Evolution of cooperative networks

No. 346
24-August-2021

Siddhi Gyan Pandey

National Institute of Public Finance and Policy
New Delhi

Evolution of cooperative networks

Siddhi Gyan Pandey*

August 14, 2021

Abstract

Situations that require individuals to mutually cooperate are often analysed as coordination games. This paper proposes a model of cooperative network formation where the network is formed through the process of the coordination game being played between multiple agents. Additionally, network effects are modelled in by the fact that the benefit to any agent from a mutually cooperative link is enhanced, over a base value, by a factor of her *trustworthiness* or reputation as observed by her partner in that link. Within this framework, evolution of cooperative networks is analysed in the presence of *altruistic* agents, through repeated interaction between *myopically best responding* agents in a finite population. Properties of networks that sustain as Nash equilibrium are also analysed.

Keywords: coordination game, network formation, game theory, social networks

1 Introduction

This paper analyses the formation of cooperative networks in contexts of cooperative sharing that have an underlying incentive structure typical of a *coordination game*. Cooperation games like the *prisoner's dilemma* and the *coordination game* have been used to represent many strategic interactions where individuals (also referred to as *players* or *agents*) may benefit from mutual cooperation but may have incentives to refrain from cooperation either as a dominant strategy or in response to the other player. Many situations of everyday cooperative interaction such as sharing of resources, combining of efforts for a joint project or venture, exchange of information and technology that is costly to acquire *etc.* have incentive structures typical of the coordination game wherein agents have incentive to cooperate only if their interacting partner also cooperates. Such a game has two pure strategy Nash equilibria; one where both agents cooperate, and one where both agents refrain from cooperation. Further, in situations where the resources or information being shared has a potentially competitive nature, it is reasonable to assume a conflict between Pareto dominance and risk dominance in the two Nash equilibria[1]. This paper attempts to study such interactions

*The author would like to thank Rajendra P Kundu, Shubhro Sarkar, Anirban Kar, and Debabrata Pal for helpful comments and suggestions.

[1]The Nash equilibria refinements of Pareto dominance and risk dominance were first introduced by Harsanyi and Selten (1988). The risk dominant equilibrium is one which has the largest basin of attraction;

between multiple individuals in contexts where *network effects* of cooperation may be of importance. Proposed here is a model of formation of cooperative networks where the outcomes of multiple pairwise cooperative interactions determine the network of cooperation. The idea is that cooperative interactions often happen in a context where the benefit or value of a link of mutual cooperation to any player may be enhanced by a factor of her own reputation as perceived by the other interacting player. The assumption that the perception of one's reputation by a cooperating partner (also referred to as *trustworthiness*) enhances one's valuation of that mutually cooperative link is supported by the idea that in most collaborative situations like exchange of information or resources, the quality of information that one is willing to give and the amount of resources that one is willing to share with a cooperative partner increases with one's trust on the partner.

The model of cooperative network formation proposed in this paper has the following features. Agents in a finite population interact with *all other agents* in situations where there are opportunities to benefit from mutual cooperation. In each interaction, agents have to decide between cooperating or refraining from cooperation. Extending and sustaining cooperation is costly; for every link of cooperation (reciprocated or non-reciprocated) that an agent i extends, a cost c_i is incurred by her. This cost is an indicator of not only the logistical costs of sharing resources or information, but also any other exogenous factors that may increase the effort or investment that i needs to make to derive benefits from the link. As mentioned before, payoffs from mutual cooperation as well as from unilateral cooperation are typical of a coordination game. Further, the benefit to any agent from a mutually cooperative link is enhanced, over a base value, by a factor of her *trustworthiness*, *i.e.* her reputation as observed by her partner in that link. Agent i's trustworthiness to j is simply the number of instance that agent j can observe, *through common mutual cooperative partners*, in which agent i extends cooperation. With this consideration, there may be situations where sustaining non-reciprocated cooperation is not sub-optimal for an agent because of the indirect utility or value that it generates by enhancing her reputation and trustworthiness in other interactions.

While cooperation is costly, we assume the presence of some *altruistic agents*, for whom the cost of cooperation is extremely low because of their preference of cooperation, *even non-reciprocated or unilateral*, over a situation of no cooperation in any interaction. The rationale for assuming the presence of altruists comes from the widely acknowledged role of altruists in evolution of cooperation in not just social sciences literature, but also in the literature of cognitive sciences and biology[2]. The rationale for assuming that a player's mutually cooperative links generate higher value if her trustworthiness to her interacting partners is higher, and the rationale for choosing to define trustworthiness or reputation as a measure of one's instances of cooperation with other agents is motivated by the idea of *indirect reciprocity* in social sciences literature[3] which states that individuals often have incentive

i.e., the more uncertainty that a player has about the actions of the other player, the more likely she is to choose the action corresponding to the risk dominant equilibrium.

[2]Lehman and Keller (2006), Carpenter and Myers (2010) and Telmo (2011) are a few examples in the wide literature analysing the nature of human altruism and its role in the evolution of cooperation under various conditions.

[3]See Semmann, Krambeck and Milinski (2004), Engelmann and Fischbacher (2009), Sylwester and Roberts (2010) and Marsh (2018) for analysis of the role of indirect reciprocity and reputation building

to cooperate in some instances without expecting reciprocity, in the expectation that their cooperative actions enhance their *reputation scores*, causing other agents to cooperate with them because of their high reputation.

With these features of altruism and indirect reciprocity or reputation building, we apply our model to analyse the process of evolution of cooperative networks starting from a situation of no cooperation, through repeated interaction of agents who *myopically best respond* in each period. The model of cooperative network formation proposed in this paper attempts to contribute to the larger literature of strategic network formation for which Aumann and Myerson (1988) acted as pioneers by being the first to model network formation explicitly as a game. The literature progressed with cooperative network formation being largely modelled as a bilateral process, since the general context of cooperation lends itself to the assumption that meaningful cooperation requires agreement of both interacting parties to cooperate. Note that this model is in agreement with this understanding; in the absence of network effects, individuals do benefit only from coordinated or mutual cooperation. However, this model goes one step further to recognize from literature the role of indirect reciprocity in evolution of cooperation and includes the possibility of extending and sustaining unilateral or non-reciprocated cooperation due to the incentives of reputation building. Thus the contribution of this model is that it allows for the possibility of forming links of *unilateral cooperation* by specifying a payoff function wherein such links can potentially contribute positively to utility. Our assumption that the value of links of cooperation is enhanced by the *embeddedness* of these cooperative links in the network conforms to a well accepted phenomenon in social sciences, with substantial research in sociology arguing that embeddedness of social relationships enhances trust and confidence in the integrity of transactions in these relationships. In the literature of social networks one of the earliest emphasis on this was laid by Granovetter (1985), who articulated the importance of embeddedness of cooperative interactions in the social structure in explaining the evolution of cooperation by observing that *the on-going networks of social relations between people discourage malfeasance.*

To analyse the evolution of cooperative networks, we assume that myopically best responding agents repeatedly interact with each other. The dynamics of this process is propelled by the presence of altruists and the incentive for reputation building that other agents have. In fact such dynamics provides a framework within which the impact of external perturbations on cooperative structures can be analysed. The process of dynamics proposed here differs from earlier dynamic models in the context of cooperative network formation in that it allows the system to converge to equilibrium through a process of best response dynamics instead of pairwise deviations. Here the process of convergence happens through independently considered and optimised unilateral deviations and the system hence converges to a network with an underlying strategy profile that is Nash equilibrium. Earlier works by Watts and Jackson in this context typically assume[4] that in the process of convergence to a steady state while links of cooperation may be severed unilaterally, they can be formed only with

in situations of cooperation in various contexts.

[4]See Watts(2001), Jackson and Watts (2001, 2002a, 2002b). The different models considered by them add incrementally to the complexity of the process, and in Jackson and Watts (2002b) the idea of reputation is incorporated by allowing players to choose whether to add or sever links based on their (prospective) partner's past behaviour.

mutual agreement. This assumption is specific to the context of these models, where links of cooperation are of one type only and are formed by mutual agreement. Thus a steady state or equilibrium for these models has the property that no single agent wants to sever a link unilaterally, and no pair of nodes wants to form a link with each other (when none exists) while simultaneously (and optionally) severing their other links. That is, the steady state or equilibrium network converged to is a *pairwise stable* network. Unilateral deviations underlying the dynamics in our proposed model is not intuitively unappealing since agents *have* incentives to extend unilateral cooperation to each other. This gives the model the features of an *agent based model* which can be used to analyse situations where agents respond to individual incentives without the need to negotiate with each other before deciding on their choice of extending cooperation.

The rest of the paper is organized as follows. The next section formally sets up the model and proceeds to analyse the evolution process. We find that under various conditions on the composition of population (depending on costs of cooperation), five different kinds of equilibria emerge. The third section discusses some necessary conditions on Nash equilibria that may emerge from different processes of convergence, and the paper concludes with a discussion on our results, the limitations of the model and further extensions that may be explored.

2 Evolution of cooperation

2.1 The model

$N = \{1, 2, 3, ...n\}$ with $n \geq 3$ is a finite set of all players.
In any time period t, each player i chooses an action a^t_{ij} for any other player j such that $(\forall i \in N)(\forall j \in N - \{i\})(a^t_{ij} \in \{\alpha, \beta\})$, where α refers to the act of cooperating while β is the act of refraining from cooperation. Any player i's strategy in time period t is denoted by $s^t_i = (a^t_{i1}, a^t_{i2} \dots a^t_{ii-1}, a^t_{ii+1}, \dots .a^t_{in})$. A strategy profile in period t is of the form $(s^t_1, s^t_2, \dots s^t_n)$. Note that the cooperative network results from every agent's action choices in a given strategy profile. The formalization of link formation and neighbourhoods is as follows.

$a^t_{ij} = \alpha$ and $a^t_{ji} = \alpha \longleftrightarrow i$ and j share a mutually cooperative tie in period t.
(also referred to as an $\alpha - \alpha$ tie)
Or, $i \in P_j(s)$ and $j \in P_i(s)$

$a^t_{ij} = \alpha \wedge a^t_{ji} = \beta \longleftrightarrow i$ is in unilateral cooperation with j in period t.
(also referred to as an $\alpha - \beta$ tie between i and j)
Or, $j \in E_i(s)$

$a^t_{ij} = \beta \wedge a^t_{ji} = \beta \longleftrightarrow i$ and j share a non cooperative tie in period t.
(also referred to as a $\beta - \beta$ tie)

Further, $T_{ij}(s) = P_j(s) \cap (P_i(s) \cup E_i(s))$ and $t_{ij}(s) = |T_{ij}(s)|$ is i's trustworthiness as perceived by j; which is the instances that j can observe, *through links of mutual cooperation*, in which

i extends cooperation to other agents. $P_{ij}(s) = P_{ji}(s) = P_i(s) \cap P_j(s)$ is the set of common mutually cooperative neighbours that agents i and j share, with $p_{ij}(s) = p_{ji}(s) = |P_{ij}(s)|$. Note that for any strategy profile and pair of nodes, $t_{ij}(s) \geq p_{ij}(s)$

The payoff matrix of the base game (*excluding network effects*) is as follows.

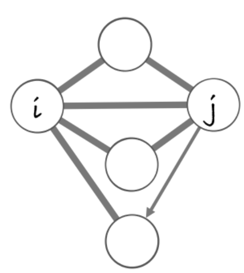
		Player j: α	Player j: β
Player i	α	(x, x)	(z, y)
	β	(y, z)	$(0, 0)$

where $x > y > 0 > z$ ensures that the payoffs are consistent with the coordination game and zero utility is attributed to the situation where both agents refrain from cooperating with each other. Finally, cooperation is costly to extend and sustain. For simplicity we assume that the cost of extending cooperation, whether reciprocated or non-reciprocated (unilateral) is the same. For any agent i, a cost c_i is incurred in every interaction where i cooperates. Unless agents are *altruistic*, it can be reasonably assumed that sustaining unilateral cooperation in absence of network effects of reputation is sub-optimal and $c_i > z$. For altruistic agents we assume $c_i < z$.

Then, the utility that agent i derives from strategy profile s is given by:

$$u_i(s) = \sum_{j \in P_i(s)} (t_{ij}(s) + 1)x + \sum_{j \in E_i(s)} z + \sum_{j \in N | i \in E_j(s)} y - \sum_{j \in P_i(s) \cup E_i(s)} c_i$$

Note that the value of every link of mutual cooperation is enhanced by a factor of one's own trustworthiness as perceived by the interacting partner. It is this utility enhancing impact of network based trustworthiness on links of mutual cooperation that brings in *network effects* into the model. For example, in the figure below (where darker links are links of mutual cooperation and the directed link is a link of unilateral cooperation), agent j gets communication (through links of mutual cooperation) about i's cooperation in two more instances, and hence i's reputation or (network based) trustworthiness as perceived by j is $t_{ij}(s) = 2$. On the other hand, j's reputation as perceived by j is $t_{ji}(s) = 3$.



The network effects of network based trustworthiness ensure that the incentive structure faced by any individual is not reduced to one where all interactions with other agents are independently considered.

At the start of analysis, note that links of unilateral cooperation contribute to utility only through enhancing trustworthiness in other interactions, while links of mutual cooperation give a direct utility and also contribute to enhancing trustworthiness in other interactions with common mutual parters. Then, any change in utility from a unilateral deviation will comprise of terms for both these effects of the deviation. For example, the following expressions give the expected change in utility from unilateral deviations comprising of single action changes. Note that the notation $\triangle u_i$ is used throughout the paper in a context specific way to express the expected change in utility from any unilateral deviation being considered.

