



UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

WORKING PAPERS IN ECONOMICS

No 603

Reciprocity Networks and the Participation Problem

Martin Dufwenberg & Amrish Patel

August 2014

ISSN 1403-2473 (print)
ISSN 1403-2465 (online)

Reciprocity Networks and the Participation Problem*

Martin Dufwenberg[†] & Amrish Patel[‡]

July 2, 2014

Abstract

Reciprocity can be a powerful motivation for human behaviour. Scholars argue that it is relevant in the context of private provision of public goods. We examine whether reciprocity can resolve the associated coordination problem. The interaction of reciprocity with cost-sharing is critical. Neither cost-sharing nor reciprocity in isolation can solve the problem, but together they have that potential. We introduce new network notions of reciprocity relations to better understand this. Our analysis uncovers an intricate web of nuances that demonstrate the attainable yet elusive nature of a unique outcome.

Keywords: Discrete public good, participation, reciprocity networks, coordination, cost-sharing

JEL codes: C72, D03, H41.

*We thank Phil Woods for able research assistance and seminar participants in Gothenburg for useful comments. Patel acknowledges generous funding from the Jan Wallander and Tom Hedelius Foundation.

[†]Bocconi University (Decision Sciences and IGIER), University of Gothenburg, University of Arizona, CESifo; martin.dufwenberg@unibocconi.it

[‡]University of Gothenburg; amrish.patel@economics.gu.se

1 Introduction

You receive an email from your boss stating that he must nominate exactly two staff as trade union representatives as soon as possible. All staff would like that someone represents them on the union, but the question is who? If no-one else volunteers, or if at least two others volunteer, there is no sense in you volunteering. If only one person volunteers then the outcome depends on you. You click reply and hesitate... how do your colleagues choose?

The story illustrates the participation problem in the private provision of discrete public goods. Grander scale examples might include international economic agreements, climate treaties, open source software, and political representation. A useful framework for analysing the issue is found in the seminal work of Palfrey & Rosenthal (1984) (P&R). They identify two types of equilibria, that where no-one participates and that where exactly those needed for provision do so. This dichotomy exposes two fundamental problems, concerning efficiency (since all individuals would be better off with provision than without) and coordination (since there are multiple provision equilibria). The problems can neither be solved by evoking the refinement of strict equilibrium nor by using refunds, that is returning costs if there are insufficient participants. However, P&R show that taken together, these two factors can achieve efficiency: with refunds, all strict equilibria are efficient.

The important insights above presume that individuals maximise material payoffs. But scholars argue that in the context of public goods provision it is natural for additional motives such as reciprocity to come into play (Sugden 1984).¹ Reciprocity can be modeled as the desire to be kind to those who are kind to you and unkind to those who are unkind to you (Rabin 1993, Dufwenberg & Kirchsteiger 2004 (D&K), Falk & Fischbacher 2006). In settings where it is beneficial that everyone contributes, the application of reciprocity theory is straightforward and leads to conditional cooperation: players desire to contribute increases as others' contributions go up. Empirical support for such behaviour has been provided (Keser & van Winden 2000, Fischbacher, Gächter & Fehr 2001, Fischbacher & Gächter 2010) via experimental public good games (Ledyard 1995). However, in discrete public goods games it isn't useful, but rather wasteful, to contribute if many others do so anyway, so the implications of reciprocity are not obvious. Will reci-

¹Reciprocity is not the only important social preference in such environments. Makris (2009) considers altruism and Rothenhäusler et al. (2013) a form of guilt.

reciprocity make it easier for you to decide how to reply to your boss' email? Will it solve the efficiency and coordination problems? Answers require systematic analysis.

We use the workhorse model of P&R and conceptualise reciprocity as in the D&K model. Cost-sharing, where individuals reduce each others' cost-burden if the number of participants exceeds that needed for provision, proves crucial. P&R's insights are robust to reciprocity and cost-sharing, taken individually. However, with *both* reciprocity and cost-sharing in the picture, conclusions are dramatically different. There may exist a unique efficient refined equilibrium so the coordination problem can be solved. But results on this potential are marvelously intricate, with possibilities as well as dead ends in the form of multiplicity or non-existence.

To untangle structure, we introduce and study new network notions of reciprocity relationships that describe the players' attitudes towards each other. Methodologically speaking, we are thus connected to an infant literature on games with networks of social preferences (Leider et al. 2009, Broulès and Bramoullé 2013). Among other things, our approach demonstrates the coordinating power of "reciprocity cliques" in a world of non-reciprocal players and how "reciprocity alienation" can impede coordination.

Our uniqueness results may seem surprising for those who believe that "in standard examples, the notion of reciprocal preferences tends to increase the number of equilibria" (Sobel 2005, p. 410). We examine an economically important class of games that under standard preferences seems intrinsically plagued with multiple equilibria, yet reciprocity potentially offers a solution.

Section 2 introduces our model. Section 3 presents results, on reciprocity in P&R (3.1) and on reciprocity and cost-sharing (3.2-3). Section 4 explores welfare, before we conclude in Section 5.

2 Model

We first recall P&R's classic participation game. Let $N = \{1, 2, \dots, n\}$ be the set of players, with $n \geq 3$. Player i 's strategy set is $S_i = \{0, 1\}$ where 1 corresponds to participating in provision and 0 to not doing so. We focus on pure strategies throughout. Let $S = \times_{i \in N} S_i$. Write $s = (s_1, s_2, \dots, s_n)$ for a profile of strategies.

