so the derivate of G is identical for all $z \in \mathcal{Z}_\theta$. *Hence*

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

where $\partial G/\partial\theta \neq 0$, since $z \in \mathcal{Z}_\theta$. *So at a stationary point* $\partial W/\partial\theta = 0$, the second term on the right-hand side must be zero. As a result, the second derivative is

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

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

Proof of Proposition 2. From (9), aggregate welfare satisfies

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

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,

$$\frac{A + w_1(z) e^{\lambda\psi_1(z)} - w_0(z) e^{\lambda\psi_0(z)}}{B + e^{\lambda\psi_1(z)} - e^{\lambda\psi_0(z)}} \leq \frac{A}{B} \quad \Leftrightarrow \quad \gamma(z) \leq \frac{A}{B},$$

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 Proposition, since $\psi_1(z) > \psi_0(z)$ and $w_1(z) > w_0(z)$, by inspection $\gamma(z) \to w_1(z)$ 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 $\tilde{Z}$. 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 proposition, in states where $w_1(z) > W^*_\lambda$ it must be the case that $\theta^*_z = 1$ for all λ large enough, and similarly $\theta_{z'} = 0$ when $w_1(z') < W^*_\lambda$. 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 Proposition 3. For a Cournot-contributions game with symmetric linear costs, the potential and welfare of state z depend only upon the sum of contributions $Z \equiv \sum_i z_i$. In fact,

$$\gamma(z) = \hat{\gamma}(Z) \quad \text{where} \quad \hat{\gamma}(Z) = \frac{L[nF(Z) - cZ] + cZ}{L - 1} \quad \text{and where} \quad L = \mathrm{e}^{\lambda F(Z)}.$$

To check concavity, differentiate twice to obtain

$$\hat{\gamma}'(Z) = \frac{nLF'(Z)[L - 1 - \lambda F(Z)]}{(L-1)^2} - c, \quad \text{and} \quad \hat{\gamma}''(Z) = \frac{\lambda n L[F'(Z)]^2}{L-1}$$
$$+ \frac{nL[L - 1 - \lambda F(Z)]\{\lambda [F'(Z)]^2 + F''(Z)\}}{(L-1)^2} - \frac{2\lambda n L^2 [F'(Z)]^2 [L - 1 - \lambda F(Z)]}{(L-1)^3}.$$

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

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

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

$$\frac{F(Z)F''(Z)}{[F'(Z)]^2} \leq \frac{2r(\mathrm{e}^r - 1) - r^2(\mathrm{e}^r + 1)}{(\mathrm{e}^r - 1)(\mathrm{e}^r - 1 - r)}. \tag{14}$$

A sufficient condition for concavity of $\hat{\gamma}(Z)$ is for the above inequality to hold for all $r \geq 0$. Noting that $F(\cdot)$ is ρ-concave if and only if $-F(Z)F''(Z)/[F''(Z)]^2 \geq \rho - 1$, and that the right-hand side of (14) is bounded below by -0.285, this completes the proof.

Nuffield College and Department of Economics, University of Oxford
Trinity College and Department of Economics, University of Oxford

Submitted: 3 July 2006
Accepted: 22 February 2008

References

Admati, A. R. and Perry, M. (1991). 'Joint projects without commitment', *Review of Economic Studies*, vol. 58(2) (April), pp. 259–76.

Anderson, S. P. and Renault, R. (2003). 'Efficiency and surplus bounds in Cournot competition', *Journal of Economic Theory*, vol. 113(2) (December), pp. 253–64.

Avriel, M. (1972). '*r*-convex functions', *Mathematical Programming*, vol. 2(1) (February) pp. 309–23.

Bagnoli, M. and Lipman, B. L. (1989). 'Provision of public goods: fully implementing the core through private contributions', *Review of Economic Studies*, vol. 56(4) (October), pp. 583–601.

Bergstrom, T. C., Blume, L. and Varian, H. R. (1986). 'On the private provision of public goods', *Journal of Public Economics*, vol. 29(1) (January), pp. 25–49.