If, from strategy profile s, agent i unilaterally withdraws cooperation from an $\alpha - \alpha$ link with j, then $\triangle u_i = c_i + y - (t_{ij}(s) + 1)x - p_{ij}(s)x$ ...(i)

If, from strategy profile s, agent i unilaterally offers cooperation to j (previously engaged in a $\beta - \beta$ link), then $\triangle u_i = -c_i + z + p_{ij}(s)x$ ...(ii)

If, from strategy profile s, agent i reciprocates cooperation to j (previously engaged in a $\beta - \alpha$ link), then $\triangle u_i = -c_i - y + (t_{ij}(s) + 1)x + p_{ij}(s)x$ ...(iii)

If, from strategy profile s, agent i withdraws cooperation from an $\alpha - \beta$ link with j, then $\triangle u_i = c_i - z - p_{ij}(s)x$ ...(iv)

2.2 Evolution

The process of network formation starts after agents start interacting from an initial state of zero cooperation. Agents repeatedly interact over discrete time periods with everyone else, **myopically best responding to the strategy profile resulting from the period.** That is, the system progresses with agents unilaterally deviating in every period to best respond to the last period's actions of other players. It converges to a *Nash equilibrium*, or simply an equilibrium when two subsequent periods see no unilateral deviation by any agent. In this section, we attempt to characterize all possible Nash equilibria that can emerge from this process, in absence of external perturbations or errors in rational decision making. In analysing the process of convergence to equilibrium, the following categorization of the population is useful.

$A = \{i \in N | c_i \leq z\}$ and $|A| = a$, where $2 \leq a \leq n - 2$
$F = \{i \in N | c_i > z \wedge c_i \leq ax - y\}$ and $|F| = f$
$D = \{i \in N | c_i > z \wedge ax - y < c_i\}$ and $|D| = d$

Note that agents in A are *altruists* who prefer to extend cooperation, even unilaterally, over being in a situation of mutual non cooperation with other agents. The rest of the population is characterized on the following basis. The set F consists of all the agents for whom cost of cooperation is low enough that if there exists a clique (completely connected set of agents) comprising of a or more agents who are all extending cooperation to them, then they have incentive to reciprocate cooperation to all the agents in this clique. Agents in D, on the other hand, do not have incentive to reciprocate cooperation in such a situation. This behaviour of non altruistic agents in F and D is discussed in detail later.

The presence of altruists is crucial to starting off the process of cooperative link formation in this evolution process; in fact, there must be at least two such agents. However, assuming $a > n - 2$ (or $a \geq n - 1$) means that the entire population except one individual consists of altruists, contradicting the incentive mechanism underlying situations that the model attempts to analyse. Further, the number of altruists are assumed to be atleast two so that analysis can focus on non trivial cases[5]. It is also assumed that when individuals are indifferent between cooperating and not cooperating, they choose to cooperate. Although this is a simplifying assumption for the purpose of the analysis (since it allows us to infer precisely the resulting strategy profile at the end of a period and best responses for the next period can be unambiguously analysed) it is also a reasonable assumption about social behaviour of humans, given that strong tendencies of reciprocity in human social behaviour is well documented. As the start of analysis, note the following lemmas.

Lemma 1: $\big(\forall t \geq 1\big)\big(\forall i \in A\big)\big[(\forall j \in N - \{i\})(a^t_{ij} = \alpha)$ is the best response to period $t - 1\big]$

That is, the unique best response for altruists in every period starting from the first period onwards is to extend cooperation to everyone else in the population.

This lemma follows directly from the fact that in any interaction, altruists prefer being in non-reciprocated cooperation over being in a situation of no cooperation. Therefore altruists do not depend on network effects of trustworthiness enhancement to engage in cooperation, which means that they extend cooperation to all other agents in the very first period that interactions begin, and do not withdraw cooperation from anyone through the entire process of convergence to equilibrium.

The next lemma is applicable to agents in F, for whom $z < c_i \leq ax - y$.

Lemma 2: Suppose $\exists M \subset N$ such that $|M| \geq a$ and $(\forall j \in M)(\forall k \in M - \{j\})(a^t_{jk} = \alpha)$. Then, $\big(\forall i \in F\big)\big[\big(\forall j \in M\big)\big(a^t_{ji} = \alpha\big) \longrightarrow \big(\forall j \in M\big)\big(a^{t+1}_{ij} = \alpha$ in i's best response in $t + 1\big)\big]$

That is, if there is a set of agents atleast as many as a, who are engaged in ties of mutual cooperation with each other and offering cooperation to $i \in F$ in a period t, then i's best response in period $t + 1$ involves cooperating with *all* agents in this set.

To prove this lemma, we consider an arbitrary $M \subset N$ of size atleast as much as a such that all agents in M are engaged in ties of mutual cooperation with each other and offering cooperation to $i \in F$ in period t. Of all of these agents, i may be refraining from cooperation with some of them in period t (call this set $M_1 \subseteq M$) and i may be already cooperating with some of them in period t (call this set $M_2 \subseteq M$). Then, we consider an arbitrary unilateral deviation by i wherein i changes her actions with respect to some of these agents.

[5] If $a = 1$ and for all non altruistic agents, $c_i > x - y$, then no links of mutual cooperation are formed through repeated interaction of these agents because non altruistic agents have no incentive to reciprocate to the single altruist. Further, we want to be able to assume $c_i > x - y$ for non altruistic agents because it is for this cost range that network effects play a role in sustaining cooperation.

In the most general form of deviation, i can reciprocate cooperation to k_1 agents in M_1 and withdraw cooperation from k_2 agents in M_2. To prove our lemma, we show that the change in utility expected from this unilateral deviation is *maximised* when k_1 takes the maximum possible value and k_2 takes the minimum possible value. This means that i's best response to period t would involve reciprocating cooperation to *all* agents in M_1 and continuing to cooperate with *all* agents in M_2, thus cooperating with all agents in M in period $t+1$. The formal proof is as follows.

Proof: Suppose $(\exists M \subset N)(|M| \geq a \wedge (\forall i \in M)(\forall j \in M - \{i\})(a^t_{ij} = \alpha))$
and $(\exists i \in F)(\forall j \in M)(a^t_{ji} = \alpha)$
Denote $|M|$ with m. Without loss of generality, we assume:
$M_1 \subseteq M = \{j \in M | a^t_{ij} = \beta\}$ $\quad |M_1| = m_1 \leq m$
$M_2 \subseteq M = \{j \in M | a^t_{ij} = \alpha\}$ $\quad |M_2| = m_2 \leq m$

Consider the arbitrary unilateral deviation in which i *switches* to reciprocate cooperation to k_1 agents in M_1 and withdraw cooperation from k_2 agents in M_2. To calculate the change in utility from this deviation, we first observe that expression for total utility that i gets from a strategy profile s is additively separable as a sum of net utility that i receives from every single link that she participates in. Consider the following calculation of the change in utility expected from this unilateral deviation. We make use of following notation.
$u_{ij}(s^0)$ is the *net* utility that i received from a link with j before the unilateral deviation.
$u_{ij}(s')$ is the *net* utility that i gets from a link with j after the unilateral deviation.
u^0_i is i's utility before the deviation and u'_i is i's utility after the deviation.

Now, $u^0_i = m_1 y - m_2 c_i + \sum_{j \in M_2} (t_{ij}(s^0) + 1)x + \sum_{j \in N-M} u_{ij}(s^0)$,
where $\sum_{j \in M_2} (t_{ij}(s^0) + 1)x \geq m_2^2 x$ because because i has a *minimum* trustworthiness of $(m_2 - 1)$ for each of m_2 the cooperative links that she is engaged in, since all the m_2 agents are linked with each other in ties of mutual cooperation.

Further, $u'_i = (m_1 - k_1 + k_2)y - (m_2 - k_2 + k_1)c_i + \sum_{j \in M | a'_{ij} = \alpha} (t_{ij}(s') + 1)x + \sum_{j \in N-M} u_{ij}(s')$,
where $\sum_{j \in M | a'_{ij} = \alpha} (t_{ij}(s') + 1)x \geq (m_2 - k_2 + k_1)^2 x$ by the same argument as above.

Thus $\triangle u_i \geq (k_1 - k_2)[x(2m_2 + k_1 - k_2) - (y + c_i)] + \sum_{j \in N-M} u_{ij}(s') - \sum_{j \in N-M} u_{ij}(s^0)$

Consider the expression $(k_1 - k_2)[x(2m_2 + k_1 - k_2) - (y + c_i)]$.
This equals $(k_1 - k_2)[m_2 x - (y + c_i) + k_1 x + (m_2 - k_2)x]$, where $k_1 x + (m_2 - k_2)x > 0$ because $m_2 \geq k_2$. If $m_2 x - (y + c_i) > 0$, this expression is strictly increasing in k_1 and strictly decreasing in k_2, hence is maximised by $k_1^* = m_1$ and $k_2^* = 0$. If $m_2 x - (y + c_i) < 0$ This may be (locally) maximized in two ways :
$k_1 = 0$ and $k_2 = m_2$ gives a locally maximised value equal to $-m_2[m_2 x - (y + c_i)]$
$k_1 = m_1$ and $k_2 = 0$ gives a locally maximised value equal to $m_1[mx + m_2 x - (y + c_i)]$

Suppose $-m_2[m_2x - (y + c_i)] \geq m_1[mx + m_2x - (y + c_i)]$.

$\rightarrow m(y + c_i) \geq [m_1 m + m_2(m_1 + m_2]x$

$\rightarrow m(y + c_i) \geq m^2 x \longrightarrow mx - (c_i + y) \leq 0$

This is a contradiction because $m \geq a \wedge ax - c_i - y > 0 \longrightarrow mx - c_i - y > 0$.

Thus $k_i^* = m_1$ and $k_2^* = 0$ maximises $(k_1 - k_2)[x(2m_2 + k_1 - k_2) - (y + c_i)]$. Further, these values also ensure that $\sum\limits_{j \in N-M} [u_{ij}(s') - u_{ij}(s^0)]$ is non negative and maximised[6].

Thus, agent i in F reciprocates cooperation to all the agents in M with whom she was earlier not cooperating, and does not withdraw cooperation from any agents in M with whom she is cooperating. This means that in i's best response, she reciprocates cooperation to all agents in M, and lemma 2 is proved. □

Thus we find in the presence of a set of $m \geq a$ agents who are completely connected by ties of mutual cooperation and offering cooperation to $i \in F$, i's best response should involve reciprocating cooperation to *all* of them. This also means that if $i \in F$ is already engaged in mutual cooperation with *all* agents in M, she has no incentives to withdraw cooperation from *any* of them. This lemma is useful in considering optimal choice of agents in F in subsequent analysis.

Given the way that the sets A, F and D are defined, the first two periods after the initial period play out in the following way: in the first period, all agents in A extend cooperation to all other agents and end up sharing links of mutual cooperation with each other. All other agents continue to refrain from cooperation. In the second period, all agents in F reciprocate cooperation to all agents in A, while refraining from cooperation with respect to all other agents (including other agents in F). Agents in D continue refraining from any cooperation even in their interactions with altruists. In the third period, agents in F may consider extending cooperation to other agents in F (while still refraining with respect to agents in D), and agents in D may now consider reciprocating cooperation to the altruists *along with extending cooperation to agents in* F for enhanced trustworthiness. Whether or not agents in F will actually cooperate with each other and whether or not agents in D will engage in any cooperative activity depends on various conditions.

Thus we find that from the third period onwards, the best response dynamics becomes contingent on various cost conditions and size of the three types of subsets of the population. These conditions can be broadly divided into two categories: when $z \leq -y$ and when $z > -y$.

When $z \leq -y$, we find that agents in D have no incentive to cooperate in any interaction under any circumstance. Any agent in $i \in F$ extends cooperation to other agents in F depending on the relation between her cost of cooperation c_i and the expression $ax + z$. If $ax + z$ is strictly less than the cost of cooperation for all agents in F, then no further cooperation is extended and the system witnesses a convergence to equilibrium with no further activity after the second period. However, if there are some agents in F for whom cost of cooperation is lesser than $ax + z$, then these agents extend cooperation to all other

[6] Forming new links of mutual cooperation without breaking off any links of mutual cooperation can only increase the payoff from other links through reputation gains.

agents in F, who find it in their interest to reciprocate in the next period. This may generate incentives for even more agents in F to cooperate with each other, till the point where only a subset of F may remain uncooperative towards each other. Alternatively, it may even happen that all agents in F cooperate with each other in equilibrium. This is also the equilibrium that is converged to when $ax + z$ is less than cost of cooperation *for all agents in* F. Note however that through this process at no point do agents in F have an incentive to cooperate with agents in D and at no point do the latter have incentive to reciprocate cooperation to altruists or to cooperate with anyone else.

When $z > -y$, agents in F have incentive to cooperate with each other in the third period. Further, agents in D *may* also have incentive (in the third period) to engage in cooperative activity by reciprocating to altruists along with extending cooperation to agents in F for the purpose of reputation building. We find that whether or not an agent i in D finds it optimal to do so depends on the sign of the expression $a(ax - c_i - y) + f(ax + z - c_i)$. If this expression is negative for all agents in D, agents in D continue to refrain from cooperation and we have an equilibrium where agents in $A \cup F$ engage in mutual cooperation with each other while altruists sustain unilateral cooperation with agents in D. However, if $a(ax - c_i - y) + f(ax + z - c_i)$ is positive for some agents in D, these agents reciprocate cooperation to *all* altruists and extend cooperation to *all agents in* F, who reciprocate in the subsequent period. This further gives these agents in D the incentive to cooperate with each other as well, and may even give incentive to some *other* agents in D to begin engaging in cooperative activity by reciprocating cooperation to altruists while extending cooperation to agents in F and those agents in D who are engaged in mutually cooperative links with agents in $A \cup F$. This process may continue and reach an equilibrium where only a few agents in D remain refraining from cooperation. Alternatively, it may converge to an equilibrium where all agents engage in mutual cooperation with all other agents, yielding a complete network of mutual cooperation. This is also the equilibrium that is converged to when $a(ax - c_i - y) + f(ax + z - c_i)$ is positive for all agents in D.