The threshold number of participants for provision of the public good is w , where $1 < w < n$. Let the cost of participation be $x \in (0, 1)$ and each

player receive an additional payoff of 1 if the good is provided. Let π_i be player i 's material payoff function.

P&R introduce two variants of their game differing in whether or not costs are refunded when there are fewer than w participants. To define payoff functions, let m denote the number of players that participate. With no refunds, player i 's material payoff function is defined by

$$\text{No Refunds: } \pi_i(s) = \begin{cases} 1 & \text{if } m \geq w \text{ and } s_i = 0, \\ 1 - x & \text{if } m \geq w \text{ and } s_i = 1, \\ 0 & \text{if } m < w \text{ and } s_i = 0, \\ -x & \text{if } m < w \text{ and } s_i = 1. \end{cases}$$

Any profile where $m = 0$ or $m = w$ is a Nash equilibrium, and in fact also a strict equilibrium.

P&R show that refunds imply the inefficient no participation outcome is no longer a strict equilibrium. With refunds, player i 's material payoff function is defined by

$$\text{Refunds: } \pi_i(s) = \begin{cases} 1 & \text{if } m \geq w \text{ and } s_i = 0, \\ 1 - x & \text{if } m \geq w \text{ and } s_i = 1, \\ 0 & \text{if } m < w. \end{cases}$$

While any profile where $m = w$ remains a strict equilibrium, $m = 0$ doesn't because deviation to participation would be costless. All strict equilibria are thus efficient.

We will additionally consider a counterpart to refunds for $m > w$, cost-sharing, where costs in excess of the provision cost are returned in equal shares to the participating players when $m > w$. Player i 's material payoff function is then defined by

$$\text{Cost-sharing: } \pi_i(s) = \begin{cases} 1 & \text{if } m \geq w \text{ and } s_i = 0, \\ 1 - \frac{w}{m}x & \text{if } m \geq w \text{ and } s_i = 1, \\ 0 & \text{if } m < w. \end{cases}$$

In some contexts cost-sharing is very natural. For instance, in our opening example 2 workers were needed to be union representatives. Suppose that 4 workers volunteer. Distributing the union tasks between the 4 seems more reasonable than making all 4 work on every task.

P&R do not study cost-sharing,² presumably because it only affects payoffs when $m > w$ and such profiles are neither equilibria nor attractive profiles to deviate to given the waste involved. Nash and strict equilibria are thus unaffected if players are motivated by material payoffs only. However with reciprocity motivation cost-sharing will matter.

We next incorporate preferences for reciprocity. Following D&K,³ when player i plays strategy s_i and holds point belief $b_{ij} \in S_j$ about player j 's strategy, player i 's kindness to player j is

$$\kappa_{ij} \left(s_i, (b_{ij})_{j \neq i} \right) = \pi_j \left(s_i, (b_{ij})_{j \neq i} \right) - \frac{1}{2} \left[\max_{s'_i} \pi_j \left(s'_i, (b_{ij})_{j \neq i} \right) + \min_{s'_i} \pi_j \left(s'_i, (b_{ij})_{j \neq i} \right) \right].$$

If $\kappa_{ij}(\cdot) > 0$ player i is kind to j and if $\kappa_{ij}(\cdot) < 0$ player i is unkind to j . When player i holds point belief b_{ij} about j 's strategy and (point) belief c_{ijk} about j 's (point) belief about k 's strategy, player i 's perceived kindness of player j towards player i is

$$\lambda_{iji} \left(b_{ij}, (c_{ijk})_{k \neq j} \right) = \pi_i \left(b_{ij}, (c_{ijk})_{k \neq j} \right) - \frac{1}{2} \left[\max_{b'_{ij}} \pi_i \left(b'_{ij}, (c_{ijk})_{k \neq j} \right) + \min_{b'_{ij}} \pi_i \left(b'_{ij}, (c_{ijk})_{k \neq j} \right) \right].$$

If $\lambda_{iji}(\cdot) > 0$ player i perceives that j is kind to him... etc. Player i 's utility is the sum of his material and reciprocity payoffs,

$$\begin{aligned} U_i \left(s, \left(b_{ij}, (c_{ijk})_{k \neq j} \right)_{j \neq i} \right) & \quad (1) \\ = \underbrace{\pi_i(s)}_{\text{material payoff}} + \underbrace{\sum_{j \in N \setminus \{i\}} \left(Y_{ij} \cdot \kappa_{ij} \left(s_i, (b_{ij})_{j \neq i} \right) \cdot \lambda_{iji} \left(b_{ij}, (c_{ijk})_{k \neq j} \right) \right)}_{\text{reciprocity payoff}}, \end{aligned}$$

where $Y_{ij} \geq 0$ is i 's reciprocity sensitivity towards j . If $Y_{ij} > 0$, a preference for reciprocation is captured by i 's utility increasing when $\kappa_{ij}(\cdot)$ and $\lambda_{iji}(\cdot)$ are non-zero with matching signs.

We now state the two solution concepts we use. Note that we may write $s = (s_i, s_{-i})$ where s_{-i} is a profile of strategies for all players except i .

²Cost-sharing has however been investigated in related games. For example, Shinohara (2009) where players have heterogenous valuations of the public good and Makris (2009) where there is incomplete information over whether or not players are altruistic.

³D&K's theory of reciprocity applies to a general class of extensive game forms. Our focus on a specific game affords us two simplifications: (i) we need not condition strategies and beliefs on histories as the game has simultaneous moves; (ii) we need not discuss D&K's notion of "inefficient strategies" (p. 276) as all strategies are efficient.

