Social preferences and the provision of public goods

Sudipta Sarangi[1], Sakshi Upadhyay[2]

[1]Department of Economics, 880 West Campus Drive Virginia Tech, Blacksburg, VA, 24061. email: ssarangi@vt.edu

[2] Department of Economics, 880 West Campus Drive Virginia Tech, Blacksburg, VA, 24061. email: usakshi@vt.edu

November 24, 2020

Abstract

This paper surveys the literature on social preferences and their incorporation in a public goods game. The Nash equilibrium in a typical public goods game is to free ride; however experimental evidence shows that players frequently cooperate. When an individual's utility function accounts for social preferences, one can explain cooperative behavior and contribution to public goods. This chapter surveys a number of different models of social preferences and their impact on public goods provision, both theoretically and empirically. Finally, we examine how provisioning of the public goods is affected by social preferences when agents play this game in networks and coalitions.

Keywords: Social Preferences, Public goods, Coalitions, Networks

JEL Classification: D91, H41, D85, C72

"*How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it, except the pleasure of seeing it*" (Smith, 1976, pg.9)

1 Introduction

Public goods are characterized by *non-rivalry*, meaning that more than one person can simultaneously benefit from them, and *non-exclusivity*, meaning that it is difficult to prevent any individual from enjoying their benefits. They simultaneously benefit many people and their creation requires the coordinated actions of people who will subsequently enjoy its benefits. Environmental protection, research and innovation, vaccination, health care services, highways, and public parks are just a few important examples.

Despite receiving benefits from public goods, individuals tend to free ride on the contributions of others in a group. Given that these goods are non-rival and non-excludable, it is evident that once the goods has been produced, every agent can consume it regardless of their contribution. Unless there exists mechanisms to make individuals act in their common interests, rational or self interested individuals will not act to achieve their common or group interests (Olson, 2012). There are various mechanisms to ensure cooperation or reduce free riding. We discuss punishment, commitment and communication as some of the mechanism which help increase cooperation in a society.

An alternative explanation for the cooperation seen can be provided by social preferences. Social norms and preferences also matter in the provision of public goods. For example, local resources can be managed well when users care about others and can organize and enforce their own rules, instead of following externally imposed norms (Fehr and Fischbacher, 2002).

Importance of social norms and preferences can also be seen in our daily lives. Consider a family with grandparents/parents and their children. Parents or grandparents might invest in infrastructure, environment, technology to mitigate climate change so that it can benefit their future generations. They have an added incentive to invest in highways, public schools, public parks, environment friendly vehicles or practices because of the concern for their children. Another way to understand this idea is that your family member might contribute more to public goods that will benefit you in future as compared to a stranger.

In the following subsection, we start with a simple public goods model and show how free riding is an equilibrium in this game. Our goal in this chapter is to examine the reasons for the absence of free-riding, in particular by focusing on the role of social preferences. In Section 2 we provide evidence on absence of free riding and also discuss mechanisms available in literature to ensure cooperation. Section 3 describes how social preferences can explain cooperation. The section also entails a model on social preferences and the equilibrium after incorporating social preferences. In Section 4 we will discuss models of various types of social preferences available in the literature. These theoretical models are also supplemented with experimental evidences. Section 5 discusses how social preferences influences public goods provisioning in a coalition or network framework. Section 6 concludes the chapter.

1.1 A simple model of public goods

To fix ideas formally, we now present the public goods model in Fehr and Schmidt (1999). We will use this model to arrive at a fundamental result in public goods which will also be the first Proposition of this chapter. Let there be $n \geq 2$ individuals in a society who simultaneously decide on their contribution levels $g_i \in [0, y]$, $i \in [1, 2, \cdots, n]$ to the public good. Each player has an endowment of y. The monetary payoff of player i is given by (Fehr and Schmidt, 1999, p.836, equation 11):

$$x_i(g_1, g_2, \cdots, g_n) = y - g_i + a \sum_{j=1}^{n} g_j, \qquad 1/n < a < 1 \tag{1}$$

Here a denotes the constant marginal return to public good $G = \sum_{i=1}^{n} g_i$. Since $a < 1$, contributing to G leads to loss of $1 - a$. The dominant strategy of an individual i is to choose $g_i = 0$.

Definition 1 *A strategy* g_i^* *is a Nash equilibrium of this game if for all* $i \in [1, 2, \cdots n]$, $x_i(g_i^*, g_{-i}) \geq x_i(g_i^{'}, g_{-i})$ *for all* $g_i' \in (0, y]$.

In this game, we have $g_i^* = 0$ as the Nash equilibrium strategy for any player $i \in [1, .. \cdots, n]$. Strategy other than $g_i * = 0$ is denoted by $g_i^{'}$. Strategy of players other than i is given by g_{-i}.

Thus, the standard model predicts $g_i = 0$ for all $i \in [1, 2, \cdots, n]$. However, since $a > 1/n$, aggregate monetary payoff is maximised at $g_i = y$. This observation leads us to fundamental result about public goods in economics summarized in our first Proposition.

Proposition 1 *Suppose the payoff function is given by equation 1 and satisfies* $1/n < a < 1$, *then in Nash equilibrium* $g_i = 0$ *for all* $i \in [1, 2, \cdots, n]$.

Not contributing to the public goods is termed as ‘free-riding’. Kim and Walker (1984) summarize the ‘free rider’ problem in their theoretical model: “If the method of voluntary contributions is used to determine the level at which a public goods will be provided, then the resulting provision level will be far below the optimal level, and many individuals will contribute nothing at all.” Free riding is also evident in their experimental results.

In practice however, we may not always see free-riding. In the next section we provide experimental and empirical evidence on lack of free riding. Various mechanisms to ensure cooperation or avoid free riding will also be discussed in the next section. In a later section we examine how social preferences can be used to explain cooperation or absence of free riding.

2 Evidence on free riding behavior

Free riding has been a widely accepted notion in the literature of public goods games. Previous theory suggests that players try to get the benefit from a public goods without contributing towards it. However those results are in sharp contrast to the existence of cooperative behavior among individuals in real-life public goods games. This behaviour has been substantiated by data from national surveys as shown in Andreoni (1988*a*), who states: "Around 85% of households make donations to charity, 50% of tax returns include charitable deductions". Another related evidence of cooperative behavior can be found in voting in elections. Individuals tend to vote in elections, even though economic theory predicts that free riding will be higher as decisive power of one vote is low. Countries joining International Environment Agreements (IEA) to solve environmental issues is also an example of cooperation. Group of 77 (G77), United Nations Framework Convention on Climate Change (UNFCCC), Kyoto Protocol are some of the existing IEA's.

The contrast between the theoretical predictions and real life evidences motivated testing of 'free-rider' hypothesis in the lab. Ledyard (1995) surveys the experimental literature on public goods before 1995. Some of the prominent papers included in the survey are Marwell and Ames (1981), Isaac, Walker and Thomas (1984), Isaac and Walker (1988), Andreoni (1988*b*). One of the findings from these experiments suggests that individuals contribute more than the Nash equilibrium prediction in a public goods game. As we saw in the previous section,the Nash equilibrium in a public good game is to free-ride, however it is optimal to contribute the full amount. On average contributions were about 40-60 % of the optimal level in these experiments. However the contributions varied over individuals. The other common observation is that contributions start at 40-60 % of the optimal level but over the periods decline to 'free riding' outcome.

This decay in contribution levels were further analyzed in (Andreoni, 1988*b*) through 'learning' and 'strategies' hypothesis. According to the learning hypothesis, repeated periods allow individuals to learn the incentives from the game which can explain the fall in contribution levels. At the same time learning also allow players to signal future moves to each other. This leads to the strategy hypothesis, where in a repeated games a rational player will develop multi period strategies that can lead to cooperative behavior. However, Andreoni finds no significant support for either of these hypothesis which could have explained the decay experienced in these games. We now state our first observation from findings in this section.