Bergstrom, T. C., Blume, L. and Varian, H. R. (1992). 'Uniqueness of Nash equilibrium in private provision of public goods: an improved proof', *Journal of Public Economics*, vol. 49(3) (December), pp. 391–92.

Blume, L. E. (1993). 'The statistical mechanics of strategic interaction', *Games and Economic Behavior*, vol. 5(3) (July), pp. 387–424.

Blume, L. E. (1995). 'The statistical mechanics of best-response strategy revision', *Games and Economic Behavior*, vol. 11(2) (November), pp. 111–45.

Blume, L. E. (1997). 'Population games', in (W. B. Arthur, S. N. Durlauf, and D. A. Lane, eds.), *The Economy as an Evolving Complex System II*, pp. 425–60, Reading MA: Addison-Wesley.

Blume, L. E. (2003). 'How noise matters', *Games and Economic Behavior*, vol. 44(2) (August), pp. 251–71.
Blume, L. E. and Durlauf S. N. (2001). 'The interactions-based approach to socioeconomic behaviour', in (S. N. Durlauf and H. P. Young, eds.), *Social Dynamics*, pp. 15–44, Cambridge MA: MIT Press.
Blume, L. E. and Durlauf S. N. (2003). 'Equilibrium concepts for social interaction models', *International Game Theory Review*, vol. 5(3) (September), pp. 193–209.
Brock, W. A. and Durlauf, S. N. (2001). 'Discrete choice with social interactions', *Review of Economic Studies*, vol. 68(2) (April), pp. 235–60.
Cabrales, A. (1999). 'Adaptive dynamics and the implementation problem with complete information', *Journal of Economic Theory*, vol. 86(2) (June), pp. 159–84.
Caplin, A. and Nalebuff, B. (1991). 'Aggregation and social choice: a mean voter theorem', *Econometrica*, vol. 59(1) (January), pp. 1–23.
Carlsson, H. and van Damme, E. (1993). 'Global games and equilibrium selection', *Econometrica*, vol. 61(5) (September), pp. 989–1018.
Cooper, R., DeJong D. V., Forsythe R. and Ross, T. W. (1992). 'Communication in coordination games', *Quarterly Journal of Economics*, vol. 107(2) (May), pp. 740–71.
Cowan, S. G. (2007). 'The welfare effects of third-degree price discrimination with nonlinear demand functions', *RAND Journal of Economics*, vol. 38(2) (Summer), pp. 419–28.
Cowan, S. G. and Vickers, J. S. (2007). 'Output and welfare effects in the classic monopoly price discrimination problem', Economics Discussion Paper No. 355, University of Oxford.
Giuri, P., Ploner, M., Rullani, F. and Torrisi, S. (2006). 'Skills, division of labor, and performance in collective inventions: evidence from open-source software', LEM Working Paper No. 2004/19, Sant' Anna School of Advanced Studies.
Granovetter, M. (1978). 'Threshold models of collective behavior', *American Journal of Sociology*, vol. 83(6) (May), pp. 1420–43.
Grimmett, G. R. and Stirzaker, D. R. (2001). *Probability and Random Processes*, 3rd edn. Oxford: Oxford University Press.
Harsanyi, J. C. and Selten, R. (1988). *A General Theory of Equilibrium Selection in Games*, Cambridge: MIT Press.
Hausman, J. A. and McFadden, D. (1984). 'Specification tests for the multinomial logit model', *Econometrica*, vol. 52(5) (September), pp. 1219–40.
Hausman, J. A. and Wise, D. A. (1978). 'A conditional probit model for qualitative choice: discrete decisions recognizing interdependence and heterogeneous preferences', *Econometrica*, vol. 46(2) (March), pp. 403–26.
Heckathorn, D. D. (1993). 'Collective action and group heterogeneity: voluntary provision versus selective incentives', *American Sociological Review*, vol. 58(3) (June), pp. 329–50.
Heckathorn, D. D. (1996). 'The dynamics and dilemmas of collective action', *American Sociological Review*, vol. 61(2) (April), pp. 250–77.
Holmström, B. R. (1982). 'Moral hazard in teams', *Bell Journal of Economics*, vol. 13(2) (Autumn), pp. 324–40.
Johnson, J. P. (2002). 'Open source software: private provision of a public good', *Journal of Economics and Management Strategy*, vol. 11(4) (Winter), pp. 637–62.
Kandori, M., Mailath, G. J. and Rob, R. (1993). 'Learning, mutation and long-run equilibria in games', *Econometrica*, vol. 61(1) (January), pp. 29–56.
Macy, M. W. (1990). 'Learning theory and the logic of critical mass', *American Sociological Review*, vol. 55(6) (December), pp. 809–26.
Macy, M. W. (1991). 'Chains of cooperation: threshold effects in collective action', *American Sociological Review*, vol. 56(6) (December), pp. 730–47.
Marschak, J. (1955). 'A theory of teams', *Management Science*, vol. 1(2) (January), pp. 127–37.
Marwell, G. and Oliver, P. E. (1993). *The Critical Mass in Collective Action: A Micro-Social Theory*, Cambridge: Cambridge University Press.
Marx, L. M. and Matthews, S. A. (2000). 'Dynamic voluntary contribution to a public project', *Review of Economic Studies*, vol. 67(2) (April), pp. 327–58.
McFadden, D. (1974). 'Conditional logit analysis of quantitative choice behavior', in (P. Zarambka, ed.), *Frontiers in Econometrics*, pp. 105–42, New York: Academic Press.
McKelvey, R. D. and Palfrey, T. R. (1995). 'Quantal response equilibria for normal form games', *Games and Economic Behavior*, vol. 10(1) (July), pp. 6–38.
Monderer, D. and Shapley, L. S. (1996). 'Potential games', *Games and Economic Behavior*, vol. 14(1) (May), pp. 124–43.
Myatt, D. P. and Wallace, C. (2008*a*). 'An evolutionary analysis of the volunteer's dilemma', *Games and Economic Behavior*, vol. 62(1) (January), pp. 67–76.
Myatt, D. P. and Wallace, C. (2008*b*). 'When does one bad apple spoil the barrel? An evolutionary analysis of collective action', *Review of Economic Studies*, vol. 75(2), pp. 499–527.
Oliver, P. E. and Marwell, G. (2001). 'Whatever happened to critical mass theory? A retrospective and assessment', *Sociological Theory*, vol. 19(3) (November), pp. 293–311.