Definition 1 (SRE) Profile s is a sequential reciprocity equilibrium (SRE) if for all $i \in N$

1. $U_i \left(s, \left(b_{ij}, (c_{ijk})_{k \neq j} \right)_{j \neq i} \right) \geq U_i \left((s'_i, s_{-i}), \left(b_{ij}, (c_{ijk})_{k \neq j} \right)_{j \neq i} \right)$ for all s'_i ,
2. $c_{ijk} = b_{jk} = s_k$ for all $j \neq i$ and $k \neq j$.

The first condition requires that player i is best-responding given others' strategies and his beliefs. The second demands that players hold correct beliefs. If for all $i, j \in N$, $Y_{ij} = 0$, then Definition 1 describes a Nash equilibrium (+ correct beliefs) in a game where utility equals material payoffs.

While a unilateral deviation from a SRE cannot increase the deviant's payoff, such a deviation from a SSRE leads to an actual loss:

Definition 2 (SSRE) Profile s is a strict sequential reciprocity equilibrium (SSRE) if for all $i \in N$

1. $U_i \left(s, \left(b_{ij}, (c_{ijk})_{k \neq j} \right)_{j \neq i} \right) > U_i \left((s'_i, s_{-i}), \left(b_{ij}, (c_{ijk})_{k \neq j} \right)_{j \neq i} \right)$ for all s'_i ,
2. $c_{ijk} = b_{jk} = s_k$ for all $j \neq i$ and $k \neq j$.

If for all $i, j \in N$, $Y_{ij} = 0$, Definition 2 describes a strict equilibrium (+ correct beliefs) in a game where utility equals material payoffs.

3 Results

Does reciprocity matter in discrete public good games? The answer is essentially no in P&R's setting (3.1). The power of reciprocity is unlocked if one considers cost-sharing, as we show first relying on examples (3.2) followed by some formal statements that generalise those examples (3.3).

3.1 Reciprocity in P&R

The following proposition characterises reciprocity equilibria in P&R's game.

Proposition 1 (Reciprocity in P&R) (i) *With no refunds, the set of SRE equals the set of SSRE and is independent of Y_{ij} .* (ii) *With refunds, the set of SSRE is independent of Y_{ij} .*⁴

It follows from the proposition, and the insight that P&R's model may be seen as the special case where for all $i, j \in N$, $Y_{ij} = 0$, that P&R's results are robust to the incorporation of reciprocity. The intuition is as follows.

Without refunds (i), a single participant cannot unilaterally provide the good. Given this, equilibrium expectations imply that if no one else participates, i 's kindness to j is zero regardless of his participation decision. Hence i has no reciprocity incentive to participate and the inefficient no-participation outcome remains an equilibrium despite reciprocity. Furthermore, since deviation incurs a material cost, no participation is also a SSRE.

With refunds (ii), the set of SRE does now depend on Y_{ij} .⁵ However, this is of little relevance to us since reasoning analogous to the previous paragraph demonstrates the inefficient SRE still exists. We thus have the same motivation to focus on SSRE in Proposition 1(ii) as P&R had to focus on strict equilibria. To see that inefficient profiles cannot be SSRE, consider $m \neq w$. There must be some i whose choice cannot affect others' material payoffs. Equilibrium beliefs then imply that i 's kindness to j is zero and that either he cannot have a strict best-response, or he has a material incentive to deviate. All that remains are the efficient profiles, $m = w$.

3.2 Reciprocity & Cost-sharing: Examples

Reciprocity has a potentially dramatic effect under cost-sharing. Not only does the set of SSRE now depend on Y_{ij} , but uniqueness becomes a real possibility. The coordination problem can thus be solved.

Our results will depend on preferences, specifically the distribution of players' reciprocity sensitivities, the "reciprocity network". It is useful to represent this using weighted directed graphs. Vertices represent players and their labels correspond to player labels. A directed edge originating from player i 's vertex and ending at player j 's vertex implies $Y_{ij} > 0$. The edge

⁴All proofs are found in the appendix; we provide only the main intuition in the text.

⁵For those interested, the set of SRE is a superset of that where for all $i, j \in N$, $Y_{ij} = 0$. The only difference arises for profiles where $m = w - 1$ which are SRE if and only if non-participants are sufficiently reciprocal towards one another.

label is Y_{ij} . Figure 1, for example, represents a 3-player network where for all $i, j \in N$, $Y_{ij} = 0$ except $Y_{12} = Y_{21} = 5$.

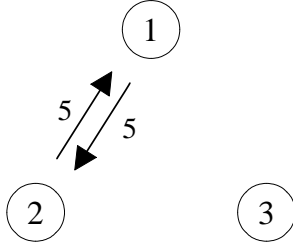


Figure 1: A reciprocity network

We now define our first network notion of reciprocity and illustrate uniqueness.

Definition 3 (Reciprocity clique of strength ρ) *A set of players $L \subseteq N$ is a reciprocity clique of strength ρ if for all $i, j \in L$, $Y_{ij}, Y_{ji} \geq \rho > 0$.*

A reciprocity clique may represent a form of friendship.⁶ It seems unlikely in hierarchical relationships such as those between a firm and workers (cf. D&K 2000) but plausible among peers. A sufficiently strong reciprocity clique may overcome the coordination problem as we illustrate in Example 1.