Observation 1: *Empirical and experimental evidence show that individuals cooperate and contribute to public goods as opposed to the theoretical predictions.*

These experimental results motivated research on importance of institutional environment which can help in achieving the optimal outcome or reduce free riding. In next subsection we discuss these mechanisms which can help further increase cooperation in a public goods game.

2.1 Mechanisms that avoid free riding

This section summarizes the institutional environments which have been used in the literature to reduce incidence of 'free-riding'. Institutional environment refers to the context or setting in which individuals would make their decisions. The payoff function remains the same as in equation 1, however we look into different settings under which the public goods game is played. We discus three such institutional environments: communication, commitment and punishment.

Communication

Communication between the participants regarding their strategies or intentions can help in increasing contributions to a public goods game. Isaac and Walker (1988) was the first paper to test face to face communication as a means to reduce 'free-riding'. According to the authors 'The role of communication is to a) help the group understand the group profit implications for different allocations and b) build credibility to the expected decisions of group members". Communication thus helps in learning the optimal strategy (contributing to public goods). Ostrom (1997) also finds that face-to-face communication can sustain cooperation even through the last period. Communication enforces no verbal agreement and hence can be thought of as 'cheap talk' (Ostrom, 2000). Her paper summarizes the finding on collective action and one of the findings suggests that when communication is implemented by allowing subjects to signal promises through their computer terminals, much less cooperation is observed as compared to the case when subjects are allowed face-to-face communication.

Communication enhances cooperation, however, the effectiveness of communication depends on it's structure and the level of private information amongst players. Palfrey, Rosenthal and Roy (2017) provide answer to this problem both theoretically and empirically. They find theoretical bounds on efficiency gains that can be attained through different modes of communication by using Bayesian-mechanism design. The bounds depend upon the distribution of private information(value of endowed unit of output) and on richness of the message space (communication structure). The authors choose three forms of pre-play communication : binary message (intention to contribute or not), practice game (announce their contribution against different contribution costs) and natural language communication (exchange of chat messages) in order to test their theoretical bounds. The results from their experiment find efficiency and public goods provisioning to be significantly higher in case of natural language communication as efficiency bounds predicted by the theoretical model were only achieved in this treatment. This might be because "unrestricted chats gives subjects an opportunity to understand each other's intentions and messages". Natural language communication can be thought of as a more personal form of communication which gives more scope to convey an individual's message and intentions than restricted message or any other form.

Commitment

Commitment can also be used as a strategy to enhance cooperation. "Commitment is a means by which players can assure one another that they are not going to free ride on others' contributions, so that group members can contribute without fearing that they will be free ridden" (Kurzban et al., 2001). Chen (1996) was one of the first paper to use 'pledge to contribute' as a commitment. The authors find that group based pledge (subjects make a pledge before making a contribution, are given feedback and have to then contribute a proportion of the mean pledge) and face to face communication have similar results in enhancing cooperation. Through commitment, individuals can eliminate free-riding. However, once an individual makes a commitment, he/she is more vulnerable of being free ridden (Kurzban et al., 2001). This is because individuals might use commitment by others as an opportunity to 'free-ride' on their contribution. To respond to this issue authors design a mechanism, where players can commit to cooperating to a small degree and then observe other player's reciprocal contributions. The mechanism allows participants to signal their commitment without exposing them to be 'free-ridden'. They test for efficiency of different 'pledge for contribution'. The study finds that 'increase only' pledge is effective in increasing cooperation. This mechanism works as a commitment strategy by not letting the players reverse their contributions and allows players to reduce their extent of free riding by limiting their commitments.

Punishment

People who cooperate might be willing to 'punish' the free riders. Ostrom, Walker and Gardner (1992) was the one of the first papers to test the impact of punishment in a public goods framework. The authors allow for costly punishments in a repeated common pool resource game and find that participants punish free riders in their experiment. However in their paper the same subjects interacted for multiple periods, thus giving them an incentive to cooperate and punish free riders. To rule out these incentives, Fehr and Gachter (2000) in their experiment have a punishment and non punishment treatment crossed with stranger (group composition changes every period)and partner treatment (group composition is fixed). The authors find that in both the treatments the punishment is heavier the more negatively individual deviates from the contributions of group members. The average contribution goes up in both the stranger and partner treatment when punishment is allowed and approaches to full cooperation in partner treatment

Previous experiments which studied the role of punishments could not elicit much about the robustness of punishment schemes. Nikiforakis and Normann (2008) in their paper provide a comparative statistics of punishment in public goods games. They find that contributions to public goods increase monotonically in the effectiveness of punishment (factor by which the punishment reduces the punished player's income). Higher effectiveness leads us near to social optimal outcome.

Individuals do not contribute in a public goods game, due to the chance of being 'free-ridden' by others. All the mechanisms discussed above change the environmental setting of a game in a manner which increases the incentive to cooperate. The success of the mechanism depends upon how effective it is in reducing chances of being 'free-ridden'.

3 Social preferences: An alternative explanation

While the experimental literature has provided us with examples of several mechanisms that can lead to free-riding and reduce cooperation, we now focus on an alternative approach to explain these findings: the presence of *social preferences*. Theories of other regarding preferences/social preferences are based on assumption (and observation) that people care about the well being of others. In his paper Andreoni (1995), shows that on average about half of all cooperation is due to subjects who understand free riding but cooperate due to kindness. The author also suggests that decline in cooperation observed in multiple trials of public goods experiment might not be due to learning but maybe a result of frustrated attempts at kindness.

According to Fehr and Fischbacher (2002): 'An individual exhibits social preference if the person cares about material resources allocated to relevant reference agents'. The relative reference agent can vary according to different domains, thus resulting in various types of social preferences. The authors empirically also show that it is difficult to understand concepts of competition on market outcomes, laws governing cooperation and collective action, optimal contracts and property rights, social norms and market failures without incorporation of social preferences.

Nash equilibrium strategy of players in a public goods game is to contribute nothing. However past literature suggests clear evidences of cooperation among players. Players' incentive to contribute positively can be predicted theoretically by including social preferences in their payoff functions. Examples of such social preferences include responsibility of the older generation (grandparents/ parents) towards their future generation. Such responsibility drives elders to contribute positively towards any public goods or service which will guarantee a secure future for their children. We illustrate such cooperative behavior using a model of social preferences from Fehr and Schmidt (1999). In the later subsections we introduce different models of social preferences.

3.1 A simple model of public goods with social preferences

In order to show how the results in a public goods model (Proposition 1) change after incorporation of social preferences we use the inequity aversion model of Fehr and Schmidt (1999). In this model, in addition to purely selfish individuals, the authors assume presence of subjects who dislike inequity both when they are worse off than other players and also when they are better off than other players.

Consider a set of n players indexed by $i \in [1, 2, , \cdots, n]$ and let $x = x_1, x_2, \cdots, x_n$ denote vector of monetary payoffs. The utility function of $i \in [1, 2, , \cdots, n]$ is given by:

$$U_i(x) = x_i - \alpha_i\left(\frac{1}{n-1}\sum_{j \neq i} max|x_j - x_i, 0|\right) - \beta_i\left(\frac{1}{n-1}\sum_{j \neq i} max|x_i - x_j, 0|\right) \quad (2)$$

The second term in equation 2 measures loss from disadvantageous inequality, the third term measures loss from advantageous inequality. The two parameters α_i and β_i measure player i's utility loss from disadvantageous inequality and from advantageous inequality. The authors assume that $\beta_i \leq \alpha_i$ and $0 \leq \beta_i < 1$. $\beta_i \leq \alpha_i$ implies, players suffer more from inequality that is to their disadvantage, i.e. the subject is loss averse in social comparisons. $\beta_i \geq 0$, rules out the subjects who like to be better than others.

We now substitute equation 1 in equation 2 to see how public goods provisioning changes due to presence of inequity aversion. For this result we focus on Proposition 4c of Fehr and Schmidt (1999) which discusses positive contribution levels of individuals.[1] Player i who does not contribute ($g_i = 0$) is a 'free-rider'. Let number of free-riders be represented by k. Recall from equation 1, g_i and a denote the contribution levels and marginal return to public good respectively.