Thus we find that five different kinds of equilibria emerge under six categorizations of cost conditions. We elaborate our analysis below; the structure of our analysis of the evolution and convergence process is as follows. Lemma 1 ensures that agents in A best respond by cooperating in all their interactions with other agents in every period; we need not consider unilateral deviations for these agents. Then, in each period a general unilateral deviation from the strategy profile of the previous period is considered for arbitrary agents in sets F and D. A general unilateral deviation can be a combination of many actions, and the specificities of those actions will yield an expected change in utility. To arrive at *optimal unilateral deviation* or best response of an agent, we find the specific actions that will maximise this expected change in utility. The best responses of all agents in any period give the strategy profile for that period, and in the next period the process of calculating best responses is repeated again. The following subsections discuss the formal proofs of the process described above.

2.2.1 First period

Consider $i \in N - A$. If i extends α to m agents in $N - \{i\}$, where $0 \leq m \leq n - 1$ then $\triangle u_i = -mc_i + mz$

Since $(\forall i \in N - A)(c_i > z)$, $\triangle u_i$ is negative for any $m > 0$ and is maximised at $m^* = 0$

Thus we have $(\forall i \in N - A)(\forall j \in N - \{i\})(a^1_{ij} = \beta)$

Also by lemma 1, $(\forall i \in A)(\forall j \in N - \{i\})(a^1_{ij} = \alpha)$. This remains true of every subsequent period as well.

2.2.2 Second period

Consider $i \in F$. By lemma 2, these agents must reciprocate cooperation to *all* altruists, and we have $(\forall i \in F)(\forall j \in A)(a^2_{ij} = \alpha)$. Further, offering α to other agents in $N - A$ yields neither direct nor reputation related benefit to i, since $c_i > z$ and there are no benefits of reputation gain (agents in $N - A$ are not connected in $\alpha - \alpha$ ties to agents in A at the end of Period 1). Thus we have $(\forall i \in F)\big[(\forall j \in N - A - \{i\})(a^2_{ij} = \beta) \wedge (\forall j \in A)(a^2_{ij} = \alpha)\big]$

Consider $i \in D$. For these agents, a general unilateral deviation may involve a combination of the following actions: (i) switching to reciprocate α to k agents in A $(0 \leq k \leq a)$, and (ii) switching to offer α to m agents in $N - A - \{i\}$ $(0 \leq m \leq f + d - 1)$

For the same reasons as for agents in F, best response can only include reciprocating α to k agents in A $(0 \leq k \leq a)$. Such a unilateral deviation yields expected change in utility $\triangle u_i = -kc_i - ky + k(k)x = k(kx - c_i - y)$

Now, $(\forall i \in D)(c_i > ax - y) \rightarrow (\forall k \geq 1)(kx - c_i - y \leq ax - c_i - y < 0)$. Thus this change in utility is maximised by $k^* = 0$ and agents in D do not deviate in this period. Thus we have $(\forall i \in D)(\forall j \in N - \{i\})(a^2_{ij} = \beta)$

Figure 1 illustrates the strategy profile that emerges at the end of the second period, where links of mutual cooperation are represented as thicker links and links of unilateral links are directed links. Also, agents in A are denoted by white nodes, agents in F by grey and agents in D by black nodes. Subsequent diagrams also adopt this representation.

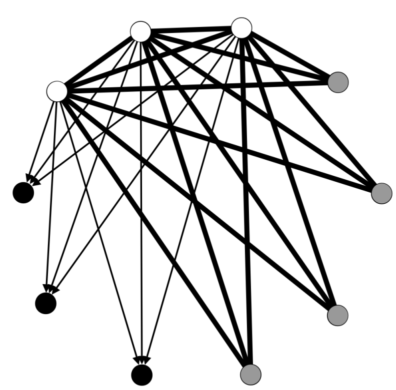


Figure 1: Strategy profile at the end of second period

2.2.3 Third period

In the third period, first note that agents in A and F do not switch to withdraw cooperation from any interaction, by lemma 1 and 2 respectively. Further, any agent in F *may* now have incentive to switch to offer cooperation to other agents in F, since it can potentially increase utility by enhancing trustworthiness in interactions with agents in A. Note that extending cooperation to agents in D can not be optimal for agents in F, because it yields neither direct benefits nor benefits of reputation enhancement (since $c_i > z$ and these agents are not connected in $\alpha - \alpha$ ties to agents in A at the end of Period 2)

On the other hand, any agent in D whose high cost of cooperation did not permit her to reciprocate cooperation to any altruists in the second period *might* now consider the possibility of doing so in combination with extending cooperation to agents in F, since it enhances her reputation and may make the deviation worthwhile. However, they have no incentive to extend cooperation to other agents in D. Let us consider the choices faced by agents in F and D.

Suppose $i \in F$ unilaterally deviates in the third period to extend α to m agents in $F - \{i\}$ $(0 \leq m \leq f - 1)$. Then the expected change in utility is

$\triangle u_i = -mc_i + mz + axm = m(ax - c_i + z)$

Whether or not this is positive depends on the relation between c_i and $ax + z$. If this change in utility is positive for any agent i, then she extends cooperation to all other agents in F in the third period; otherwise she refrains from cooperation with other agents in F. Thus we have a decision making criterion for agents in F in the third period. Various possibilities arising from this are discussed subsequently. Now we consider agents in D.

Suppose $i \in D$ unilaterally deviates in the third period by switching to reciprocate α to k agents in A $(0 \leq k \leq a)$ and switching to offer α to m agents in F $(0 \leq m \leq f)$. Then the expected change in utility is

$\triangle u_i = k(k + m)x - k(c_i + y) + m(z - c_i)$

Let (k^*, m^*) be the values of k and m that maximise $\triangle u_i$. Note that this maximised value must be non negative, since $\triangle u_i = 0$ for $k = 0$ and $m = 0$.

Claim 1: $(\forall i \in D)$[Either $(k^*, m^*) = (0, 0)$ or $(k^*, m^*) = (a, f)$].

Proof: If $k^* = 0$ and $m^* \geq 1$, then $\triangle u_i = -m^*(c_i - z) < 0$ (since $c_i > z$)

Thus $k^* = 0 \longrightarrow m^* = 0$ and $m^* \geq 1 \longrightarrow k^* \geq 1$

If $m^* = 0$ and $k^* \geq 1$, then $\triangle u_i \geq 0 \longrightarrow (k^*)^2 x - k^*(c_i + y) \geq 0$

$\longrightarrow k^* x - y \geq c_i$

Now, $a \geq k^* \longrightarrow ax - y \geq k^* x - y$

$\longrightarrow ax - y \geq c_i$. This contradicts with $(\forall i \in D)(c_i > ax - y)$.

This implies that $\triangle u_i|_{m^*=0, k^* \geq 1} < 0$

Thus $m^* = 0 \longrightarrow k^* = 0$ and $k^* \geq 1 \longrightarrow m^* \geq 1$

Hence $m^* = 0 \longleftrightarrow k^* = 0$ and $k^* \geq 1 \longleftrightarrow m^* \geq 1$ $\quad \ldots (v)$

Suppose $m^* \geq 1 \wedge k^* < a$
Then, $\triangle u_i(k^*, m^*) \geq 0 \longrightarrow k^*(k^* + m^*)x - k^*(c_i + y) + m^*(z - c_i) \geq 0$
$\rightarrow (k^* + m^*)x - \frac{m^*}{k^*}(c_i - z) \geq (c_i + y)$ $\ldots(vi)$
Also, $\triangle u_i(k^*, m^*) \geq \triangle u_i(a, m^*)$ $\quad((k^*, m^*)$ maximises $\triangle u_i$ and $k^* < a)$
$\rightarrow k^*(k^* + m^*)x - k^*(c_i + y) + m^*(z - c_i) - a(a + m^*)x + a(c_i + y) - m^*(z - c_i) \geq 0$
$\rightarrow k^*(k^* + m^*)x - a(a + m^*)x + (a - k^*)(c_i + y) \geq 0$
$\rightarrow x[(k^* - a)(a + m^* + k^*)] \geq (k^* - a)(c_i + y)$
Now, $(k^* - a) \leq 0 \rightarrow x[(a + m^* + k^*)] \leq (c_i + y)$ $\ldots(vii)$

From (vi) and (vii), we have:
$x[(a + m^* + k^*) \leq (k^* + m^*)x - \frac{m^*}{k^*}(c_i - z)$
$\rightarrow ax + \frac{m^*}{k^*}(c_i - z) \leq 0$. This contradicts with $ax > 0 \wedge c_i > z$.

Hence $m^* \geq 1 \longrightarrow k^* = a$ $\ldots(viii)$

Suppose $k^* = a$ and $m^* < f$
Then, by definition we have $\triangle u_i(a, m^*) \geq \triangle u_i(a, f)$
$\rightarrow a(a + m^*)x - a(c_i + y) + m^*(z - c_i) - a(a + f)x + a(c_i + y) - f(z - c_i) \geq 0$
$\rightarrow a(m^* - f)x + (m^* - f)(z - c_i) \geq 0$
$\rightarrow (m^* - f)(ax + z - c_i) \geq 0$
Since $m^* < f$, this implies that $(ax + z - c_i) \leq 0$. $\ldots(ix)$

Also, $\triangle u_i(k^* = a, m^*) > 0 \rightarrow a(a + m^*)x - a(c_i + y) + m^*(z - c_i) \geq 0$
$\rightarrow a^2x - am^*x - ac_i - ay + m^*z - m^*c_i \geq 0$
$\rightarrow a(ax - c_i - y) + m^*(ax + z - c_i) \geq 0$
From (ix), $m^*(ax + z - c_i) \leq 0$. Thus $a(ax - c_i - y) \geq 0$.
$\rightarrow ax - y \geq c_i$. This contradicts with $(\forall i \in D)(c_i > ax - y)$.
Hence $k^* = a \longrightarrow m^* = f$ $\ldots(x)$

From $(v), (viii)$ and (x) we can deduce:
$(\forall i \in D)[$ Either $(k^*, m^*) = (0, 0)$ or $(k^*, m^*) = (a, f)]$.

Claim 2: $\big(\forall \mathbf{i} \in \mathbf{D}\big)\big((\mathbf{k}^*, \mathbf{m}^*) = (\mathbf{a}, \mathbf{f})$ **if and only if** $\mathbf{a}[\mathbf{ax} - \mathbf{c_i} - \mathbf{y}] + \mathbf{f}[\mathbf{ax} + \mathbf{z} - \mathbf{c_i}] \geq \mathbf{0}\big)$

Proof: If $a[ax - c_i - y] + f[ax + z - c_i] < 0$, then the change in utility for agent i from choosing $k = a$ and $m = f$ is negative, thus $(k^*, m^*) \neq (a, f)$. Further, if $(k^*, m^*) \neq (a, f)$, then it must be the case that $(k^*, m^*) = (0, 0)$ and $\triangle u_i(a, f) < \triangle u_i(0, 0)$. This implies that $a[ax - c_i - y] + f[ax + z - c_i] < 0$.

Note that by assumption, $(\forall i \in D)(a[ax - c_i - y] < 0)$. The possibility of an optimal deviation of $(k^*, m^*) \neq (0, 0)$ is made feasible only when $z > -y$ such that $c_i < ax + z$ and hence $f[ax + z - c_i] > 0$.

Thus we have a decision making criterion for agents in D in the third period. Clearly, the analysis of the evolution process from third period onwards becomes contingent on various cost conditions. Broadly, we divide all possible cases into two categories, when $z \leq -y$ and when $z > -y$.

2.2.4 When $\mathbf{z} \leq -\mathbf{y}$

Case 1: When $z \leq -y$ and $(\forall i \in F)(ax + z < c_i)$

First note that for any agent $i \in D$, $z \leq -y \longrightarrow ax + z \leq ax - y < c_i$. This implies that $f[ax + z - c_i] < 0$, which further implies that $a[ax - c_i - y] + f[ax + z - c_i] < 0$. Thus when $z \leq -y$, the only optimal deviation for agents in D in Period 3 is given by $(k^*, m^*) = (0, 0)$, that is, agents in D do not deviate from the third period and we have $(\forall i \in D)(\forall j \in N - \{i\})(a^3_{ij} = \beta)$.

Secondly, for any agent $i \in F$ we have $ax+z < c_i$, which means that if i deviates to extend cooperation to m other agents in F, the expected change in utility $\triangle u_i = m(ax - c_i + z)$ is negative for all non zero values of m and is maximised by $m^* = 0$. Thus agents in F also do not deviate in Period 3, and we have $\big(\forall i \in F\big)\big[(\forall j \in A)(a^3_{ij} = \alpha) \wedge (\forall j \in D \cup F - \{i\})(a^3_{ij} = \beta)\big]$. Thus we reach an equilibrium after the second period itself, and we have the following proposition.

Proposition 1 When $z \leq -y$ and $(\forall i \in F)(ax + z < c_i)$, the system converges to an equilibrium strategy profile s^* where:

$(\forall i \in A)(\forall j \in N - \{i\})(a^*_{ij} = \alpha)$

$(\forall i \in F)[(\forall j \in A)(a^*_{ij} = \alpha) \wedge (\forall j \in N - A - \{i\})(a^*_{ij} = \beta)]$

$(\forall i \in D)(\forall j \in N - \{i\})(a^*_{ij} = \beta)$

Denote this equilibrium as e_1.

Case 2: When $z \leq -y$, $(\exists i \in F)(ax + z < c_i)$ and $(\exists j \in F)(ax + z \geq c_j)$

Optimal behaviour of agents in D remains the same as in Case 1, which is that they continue to refrain from cooperating with everyone.