Example 1: Let $n = 3$, $w = 2$ and the reciprocity network be as in Figure 1. Full participation cannot be a SSRE as player 3 would deviate. Only player 3 and one of the clique members participating cannot be either since the other clique member's reciprocity gain from reducing the cost-share of the participating member is greater than his material cost of participation. This leaves only the clique participating as the unique SSRE. The coordination problem is solved.

The existence of a clique is not generally sufficient for uniqueness as Example 2 will show. The example will also illustrate our second network notion of reciprocity.

⁶Reciprocating unkindness is probably not a particularly common feature of enduring friendships however!

Definition 4 (Reciprocity alienation) *A set of players $L \subset N$ is reciprocity alienated if for all $i \in L$ and all $j \in N \setminus L$, $Y_{ij} = Y_{ji} = 0$.*

Players being reciprocity alienated neither implies nor is implied by them having selfish preferences. For example, players in L having positive reciprocity sensitivities towards one another does not contradict alienation; and, a player not in L with a positive reciprocity sensitivity towards a player in L implies L is not alienated. Intuitively, the notion implies that players inside the alienated group provide no reciprocity incentives to those outside the group and vice-versa.

Our next example adds a non-reciprocal player to Example 1 to illustrate the coordination problems caused by alienation and that a clique is not sufficient for coordination.

Example 2: Let $n = 4$, $w = 2$ and the reciprocity network be as in Figure 2.

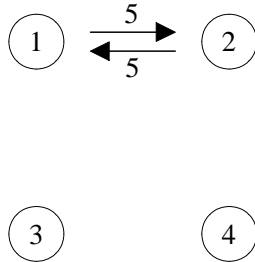


Figure 2: Reciprocity Alienation

Uniqueness is now impossible despite the reciprocity clique. A SSRE with only 1 and 2 participating exists; however, there also exists one where only players 3 and 4 participate. Alienation implies there are no reciprocity incentives acting across the pairs to break the equilibria.

Although both examples have cliques and multiple sets of alienated players, there is uniqueness in the first example and multiplicity in the second. The difference arises because Example 2 has multiple sets of alienated players of size w , whereas Example 1 does not. Section 3.3 shows that multiple alienated sets of this size prevents uniqueness more generally. Example 3 illustrates how a clique can recover uniqueness.

Example 3: Let $n = 4$, $w = 2$, $x = 0.5$ and the reciprocity network be as in Figure 3.

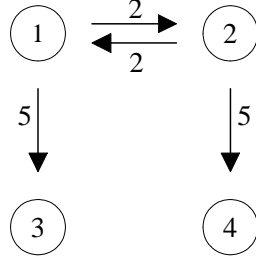


Figure 3: Preventing alienation

Players 3 and 4 only participating is now not a SSRE as player 1(2) has an incentive to deviate: the gain from reciprocating the kindness of player 3(4) is greater than the material cost. All profiles except where only the clique participate can be excluded, uniqueness thus re-emerges.

Example 3 illustrates how a clique being reciprocal towards non-clique players avoids alienation and can imply uniqueness. Section 3.3 shows how general reciprocity networks with this property may achieve coordination.

The final example concerns networks with high reciprocity sensitivities. Given that multiple equilibria exist with standard preferences one may conjecture that sufficiently high reciprocity sensitivity solves the coordination problem. This is not necessarily true, however.

Example 4: Let $n = 4$, $x = 0.5$ and the reciprocity network be as in Figure 4.

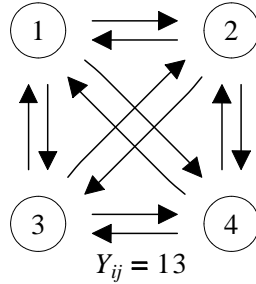


Figure 4: High reciprocity sensitivity

Consider two cases, first $w = 3$. Three participants cannot be a SSRE as the non-participant deviates since his reciprocity gain from doing so outweighs

his material cost. Full participation is the unique SSRE. Now suppose $w = 2$ and 3 players are participating. Participants have no incentive to deviate as their reciprocity gain from reciprocating the kindness of fellow participants is greater than the material cost. Non-participants do not deviate either since equilibrium beliefs imply a particular participant's kindness towards them is zero. Thus, there are multiple SSRE.

More generally high reciprocity gives uniqueness if and only if $w = n - 1$; see Section 3.3. Example 4 also illustrates that there may be more than w participants in a unique SSRE. We show that any number between w and n can be consistent with uniqueness in the next section.

To summarise the insights from Examples 1-4.

1. Existence of a unique SSRE is possible with cost-sharing and reciprocity (Examples 1, 3 and 4).
2. Reciprocity alienation can prevent uniqueness (Example 2).
3. A sufficiently strong reciprocity clique with positive reciprocity sensitivities towards players outside the clique can imply a unique SSRE (Example 3).
4. High reciprocity sensitivities do not necessarily imply uniqueness or multiplicity of SSRE (Example 4).

Readers looking mainly for key intuitions exhibited by typical examples may now skip to Section 4. Readers interested in seeing some more general formal statements, plus further details, should continue to Section 3.3.

3.3 Reciprocity & Cost-sharing: Details

First we describe the SSRE of our model. For strategy profile s , let M denote the associated set of participating players and m be the cardinality of M .