Proposition 2 *(Fehr and Schmidt, 1999, p.839 Proposition 4(c)) If* $k/(n-1) < (a + \beta_j - 1)/(\alpha_j + \beta_j)$ *for all players* $j \in [1, 2, \cdots, n]$ *with* $a + \beta_j > 1$*, then other equilibria with positive contribution levels does exist. In this equilibria all* k *players with* $a + \beta_i < 1$ *must choose* $g_i = 0$ *while all other players contribute* $g_i = g \in [0, y]$*. Note further that* $(a + \beta_j - 1)(\alpha_j + \beta_j) < a/2$

[1] The case of free-rider ($g_i = 0$) is studied in part a and b of the Proposition 4 of the original paper.

We first discuss the author's intuition behind the proof and then move towards sketch of the proof. If there are sufficiently many players with $a + \beta_j > 1$, they can sustain cooperation among themselves even when other players are free riding. This only holds when contributors are not affected much by the disadvantageous inequality. This is because if α_j increases, it is less likely to be the case that: $k/(n-1) < (a + \beta_j - 1)/(\alpha_j + \beta_j)$

Sketch of the proof

- Following from the author's Proposition 4a, the dominant strategy of k free-riders, with $a + \beta_i < 1$, is $g_i = 0$ (not contribute). This is because free-rider's return from public good (a) and non pecuniary benefit from reducing inequality(β_i) is less than 1.

- The remaining $n - k$ or j players with $a + \beta_j > 1$ contribute positively with $g_i = g \in [0, y]$. j's payoff is given by:

$$U_j(g) = y - g + (n-k)ag - \alpha_j \left(\frac{1}{n-1} kg \right) \quad (3)$$

  Any individual who contributes is deprived of the advantageous utility which reduces the third term in equation 2 to zero, thereby forming equation 3.

  Suppose player deviates from contributing g to $g - \triangle$, such that $\triangle > 0$.The deviation strategy towards contributing less than g will not payoff if and only if $U(g - \triangle) \leq U(g)$. Simplifying this inequality leads us to the following condition: $k/(n-1) \leq (a + \beta_j - 1)/(\alpha_j + \beta_j)$.

- Following from author's Proposition 4b, if there are only few players with $a+\beta_i > 1$, they would suffer too much loss from the disadvantageous inequality caused by the free riders. The proof given by the authors shows that if a potential contributor knows that the number of free riders, k, is larger than $a(n-1)/2$, then he will not contribute either.

4 Types of social preferences

As seen in the previous section, incorporation of inequity aversion in the standard utility functions, predicts cooperation in a public goods game. Depending on the assumptions of model and the type of social preference, the model can look different. For instance we can have altruism as a social preference, incorporated into standard utility function. However the mechanism to arrive at the equilibrium will be similar and will lead to positive contributions being made to the public goods game. In the next subsection we will discus other papers on fairness and inequity aversion. In the later subsections we explain models with different social preferences and their outcomes.

4.1 Fairness and inequity aversion

Inequity aversion implies that individuals care for equitable distribution of resources or equal outcomes. These models consider an individual 'fair' if the individual is willing to give up their payoff to help others. A model of fairness is represented in equation 2. The second and third term which measures the individual loss from disadvantageous and advantageous inequality respectively are measure of fairness in their model.

Rabin (1993) was one of the first to develop game theoretic solution concept "fairness equilibria". An outcome is considered to be fair if the intention behind the action is kind , whereas if the intention is hostile, the action is considered to be unfair. The model is applicable to all finite-strategy games involving two players. Each player's expected subjective utility depends on: his strategy, his beliefs about other player's strategy choices and his beliefs about other player's beliefs about his strategy.

Let $a_1 \in S_1$ and $a_2 \in S_2$ represent strategies chosen by two players; $b_1 \in S_1$ and $b_2 \in S_2$ represent player 2's belief about strategy player 1 is choosing, and player 1's belief about what strategy player 2 is choosing. $c_1 \in S_1$ and $c_2 \in S_2$ represent player 1's belief about what player 2 believes player 1's strategy is, and player 2's beliefs about what player 1 believes player 2's strategy is.

Each player i chooses a_i to maximize expected utility:

$$U_i(a_i, b_j, c_i) = \pi_i(a_i, b_j) + \bar{f}_j(b_j, c_i).[1 + f_i(a_i, b_j)] \tag{4}$$

- $\pi_i(a_i, b_j)$ is individual i's material payoff.
- Player i's kindness to player j is measured by $f_i(a_i, b_j)$. The function measures how much more than or less than player j's equitable payoff[2] player i believes he is giving to player j. When $f_i = 0$, player i is giving j her equitable payoff. If $f_i > 0$, player i is giving j more than her equitable payoff. When $f_i < 0$, player i is giving j less than her equitable payoff.
- $\bar{f}_i(b_j, C_i)$ measures player i's belief about how kind player j is being to him. If player i believes that player j is treating him badly ($\bar{f}_i(b_j, C_i) < 0$), then i chooses a_i such that $f_i(a_i, b_j)$ is low or negative. The opposite situation occurs when $\bar{f}_i(b_j, C_i) > 0$.

The above game by Rabin is a psychological game of the type described by Geanakoplos, Pearce and Stacchetti (1989). The equilibrium concept in these games is called psychological Nash equilibrium which is an analog of Nash equilibrium. Psychological Nash equilibrium concept imposes an additional condition that all higher order beliefs match actual behavior. Rabin uses psychological Nash equilibrium to arrive at the *fairness equilibrium*, which we describe in the next definition.

Definition 2 *(Rabin, 1993, p.1288, Definition 3) The pair of strategies* $(a_1, a_2) \in (S_1, S_2)$ *is fairness equilibrium if for* $i = 1, 2, j \neq i$:

- $a_i \in arg\ max_{a \in S_i} U_i(a, b_j, c_i)$
- $c_i = b_i = a_i$

According to the above definition, individual i's strategy (a_i) should maximize her payoff. The strategy should also be equal to player j's belief about player i's strategy (b_i) and player i's belief about what player j believes player i's strategy is

[2] equitable payoff is the average of highest and lowest payoff of player j

(c_i). Thus individuals actions and their higher order beliefs both match their actual behavior.

A mutual max(min) outcome is one where player's mutually maximize (minimize) each other's payoffs. We now discuss one of the Propositions which talks about two types of Nash equilibrium being 'fairness equilibrium'.

Proposition 3 *(Rabin, 1993, p.1290, Proposition1) Suppose that* (a_1, a_2) *is a Nash equilibrium, and either a mutual-max outcome or a mutual-min outcome. Then* (a_1, a_2) *is a fairness equilibrium.*

The proof is intuitive. First, suppose (a_1, a_2) is mutual-max outcome, then both f_1 and f_2 are non-negative. This implies players have positive regard for each other. Since both players are choosing strategy that maximizes their payoff and payoff of other player, this must maximize their own utility. Now suppose (a_1, a_2) is mutual-min outcome, then f_1 and f_2 will be non -positive, both players would like to decrease well-being of others. Simultaneously player also maximizes his own utility by maximizing his material well-being.

We can draw prediction by applying a prisoner dilemma game into a public goods framework with only 2 players. The Nash equilibrium in a prisoner's dilemma game is to defect, which can be interpreted as no cooperation in a public goods game. Incorporation of reciprocal motives, in a public goods game can lead to full cooperation as one of the equilibrium. The implications will be difficult if there are more than two people (which is usually the case). The payoff function incorporates the stylized facts evident in many experiments: people are willing to sacrifice their own well being to help those who are kind, people are willing to sacrifice their own well being to punish those who are unkind. However, Rabin (1993) model can only be applied to two persons game.