Consider $i \in F$. We saw that if i offers α to m agents in $F - \{i\}$ $(0 \leq m \leq f - 1)$, then $\triangle u_i = m(ax - c_i + z)$. Now, $(ax - c_i + z) \geq 0 \longrightarrow m^* = f - 1$ and $(ax - c_i + z) < 0 \longrightarrow m^* = 0$.

Let $F_1 \subset F = \{i \in F | (ax - c_i + z) \geq 0\}$ and $F_2 \subset F = \{j \in F | (ax - c_j + z) < 0\}$. Then we have the following at the end of third period.

$\big(\forall i \in F_1\big)\big[(\forall j \in A \cup F - \{i\})(a^3_{ij} = \alpha) \wedge (\forall j \in D)(a^3_{ij} = \beta)\big]$
$\big(\forall i \in F_2\big)\big[(\forall j \in A)(a^3_{ij} = \alpha) \wedge (\forall j \in N - A - \{i\})(a^3_{ij} = \beta)\big]$

Figure 2 displays a possible strategy profile at the end of the third period.

Now, in the fourth period the incentive structures for $i \in A \cup D$ remain unchanged in this case and they do not deviate from their third period strategies. For agents in F, since agents in D are still unconnected with $\alpha - \alpha$ ties to agents in A, their best response action towards them also remains unchanged (no direct or reputation related utility gain possible).

First consider $i \in F_1$. By lemma 2, she has no incentive to change their actions with respect to $j \in A \cup F_1$ in the fourth period. If i switches to offer β to $0 \leq k \leq f_2$ agents in F_2, expected change in utility $\triangle u_i = kc_i - kz - akx = -k(ax - c_i + z)$ is negative for any $k \geq 0$.

Thus agents in F_1 do not deviate in the fourth period.

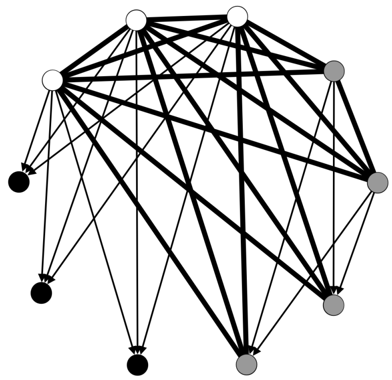


Figure 2: Third period when $z \leq -y$, $(\exists i \in F)(ax + z < c_i)$ and $(\exists j \in F)(ax + z \geq c_j)$

Now consider $i \in F_2$. By lemma 2, i's best response actions towards agents in A remain unchanged in the fourth period. An optimal unilateral deviation may involve (i) reciprocating α to $0 \leq k \leq f_1$ agents in F_1 and (ii) offering α to $0 \leq m \leq f_2 - 1$ agents in F_2, giving expected change in utility

$\triangle u_i = -kc_i - mc_i - ky + mz + kx(a + k) + akx = k[x(2a + k - c_i - y)] + m(z - c_i)$

Now, $c_i > z \longrightarrow m^* = 0$ and $c_i < ax - y \longrightarrow c_i < 2ax + kx - y \longrightarrow k^* = |F_1| = f_1$

Thus $\big(\forall i \in F_2\big)\big[(\forall j \in A \cup F_1)(a^4_{ij} = \alpha) \wedge (\forall j \in D \cup F_2)(a^4_{ij} = \beta)\big]$

In the fifth period as already argued, agents in $A \cup D \cup F_1$ do not deviate from their strategies from the previous period. Consider $i \in F_2$. By lemma 2, i's best response actions towards agents in $A \cup F_1$ remain unchanged. Her actions towards $j \in D$ also remain unchanged, as previously argued. An optimal unilateral deviation may involve offering α to $0 \leq k \leq f_2 - 1$ agents in $F_2 - \{i\}$, giving change in utility

$\triangle u_i = -kc_i + kz + kx(a + f_1) = k(x(a + f_1) - c_i + z)$

Now, $((a + f_1)x - c_i + z) \geq 0 \longrightarrow k^* = f_2 - 1$ and $((a + f_1)x - c_i + z) < 0 \longrightarrow k^* = 0$.

Let $F_{21} \subset F_2 = \{i \in F_2 | ((a+f_1)x - c_i + z) \geq 0\}$ and $F_{22} \subset F_2 = \{i \in F_2 | ((a + f_1)x - c_i + z) < 0\}$. Then we have:

$\big(\forall i \in F_{21}\big)\big[(\forall j \in A \cup F - \{i\})(a^5_{ij} = \alpha) \wedge (\forall j \in D)(a^5_{ij} = \beta)\big]$

$\big(\forall i \in F_{22}\big)\big[(\forall j \in A \cup F_1)(a^5_{ij} = \alpha) \wedge (\forall j \in D \cup F_2)(a^5_{ij} = \beta)\big]$

The argument behind this process is as follows. In this case there *are* some agents in F for whom cost conditions are such that they benefit from unilaterally extending cooperation to *all* other agents in F, because of the reputation related utility that it brings them by raising their trustworthiness in their interactions of mutual cooperation with agents in A. When these agents (whose cost conditions allow them to invest in reputation building) offer links to all other agents in F in the third period, they end up with links of mutual cooperation with each other and unilateral cooperation with remaining agents in F. In the fourth period, these links are reciprocated and now some of the remaining agents may also find it profitable to extend links of cooperation to each other since the benefits of reputation building accrue on not only links of mutual cooperation with agents in A, but also some links within F. This process continues such that in every period, more agents in F become connected in $\alpha - \alpha$ ties with each other. In fact, this process may converge to an equilibrium where *all* agents in F mutually cooperate with each other (equilibrium e_3 defined below). It may also, however, converge to an equilibrium where some agents in F remain uncooperative towards each other. In both cases, agents in D have no incentives to extend any links of cooperation (equilibrium e_2 defined below).

Suppose this process converges to an equilibrium such that there are still some agents in F who are not connected in $\alpha - \alpha$ ties with each other and we have the following equilibrium.

$(\forall i \in A)(\forall j \in N - \{i\})(a^*_{ij} = \alpha)$
$(\forall i \in D)(\forall j \in N - \{i\})(a^*_{ij} = \beta)$
$F_k = \{i_1, i_2, i_3, ...i_k\} \subset F$ is such that:
$(\forall i \in F - F_k)[(\forall j \in A \cup F - \{i\})(a^*_{ij} = \alpha) \wedge (\forall j \in D)(a^*_{ij} = \beta)]$
$(\forall i \in F_k)[(\forall j \in A \cup F - F_k)(a^*_{ij} = \alpha) \wedge (\forall j \in F_k \cup D - \{i\})(a^*_{ij} = \beta)]$

Since this strategy profile is an equilibrium, agents in F_k must have no incentive to engage in any further cooperative activity. Suppose $i \in F_k$ deviates to offer α to m agents in $F_k - \{i\}$, then the expected change in utility from this deviation should be negative:

$\triangle u_i = -mc_i + mz + mx(a + f - k) < 0 \rightarrow x(a + f - k) + z < c_i$
$\rightarrow (\forall i \in F_k)(x(a + f - k) + z < c_i)$

Alternatively, this process may converge to an equilibrium where all agents in F are linked in ties of mutual cooperation to each other. Note that through the entire process, agents in F do not extend cooperation to agents in D and agents in D continue refraining from cooperation with respect to everyone else. Thus we have the following proposition.

Proposition 2 When $z \leq -y$, $(\exists i \in F)(ax + z < c_i)$ and $(\exists j \in F)(ax + z \geq c_j)$, the system converges to either of the following equilibria.

1) Suppose there exists an integer $m < f$ and a subset $F_m = \{i_1, i_2, i_3, \ldots i_m\} \subset F$ such that $(\forall i \in F_m)(x(a + f - m) + z < c_i)$. Let k be the largest possible value of this integer and $F_k \subset F = \{i_1, i_2, i_3, \ldots i_k\}$ be the largest possible set of this kind. Under these conditions the system converges to the following equilibrium s^* :

$(\forall i \in A)(\forall j \in N - \{i\})(a^*_{ij} = \alpha)$

$(\forall i \in F - F_k)[(\forall j \in A \cup F - \{i\})(a^*_{ij} = \alpha) \wedge (\forall j \in D)(a^*_{ij} = \beta)]$

$(\forall i \in F_k)[(\forall j \in A \cup F - F_k)(a^*_{ij} = \alpha) \wedge (\forall j \in F_k \cup D - \{i\})(a^*_{ij} = \beta)]$

$(\forall i \in D)(\forall j \in N - \{i\})(a^*_{ij} = \beta)$

(Denote this equilibrium as e_2)

Or

2) If for any integer $m < f$ and subset $F_m = \{i_1, i_2, i_3, \ldots i_m\} \subset F$ we have $(\forall i \in F_m)(x(a + f - m) + z \geq c_i)$, then the system converges to a Nash equilibrium s^* where:

$(\forall i \in A)(\forall j \in N - \{i\})(a^*_{ij} = \alpha)$

$(\forall i \in F)[(\forall j \in A \cup F - \{i\})(a^*_{ij} = \alpha) \wedge (\forall j \in D)(a^*_{ij} = \beta)]$

$(\forall i \in D)(\forall j \in N - \{i\})(a^*_{ij} = \beta)$

(Denote this equilibrium as e_3)

Case 3: When $z \leq -y$ and $(\forall i \in F)(ax + z \geq c_i)$.

In this case as discussed before, *all* agents in F have incentive to extend cooperation to each other in the third period, since the added benefit from reputation enhancement and increased trustworthiness in their interactions with altruists more than justifies the cost. Incentives towards cooperation with agents in D remain unchanged, and we have the following proposition.

Proposition 3 When $z \leq -y$ and $(\forall i \in F)(ax + z \geq c_i)$, the system converges to an equilibrium strategy profile identical to equilibrium e_3.

2.2.5 When $\mathbf{z > -y}$

$z > -y$ implies the following. Firstly, for any agent $i \in F$, $z > -y$ and $c_i < ax - y$ implies that $c_i < ax + z$. This means that all agents in F have incentive to extend cooperation to each other in the third period, since extending cooperation to m other agents in F gives change in utility $m(ax - c_i + z)$.

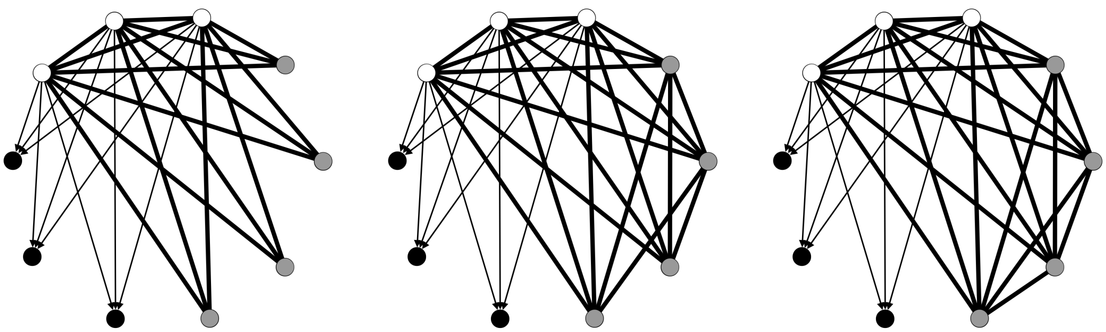


Figure 3: Three kinds of equilibria that may emerge under $z \leq -y$

Secondly, recall that in the third period if agent $i \in D$ unilaterally deviates by switching to reciprocate α to k agents in A ($0 \leq k \leq a$) and switching to offer α to m agents in F ($0 \leq m \leq f$). Then the expected change in utility is

$\triangle u_i = k(k+m)x - k(c_i + y) + m(z - c_i)$

which is maximised by (k^*, m^*), where

$(k^*, m^*) = (0, 0)$ if and only if $a[ax - c_i - y] + f[ax + z - c_i] < 0$, and

$(k^*, m^*) = (a, f)$ if and only if $a[ax - c_i - y] + f[ax + z - c_i] \geq 0$.

We know that for agents in D, $c_i > ax - y$ implies that the first term on the right hand side of above expression is always negative. However, $z > -y$ opens up the possibility that the second term $f[ax + z - c_i]$ may be positive, and even larger in magnitude than the first term. Thus we have the following cases.

Case 4: When $z > -y$ and $(\forall i \in D)\big[a(ax - c_i - y) + f(ax + z - c_i) < 0\big]$

In this case as discussed before, agents in F link with each other in ties of mutual cooperation in the third period. However, agents in D *still do not have incentive to engage in any cooperative activity.* Thus we have the following proposition.

Proposition 4 When $z > -y$ and $(\forall i \in D)\big[a(ax - c_i - y) + f(ax + z - c_i) < 0\big]$, the system converges to a strategy profile identical to equilibrium e_3.

Although the equilibrium converged to is the same in this case as in the previous case, the incentive mechanism driving and determining the actions of agents in D is subtly different. Note that $z > -y$ implies for agents in D, while in the previous case (Case 3) it is not *possible* to benefit from reciprocating cooperation to altruists, in this case it is possible but restricted by different conditions. In Case 3, $z \leq -y$ prohibits any reputation building for these agents. In this case, even though $z > -y$, it is the unavailability of *enough* opportunities that prohibits reputation building. A larger f, in some sense, would have changed the situation and allowed these agents to reciprocate cooperation to altruists as well as extend unilateral cooperation to agents in F. This constraint is eased for some agents in D in the next case examined.

Case 5: When $z > -y$, $(\exists i \in D)\big[a(ax - c_i - y) + f(ax + z - c_i) < 0\big]$

and $(\exists j \in D)\big[a(ax - c_j - y) + f(ax + z - c_j) \geq 0\big]$

To analyse this case, consider the following partition of D.