Proposition 2 (Cost-sharing equilibria) *With cost-sharing, all SSRE involve provision of the public good. Profile s where $m = w$ is a SSRE if and only if for all $i \in M$,*

$$1 - x + \frac{(1-x)^2}{2} \sum_{j \in M \setminus \{i\}} Y_{ij} > \frac{x}{2(w+1)} \sum_{j \in N \setminus M} Y_{ij}, \quad (2)$$

and for all $i \in N \setminus M$,

$$\sum_{j \in M} Y_{ij} < 2w. \quad (3)$$

Profile s where $m > w$ is a SSRE if and only if for all $i \in M$,

$$\sum_{j \in M \setminus \{i\}} Y_{ij} > \frac{2m(m-1)^2}{wx}. \quad (4)$$

Proposition 2 states that $m \geq w$ in all SSRE, consistent with refunds ruling out profiles with $m < w$ (Section 3.1). The conditions for a SSRE with $m = w$ are understood as follows. Participant non-deviation (2) requires that the gain from reciprocating the kindness of other participants and the material gain from provision be greater than the loss of not reciprocating the unkindness of non-participants. Non-participant non-deviation (3) demands the material savings be greater than the reciprocity cost of not reciprocating participant kindness. As non-participants cannot affect each other's material payoffs they provide no reciprocity incentives to one another.

For a SSRE with $m > w$ a non-deviation condition is not needed for non-participants as they have no incentive to deviate: each non-participant believes other players believed they could not influence his material payoff. Participant non-deviation (4) requires the gain from reciprocating the kindness of fellow participants be greater than the material cost.

Note that Proposition 2 does not guarantee existence of SSRE. To see one instance of non-existence suppose that for all $i, j \in N$, $Y_{ij} \geq 2w$, so that (3) does not hold. Then take x sufficiently small such that (4) does not hold.

Cliques Coordinating a Selfish World

Examples 1 and 3 demonstrated the potential of a reciprocity cliques to coordinate behaviour. The next two results describe a fairly rich set of circumstances where this is true.

Proposition 3 (Clique with w members) *Suppose that all players have selfish preferences except a reciprocity clique M' of strength $\rho > \max\{0, 2[wx(n-w) - (1-x)(w+1)] / (w^2-1)(1-x)^2\}$ with $|M'| = w$, and that for all $i \in M'$ and $j \in N \setminus M'$, $Y_{ij} = 2w$. In this case there exists a unique SSRE profile s where $M = M'$.*

The proposition generalises Example 3: a sufficiently strong clique with w members and positive reciprocity sensitivity towards non-clique players (to avoid alienation) can coordinate behaviour in a selfish world. See Example 3 for the intuition.

We have seen examples of coordination by cliques with w members (Examples 1 and 3, Proposition 3) and n members (Example 4) participating, but can cliques of other sizes coordinate behaviour and provide the good in a selfish world? The following proposition provides an answer.

Proposition 4 (Clique with $> w$ members) *Suppose that all players have selfish preferences except $M'' = \{1, 2, \dots, m''\}$ where $m'' > w$. Also suppose that for all $i \in M'' \setminus \{m''\}$, $Y_{i(i+1)} = Y_{m''1} = 1 + 2[m''(m''-1)^2 - w(m''-2)(w+1)] / wx$, that for all $k \in M'' \setminus \{i, i+1\}$ and $l \in M'' \setminus \{1, m''\}$, $Y_{ik} = Y_{m''l} = 2(w+1)/x$, and that for all $i \in M''$ and $j \in N \setminus M''$, $Y_{ij} = 2$. In this case there exists a unique SSRE profile s where $M = M''$.*

Note that M'' is a reciprocity clique, however unlike that in Proposition 3 this clique has heterogenous intra-clique reciprocity sensitivities. To see the necessity of this, consider a clique of size $m > w + 1$ with homogenous intra-clique reciprocity sensitivities. The level of reciprocity required for a SSRE with the entire clique participating is more than that required for a SSRE with a proper subset of the clique providing the good (condition (4)). Uniqueness is thus incompatible with the entire clique providing.

By contrast, a clique with heterogenous reciprocity sensitivities can coordinate behaviour and provide the good as stated in Proposition 4. Each clique member is relatively reciprocal to all other players, but is strongly reciprocal towards one other clique member towards whom no-one else is strongly reciprocal. A proper subset of the clique providing the good now cannot be an equilibrium as at least one participant will not have his strongly reciprocal player participating and thus he has an incentive to deviate.

Coordination Difficulties

Our insights on reciprocity cliques may make coordination seem easy. In general, however, coordination via reciprocity networks is far from trivial. We have already noted equilibrium existence problems. Here follow three negative observations of reciprocity networks where instead the coordination problem is very much alive, beginning with alienation.

Observation 1 (Reciprocity alienation) *If there exist two sets of reciprocity alienated players L' and L'' , such that $L' \neq L''$ and $|L'| = |L''| = w$, then there are multiple SSRE.*

To see the logic behind the statement, simply note that provision by an alienated group of size w is a SSRE and multiplicity follows immediately if there are two such groups (e.g. Example 2).

Having reciprocity incentives throughout the set of players is clearly important to avoid this outcome. The conditions in Observation 1 are satisfied more easily than it may seem. For example, at least $w + 1$ players with standard preferences towards whom no-one is reciprocal is sufficient. Sufficiently high reciprocity among all players is however not enough to prevent multiplicity as our next observation shows.

Observation 2 (High reciprocity) *If for all $i, j \in N$, $Y_{ij} > 2(w + 1)/x$, then there exists a unique SSRE if and only if $w = n - 1$.*

To see this, reason as follows. When $w = n - 1$ candidate SSRE with $m = w$ are broken as non-participants have an incentive to reciprocate the kindness of participants (3), leaving only a SSRE with $m = n$, (2). When $w < n - 1$, (4) implies existence of a SSRE with $m = w + 1$ is guaranteed if for all $i, j \in N$, $Y_{ij} > 2(w + 1)/x$, however there are many such equilibria (see e.g. Example 4).