Fehr and Schmidt (1999) models fairness as a self centered inequity aversion, i.e. individuals are willing to give up some payoff to move in the direction of equitable outcome. Individuals in these models are concerned about their relative utility or payoff as compared to others. Unlike Rabin (1993), Fehr and Schmidt (1999) do not model intentions explicitly and use standard game theory in order to analyze n-person public goods game. The authors assume that subjects suffer more from inequity due to their material disadvantage than from inequity due to their material advantage (see section 3.1). In presence of inequity averse people, the authors can explain "fair" and "cooperative" as well as "competitive" and "non cooperative" behavioral patterns. The model also accounts for interaction between distribution of preferences in a given society. For instance, presence of 'free-riders' in the society induces many inequity-averse individuals to behave in a selfish manner. This happens because if there are only few individuals who have $\alpha + \beta_i > 1$, they suffer too much loss from disadvantageous inequality caused by free riders. This is Proposition 4 (b) in their paper, discussed briefly in our section 3.1 .

The experimental evidence on fairness and inequity aversion is not obvious. Dannenberg et al. (2007) test for inequity aversion using model from Fehr and Schmidt (1999). The experiment is a a two-step procedure using within-subject design. In the first step subjects played selected games to estimate their individual other regarding preferences. In the second step, subjects with preferences (fair and selfish) according to Fehr and Schmidt (1999) were matched into pairs and interacted with possibility of punishment. They find significant effect of advantageous inequity aversion (third term in equation 2) on individual's contribution to public goods. Another paper, Blanco, Engelmann and Normann (2010) also uses within-subject design to assess predictive power of Fehr and Schmidt (1999) model. They find that inequity aversion can explain individual's behavior in a public goods game at an aggregate level, however not at the individual level. Aggregate level tests compare the distribution of outcomes across different experiments that were run with different samples and thereby check for consistency. Individual level analysis on the other hand uses within subject design to test for decisions in different experiments with same sample. The model of inequity aversion was based on relative payoff of individuals. In our next section we discuss models of altruism focusing on absolute payoff of individuals.

4.2 Altruism

Standard utility function as defined in equation 1 focus on individual's monetary payoff. Models of giving or donating to a charity have been based on 'altruism', where an individual is assumed to contribute to the pubic goods because they simply demand more of public goods. However these models have low predictive power and were not able to incorporate the empirical findings, this lead to development of models with 'impure altruism'. In the models with 'impure altruism' individuals are assumed to contribute to public goods because of two reasons: 1) altruism: people demand more of a public goods, 2) people get some private goods benefits from the gift per se which is called 'warm -glow'. The second motive is also termed as 'egoistic motive'.

Andreoni (1989), Andreoni (1990) presents model of giving that incorporates warm glow in a public goods game. Suppose there is one private good and one public good. Individuals are endowed with wealth w_i which they can allocate between consumption of private good x_i and their gift to the public good g_i. Let n be total number of individuals and $G = \sum_{i=1}^{n} g_i$. In order to explain how a utility function transforms in case of impure altruism, we use the utility function from (Andreoni, 1990, p.465, equation 1) as stated below:

$$U_i = U_i(x_i, G, g_i), \qquad i = 1, 2, \cdots, n \tag{5}$$

Here, U_i is assumed to be strictly quasi concave. Notice that g_i enters twice in the utility function, once as part of public good G, and as private good g_i. This captures the fact that an individual's contribution/gift (g_i) has properties of a private good that are independent of it's properties as a public good.

If the utility function is of the form $U_i = U_i(x_i, G)$ then preferences are purely altruistic. This is because, individual does not get any private goods benefit from the contribution. In contrast, if the utility function is of the form $U_i = U_i(x_i, g_i)$ then preferences are purely egoistic and individual is only motivated to give because of warm glow. Individual only derives private goods benefit from contributing to the public goods.

Let gift/contribution of all the other players except i be denoted by $G_{-i} =$

$\sum_{j \neq i} g_j$, individual donations/contributions can be found by solving:

$$\begin{aligned} \max_{x_i, g_i, G} \quad & U_i(x_i, G, g_i) \\ \text{s.t.} \quad & x_i + g_i = w_i \\ & G_{-i} + g_i = G \end{aligned}$$

Under the Nash equilibrium, G_{-i} is treated exogenously, thus we can re-write $g_i = G - G_{-i}$. Substituting the budget constraints given above into utility function (equation 5) we get:

$$\max_G \quad U_i(w_i + G_{-i} - G, G, G - G_{-i}) \tag{6}$$

Differentiating equation 6 w.r.t G and solving leads a donation function that is given by the following:

$$G = f_i(w_i + G_{-i}, G_{-i}) \tag{7}$$

$$g_i = f_i(w_i + G_{-i}, G_{-i}) - G_{-i} \tag{8}$$

The first argument in equation 7 is from the public dimension of the utility function. The second argument is from the private goods dimension of the utility function. The partial derivative of f_i with respect to first argument is denoted by $f_{i\alpha}$. This is i's marginal propensity to donate for altruistic reasons. f_{ie} represents the partial derivative of f_i with respect to second argument. This is i's marginal propensity to donate due to egoistic reasons. Thus model incorporates both altruistic and egoistic reasons for contributing to the public goods. In the model, $0 < f_{i\alpha} < 1$ and $f_{ie} > 0$. From the equations above we can say that individual's contribution g_i is increasing in both egoistic and altruistic motives.

Incorporation of these motives can lead to positive contribution ($g_i > 0$) unlike the standard model, which will predict no contribution ($g_i = 0$). The predictions from their model are also consistent with various empirical findings mentioned by Andreoni. Including private provisioning of public goods or impure altruism also increase the predictive power of the models. For instance, pure altruism model predicts that an increase in amount of public good provided (G), implies a dollar-for-dollar decrease in individual's own contribution (g_i). If there is a dollar-for-dollar decrease in g_i for any increase in G, we call such crowding out of g_i as complete. However empirically, the magnitude of such crowding out is found to be incomplete or proportionately less than the magnitude of change in G. Such findings are consistent with the theoretical predictions from impure altruism models.

Andreoni, Harbaugh and Vesterlund (2010) provide evidence of altruistic preferences in various games: prisoner's dilemma, public goods game, dictator game, trust games and gift exchange games. The survey also suggests the formation of altruistic can be due to cultural norms, psychological development, socialization and neural foundations.

Model of impure altruism also predicts that individuals will reduce their contribution to the public goods when other individuals increase their contributions. However, this observation is in contrast to various other outcomes in a public goods game. For instance conditional cooperation (discussed in the next section) is observed in a public goods game, where individuals cooperate if they see others contributing. Reciprocity observed in many games also contradicts the assumption of altruism." An altruistic person's kindness does not depend on behavior of others, whereas kindness of a strong reciprocator is conditional on the perceived kindness of other players" Fehr, Fischbacher and Gächter (2002). We next discuss reciprocity and conditional cooperation.

4.3 Reciprocity and conditional cooperation

Theories of reciprocity and conditional cooperation incorporate individual's willingness to cooperate if others are cooperating as well. In these model, individuals are also concerned about intentions behind other's decision. We can also apply the concept of 'conditional cooperation' to the section of inequity aversion, wherein individuals contribute if they believe others will contribute due to his/her concern for equity in payoffs (Chaudhuri, 2011). However, models of reciprocity or conditional cooperation capture intentions or beliefs of individuals as compared to models of inequity aversion.