$D_1 = \{i \in D | a[ax - c_i - y] + f[ax + z - c_i] \geq 0\}$ $\qquad |D_1| = d_1$

$D_2 = \{i \in D | a[ax - c_i - y] + f[ax + z - c_i] < 0\}$ $\qquad |D_2| = d_2$

As already argued, agents in $A \cup F$ are all linked to each other in ties of mutual cooperation at the end of the third period (since $z > -y$) and given above partitioning based on costs, agents in D_1 have incentive to reciprocate cooperation to *all altruists* while extending unilateral cooperation to *all agents in F*. Thus we have

$(\forall i \in D_1)[(\forall j \in A \cup F)(a^3_{ij} = \alpha) \wedge (\forall j \in D - \{i\})(a^3_{ij} = \beta)$

$(\forall i \in D_2)(\forall j \in N - \{i\})(a^3_{ij} = \beta)$

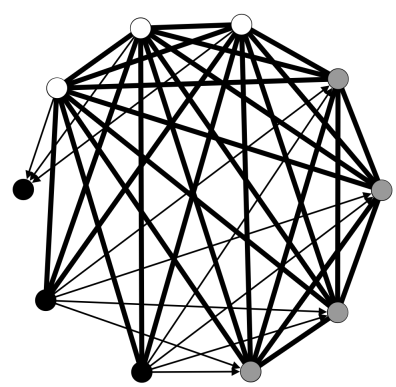


Figure 4: A possible strategy profile at the end of the third period

In the fourth period, agents in $A \cup F$ have no incentive to deviate (by lemma 1 and 2) to withdraw cooperation in any interaction. Further, agents in F may consider reciprocating cooperation to agents in D_1. If $i \in F$ reciprocates α to $0 \leq k \leq d_1$ agents in D_1, the expected change in utility is

$\triangle u_i = -kc_i - ky + kx(a) = k(ax - c_i - y) > 0$ for all $k \geq 1$.

Thus $k^* = d_1$ maximises this change in utility and agents in F reciprocate cooperation to *all agents in* D_1.

Now consider $i \in D_1$. In the fourth period, these agents have no incentives to offer α to agents in D_2. Suppose i deviates to (i) offer β to $0 \leq k \leq a$ agents in A, (ii) offer β to $0 \leq l \leq f$ agents in F, and (iii) offer α to $0 \leq m \leq d_1 - 1$ agents in $D_1 - \{i\}$. The change in utility expected is

$\triangle u_i = -kx(a + f) + x(a - k)(m - l - k) + (k + l - m)c_i + ky + (m - l)z$

Let k^*, l^* and m^* be the values of k, l and m that maximise $\triangle u_i$.

Since i is maximising utility by offering α all agents in $A \cup F$ at the end of the third period, i has no incentive to switch to offering β to any agent in $A \cup F$ *in isolation*.

Then, $m^* = 0 \longrightarrow k^* = 0$ and $l^* = 0$ $\qquad \ldots(xi)$

Secondly, $\dfrac{\triangle(\triangle u_i)}{\triangle m} = \dfrac{-\triangle(\triangle u_i)}{\triangle l} = x(a - k) - c_i + z$

Now, $k^* = 0 \rightarrow \dfrac{\triangle(\triangle u_i)}{\triangle m} => 0$ and $\dfrac{\triangle(\triangle u_i)}{\triangle l} < 0$ $\qquad$ (since $ax - c_i + z > 0$)

Thus $k^* = 0 \longrightarrow m^* = d_1 - 1$ and $l^* = 0$ $\quad \ldots(xii)$

From (xi), we have $k^* \geq 1 \longrightarrow m^* \geq 1$

$m^* \geq 1 \longrightarrow \dfrac{\triangle(\triangle u_i)}{\triangle m} > 0$ and $\dfrac{\triangle(\triangle u_i)}{\triangle l} < 0$

$\dfrac{\triangle(\triangle u_i)}{\triangle l} < 0 \rightarrow l^* = 0$

Then, $\triangle u_i = -k^* x(a+f) + x(a-k^*)(m^*-k^*) + (k^*-m^*)c_i + k^* y + m^* z$

Substituting $f + m^* = \mu$ and $a - k^* = \kappa$ gives:

$\triangle u_i = \kappa x(\kappa+\mu) - c_i(\kappa+\mu) + \mu z - \kappa y + \{c_i(a+f) - fz + ay - ax(a+f)\}$

where the last term of expression is a constant. The non-constant part of the above expression is maximised by $\kappa = a$ and $\mu = f$ (this is the same expression as $\triangle u_i$ being considered in the third period for agents in D)

$a - k^* = \kappa = a \longrightarrow k^* = 0$. This is a contradiction.

Hence $k^* \geq 1$ leads to a contradiction and agents in D_1 must not withdraw cooperation from any altruist. From (xii), this implies that $k^* = 0, l^* = 0$ and $m^* = d_1 - 1$ maximise the expected increase in utility for i. That is, $i \in D_1$ remains cooperating with agents in F, *and* extends cooperation to other agents in D_1. Further, this expected increase in utility is positive:

$\triangle u_i = ax(d_1-1) - (d_1-1)c_i + (d_1-1)z = (d_1-1)[ax - c_i + z] > 0$

Thus we have $(\forall i \in D_1)[(\forall j \in N - D_2 - \{i\})(a_{ij}^4 = \alpha) \wedge (\forall j \in D_2)(a_{ij}^4 = \beta)$

Now consider $i \in D_2$. In the fourth period, these agents do not have incentives to offer α to other agents in D_2. If $i \in D_2$ reciprocates α to k agents in A and offers α to m agents in $F \cup D_1$:

$\triangle u_i = -(k+m)c_i - ky + mz + k(a+m)x = k(ax - c_i - y) + m(kx - c_i + z)$

where $ax - c_i - y < 0$

Now, $m^* = 0 \rightarrow \triangle u_i|_{k^* \geq 1} < 0 \rightarrow k^* = 0$

Further, $k^* = 0 \rightarrow \triangle u_i|_{m^* \geq 1} < 0 \rightarrow m^* = 0$. Thus $k^* = 0 \leftrightarrow m^* = 0$ $\quad \ldots(xiii)$

Suppose $m^* \geq 1$ and $k^* < a$

$\rightarrow \triangle u_i|_{m^*,k^*} \geq 0$ and $\triangle u_i|_{m^*,k^*} \geq \triangle u_i|_{m^*,a}$

$\triangle u_i|_{m^*,k^*} \geq 0 \rightarrow c_i \leq \dfrac{k^* ax + k^* m^* x + m^* z - ky}{k^* + m^*}$

$\triangle u_i|_{m^*,k^*} \geq \triangle u_i|_{m^*,a} \rightarrow c_i \geq x(m^*+a) - y$

$\longrightarrow k^* x(a+m^*) + m^* z - k^* y \geq x(a+m^*)(k^*+m^*) - y(k^*+m^*)$

$\rightarrow z \geq x(a+m^*) - y$. This is a contradiction since $z < 0$ and $x(a+m^*) - y > 0$.

Thus $m^* \geq 1 \rightarrow k^* = a$ $\quad \ldots(xiv)$

Now, $k^* = a \longrightarrow \triangle u_i(k^*, m) = a(ax - c_i - y) + m(ax - c_i + z)$

$\dfrac{\triangle(\triangle u_i(k^*, m))}{\triangle m} = ax - c_i + z > 0 \longrightarrow m^* = d_1 + f$

Thus $k^* = a \longleftrightarrow m^* = d_1 + f$ $\quad \ldots(xv)$

From $(xiii)$ and (xv) we can deduce that either $(k^*, m^*) = (a, d_1+f)$ or $(k^*, m^*) = (0, 0)$.

Further, $(k^*, m^*) = (a, d_1+f)$ if and only if $a(ax - c_i - y) + (f + d_1)(ax - c_i + z) \geq 0$

Consider the following partition of D_2

$D_{21} = \{i \in D_2 | a[ax - c_i - y] + (f + d_1)[ax + z - c_i] \geq 0\}$
$D_{22} = \{i \in D_2 | a[ax - c_i - y] + (f + d_1)[ax + z - c_i] < 0\}$

Then, agents in D_{21} reciprocate cooperation to altruists in A while extending cooperation to all agents in $F \cup D_1$, while agents in D_{22} continue refraining from cooperation.
$(\forall i \in D_{21})[(\forall j \in A \cup F \cup D_1)(a^4_{ij} = \alpha) \wedge (\forall j \in D_2 - \{i\})(a^4_{ij} = \beta)$
$(\forall i \in D_{22}(\forall j \in N - \{i\})(a^4_{ij} = \beta)$

The argument behind best response dynamics in the fourth period is as follows. At the end of the third period, we find that agents $i \in D_1$ reciprocate α to all agents in A and extend α to all agents in F in the third period, while all $j \in D_2$ refrain from any cooperation. In the fourth period, all agents in F reciprocate α to all agents in D_1, all $i \in D_1$ extend α to each other, and some agents in D_2 have incentives to reciprocate α to all agents in A while unilaterally extending α to all agents in F and D_1. Call this set of agents D_{21}. Then, we know that in the fifth period agents in F will reciprocate α to these agents in D_{21} and agents in D_{21} will have incentive to extend α to each other. Thus, in every period more and more agents in D find incentives to enter ties of cooperation by reciprocating cooperation to the altruists along with extending cooperation to not only agents in F but also the growing set of agents from D who are linked to everyone in $A \cup F$ and to each other. This process may converge to an equilibrium with a complete network of mutual cooperation ties, or it may converge to an equilibrium where some agents in D continue to refrain from cooperation.

Suppose this process converges to an equilibrium where:
$(\forall i \in A)(\forall j \in N - \{i\})(a^*_{ij} = \alpha)$

$D_k = \{i_1, i_2, i_3, \ldots i_k\} \subset D$ is such that:
$(\forall i \in D - D_k)[(\forall j \in N - D_k - \{i\})(a^*_{ij} = \alpha) \wedge (\forall j \in D_k)(a^*_{ij} = \beta)]$
$(\forall i \in D_k)(\forall j \in N - \{i\})(a^*_{ij} = \beta)$

$(\forall i \in F)[(\forall j \in N - D_k - \{i\})(a^*_{ij} = \alpha) \wedge (\forall j \in D_k)(a^*_{ij} = \beta)]$

Then, it must be the case that agents in D_k have no incentives to simultaneously reciprocate α to agents in A while extending α to agents in $F \cup D - D_k$. This implies:
$-(a + f + d - k)c_i + (f + d - k)z - ay + ax(a + f + d - k) < 0$
$\rightarrow a(n - k)x - ay + (f + d - k)z < (n - k)c_i$
$\rightarrow ax + z - \dfrac{a(y + z)}{n - k} < c_i$

Alternatively, the process may converge to a complete equilibrium. Thus we have the following proposition.

Proposition 5 When $z > -y$, $(\exists i \in D)\big[a(ax - c_i - y) + f(ax + z - c_i) < 0\big]$ and $(\exists j \in D)\big[a(ax - c_j - y) + f(ax + z - c_j) \geq 0\big]$, the system may converge to either of the following equilibria:

1) Suppose there exists an integer $m < n - 1$ and a subset $D_m = \{i_1, i_2, i_3 \ldots i_m\} \subset D$ such that $(\forall i \in D_m)(ax + z - \dfrac{a(y+z)}{n-m} < c_i)$. Let k be the largest possible value of this integer and $D_k = \{i_1, i_2, i_3 \ldots i_k\} \subset D$ be the largest possible set of this kind. Under these conditions, the system converges to the following equilibrium s^* where:

$(\forall i \in A)(\forall j \in N - \{i\})(a^*_{ij} = \alpha)$

$(\forall i \in D - D_k)[(\forall j \in N - D_k - \{i\})(a^*_{ij} = \alpha) \wedge (\forall j \in D_k)(a^*_{ij} = \beta)]$

$(\forall i \in D_k)(\forall j \in N - \{i\})(a^*_{ij} = \beta)$

$(\forall i \in F)[(\forall j \in N - D_k - \{i\})(a^*_{ij} = \alpha) \wedge (\forall j \in D_k)(a^*_{ij} = \beta)]$

(Denote this equilibrium as e_4)

Or

2) If for any integer $m < n - 1$ and subset $D_m = \{i_1, i_2, i_3, \ldots i_m\} \subset D$ we have $(\forall i \in D_m)(ax + z - \dfrac{a(y+z)}{n-m} \geq c_i)$, then the system converges to a Nash equilibrium s^* where $(\forall i \in N)(\forall j \in N - \{i\})(a^*_{ij} = \alpha)$

(Denote this equilibrium as e_5)

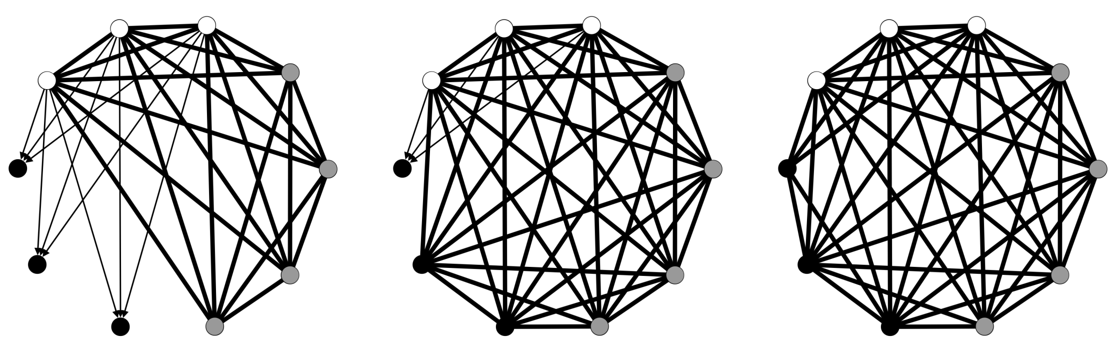


Figure 5: Three kinds of equilibria that may emerge under $z > -y$

Case 6: When $z > -y$ and $(\forall i \in D)\big[a(ax - c_i - y) + f(ax + z - c_i) \geq 0\big]$

From the arguments made above, we have the following strategy profile at the end of third period:

$(\forall i \in A)(\forall j \in N - \{i\})(a^3_{ij} = \alpha)$

$(\forall i \in F)\big[(\forall j \in A \cup (F - \{i\}))(a^3_{ij} = \alpha) \wedge (\forall j \in D)(a^3_{ij} = \beta)\big]$

$(\forall i \in D)[(\forall j \in A \cup F)(a^3_{ij} = \alpha) \wedge (\forall j \in D - \{i\})(a^3_{ij} = \beta)$

In the fourth period, consider $i \in F$. By lemma 2, i does not withdraw cooperation from any agent in $A \cup F - \{i\}$. An optimal unilateral deviation can only involve i switching to reciprocate α to $0 \leq k \leq d$ agents in D. This yields change in utility

$\triangle u_i = -kc_i - ky + k(a+1)x + akx = k[(2a+1)x - c_i - y]$.