A direct implication of Observation 2 is that for a high reciprocity network, provision of a public good which requires almost full participation may be easier than one that requires low participation. Our next observation points out low participation public goods are also more difficult to provide when preferences are homogenous.

Observation 3 (Homogenous reciprocity) *If $w < n - 1$ and for all $i, j \in N$, $Y_{ij} = Y$, then there exist multiple SSRE.*

To see the logic, reason as follows. With homogenous preferences a unique SSRE cannot have less than n participants otherwise another SSRE would exist where at least one participant and non-participant exchange strategies. Furthermore, full participation cannot be a unique SSRE since the level of reciprocity required for equilibrium is increasing in the number of participants, (4), so there would also exist equilibria with fewer participants. Preference heterogeneity is thus often key to coordination via reciprocity networks.

4 Welfare

We have so far implicitly defined efficiency in terms of material payoffs. This is not obvious given that utility functions are assumed to be a combination of both material and reciprocity payoffs. While some scholars have argued that welfare should be defined on full utility functions (e.g. Rabin 1993, p. 1294), more generally the jury remains out on whether non-material payoffs in social preference models should be included in normative analysis (Bernheim & Rangel 2005).

To illustrate one problem with including reciprocity payoffs in efficiency definitions, recall P&R with neither refunds nor cost-sharing. Let $n = 4$, $w = 2$, $x = 0.5$, player 1 be alienated and the others form a clique of strength 3. Including reciprocity payoffs in efficiency implies player 1 only participating is efficient as clique members get strictly lower utility in all other profiles due to the absence of unkindness reciprocation. Reciprocation of unkindness hardly seems "normatively desirable".

A more thorough examination of whether and how one should incorporate reciprocity payoffs into welfare criteria is clearly beyond the scope of our paper. Nonetheless, we now explain the main implications of defining Pareto efficiency on full utility functions in our game.

With neither refunds nor cost-sharing it remains the case that the w participant SRE is Pareto superior to the no-participation SRE: all players receive strictly higher material payoffs and participants reciprocate kindness towards one another, so receive strictly higher reciprocity payoffs.

Although the inefficiency problem remains, P&R's refunds may no longer recover efficiency. For example, take P&R with only a refund, let $n = 3$, $w = 2$, $x = 0.5$ and all $Y_{ij} = 0$ except $Y_{21} > 8$. Despite players 1 and 2 only participating and players 1 and 3 only participating both being SSRE, the

former now Pareto dominates the latter as 1 receives the same utility and 2 and 3 receive higher utility.

While the set of SSRE may be (partially) Pareto rankable, transitivity of the Pareto criterion implies that there exists some SSRE undominated by other SSRE. This SSRE is in fact efficient overall: if $m < w - 1$, participants in the SSRE are now worse off, and if $m = w - 1$, participants in the SSRE who are now non-participants are worse off.

In Section 3 the set of efficient profiles was fixed and the reciprocity network determined equilibrium behaviour. With reciprocity included in the efficiency definition, the reciprocity network simultaneously determines which strategy profiles are efficient and equilibrium behaviour. Is a unique and efficient SSRE still attainable with cost-sharing?

Recall Example 3 from Section 3.3 to see that the answer is yes. Example 3: $n = 4$, $w = 2$, $x = 0.5$ and for all $i, j \in N$, $Y_{ij} = 0$ except $Y_{13} = Y_{24} = 5$ and $Y_{12} = Y_{21} = 2$, implying a unique SSRE with the clique participating. This is efficient since any other profile either involves non-provision or player 3 and/or 4 participating, hence player 3 and/or 4 being strictly worse off. Since Proposition 3 is established by generalising the logic of this example, a unique and efficient SSRE remains possible.

5 Conclusion

Societies around the world have faced and continue to face great challenges. A fundamental prerequisite to overcome many of these is the coordination of social participation. Palfrey & Rosenthal (1984) (P&R) provided an adequate and specific framework for exploring related issues: discrete-level public goods games, which capture the key problems of private provision of public goods in condensed form.

Reciprocity is an important form of human motivation which many scholars have suggested may matter in public goods provision settings. We have explored how the application of a formal model of reciprocity (Dufwenberg & Kirchsteiger 2004; cf. Rabin 1993) affects predictions in P&R's context, with extensions.

A new potential source of coordination power was uncovered: the networks of reciprocity incentives embedded amongst citizens. These incentives, however, coordinate behaviour only if two key ingredients are present. The first is cost-sharing. Since cost-sharing can potentially be manipulated by

institution designers, understanding its effects seems particularly important. While there is some experimental work in continuous contribution games (Marks & Croson 1998, Spencer et al. 2009), the conditions under which cost-sharing can coordinate behaviour are poorly understood. Our work suggests that empirical exploration of its interaction with reciprocity may be worthwhile.

The second key ingredient we identified for coordination is the nature of the reciprocity network. Properties of the network can both help coordination and hinder it. We are then left with a pressing empirical question: what does the reciprocity network look like in reality? Given its importance and the absence of empirical research on its nature, addressing this may be of interest for public good scholars.

Societies will probably always have new challenges waiting for them. We hope that by understanding how social participation can be coordinated, resolutions are less of a struggle, leaving more time to enjoy the public goods produced.