Oliver, P. E., Marwell, G. and Teixeira, R. (1985). 'A theory of the critical mass. I. Interdependence, group heterogeneity, and the production of collective action', *American Journal of Sociology*, vol. 91(3) (November), pp. 522–56.

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

Palfrey, T. R. and Rosenthal, H. (1984). 'Participation and the provision of discrete public goods: a strategic analysis', *Journal of Public Economics*, vol. 24(2) (July), pp. 171–93.

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

Sandholm, W. H. (2002). 'Evolutionary implementation and congestion pricing', *Review of Economic Studies*, vol. 69(3) (July), pp. 667–89.

Sandholm, W. H. (2005). 'Negative externalities and evolutionary implementation', *Review of Economic Studies*, vol. 72(3) (July), pp. 885–915.

Sandholm, W. H. (2007). 'Pigouvian pricing and stochastic evolutionary implementation', *Journal of Economic Theory*, vol. 132(1) (January), pp. 367–82.

Shibata, H. (1971). 'A bargaining model of the pure theory of public expenditure', *Journal of Political Economy*, vol. 79(1) (January), pp. 1–29.

van de Kragt, A. J. C., Orbell, J. M. and Dawes, R. M. (1983). 'The minimal contributing set as a solution to public goods problems', *American Political Science Review*, vol. 77(1) (March), pp. 112–22.

Warr, P. (1983). 'The private provision of a public good is independent of the distribution of income', *Economics Letters*, vol. 13(2–3) (October), pp. 207–11.

Young, H. P. (1993). 'The evolution of conventions', *Econometrica*, vol. 61(1) (January), pp. 57–84.

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

Young, H. P. (2001). 'The dynamics of conformity', in (S. N. Durlauf and H. P. Young, eds.), *Social Dynamics*, pp. 133–53. Cambridge MA: MIT Press.