Since $c_i < ax - y < (2a+1)x - y$, this is positive for all $k \geq 1$ and maximised by $k^* = d$

Thus agents in F reciprocate cooperation to all agents in D and we have

$(\forall i \in F)(\forall j \in N - \{i\})(a^4_{ij} = \alpha)$

Consider $i \in D$ in the fourth period. Suppose i deviates to (i) offer β to $0 \leq k \leq a$ agents in A, (ii) offer β to $0 \leq l \leq f$ agents in F, and (iii) offer α to $0 \leq m \leq d-1$ agents in $D - \{i\}$. The change in utility expected is:

$\triangle u_i = -kx(a+f) + x(a-k)(m-l-k) + (k+l-m)c_i + ky + (m-l)z$

First we note that since i is maximising utility by offering α all agents in $A \cup F$ at the end of the third period, i has no incentive to switch to offering β to any agent in $A \cup F$ *in isolation*, that is: $m^* = 0 \longrightarrow k^* = 0$ and $l^* = 0$ $\quad \ldots (xvi)$

Secondly, $\dfrac{\triangle(\triangle u_i)}{\triangle m} = -\dfrac{\triangle(\triangle u_i)}{\triangle l} = x(a-k) - c_i + z$

Now, $k^* = 0 \longrightarrow \dfrac{\triangle(\triangle u_i)}{\triangle m} > 0$ and $\dfrac{\triangle(\triangle u_i)}{\triangle l} < 0$ (since $ax - c_i + z > 0$)

Thus $k^* = 0 \longrightarrow m^* = d_1 - 1$ and $l^* = 0$ $\quad \ldots (xvii)$

From (xvi), we have $k^* \geq 1 \longrightarrow m^* \geq 1$

$m^* \geq 1 \longrightarrow \dfrac{\triangle(\triangle u_i)}{\triangle m} > 0$ and $\dfrac{\triangle(\triangle u_i)}{\triangle l} < 0$

$\dfrac{\triangle(\triangle u_i)}{\triangle l} < 0 \longrightarrow l^* = 0$

Then, $\triangle u_i = -k^*x(a+f) + x(a-k^*)(m^*-k^*) + (k^*-m^*)c_i + k^*y + m^*z$

Substituting $f + m^* = \mu$ and $a - k^* = \kappa$ gives:

$\triangle u_i = \kappa x(\kappa+\mu) - c_i(\kappa+\mu) + \mu z - \kappa y + \{c_i(a+f) - fz + ay - ax(a+f)\}$

where the last term of the expression is a constant and the non-constant part of the expression is maximised by $\kappa = a$ and $\mu = f$ (this is the same expression as $\triangle u_i$ being considered in the third period for agents in D)

$a - k^* = \kappa = a \rightarrow k^* = 0$. This is a contradiction.

Hence $k^* \geq 1$ leads to a contradiction and from $(xvii)$, $k^* = 0, l^* = 0$ and $m^* = d - 1$ maximise the expected increase in utility for i. Further, this expected increase in utility is positive:

$\triangle u_i = ax(d-1) - (d-1)c_i + (d-1)z = (d-1)[ax - c_i + z] > 0$

That is, agents in D do not withdraw cooperation from any agent in $A \cup F$ and instead extend cooperation to other agents in D as well, and we have

$(\forall i \in D)(\forall j \in N - \{i\})(a^4_{ij} = \alpha)$

Thus in this case, we have a complete network of mutual cooperation at the end of the fourth period. Suppose, in the fifth period, an arbitrary agent $i \in N$ considers a unilateral deviation involving switching to offer β to $0 \leq k \leq n-1$ agents. This yields the following expected change in utility:

$\triangle u_i = k(c_i + y) - kx(n-1) - kx(n-1-k) = k[-x(2n-2-k) + c_i + y]$

$\triangle u_i \geq 0 \rightarrow c_i \geq x(2n-2-k) - y > (n-1)x - y$

That is, an agent who deviates must have $c_i > (n-1)x - y$. By assumption, $(\forall i \in F)(c_i < ax - y < (n-1)x - y)$. Further, we have $(\forall i \in D)(c_i < ax + z)$. Suppose $(\exists i \in D)((n-1)x - y < c_i < ax + z)$. This implies that $(n-1-a)x - y < z$. Now, $n-1-a \geq 1$ and $x > y$ means $z > (n-1-a)x - y > 0$. This is a contradiction, since $z < 0$ by assumption.

Hence we must have $(\forall i \in N)(c_i < (n-1)x - y)$, and no agent has incentive to unilaterally deviate, giving us the following proposition.

Proposition 6 When $z > -y$ and $\big(\forall i \in D\big)\big(a[ax - c_i + y] + f[ax + z - c_i] \geq 0\big)$, the system converges to an equilibrium strategy profile s^* where: $(\forall i \in N)(\forall j \in N - \{i\})(a^*_{ij} = \alpha)$

From this evolution analysis, we find that equilibria of varying degrees of cooperation emerge under different conditions on the costs of cooperation of agents and distribution of altruists in the population. The presence of altruists is not only crucial to starting off the process of cooperation, but the number of altruists in the population has a positive relation with the degree of cooperation observed in the equilibrium network that evolves. For instance, a relatively larger number of altruists in the population would mean that many more agents are likely to fall in subgroup F, which is the group of agents that reciprocate to the altruists in the second period. Further, if the number of altruists is large then the likelihood of agents in F extending cooperation to each other in the third period is also higher, since each cooperative link extended to another agent in F enhances trustworthiness in a significant number of cooperative ties (*i.e* mutually cooperative ties with all altruists).

Note that while altruism is essential for any cooperation to evolve from a situation of zero cooperation, the incentive to build reputation also plays an important role in explaining cooperative activity initiated by non altruistic agents. Apart from their interactions with altruistic agents, where reciprocity has a role to play, all other cooperative ties that non altruistic agents form are first initiated to build trustworthiness or reputation to support existing ties of mutual cooperation. Also observe that the symmetric nature of the evolution process in absence of any perturbations ensures that links of cooperation once formed are not broken through the evolution process. This is because a non altruistic agent decides to extend unilateral cooperation to another agent only when the benefits of trustworthiness that it yields are more than the cost to utility. In the next period, this may convert to a link of mutual cooperation through reciprocation by the other interacting partner, or it may remain a link of unilateral cooperation. In both cases, it still serves its purpose of enhancing reputation and continues to justify its cost, *unless the number of common mutually cooperative partners between the two interacting agents falls*. That is, cooperation extended in one period is withdrawn in any future period only if the *embeddedness*[7] of the cooperative

[7] The embeddedness of an edge in a network is the number of common neighbors the two endpoints have.

tie falls, which would happen if *other* cooperative ties were to be withdrawn. Because of the symmetric nature of the process, in absence of exogenous increases in costs of cooperation and in absence of mistakes or errors in rational decision making, there is no independent reason for cooperation to be withdrawn.

However, we can realistically expect a variety of exogenous perturbations in a process of repeated interaction between agents. For instance, some agents might make errors in decision making, while others might experience changes in their cost of cooperation due to external factors. In fact, we can even imagine altruists' attitude to remain altruistic only for a few periods of repeated interaction. This raises the possibility of a wider range of Nash equilibria to emerge from the given incentive mechanism of network formation. In particular we can have a variety of strategy profiles that do not emerge as equilibria from the undisturbed and symmetric process of evolution of cooperation discussed above, but once manifested through alternative processes, sustain as Nash equilibrium. Then, regardless of the process through with equilibria emerge, they will always exhibit some characteristic features. These are discussed in the next section.

3 Nash equilibria

We begin by revisiting certain observations from our analysis of the evolution process.

1. The complete non cooperation network such that $(\forall i, j \in N)(a_{ij} = \beta \wedge a_{ji} = \beta)$ is a Nash equilibrium *if and only if* $(\forall i \in N)(c_i > z)$.

   This means that in absence of altruists, the strategy profile where no one cooperates with anyone else is always a Nash equilibrium. This reminds us of the importance of altruists for starting off a process of cooperation network formation through repeated interactions of agents.

2. A complete network of mutual cooperation where $(\forall i, j \in N)(a_{ij} = \alpha \wedge a_{ji} = \alpha)$ is a Nash equilibrium *if and only if* $(\forall i \in N)(c_i < (n-1)x - y)$

   The interpretation of the condition is simply that from a complete network of mutual cooperation, unilaterally withdrawing cooperation simultaneously from *all* partners must not be profitable to any agent. The necessity of this condition for the complete network to be a Nash equilibrium emerges in our discussion of Case 6 in the previous section. Further, this condition is also sufficient to ensure that the complete network is a Nash equilibrium. To see this, suppose this condition is satisfied and a complete network of mutual cooperation is not a Nash equilibrium. Then there exists an agent i who can make a profitable unilateral deviation to withdraw cooperation from $1 \leq k \leq n-1$ interactions, or, $\triangle u_i = k[c_i + y - x(2(n-1) - k)] > 0$ for some $k \geq 1$. This implies that $c_i + y - x(2(n-1) - k) > 0$ for some $k \geq 1$.

   Now, $k \leq n-1 \longrightarrow -k \geq -(n-1) \longrightarrow 2n - 2 - k \geq n - 1$

   $\longrightarrow c_i + y - (2n-2-k)x \leq c_i + y - (n-1)x$

   Further, $c_i + y - (n-1)x < 0 \longrightarrow c_i + y - (2n-2-k)x < 0$. This is a contradiction with $\triangle u_i > 0$. Thus the complete network of mutual cooperation must be a Nash equilibrium when the condition $(\forall i \in N)(c_i < (n-1)x - y)$ is satisfied.

3. Finally, we find that in all Nash equilibria that emerge through the undisturbed process of evolution of cooperation, links of unilateral cooperation are sustained only by altruistic agents.

   Regardless of the process through which equilibria manifest, we find that links unilateral cooperation can not be sustained by non altruistic agents in absence of cost heterogeneity. In presence of cost heterogeneity, however, non altruistic agents may also sustain unilateral cooperation in Nash equilibrium. These observations are discussed next.

Proposition 7 If all agents in the population are homogeneous and non altruistic, such that $(\forall i \in N)(c_i = c > z)$ and s^* is a Nash equilibrium profile, then $(\nexists i, j \in N)(a^*_{ij} = \alpha \wedge a^*_{ji} = \beta)$

Proof: Suppose all agents are homogeneous and s^* is a Nash equilibrium profile such that $(\exists i, j \in N)(a^*_{ij} = \alpha \wedge a^*_{ji} = \beta)$.

A unilateral deviation makes i worse off. From (iv), this implies:

$c - z - p_{ij}(s)x < 0 \longrightarrow c < p_{ij}(s)x + z$

A unilateral deviation makes j worse off. From (iii), this implies:

$-c - y + (t_{ji}(s) + 1)x + p_{ij}(s)x < 0 \longrightarrow c > (t_{ji}(s) + p_{ij}(s) + 1)x - y$

$\longrightarrow (t_{ji}(s) + p_{ij}(s) + 1)x - y < p_{ij}(s)x + z$

$\longrightarrow (t_{ji}(s) + 1)x - y < z$

Now, $t_{ji}(s) \geq 0$ and $x > y$ implies that $(t_{ji}(s) + 1)x - y > 0$. Also, $z < 0$ by assumption. Thus proposition 7 is proved by contradiction. □

The argument here is that if agent i finds it optimal to sustain a link of unilateral cooperation towards agent j, it must be because of the reputation related benefits that this brings. That is, i must have a sufficiently large number of common mutually cooperative partners with j so that the cost of sustaining unilateral cooperation with j is more than matched by the benefits from increased trustworthiness to all these common cooperative partners. But then j also has enough common mutually cooperative partners with i such that the benefit from reciprocating cooperation to i outweighs the cost. Thus, in any case where agent i has incentive to extend unilateral cooperation to agent j, the latter also has incentive to reciprocate cooperation to i. In other words, in any situation where sustaining unilateral cooperation is not sub optimal, incentives to reciprocate cooperation are stronger than incentives to remain a beneficiary of unilateral cooperation. Note that while unilateral cooperation does not exist as a feature of equilibria, the act of extending unilateral cooperation or reputation building still has a role to play in best response dynamics and convergence to equilibrium in case of any perturbations to the given network. Further, unilateral cooperation may feature into Nash equilibrium networks in presence of heterogeneous agents and may even be sustained by non altruists. Following example illustrates one such case.

Example 1 Suppose $n = 7$, $x = 7$, $y = 4$, $z = -1$, $c = (c_1, c_2, ...c_7) = (5, 15, 15, 15, 15, 15, 10)$. Note that there are no altruists in the population. The diagram below illustrates a possible network emerging from a Nash equilibrium, with agent 1 sustaining unilateral cooperation with agents $2, 3, 4, 5$ and 6.