Reciprocal motivation is modelled in Rabin (1993), however the model does not apply to sequential games. Falk and Fischbacher (2006) extend the notion of reciprocity in a sequential game. The authors present a formal theory of reciprocity (equation 9), where the players utility now depends upon an individual's payoff and also on the kindness (how kind a person perceives action by another player) and the reciprocation term (response to the experienced kindness). We now represent their utility function in a 'reciprocity game' (Falk and Fischbacher, 2006, p.301, Definition 3):

$$U_i(f, s_i", s_i') = \pi_i(f) + \rho_i \sum_{n \to f} \psi_j(n, s_i", s_i')\sigma_i(n, s_i", s_i') \qquad n \in N_i \tag{9}$$

The game is a two-player extensive form game with finite number of stages. Let $i \in \{1, 2\}$ be a player in the game and let player j be the other player. N denotes the set of nodes and N_i is the set of nodes where player i has the first move. $n \in N$ is one of the node in the game. S_i and S_j is behavioral strategy space of player i and j respectively. $s_i \in S_i$, $s_j \in S_j$ are behavior strategy of player i and j respectively. s_i' denotes *first order belief* of player i and captures i's belief about the behavior strategy player j will choose. $s_i"$ denotes the *second order belief* of player i and captures i's belief about j's belief about which strategy player i will choose. F denotes the set of end nodes of the game. The models fixes, f as an end node that follows (directly or indirectly) node n.

The first term in equation 9 : $\pi_i(f)$ is individual i's material payoff. The second term is the reciprocity utility and comprises of:

- Reciprocity parameter: ρ_i is a positive constant and common knowledge. It captures strength of player i's reciprocal preferences. A high ρ_i implies reciprocal utility is more important as compared to the other utility. If $\rho_i = 0$, then utility equals material payoff $\pi_i(f)$.

- Kindness term: $\psi_j(n, s_i", s_i')$ which measures how kind i perceives action by another player j. It depends upon the consequence or outcome of that action and underlying intention. ψ_j is product of outcome term ($\triangle_j$) and intention factor (v_j). Outcome term measures the output, $\triangle_j > 0$ expresses advantageous outcome for i, $\triangle_j < 0$ expresses disadvantageous outcome for i. Intention factor measures the intention behind the outcome. $v_j = 1$ captures a situation where $\triangle_j$ is result of an action which j completed intentionally and $v_j < 1$ implies j's action was not fully intentional.

- Reciprocation term: $\sigma_i(n, s_i", s_i')$ expresses response to experienced kindness, how much i alters payoff of j with his move in node n. A rewarding action implies a positive reciprocation term, whereas a punishment implies a negative reciprocation term.

- Product of kindness term and reciprocation term measures the reciprocal utility in a particular node. If kindness term in a particular node n is positive, then individual i's utility increases if he/she chooses an action in that node which increases j's payoff. The opposite holds when kindness term is negative and i has an incentive to reduce j's payoff. The model measures kindness in each node where i has the move, hence the overall reciprocity utility is the sum of the reciprocity utility in all nodes (before the considered end node), weighted with the reciprocity parameter (ρ_i).

The authors discuss the intuition of their theoretical prediction for public goods game. According to authors: 'the strategic structure of a prisoner's dilemma is very similar to a public goods game'. In a sequential prisoner's dilemma player 1 either cooperates or defect and after observing player 1's outcome, player 2 also chooses to cooperate or defect. The sub-game perfect solution is for both the players to defect. The above model predicts that if player 2 is sufficiently reciprocally motivated there is a positive probability that player 2 rewards player 1's cooperation with cooperation. Also, player 2 will defect if player 1 defects. Conditional cooperation between subjects can also be seen through the payoff function (equation 9), subjects contribute more if the contributions of other group members are also higher. In terms of model parameters: if player j is cooperative then kindness term ψ_j will be positive and if player i expresses response to experienced kindness then, σ_i is also positive. Since $\rho_i > 0$, i's contribution increases in response to higher contribution by j.

The authors also substantiate their theoretical predictions with experimental results from Fischbacher, Gächter and Fehr (2001). In their experiment, subjects could conditionally indicate how many tokens they wanted to contribute to public goods. The best strategy is to contribute nothing irrespective of others contributions, however subjects' average contribution was increasing in the mean contributions of others. Using this conditional-cooperation strategy, more than half of the subjects were classified as "conditional cooperators" and the rest were classified as free riders.

Fehr and Gachter (2000) also find evidence of reciprocity in their public goods experiment, the more a subject free rides relative to others the more he/she is punished. In order to test multiple preferences, Croson (2007) conducts an experiment to test theories of commitment, theories of altruism and theories of reciprocity in a public goods game. Almost all subjects demonstrate a positive correlation between their own contribution and belief of others' contributions, consistent with theory of reciprocity.

4.4 Heterogeneous social preferences

Positive contribution towards public goods game can be explained and sustained through incorporation of other regarding preferences. We have discussed theoretical models of three types of social preferences along with their experimental evidences. These models incorporate heterogeneity, thereby allowing for presence of different equilibria.

Experimental evidences in public goods game, further add to the above observation. Chaudhuri (2011) in his survey summarizes the advances made in the literature since Ledyard (1995) by agreeing upon presence of distinct type of players. These players differ in social preferences and/or their beliefs about others, which can explain their behavior being contrary to standard theoretical prediction of free riding. Gunnthorsdottir, Houser and McCabe (2007) in their voluntary contribution mechanism (VCM) public goods experiment classify the subjects into 'free riders' (contributes 30 % or less of his/her endowment) and 'cooperators' (contributes more than 30 %) based on their first round contribution.

Heterogeneous preferences can also explain the decline in cooperation in these experiments due to the presence of free riders (Fischbacher and Gachter, 2010).

The decline in cooperation over periods is suggested due to"presence of imperfect conditional cooperators", those who match others contribution but only partially. Interaction of "imperfect conditional cooperators" with free riders leads to an increase in free rider behavior.

5 Extensions

In this section we discuss extensions of the model of social preferences in public goods game. We explain how effectiveness of 'Coalition' and 'Networks' increases through incorporation of social preferences.

5.1 Coalition Formation

Coalitions, subgroups of individuals who agree to act collectively to produce a public goods, represent a possible solution to the public goods problem. Coalitions such as International Environmental Agreements (IEA) where countries cooperate for an environmental cause are also observed in practice. Agents in a coalition first decide whether or not to join a coalition, then members decide how much to contribute. Social preferences also influences coalition size and their inclusion can lower the threshold for contributing to the public goods.

Kolstad (2014) assumes homogeneous Charness and Rabin (2002) preferences. Let there be $i = 1, 2, \cdots, N$ countries, each with potential to emit w_i. Each country chooses level of abatement (a public good) given by g_i or level of emissions $x_i = w_i - g_i$.

Welfare is positively affected by

1. Direct benefits of emitting: x_i. If the country has more emissions, the production cost reduces.

2. Aggregate level of abatement: G. Higher abatement leads to lower pollution levels or reduces the environmental damage.

National welfare (u_i) depends upon egoistic component/self centered (π_i) and pro-social or altruistic component (α_i). Altruistic component (α_i) depends on the vector of egoistic payoffs of other countries. The payoff is given as follows:

$$u_i(x_i, G) = \lambda_i \pi_i(x_i, G) + (1 - \lambda_i)\alpha_i(\pi) \tag{10}$$

Here, $\lambda_i \in [0, 1]$ reflects extent to which country is selfish or altruistic. Egoistic component can be described as following in a public goods game framework:

$$\pi_i = x_i + aG, \qquad \text{where} \quad x_i + g_i = w_i, G = \sum g_i \tag{11}$$

$$= w_i - g_i + aG, \quad \text{where} \quad G = \sum g_i; 0 \leq g_i \leq w_i \tag{12}$$

Here, w_i is the maximum possible emissions for country i, g_i is level of abatement for country i and G is aggregate abatement over all the countries. a represents the marginal per capita return (MPCR) and indicates how much an investment in abatement returns privately. The authors assume $a \in (1/(N-1), 1)$. This is because, for $a = 1$, the individual will be indifferent between abating and non abating (welfare from emitting and abatement have the same return). Small values of a are also excluded because coordination might not be enough for abatement.