6 Appendix

Proof of Proposition 1 (Reciprocity in P&R)

(i) *No refunds.* Identify the set of SRE and the set of SSRE, to prove that they are equal and independent of Y_{ij} . First consider a candidate equilibrium strategy profile with no participants. Deviation gives i a strictly lower material payoff given $w > 1$ and does not affect his reciprocity payoff since for all i, j and s_i , $\kappa_{ij}(s_i, s_{-i}) = 0$. Thus non-participation is a SRE and a SSRE. Second, consider a candidate equilibrium profile with w participants. Deviation gives i a strictly lower material payoff and no increase in reciprocity payoff. More specifically, the reciprocity payoff is unchanged for i such that $s_i = 0$ since for all s_i and j , $\kappa_{ij}(s_i, s_{-i}) = 0$; and, is strictly lower for i such that $s_i = 1$ since for all j such that $s_j = 1$, $\lambda_{iji}(1, s_{-j}) > 0$ and $\kappa_{ij}(1, s_{-i}) > \kappa_{ij}(0, s_{-i})$, and for all j such that $s_j = 0$, $\lambda_{iji}(0, s_{-j}) = 0$. Thus w participants is a SRE and a SSRE. To see that there are no more equilibria, consider a candidate equilibrium profile with any number of participants other than 0 or w . Deviation gives i such that $s_i = 1$ a strictly higher material payoff and leaves his reciprocity payoff unchanged given that for all j and s_i , $\kappa_{ij}(s_i, s_{-i}) = 0$. Thus these profiles are neither SRE nor SSRE.

(ii) *Refunds.* Identify the set of SSRE to prove that it is independent of Y_{ij} . Consider a candidate SSRE profile with w participants. Deviation gives i a strictly lower material payoff and no increase in reciprocity payoff. More specifically, i such $s_i = 0$ receives an identical reciprocity payoff since for all j and s_i , $\kappa_{ij}(s_i, s_{-i}) = 0$; and, i such that $s_i = 1$ receives a strictly lower reciprocity payoff since for j such that $s_j = 1$, $\lambda_{iji}(1, s_{-j}) > 0$ and $\kappa_{ij}(1, s_{-i}) > \kappa_{ij}(0, s_{-i})$, and for j such that $s_j = 0$, $\lambda_{iji}(0, s_{-j}) = 0$. Thus w participants is a SSRE. To see that there are no more SSRE, first consider a candidate SSRE profile with more than w participants. Deviation by i such that $s_i = 1$ increases his material payoff and gives identical reciprocity payoff given that for all j , $\lambda_{iji}(s_j, s_{-j}) = 0$. Thus this profile is not a SSRE. Second, consider a candidate SSRE profile with a number of participants strictly between zero and w . Deviation by i such that $s_i = 1$ gives identical material and reciprocity payoffs since for all s_i and j , $\kappa_{ij}(s_i, s_{-i}) = 0$. Thus this profile is not SSRE. Finally, consider the candidate SSRE profile with no participants. Deviation gives identical material and reciprocity payoffs since for all s_i and j , $\kappa_{ij}(s_i, s_{-i}) = 0$. Thus this profile is not SSRE. ■

Proof of Proposition 2 (Cost-sharing equilibria)

First show that non-provision profiles cannot be SSRE. Consider a candidate SSRE profile with no participants. Deviation gives i the same material and reciprocity payoffs since for all s_i and j , $\kappa_{ij}(s_i, s_{-i}) = 0$. Thus this is not SSRE. Consider a candidate SSRE profile with a number of participants strictly between 0 and w . Deviation by i such that $s_i = 1$ gives him the same material and reciprocity payoffs since for all s_i and j , $\kappa_{ij}(s_i, s_{-i}) = 0$. Thus this is not SSRE.

Then identify conditions for provision profiles to be SSRE. A SSRE profile with w participants requires that for all i such that $s_i = 1$, $u_i((1, s_{-i}), (s_{-i}, (s_{-j})_{j \neq i})) > u_i((0, s_{-i}), (s_{-i}, (s_{-j})_{j \neq i}))$, implying (2), and that for all i such that $s_i = 0$, $u_i((0, s_{-i}), (s_{-i}, (s_{-j})_{j \neq i})) > u_i((1, s_{-i}), (s_{-i}, (s_{-j})_{j \neq i}))$, implying (3). In a SSRE profile with strictly more than w participants, deviation by i such that $s_i = 0$ reduces his material payoff and leaves his reciprocity payoff unchanged since for all j , $\lambda_{iji}(s_j, s_{-j}) = 0$. Thus a SSRE requires only that for all i such that $s_i = 1$, $u_i((1, s_{-i}), (s_{-i}, (s_{-j})_{j \neq i})) > u_i((0, s_{-i}), (s_{-i}, (s_{-j})_{j \neq i}))$, implying (4). ■

Proof of Proposition 3 (Clique with w members)

Recall the reciprocity network in the proposition: take $M' \subset N$ such that $|M'| = w$ and (a) for all $i \in N \setminus M'$ and $j \in N$, let $Y_{ij} = 0$, (b) for all $i \in M'$ and $j \in N \setminus M'$, let $Y_{ij} = 2w$, and (c) for all $i, j \in M'$ let $Y_{ij} > \max\{0, 2[w x (n - w) - (1 - x)(w + 1)] / (w^2 - 1)(1 - x)^2\}$. Reason as follows to see that there is a unique SSRE. First consider a candidate SSRE profile with more than w participants. There must exist some $i \in N \setminus M'$ such that $s_i = 1$. But then for such i , (a) implies the LHS of the inequality in (4) is equal to zero. Thus this is not a SSRE. Second, consider a candidate SSRE profile with w participants where the set of participants is not equal to M' . There must exist some $i \in M'$ such that $s_i = 0$ and some $j \in N \setminus M'$ such that $s_j = 1$. But then for such i , (b) implies the LHS of the inequality in (3) equals $2w$. Thus this is not SSRE. Consider the only remaining candidate SSRE where the set of participants equals M' . For all $i \in M'$ inequality (2) is satisfied given (b) and (c). The inequality in (3) is satisfied given (a). Thus this is the unique SSRE. ■