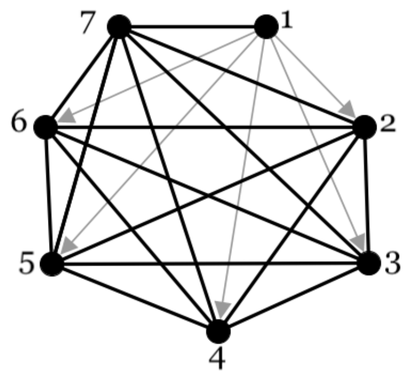


Figure 6: Agent 1 sustains unilateral cooperation with multiple agents because it adds to her trustworthiness in her cooperative tie with agent 7

Consider $i \in \{2,3,4,5,6\}$ making an arbitrary unilateral deviation by withdrawing cooperation from $0 \leq k \leq 5$ links, without reciprocating cooperation to 1. Then her expected change in utility is $\triangle u_i = kc_i + ky - 5kx - (5-k)kx = k(-51+7k) \leq 0$ for all possible values of k.

If i reciprocates cooperation to 1, and withdraws cooperation from $0 \leq k \leq 4$ links (while retaining cooperation with 7), her expected change in utility is $\triangle u_i = -c_i - y + 2x + x + kc_i + ky - 5kx - (5-k)kx = k(-51+7k) + 2 \leq 0$ for all possible values of k.

If i reciprocates cooperation to 1, withdraws cooperation from 7, and withdraws cooperation from $0 \leq k \leq 4$ other links, her expected change in utility is $\triangle u_i = -c_i - y + x + c_i + y - 5x + kc_i + ky - 5kx - (5-k)(k+1)x = -4x + kc_i + ky - 5kx - (5-k)kx - (5-k)x = k(-51+7k) + 7k - 63 < 0$ for all possible values of k. Thus i has no incentive to deviate.

Consider agent 7 making an arbitrary deviation by withdrawing cooperation from $0 \leq k \leq 5$ agents in $\{2,3,4,5,6\}$. If she does not withdraw cooperation from 1, $\triangle u_7 = kc_7 + ky - 5kx - (5-k)kx = -56k + 7k^2 = k(7k-56) \leq 0$ for all possible values of k. If she simultaneously withdraws cooperation from 1, $\triangle u_7 = k(7k-56) + c_7 + y - x = k(7k-56) + 7 \leq 0$ for all possible values of k. Thus agent 7 has no incentive to deviate.

Consider agent 1 making an arbitrary deviation by withdrawing cooperation from $0 \leq k \leq 5$ agents in $\{2,3,4,5,6\}$. If she does not withdraw cooperation from 7, $\triangle u_1 = kc_1 - kz - kx = 5k + k - 7k = -k \leq 0$ for all possible values of k. If she simultaneously withdraws cooperation from 7, $\triangle u_1 = c_1(1+k) - kz - 6x = 5 + 5k + k - 42 = -37 + 6k < 0$ for all values of k. Thus agent 1 also has no incentive to deviate and the strategy profile underlying the given network is a Nash equilibrium where a non altruistic agent sustains multiple links of unilateral cooperation.

In example 1, note that the cost of cooperation for the agent sustaining unilateral cooperation is much lower than the cost for agents who do not reciprocate cooperation for these interactions. In fact, in any Nash equilibrium where agent i sustains a link of unilateral cooperation with agent j, it must be the case that $c_i < c_j$. This is seen from the next proposition.

Proposition 8 If s^* is a Nash equilibrium strategy profile, then for any pair of agents i and j, $\big(a^*_{ij} = \wedge a^*_{ji} = \beta \longrightarrow c_i < p_{ij}(s^*)x + z \wedge c_j > (2p_{ij}(s^*) + 1)x - y\big)$

Proof: Suppose s^* is a Nash equilibrium strategy profile and $\big(\exists i, j \in N\big)\big(a^*_{ij} = \alpha \wedge a^*_{ji} = \beta\big)$. If i deviates to offer β to j:

$\triangle u_i = c_i - z - p_{ij}(s^*)x < 0$ (from (iv))

$\rightarrow c_i < p_{ij}(s^*)x + z$

If j deviates to offer α to i:

$\triangle u_j = -c_j - y + (t_{ji}(s^*) + 1)x + p_{ij}(s^*)x < 0$ (from (iii))

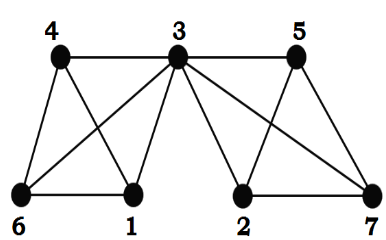
$\rightarrow c_j > (t_{ji}(s^*) + p_{ij}(s^*) + 1)x - y$

$\rightarrow c_j > (2p_{ij}(s^*) + 1)x - y$ (since $t_{ji}(s^*) \geq p_{ij}(s^*)$) □

Thus we have $c_j > (2p_{ij}(s^*) + 1)x - y$ and $c_i < p_{ij}(s^*)x + z$ when i sustains unilateral cooperation in her interaction with j in equilibrium. Now, $x > y$ and $z < 0$ implies that $(2p_{ij}(s^*) + 1)x - y > 2p_{ij}(s^*) > p_{ij}(s^*) > p_{ij}(s^*) + z$. Thus $c_j > (2p_{ij}(s^*) + 1)x - y > p_{ij}(s^*) + z > c_i$; that is, if agent i sustains a link of unilateral cooperation with another agent j in Nash equilibrium, it must be that $c_j > c_i$. Intuitively, the cost of extending cooperation is low enough for i to find it justified by the reputation related benefits of reputation gain that it unilateral cooperation brings, given the common cooperative partners that i shares with j. However, this cost is large enough for j to prohibit her from reciprocating. This highlights again the role of cost heterogeneity in explaining unilateral cooperation in Nash equilibria.

Finally, we return to our observation that a complete network of mutual cooperation is a Nash equilibrium *if and only if* $(\forall i \in N)(c_i < (n-1)x - y)$. Note that if this condition is satisfied, then there can exist multiple Nash equilibria with varying degrees of cooperation, the complete cooperation equilibrium being one of them. The example below illustrates one such equilibrium network with less than complete mutual cooperation.

Example 2 Suppose $n = 7, x = 5, y = 4, z = -1$ and for all $i \in N$, cost of cooperation is $c_i = 6$. Note that $c_i > x - y$ for all $i \in N$ implies that mutual cooperation can not be sustained in equilibrium without network effects. The cooperative network illustrated beloq(with only links of mutual cooperation shown) is supported by a Nash equilibrium strategy profile.



Consider agent 3. An arbitrary unilateral deviation can consist of withdrawing cooperation from $0 \leq k \leq 3$ agents in $\{1, 4, 6\}$ and $0 \leq m \leq 3$ agents in $\{2, 5, 7\}$ (where $k > 0 \vee m > 0$). This yields change in utility:

$\triangle u_3 = (k+m)(c+y) - kx(6-k) - mx(6-m) = 10(k+m) - 5[k(6-k) + m(6-m)]$

which is negative for any positive value of $k + m$, and zero otherwise. Thus agent 3 has no incentive to unilaterally deviate.

Consider agent 1. An arbitrary unilateral deviation can consist of a combination of withdrawing cooperation from $0 \leq k \leq 2$ agents in $\{4, 6\}$, extending cooperation to $0 \leq m \leq 3$ agents in $\{2, 5, 7\}$ and withdrawing cooperation from 3.

If 1 withdraws cooperation from 3:

$\triangle u_1 = (k+1)(c+y) - 3x(k+1) - (k+1)(2-k)x + m(z-c) = (k+1)[10+5(k-5)] - 7m$

which is negative for all combinations of positive values of k and m.

If i does not withdraw cooperation from 3:

$\triangle u_1 = k(c+y) - 3kx - (3-k)kx - mc + mz + mx = 10k - 5k(6-k) - 2m$

which is negative for all combinations of positive values of k and m.

Thus agent 1 has no incentive to deviate. Note that agents 1, 2, 4, 5, 6 and 7 are symmetrically placed in this network. Thus a symmetric argument can be made to show that other agents in $\{2, 4, 5, 6, 7\}$ also have no incentive to deviate and the strategy profile underlying the given cooperation network is a Nash equilibrium. □

Thus we find that when the cost of cooperation for every agent is less than $(n-1)x - y$, networks with varying degrees of cooperation can sustain as Nash equilibria. However, we find that presence of agents with cost higher than $(n-1)x - y$ is detrimental to cooperation.

Proposition 9 If $(\exists i \in N)(c_i > (n-1)x - y)$, then i does not engage in any links of mutual cooperation in equilibrium.

Suppose i is involved in a link of mutual cooperation with j in Nash equilibrium s^*. This implies that $c_i + y - (t_{ij}(s^*) + p_{ij}(s^*) + 1)x < 0$. Given that $c_i > (n-1)x - y$, this implies that $p_{ij}(s^*) > n/2 - 1$, or $p_{ij}(s^*) \geq n/2$. Further, for any m in the set $P_{ij}(s^*)$, $p_{im}(s^*) \geq n/2$ and for any $l \in P_{im}(s^*)$, $p_{il}(s^*) \geq n/2$ and so on.

The highest utility that i can be facing in such an equilibrium is when any link of mutual cooperation that she engages in has the maximum possible embeddedness, which is $n-2$, that is when the network is a complete network of mutual cooperation. However, $c_i > (n-1)x - y$ ensures that i has incentive to deviate from such a strategy profile by withdrawing cooperation from all agents.

Thus we see that agents whose cost of cooperation exceeds $(n-1)x - y$ do not ever find it optimal to extend cooperation in any interaction. Further, the presence of such agents in the population can be highly detrimental for cooperation to exist in equilibria. The next example illustrates an extreme situation where in the presence of a single agent of this kind we find that the unique Nash equilibrium is one with zero cooperation between all agents.

Example 3 Suppose $n = 5$, $x = 20$, $y = 10$, $z = -5$ and $c = (c_1, c_2, \ldots c_5) = (75, 55, 35, 15, 5)$. Note $c_1 > 4x - y$ implies that agent 1 does not extend cooperation in any interaction in equilibria. Since agent 1 is never linked in ties of mutual cooperation with anyone else in any equilibrium, sustaining cooperation with 1 does not yield benefits of trustworthiness or reputation enhancement for other agents. Thus in any equilibrium cooperative network, 1 is an isolate (or engaged in $\beta - \beta$ ties with everyone else).

Now consider the set $N - \{1\}$. Note that $c_2 > 3x - y$ implies that even in a situation where agents 3, 4 and 5 are all linked in ties of mutual cooperation and extending cooperation to agent 2, the latter has no incentive to reciprocate cooperation to anyone. Similarly, if agents 2, 3, 4 and 5 are all linked with mutual cooperation, agent 2 has incentive to deviate and withdraw cooperation from all interactions. That is, even in the scenario that is most favourable towards cooperation, agent 2 does not cooperate with anyone in equilibrium. Given that, there are no benefits of reputation enhancement possible from sustaining cooperation with 2; thus 2 is engaged in $\beta - \beta$ links with everyone in any equilibrium.

By the same argument, observe that $c_3 > 2x - y$ and $c_4 > x - y$ and $c_5 > z$ imply that none of the remaining agents can be extending cooperation to anyone else in equilibria. But this implies that the unique Nash equilibrium is the empty network.

In the above example, suppose agent 5 were an altruist instead. Note in that scenario even agent 4 (who has lowest cost among other agents) would not optimally reciprocate cooperation to 5 and in that case the unique Nash equilibrium would be one with agent 5 extending cooperation to all other agents and all other agents refraining from cooperating in all interactions. On the other hand, suppose the cost of cooperation for agent 1 were 55 instead of 75. Then, from proposition 9, the complete network of mutual cooperation is a Nash equilibrium. Thus the above example illustrates how the presence of a single agent with $c_i > (n-1)x - y$ can be detrimental to overall cooperation. For instance, if from a complete network in equilibrium, the cost of cooperation for a single agent were to suddenly increase then it can lead to a larger breakdown of cooperation through the network effect because when one agent exits cooperation, it reduces the trustworthiness supporting all other interactions in the previously complete network.

4 Discussion

This paper proposes a model of cooperative network formation that serves to model cooperative interactions between individuals in situations that can be formalised by the coordination game, with an important application being costly sharing of valuable resources in the context of general scarcity. In line with literature, we assume that being observed to be cooperating in many situations increases one's *reputation* or trustworthiness. Further, we assume that information about one's reputation is communicated through the cooperation network itself, and that a higher perceived trustworthiness of any agent involved in a mutually cooperative interaction yields a higher value from that interaction for the agent. This results in an incentive for *indirect reciprocity*, which is a well documented motive and tendency in literature on cooperation. Note that the indirect reciprocity is directed towards partners of partners, resulting in a tendency for *triadic closure*. Triadic closure refers to the tendency of two nodes who share one or more neighbours to become linked to each other in social networks. Agents who share neighbours in social networks are expected to form social links with each other because they have higher opportunity of interacting with each other, can develop higher trust on each other and may even have a higher incentive to link with each other to avoid any social stress. Through our model we find that reputation building in cooperative interactions can also explain triadic closure in some networks.

The network effects of trustworthiness considered in our model have the implication that the more *embedded* an interaction is in common mutual cooperative partnerships, the higher incentives that both individuals have to maintain cooperation in the interaction since it a) enhances trustworthiness for many other interactions (even if cooperation is not reciprocated); and b) yields a high value when cooperation is reciprocated, because of high trustworthiness due to high embeddedness. These network effects have an important role in not just the evolution of cooperation but also in sustaining cooperation in equilibria. Because of the importance of network effects in sustaining cooperation, we find that the presence of a few altruists can generate high levels of cooperation in equilibria whereas the presence of a few high cost agents (whose cost of cooperation is too high to *ever* cooperate) can be potentially very detrimental to the overall level of cooperation in equilibria.