Now we talk about the altruistic payoff in the utility function which is taken from Charness and Rabin (2002).

$$\alpha_i(\pi) = [\delta_i(min_{j \neq i}\pi_j) + \epsilon_i \sum_j \pi_j]/(1 - \lambda_i) \quad \text{where} \quad \delta_i, \epsilon_i \geq 0; \delta_i + \epsilon_i + \lambda_i = 1 \tag{13}$$

Here, δ_i reflects relative importance of agent i of distribution/equity and ϵ_i reflects importance of efficiency. Equity is represented in the model by a Rawlsian preference which is the minimum monetary payoff over rest of the population. Efficiency is represented by total monetary payoffs over the population. Inclusion of social preferences in the model, reduces the threshold for contributing to the public goods. This result is given by Proposition 1 in the paper:

Proposition 4 *(Kolstad, 2014, p.15, Proposition 1) Assuming the N homogeneous player public goods game with Charness and Rabin social preferences, then*

1. *Efficient (Pareto Optimal) outcomes involve all countries undertaking maximal abatement; and*

2. *The Non-cooperative Nash equilibrium involves each agent either not abating ($gi = 0$) or fully abating ($gi = wi$) according to*

$$g_i = 0 \qquad if \quad a < \bar{a}_i \tag{14}$$

$$g_i = w_i \qquad if \quad a > \bar{a}_i \tag{15}$$

$$where \qquad \bar{a}_i = (\lambda_i + \epsilon_i)/[1 + \epsilon_i(N-1)] \tag{16}$$

In case of standard preferences, the cutoff for abating and not abating is $a = 1$, with social preferences cutoffs are lower, by construction $\bar{a}_i < 1$. $\bar{a}_i$ can also be interpreted as MPCR between cooperation and non cooperation. Concerns for efficiency ($\epsilon_i > 0$), keeping δ_i constant also lowers $\bar{a}_i$. Thus inclusion of social preferences reduces the cutoff for abating or not abating. With presence of social preferences, countries find it individually rational to abate (provide public goods).

Ringius, Torvanger and Underdal (2002) identify 'fairness' as a motivation for countries in environmental negotiation. The study also analyzes various IEA's with negotiations leading to Kyoto protocol and find considerations of fairness and equity to be building characteristics of these negotiations. In their empirical analysis Lange, Vogt and Ziegler (2007) show that equity issues are considered highly important in international climate negotiations by using a world wide survey of people involved in international climate policy. Polluter pays rule (rule of equal ratio between abatement costs and emissions) and the accompanying poor losers rule (exempting due to GDP) are the most widely accepted equity principles according to this study.

Grüning and Peters (2010) in their paper incorporate fairness and justice in countries' preferences. We now illustrate their utility function to understand how coalitions/IEA incorporate social preferences. The public goods problem arises because each country can choose their level of abatement (say reducing carbon emissions) and benefits from the reduced emissions by all the other countries as well. Country j's payoff can be represented by the following quasi linear logarithmic function (consisting of benefit minus abatement cost) minus a term which measures heterogeneity by means of variance in all abatement strategies (Grüning and Peters, 2010, p.141, equation 1).

$$P_j = ln(\sum_i a_i) - a_j - \theta.\sigma(a_1, a_2, \cdots, a_N) \tag{17}$$

In the above payoff, $ln(\sum_i a_i)$, measure the benefit from abatement of all the countries. (a_i, a_j) measures the abatement levels of country i and j. $\sigma(a_1, a_2, \cdots, a_N)$ measures variance in the environmental policies of all the countries. Variance is a measure for fairness and justice in their model, since countries prefer a more egalitarian cost sharing. Variance in their model is defined as $\sum_i \frac{(a_i - \bar{a})}{N}$ where $\bar{a}$ is the global average of all countries environmental policies. A country's payoff is also assumed to be concave in it's own strategy and continuous in that of the opponents. $\theta \geq 0$ represents preference intensity for welfare loss due to cost dispersion. For instance $\theta = 0$ corresponds to the case of pure selfishness, increasing θ corresponds with stronger concern for 'fair or just' cost sharing. Equity concerns are homogeneous in this symmetric payoff function, however their model is also robust to heterogeneous countries. The authors extend their results to asymmetry in measuring equity (σ is modified by incorporating countries self interest), heterogeneous countries (countries have different θ or different abatement cost).

The authors find that stronger fairness attitudes leads to homogeneous results as countries both inside and outside IEA adjust abatement levels to each other. This is explained by Proposition 1 from their paper

Proposition 5 *(Grüning and Peters, 2010, p.143, Proposition 1) Abatements inside and outside the coalition.*

1. *For signatories, stronger fairness preferences result in smaller abatement activities. If θ exceeds the threshold level $\tilde{\theta}$, even an outsider becomes active. The stronger θ, the more abatements an outsider carries out. In the limit (for $\theta \to \infty$) there is no difference between an insider and an outsider.*

2. *The aggregate does not significantly change in θ. For $\theta < \tilde{\theta}$; fairness has a negative impact on global abatements , while $A(S; \theta)$ remains constant for all θ exceeding the threshold $\tilde{\theta}$.*

Here $A(S;\theta)$, is the aggregate abatement activity and is given by:

$$A(S;\theta) = Sa_S^*(S,\theta) + (N-S)a_o^*(S,0) \tag{18}$$

Here, $a_S^*(S,\theta)$ and $a_o^*(S,0)$ are the abatement activities of countries inside and outside IEA respectively. Countries in IEA are signatories and are represented by S. According to the above Proposition, if $\theta < \tilde{\theta}$[3]; outsiders are free riders and fairness concern leads to signatories reducing their abatements. This leads to a lower $A(S;\theta)$ or loss in environmental quality. For $\theta > \tilde{\theta}$, $A(S;\theta)$ does not change and for stronger θ countries abatement becomes similar, leading to not much difference between insider and outsider. In words, stronger fairness preferences lead to more abatement by non signatories. Fairness concerns implies countries should not deviate too much from other countries environmental policies. This deviation is measured by the variance in the payoff function (equation 17). Fairness concern thus leads to similar abatements by signatories as well.

Thus either all or almost none of the countries form an IEA. Internalization of the global environmental externality stabilizes IEA's whereas the free riding hinders larger coalitions. Thus stronger fairness preferences are needed to overcome instability of grand coalition as these preferences favor similar behavior with respect to abatement .

Sarangi and Upadhyay (2019) study the role of social preference in a two stage public goods game where, in the first stage, heterogeneous agents first choose whether or not to join a coalition then, in the next stage, the coalition votes on whether its members will contribute. The preferences are assumed to be Rawlsian, wherein the individuals care about the least well off person in the society.

Let there be $i = 1, 2, \cdots, n$ players. The individuals payoff depend on their own payoff and the payoff of the least well off person. λ_i is the weight on their won payoff. The utility function in their model is as follows:

$$\pi_i = \lambda_i(P_i) + (1-\lambda_i)(min(P_j)) \tag{19}$$

Here P_i is the monetary payoff of i and min (P_j) is the lowest monetary payoff of

[3] $a_o^*(S,0) = 0$ for $\theta < \tilde{\theta}$, see (Grüning and Peters, 2010, page 142, equation 5)

any player j[4].

They find that individuals with stronger social preferences are more likely to join the coalition and vote for the coalition to contribute to the public goods. This can be summarized from Proposition 2 in their paper given below. Let the decision to join be given by j_i, $j_i = 1$ means individual joins the coalition and $j_i = 0$ implies individual does not join the coalition.

Proposition 6 *(Sarangi and Upadhyay, 2019, p.10, Proposition 2) In the subgame perfect Nash equilibrium, if $\lambda_i \leq \gamma$ then $j_i = 1$. If $\lambda_i > \gamma$ then $j_i = 0$.*

The threshold for joining the coalition ($\lambda_i \leq \gamma$), also satisfies the threshold for contributing to the public goods in their paper. Thus incorporation of social preferences can result in larger coalition. The result is intuitive, since individuals with stronger social preferences are more likely to join the coalition and contribute to public goods.

Inclusion of social preferences, lowers the thresholds for contribution and increases the likelihood of a larger coalition/grand coalition. Accounting for social preferences in the coalition framework helps in learning about the successful development of coalitions.