Proof of Proposition 4 (Clique with $> w$ members)

Recall the reciprocity network in the proposition: take $M'' = \{1, 2, \dots, m''\}$ where $|M''| = m'' > w$ and (d) for all $i \in M'' \setminus \{m''\}$ let $Y_{i(i+1)} = Y_{m''1} = 1 + 2[m''(m'' - 1)^2 - w(m'' - 2)(w + 1)] / wx$, (e) for all $k \in M'' \setminus \{i, i + 1\}$ and $l \in M'' \setminus \{m'', 1\}$, let $Y_{ik} = Y_{m''l} = 2(w + 1) / x$, (f) for all $i \in M''$ and $j \in N \setminus M''$, let $Y_{ij} = 2w$, and (g) for all $i \in N \setminus M''$ and $j \in N$, let $Y_{ij} = 0$. Reason as follows to see that there is a unique SSRE. First consider a candidate SSRE where the set of participants is a proper subset of M'' with strictly more than w participants. Given (e) there must be some $i \in M''$ such that $s_i = 1$ and for all j for whom $s_j = 1$, $Y_{ij} = 2(w + 1) / x$. But then for such i , inequality (4) does not hold. Thus this is not SSRE. Second, consider a candidate SSRE where there are strictly more than w participants and the set of participants is not a subset of M'' . There must be some $i \in N \setminus M''$ such that $s_i = 1$. But then for such i , inequality (4) does not hold given (g). Thus this is not SSRE. Third, consider a candidate SSRE profile where there are w participants. There must exist some $i \in M''$ such that $s_i = 0$. But then for such i , inequality (3) does not hold given (d), (e) and (f). Thus this is not SSRE. The final candidate SSRE profile is where the set of participants is M'' . Inequality (4) holds given (d) and (e). This is the unique SSRE. ■

References

- [1] Bernheim, D. B. and A. Rangel (2005) "Behavioral public economics: Welfare and policy analysis with non-standard decision makers" NBER Working Paper Series, Working Paper 11518.
- [2] Bourlès, B. and Y. Bramoullé (2013) "Altruism in networks" *AMSE Working Papers*, no. 1356.
- [3] Dufwenberg, M. and G. Kirchsteiger (2000) "Reciprocity and wage undercutting" *European Economic Review* 44: 1069-78.
- [4] Dufwenberg, M. and G. Kirchsteiger (2004) "A theory of sequential reciprocity" *Games and Economic Behavior* 47: 268-98.
- [5] Falk, A. and U. Fischbacher (2006) "A theory of reciprocity" *Games and Economic Behavior* 54: 293-315.
- [6] Fischbacher, U., Fehr, E. and S. Gächter (2001) "Are people conditionally cooperative? Evidence from a public goods experiment." *Economic Letters* 71(3): 397-404.
- [7] Fischbacher, U. and S. Gächter (2010) "Social preferences, beliefs, and the dynamics of free-riding in public goods experiments" *American Economic Review* 100(1): 541-56.
- [8] Keser, C. and F. van Winden (2000) "Conditional cooperation and voluntary contributions to public goods" *Scandinavian Journal of Economics* 102 (1): 23-39.
- [9] Ledyard, J. (1995) "Public goods: A survey of experimental research" in *Handbook of Experimental Economics*, edited by Kagel, J. and A. Roth, Princeton University Press.
- [10] Leider, S., Möbius, M. M., Rosenblat, T. and Q.-A. Do (2009) "Directed altruism and enforced reciprocity in social networks" *Quarterly Journal of Economics* 124(4): 1815-51.
- [11] Makris, M. (2009) "Private provision of discrete public goods," *Games and Economic Behavior* 67: 292-99.

- [12] Marks, M. and R. Croson (1998) "Alternative rebate rules in the provision of a threshold public good: An experimental investigation" *Journal of Public Economics* 67: 195-220.
- [13] Palfrey, T. and H. Rosenthal (1984) "Participation and the provision of discrete public goods: A strategic analysis" *Journal of Public Economics* 24: 171-93.
- [14] Rabin, M. (1993) "Incorporating fairness into game theory and economics" *American Economic Review* 83: 1281-302.
- [15] Rothenhäusler, D., Schweizer, N. and N. Szech (2013) "Institutions, shared guilt and moral transgression," *WZB discussion paper*, no. SP II 2013-305.
- [16] Shinohara, R. (2009) "The possibility of efficient provision of a public good in voluntary participation games" *Social Choice and Welfare* 32: 367-87.
- [17] Sobel, J. (2005) "Interdependent preferences and reciprocity" *Journal of Economic Literature* 63: 392-436.
- [18] Spencer, M. A., Swallow, S. K., Shogren, J. F. and J. A. List (2009) "Rebate rules in threshold public good provision" *Journal of Public Economics* 93: 798-806.
- [19] Sugden, R. (1984) "Reciprocity: The supply of public goods through voluntary contributions" *Economic Journal* 94(376): 772-87.