From our analysis of the evolution process, we find that equilibria of varying degrees of cooperation emerge under different conditions on the costs of cooperation of agents and distribution of altruists in the population. The presence of altruists is crucial to starting off the process of cooperation and the number of altruists in the population has a positive relation with the degree of cooperation observed in the equilibrium network that evolves. For non altruistic agents, the incentive to build reputation plays an important role in explaining cooperative activity. Cooperative ties that non altruistic agents form with other non altruistic agents are first initiated to build trustworthiness or reputation to support existing ties of mutual cooperation.

From our analysis of Nash equilibria, we find that under if costs of cooperation for all agents in the population are below a certain threshold, cooperative networks of various kinds can manifest as equilibria. Network effects of trustworthiness ensure that cooperative activity is found in equilibria even under cost conditions that would have prohibited cooperation if each pair-wise interaction were independently conducted. We also find that in presence of heterogeneity of costs of cooperation, even non altruistic agents may sustain links of unilateral cooperation in equilibrium. Further, the presence of even a few individuals with prohibitively high costs of cooperation (such that they refrain from reciprocating cooperation even if the entire set of other agents, being completely connected in ties of mutual cooperation, were extending cooperation to them) is highly detrimental for cooperative activity in equilibria.

Various dynamic models of network formation in the context of cooperative networks have been proposed by Jackson and Watts (2001, 2002a, 2002b). A significant difference between these formulations in literature and the model proposed here is that while in the works of Jackson and Watts players choose their interaction patterns on an individual-by-individual basis (in every period a single pair of nodes is stochastically chosen to update the link between them), in our model agents decide in each period of discrete time their best response actions towards *all other players*. This limits the applicability of our model such that it is best suited to study cooperation in small populations or geographic areas where agents have repeated interactions with *everyone else* and hence choose actions for all their interactions simultaneously. We acknowledge this limitation in the model; a richer version of the model could be one that restricts agents to be able to change actions with respect to only a few other agents in any given period. Given that, it may be useful to analyse evolution of cooperation from given initial network structures (which can also influence the equilibria that emerge) instead of a situation of zero cooperation. Dynamics of myopic best responses could also be replaced by far-sightedness for a richer analysis.

Another major limitation of the model is that it assumes for simplicity a single cost of cooperation, whether reciprocated or non-reciprocated. This gives rise to the following problem. With given incentive parameters x, y and z assumed to be the same for all individuals, any heterogeneity in agents' inherent trustworthiness (independent of their network position) is reflected in their costs of cooperation. For instance, if we wanted to model the fact that two agents i and j are identically placed in the network and yet i enjoys a higher payoff because of an inherently higher trustworthiness (which causes other agents to share higher quality and quantity of resources), we can do it only by assuming a lower cost of cooperation for i than for j. This is why we interpret c_i as *an indicator of not only the logistical costs of sharing resources or information, but also any other exogenous factors that may increase the effort or investment that i needs to make to derive benefits from the link.* But note also that attitudes of altruism are also defined using cost of cooperation. Since altruists prefer to unilaterally cooperate over a situation of no cooperation, this can only be modelled by assuming $c_i < z$ for altruists. This gives rise to an unintentional interpretation of altruists as individuals with high levels of *inherent* trustworthiness. While altruists can develop a high network-based reputation by extending cooperation to all other agents, there is no reason to assume that they are inherently more trustworthy than other agents. This problem can be resolved by allowing unilateral cooperation and reciprocated cooperation to have different associated costs.

Finally, thus far we have assumed that agents have the same cost of extending cooperation towards *all* other agents, hence excluding the possibility of heterogeneity based communal identity[8]. This assumption may be dropped to yield more asymmetric and communally segregated equilibria through the evolution process. Further, as mentioned before, the best response dynamics considered in the evolution process provides a tool to analyse the impact of external perturbations on existing cooperative structures.

[8] Agents belonging to the same ethnic, linguistic, caste or religious community may lower costs of extending cooperation to each other because of logistical ease and higher implicit trust on each other

References

Alexander, R. (1987). The biology of moral systems (foundations of human behavior).

Allen, S. M., Colombo, G., & Whitaker, R. M. (2010). Cooperation through self-similar social networks. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, *5*(1), 4.

Aumann, R., & Myerson, R. (1988). Endogenous formation of links between players and coalitions: an application of the shapley value. *The Shapley Value*, 175–191.

Axelrod, R. (1984). *The evolution of cooperation.* Basic Books.

Axelrod, R. (2000). On six advances in cooperation theory. *Analyse & Kritik*, *22*(1), 130–151.

Axelrod, R., & Dion, D. (1988). The further evolution of cooperation. *Science*, *242*(4884), 1385–1390.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *science*, *211*(4489), 1390–1396.

Bala, V., & Goyal, S. (2000). A noncooperative model of network formation. *Econometrica*, *68*(5), 1181–1229.

Blau, P. M. (1960). Structural effects. *American sociological review*, 178–193.

Blau, P. M. (1994). *Structural contexts of opportunities.* University of Chicago Press.

Bloch, F., & Dutta, B. (2011). Formation of networks and coalitions. In *Handbook of social economics* (Vol. 1, pp. 729–779). Elsevier.

Bó, P. D. (2005). Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. *American economic review*, *95*(5), 1591–1604.

Bourlès, R., Bramoullé, Y., & Perez-Richet, E. (2017). Altruism in networks. *Econometrica*, *85*(2), 675–689.

Buskens, V., & Weesie, J. (2000). Cooperation via social networks. *Analyse & Kritik*, *22*(1), 44–74.

Carpenter, J., & Myers, C. K. (2010). Why volunteer? evidence on the role of altruism, image, and incentives. *Journal of Public Economics*, *94*(11-12), 911–920.

Caulier, J. (2009). Network games as tu cooperative games: The core, the shapley value and simple network games.

Curry, O., & Dunbar, R. (2011). Altruism in networks: the effect of connections. *Biology letters*, rsbl20101202.

Dutta, B., Ghosal, S., & Ray, D. (2005). Farsighted network formation. *Journal of Economic Theory*, *122*(2), 143–164.

Dutta, B., & Jackson, M. O. (2003). On the formation of networks and groups. In *Networks and groups* (pp. 1–15). Springer.

Dutta, B., & Jackson, M. O. (2013). *Networks and groups: Models of strategic formation.* Springer Science & Business Media.

Dutta, B., & Mutuswami, S. (1997). Stable networks. *Journal of economic theory*, *76*(2), 322–344.

Dutta, B., van den Nouweland, A., & Tijs, S. (1998). Link formation in cooperative situations. *International Journal of Game Theory*, *27*(2), 245–256.

Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets: Reasoning about a highly connected world.* Cambridge University Press.

Eguíluz, V. M., Zimmermann, M. G., Cela-Conde, C. J., & Miguel, M. S. (2005). Cooperation and the emergence of role differentiation in the dynamics of social networks. *American journal of sociology*, *110*(4), 977–1008.

Engelmann, D., & Fischbacher, U. (2009). Indirect reciprocity and strategic reputation building in an experimental helping game. *Games and Economic Behavior*, *67*(2), 399–407.

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425*(6960), 785.

Fu, F., Hauert, C., Nowak, M. A., & Wang, L. (2008). Reputation-based partner choice promotes cooperation in social networks. *Physical Review E*, *78*(2), 026117.

Goyal, S. (2012). *Connections: an introduction to the economics of networks*. Princeton University Press.

Goyal, S., & Vega-Redondo, F. (2005). Network formation and social coordination. *Games and Economic Behavior*, *50*(2), 178–207.

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American journal of sociology*, *91*(3), 481–510.

Granovetter, M. (2005). The impact of social structure on economic outcomes. *Journal of economic perspectives*, *19*(1), 33–50.

Gross, J. L., & Yellen, J. (2004). Handbook of graph theory crc press. *Boca Raton*.

Hanaki, N., Peterhansl, A., Dodds, P. S., & Watts, D. J. (2007). Cooperation in evolving social networks. *Management Science*, *53*(7), 1036–1050.

Jackson, M., & Wolinsky, A. (1996). A strategic model of economic and social network. *J Econ Theory*, *71*, 44–74.

Jackson, M. O. (2005). A survey of network formation models: stability and efficiency. *Group Formation in Economics: Networks, Clubs, and Coalitions*, 11–49.

Jackson, M. O. (2009). Networks and economic behavior. *Annu. Rev. Econ.*, *1*(1), 489–511.

Jackson, M. O. (2010). *Social and economic networks*. Princeton university press.

Jackson, M. O., Rogers, B. W., & Zenou, Y. (2017). The economic consequences of social-network structure. *Journal of Economic Literature*, *55*(1), 49–95.

Jackson, M. O., & Watts, A. (2002a). The evolution of social and economic networks. *Journal of Economic Theory*, *106*(2), 265–295.

Jackson, M. O., & Watts, A. (2002b). On the formation of interaction networks in social coordination games. *Games and Economic Behavior*, *41*(2), 265–291.

Jackson, M. O., & Wolinsky, A. (2003). A strategic model of social and economic networks. In *Networks and groups* (pp. 23–49). Springer.

Jackson, M. O., & Zenou, Y. (2015). Games on networks. In *Handbook of game theory with economic applications* (Vol. 4, pp. 95–163). Elsevier.

Lehmann, L., & Keller, L. (2006). The evolution of cooperation and altruism–a general framework and a classification of models. *Journal of evolutionary biology*, *19*(5), 1365–1376.

Lim, C. (2008). Social networks and political participation: How do networks matter? *Social forces*, *87*(2), 961–982.

Lindgren, K., & Nordahl, M. G. (1994). Evolutionary dynamics of spatial games. *Physica D: Nonlinear Phenomena*, *75*(1-3), 292–309.

Marsh, C. (2018). Indirect reciprocity and reputation management: Interdisciplinary findings from evolutionary biology and economics. *Public Relations Review*.

Myerson, R. B. (1977). Graphs and cooperation in games. *Mathematics of operations research*, *2*(3), 225–229.
Myerson, R. B. (1980). Conference structures and fair allocation rules. *International Journal of Game Theory*, *9*(3), 169–182.
Nava, F. (2016). Repeated games and networks.
Nowak, M. A. (2006). Five rules for the evolution of cooperation. *science*, *314*(5805), 1560–1563.
Nowak, M. A., & May, R. M. (1992). Evolutionary games and spatial chaos. *Nature*, *359*(6398), 826.
Nowak, M. A., & Sigmund, K. (1998a). The dynamics of indirect reciprocity. *Journal of theoretical Biology*, *194*(4), 561–574.
Nowak, M. A., & Sigmund, K. (1998b). Evolution of indirect reciprocity by image scoring. *Nature*, *393*(6685), 573.
Phelps, S. (2013). Emergence of social networks via direct and indirect reciprocity. *Autonomous agents and multi-agent systems*, *27*(3), 355–374.
Pievani, T. (2011). Born to cooperate? altruism as exaptation and the evolution of human sociality. In *Origins of altruism and cooperation* (pp. 41–61). Springer.
Riolo, R. L., Cohen, M. D., & Axelrod, R. (2001). Evolution of cooperation without reciprocity. *Nature*, *414*(6862), 441.
Roger, B. M. (1991). Game theory: analysis of conflict. *The President and Fellows of Harvard College, USA*.
Schroeder, D. A. (1995). *Social dilemmas: Perspectives on individuals and groups*. Greenwood Publishing Group.
Semmann, D., Krambeck, H.-J., & Milinski, M. (2004). Strategic investment in reputation. *Behavioral Ecology and Sociobiology*, *56*(3), 248–252.
Skyrms, B., & Pemantle, R. (2009). A dynamic model of social network formation. In *Adaptive networks* (pp. 231–251). Springer.
Slikker, M., & Van den Nouweland, A. (2012). *Social and economic networks in cooperative game theory* (Vol. 27). Springer Science & Business Media.
Smith, K. P., & Christakis, N. A. (2008). Social networks and health. *Annu. Rev. Sociol*, *34*, 405–429.
Sussman, R. W., & Cloninger, C. R. (2011). *Origins of altruism and cooperation*. Springer.
Sylwester, K., & Roberts, G. (2010). Cooperators benefit through reputation-based partner choice in economic games. *Biology letters*, rsbl20100209.
van den Nouweland, A. (2005). Models of network formation in cooperative games. *Group formation in economics*, 58–88.
Wang, Z., Wang, L., Yin, Z.-Y., & Xia, C.-Y. (2012). Inferring reputation promotes the evolution of cooperation in spatial social dilemma games. *PloS one*, *7*(7), e40218.
Watts, A. (2001). A dynamic model of network formation. *Games and Economic Behavior*, *34*(2), 331–341.
West, S. A., Griffin, A. S., & Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of evolutionary biology*, *20*(2), 415–432.
Young, H. P. (1993). The evolution of conventions. *Econometrica: Journal of the Econometric Society*, 57–84.

MORE IN THE SERIES

- Kundu, R. P., and Pandey, S. G., (2021). Multiple Public Goods in Networks, W.P. No. 345 (August).
- Balakina, O., Balasubramaniam, V., Dimri, A., and Sane, R., (2021). Unshrouding product-specific attributes through financial education, W.P. No. 344 (August).
- Goyal, A., Pandey, R., and Sane, R., (2021). Consumption baskets of Indian households: Comparing estimates from the CPI, CES and CPHS, W.P. No. 343 (August).

Siddhi Gyan Pandey, Jindal School of Liberal Arts and Humanities, O.P. Jindal Global University, Sonipat, Haryana.
E-mail: siddhigyanpandey@gmail.com

National Institute of Public Finance and Policy,
18/2, Satsang Vihar Marg, Special Institutional Area (Near JNU), New Delhi 110067
Tel. No. 26569303, 26569780, 26569784
Fax: 91-11-26852548
www.nipfp.org.in