5.2 Network Formation

Bramoullé, Kranton et al. (2007) provide the first network model of public goods and answer how social or geographical structure affect level and pattern of public goods. The study finds that individuals who have active social neighbors usually gain more from contribution of others (due to more links) but contribute less to public goods. This is similar to the concept of free riding observed in a general public goods game. With similar reasons, an addition of new link increases access to public goods, however reduces individual's incentive to contribute. Galeotti et al. (2010) suggest that effect of adding links to a network depend upon where is the link added.

[4]player i can also have the least payoff, in that case $\pi_i = P_i$

Galeotti et al. (2010) in their paper also examine patterns of social communication in a network. In the game, individuals choose to personally acquire information and form connections with others. Main finding of the paper, suggests that strict equilibrium of the game exhibits “law of the few”. According to “law of the few”, in a social group, small subset of individuals personally acquire information (called hub), while rest of the population forms connections with this small set of information acquirers. Individual information acquisition is a local public good game and implies that an equilibrium which entails links can lead to under provision of information acquisition. The socially optimal output is when central player in a star network acquires information and the others form links with hub. This happens when cost for forming an additional link is less than cost of acquiring information. If this is not the the case then in the social optimal outcome all players acquire information and no one forms links.

Caria and Fafchamps (2019) conduct a public good experiment on a star network (one central player and seven spokes). The design of experiment is based on the theoretical work of Bramoullé, Kranton et al. (2007). By design, the contribution of the centre player benefits all individuals located at the spokes, while the contributions of the spokes only benefits the center. Following prediction from equity and efficiency, center player should be motivated to contribute more than the spokes. Also the central player experiences ‘social pressure’ as other players also expect central player to contribute more than others. This is captured using ‘guilt aversion’ model from Battigalli and Dufwenberg (2007) where subjects experience guilt if their actions determine a payoff for other players that is lower than what these player expect.

Guilt aversion that star center i feel towards player j can be captured by (Caria and Fafchamps, 2019, p.397):

$$G_{ij}(c_i, \alpha_j, z) = max\{E_j[\pi_j] - \pi_j, 0\} = max\{r(\alpha_j^z - c_i^z), 0\} \tag{20}$$

Here, $E_j[\pi_j]$ is the expected payoff of spoke player j and π_j is the actual payoff of player j. c_i is contribution profile of star player, z indicates average contribution of all spoke players. Thus c_i^z indicates contribution of player i when seven spokes have contributed on average z. α_j^z is the expectation profile of player j from player i when spoke members contribute on average z. r indicates the rate of return to public good contributions. Thus guilt is a measure of the difference between player i's contribution and what spoke members expect i to contribute.

Each player was also asked to predict average value of contribution among other 7 players for each level of z. α_i^z records how much player i expects other 7 players to contribute when they play as center of the star and spokes on average contribute z. $\bar{\alpha}_z$ is contribution that individuals in network, on average expect from a player at the center of the star. This is arrived by taking the group expectations: average of α_i^z over all the eight players.

We now use their utility function to illustrate incorporation of social preference (guilt aversion here) in a Network (Caria and Fafchamps, 2019, p.397, equation 2).

$$u_i(c_i, \bar{\alpha}, z) = \pi_i - \frac{1}{7}\sum_{j \neq i} 7g_i * G_{ij}(c_i, \alpha_j, z) \tag{21}$$

Utility is monetary payoff minus cost of guilt central player experiences for each of the 7 spoke members. The authors assume that player i believes that each spoke has the same expectations, so that individual expectation coincide with group expectation ($\alpha_j = \bar{\alpha}$). Hence, the central player experiences same guilt towards each of the 7 spoke players. The utility function simplifies to (Caria and Fafchamps, 2019, p.397, equation 3)

$$u_i(c_i, \bar{\alpha}, z) = \pi_i - g_i * G_{ij}(c_i, \bar{\alpha}, z) \tag{22}$$

Here $G_{ij}(c_i, \bar{\alpha}, z) = max\{r(\bar{\alpha}^z - c_i^z), 0\}$. The first term in the utility function, equation 22 reflects concern for monetary payoff and second term is cost of guilt. If player i is sufficiently averse to guilt, he /she will align his/her contributions to expectations of other players to minimise guilt. For instance if player i contributes an amount lower than what other players expected, he/she will be guilty. Suppose player i increases contribution by one unit, this will decrease guilt of player i by $g_i r$.

The Contribution decisions in the game were made before assigning positions in the network. This was done in order to ask subjects how much they would like to contribute (i) if they are assigned the spoke position and (ii) if they are assigned the center position. Contribution in case (i) is denoted by denoted by s_i and in case (ii) is denoted by c_i. Each player had three notes worth 50 INR and had to decide how many notes to contribute, thus $z \in \{0, 1, 2, 3\}$. For each value of z, central player has to decide how much he would like to contribute. Vector $c_i = (c_i^0, c_i^1, c_i^2, c_i^3)$ collects four conditional decision of player i. Subjects were also asked to predict average value of contribution (c_j^z) among other seven players for each value of z. This helps to get an estimate of α_j^z and we thereby arrive at $\bar{\alpha}^z$ .The results from the experiment suggest that subjects in the center contribute as much as the average contribution, thus suggesting evidence of 'conditional cooperation'. Subjects play 'conditional cooperation' even when efficiency and equity concerns would require star player to contribute more than others. Disclosing group expectations significantly increases the contribution made by the star central player, thus confirming evidence of guilt aversion.

Altruism has been studied in a network framework (Bourlès, Bramoullé and Perez-Richet, 2017), however in the context of transfers. Structure of the network again plays a role in determining how income shocks lead to change in inequality. The consequence of change in altruism network is uncertain and also depends upon where the expansion takes place.

Zhang (2018) investigate social preferences in a networks game. Their models incorporate inequity aversion (Fehr and Schmidt, 1999) and welfare preferences(Charness and Rabin, 2002). The experiment manipulated the network structure: star or circle and the return from public good. Subjects at the core/center of a network contribute more than others in a star network. Subjects in a circular network, who earned less than others also contribute more than predicted outcome in subsequent period. This behavior suggests individuals exhibit welfare preferences rather than inequity aversion.

Cooperation is reinforced when conditional cooperators are more likely to interact. Thus, cooperation should fare better in highly clustered networks. Suri and Watts (2011) conducted a series of web experiments in which individuals play local public goods game with network topology varying across the sessions. In contrast to the earlier results, they find that network topology had no significant effect on average contribution. Players were as likely to decrease their contributions for low contributing neighbors as they were to increase their contributions in response to high contributing neighbors, thereby suggesting evidence of conditional cooperation. Positive effects of cooperation were contagious only to direct neighbors in network.

6 Conclusion

Public goods simultaneously benefit many people and are vital to individuals and societies which further fosters economic growth. A key theme in public goods research is deciding how much of a public goods to produce and how to pay for it. While public goods theory predicts free riding and inefficient outcomes, experimental results suggests existence of cooperation , with contribution rates at 40-60 percent of the efficient level. Donations to charity, payment of taxes, voting in elections and countries participating in IEA's are some of the other examples which support the claim that cooperation does exist. There are various mechanisms in the literature to reduce 'free riding'. Face to face communication, pledging the contribution, punishing the free riders are some of the effective tools to increase and sustain cooperation over the periods.

Studies on public goods highlight that human behavior is not entirely motivated by pure self interest. This has led to the formalization of other regarding preferences in the standard utility function. We have classified social preferences as impure altruism, fairness and inequity aversion, reciprocity and conditional cooperation. These models with social preferences are able to generate predictions for positive contribution and cooperation in a public goods game. However, there are variations in predictions of these models which arise from the heterogeneity of preferences of individuals. Further, preference of one individual might vary, contingent on the situation. Thereby in some situations, a individual's behaviour can be driven by fairness, while in other scenarios, he might be influenced by reciprocity.

Another possible solution to the public goods problem can be carried out through coalitions among individuals who agree to act collectively. Incorporation of social preferences in a public goods framework with coalition can then explain the existence of groups like IEA. Further incorporation of social preferences in a network framework can lead to interesting insights in a public goods game.

References

Andreoni, James. 1988*a*. "Privately provided public goods in a large economy: the limits of altruism." *Journal of Public Economics*, 35(1): 57–73.

Andreoni, James. 1988*b*. "Why free ride?: Strategies and learning in public goods experiments." *Journal of Public Economics*, 37(3): 291–304.

Andreoni, James. 1989. "Giving with impure altruism: Applications to charity and Ricardian equivalence." *Journal of Political Economy*, 97(6): 1447–1458.

Andreoni, James. 1990. "Impure altruism and donations to public goods: A theory of warm-glow giving." *The Economic Journal*, 100(401): 464–477.

Andreoni, James. 1995. "Cooperation in public-goods experiments: kindness or confusion?" *American Economic Review*, 891–904.

Andreoni, James, William T Harbaugh, and Lise Vesterlund. 2010. "Altruism in experiments." *Behavioural and Experimental Economics*, 6–13.

Battigalli, Pierpaolo, and Martin Dufwenberg. 2007. "Guilt in games." *American Economic Review*, 97(2): 170–176.

Blanco, Mariana, Dirk Engelmann, and Hans-Theo Normann. 2010. "A within-subject analysis of other-regarding preferences." *Available at SSRN 934700.*

Bourlès, Renaud, Yann Bramoullé, and Eduardo Perez-Richet. 2017. "Altruism in networks." *Econometrica*, 85(2): 675–689.

Bramoullé, Yann, Rachel Kranton, et al. 2007. "Public goods in networks." *Journal of Economic Theory*, 135(1): 478–494.

Caria, A Stefano, and Marcel Fafchamps. 2019. "Expectations, network centrality, and public good contributions: Experimental evidence from India." *Journal of Economic Behavior & Organization*, 167: 391–408.

Charness, Gary, and Matthew Rabin. 2002. "Understanding social preferences with simple tests." *The Quarterly Journal of Economics*, 117(3): 817–869.

Chaudhuri, Ananish. 2011. "Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature." *Experimental Economics*, 14(1): 47–83.

Chen, Xiao-Ping. 1996. "The group-based binding pledge as a solution to public goods problems." *Organizational Behavior and Human Decision Processes*, 66(2): 192–202.

Croson, Rachel TA. 2007. "Theories of commitment, altruism and reciprocity: Evidence from linear public goods games." *Economic Inquiry*, 45(2): 199–216.

Dannenberg, Astrid, Thomas Riechmann, Bodo Sturm, and Carsten Vogt. 2007. "Inequity aversion and individual behavior in public good games: An experimental investigation." *ZEW-Centre for European Economic Research Discussion Paper*, , (07-034).

Falk, Armin, and Urs Fischbacher. 2006. "A theory of reciprocity." *Games and Economic Behavior*, 54: 293–315.

Fehr, Ernst, and Klaus M Schmidt. 1999. "A theory of fairness, competition, and cooperation." *The Quarterly Journal of Economics*, 114(3): 817–868.

Fehr, Ernst, and Simon Gachter. 2000. "Cooperation and punishment in public goods experiments." *American Economic Review*, 90(4): 980–994.

Fehr, Ernst, and Urs Fischbacher. 2002. "Why social preferences matter–the impact of non-selfish motives on competition, cooperation and incentives." *The Economic Journal*, 112(478): C1–C33.

Fehr, Ernst, Urs Fischbacher, and Simon Gächter. 2002. "Strong reciprocity, human cooperation, and the enforcement of social norms." *Human Nature*, 13(1): 1–25.

Fischbacher, Urs, and Simon Gachter. 2010. "Social preferences, beliefs, and the dynamics of free riding in public goods experiments." *American Economic Review*, 100(1): 541–56.

Fischbacher, Urs, Simon Gächter, and Ernst Fehr. 2001. "Are people conditionally cooperative? Evidence from a public goods experiment." *Economics Letters*, 71(3): 397–404.

Galeotti, Andrea, Sanjeev Goyal, Matthew O Jackson, Fernando Vega-Redondo, and Leeat Yariv. 2010. "Network games." *The Review of Economic Studies*, 77(1): 218–244.

Geanakoplos, John, David Pearce, and Ennio Stacchetti. 1989. "Psychological games and sequential rationality." *Games and Economic Behavior*, 1(1): 60–79.

Grüning, Christine, and Wolfgang Peters. 2010. "Can justice and fairness enlarge international environmental agreements?" *Games*, 1(2): 137–158.

Gunnthorsdottir, Anna, Daniel Houser, and Kevin McCabe. 2007. "Disposition, history and contributions in public goods experiments." *Journal of Economic Behavior & Organization*, 62(2): 304–315.

Isaac, R Mark, and James M Walker. 1988. "Communication and free-riding behavior: The voluntary contribution mechanism." *Economic Inquiry*, 26(4): 585–608.

Isaac, R Mark, James M Walker, and Susan H Thomas. 1984. "Divergent evidence on free riding: An experimental examination of possible explanations." *Public Choice*, 43(2): 113–149.

Kim, Oliver, and Mark Walker. 1984. "The free rider problem: Experimental evidence." *Public Choice*, 43(1): 3–24.

Kolstad, Charles D. 2014. "International environmental agreements among heterogeneous countries with social preferences." *National Bureau of Economic Research.*

Kurzban, Robert, Kevin McCabe, Vernon L Smith, and Bart J Wilson. 2001. "Incremental commitment and reciprocity in a real-time public goods game." *Personality and Social Psychology Bulletin*, 27(12): 1662–1673.

Lange, Andreas, Carsten Vogt, and Andreas Ziegler. 2007. "On the importance of equity in international climate policy: An empirical analysis." *Energy Economics*, 29(3): 545–562.

Ledyard, O. 1995. "Public goods: some experimental results." *Handbook of Experimental Economics*, 1.

Marwell, Gerald, and Ruth E Ames. 1981. "Economists free ride, does anyone else." *Journal of Public Economics*, 15(3): 295–310.

Nikiforakis, Nikos, and Hans-Theo Normann. 2008. "A comparative statics analysis of punishment in public-good experiments." *Experimental Economics*, 11(4): 358–369.

Olson, Mancur. 2012. "The Logic of Collective Action [1965]." *Contemporary Sociological Theory*, 124.

Ostrom, Elinor. 1997. "Neither Markets nor States: Linking Transformation Process in Collective Action Arenas." *Perspectives on Public Choice: A handbook.*

Ostrom, Elinor. 2000. "Collective action and the evolution of social norms." *Journal of Economic Perspectives*, 14(3): 137–158.

Ostrom, Elinor, James Walker, and Roy Gardner. 1992. "Covenants with and without a sword: Self-governance is possible." *American Political Science Review*, 86(2): 404–417.

Palfrey, Thomas, Howard Rosenthal, and Nilanjan Roy. 2017. "How cheap talk enhances efficiency in threshold public goods games." *Games and Economic Behavior*, 101: 234–259.

Rabin, Matthew. 1993. "Incorporating fairness into game theory and economics." *American Economic Review*, 1281–1302.

Ringius, Lasse, Asbjørn Torvanger, and Arild Underdal. 2002. "Burden sharing and fairness principles in international climate policy." *International Environmental Agreements*, 2(1): 1–22.

Sarangi, Sudipta, and Sakshi Upadhyay. 2019. "To Join or not to Join: Coalition Formation in Public Good Games."

Smith, Adam. 1976. "The Theory of Moral Sentiments." *Liberty Classics.*

Suri, Siddharth, and Duncan J Watts. 2011. "Cooperation and contagion in web-based, networked public goods experiments." *ACM SIGecom Exchanges*, 10(2): 3–8.

Zhang, Yang. 2018. "Social Preferences in Network Games: Theory and Laboratory Evidence." *Available at SSRN 3059710.*