

# EERI

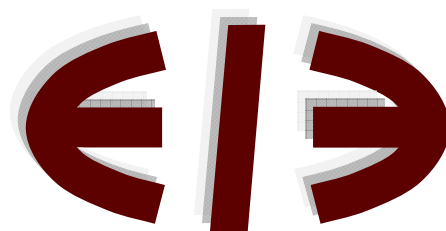
Economics and Econometrics Research Institute

## **Modeling Social Preferences: A Generalized Model of Inequity Aversion**

Hayat Khan

**EERI Research Paper Series No 21/2009**

**ISSN: 2031-4892**



**EERI**  
**Economics and Econometrics Research Institute**  
Avenue de Beaulieu  
1160 Brussels  
Belgium

Tel: +322 299 3523  
Fax: +322 299 3523  
[www.eeri.eu](http://www.eeri.eu)

# Modeling Social Preferences: A Generalized Model of Inequity Aversion

(Last Draft: May 13, 2009).

(This Draft: Jan 3, 2010)

HAYAT KHAN

*School of Economics and Finance, La Trobe University*

*Bundoora, Vic 3086*

Ph: +61 3 9479 3536

Fax: +61 3 9479 1654

Email: [h.khan@latrobe.edu.au](mailto:h.khan@latrobe.edu.au)

## ***Abstract:***

This paper argues that a generalized model of social preferences must simultaneously pass two tests; the Variety Test (explain outcomes under variety, the V-test) and the Psychological Test (conform to psychological intuition, the P-test). It is shown that none of the models proposed to date unconditionally passes these tests. The paper extends the Fehr and Schmidt model of inequality-aversion to a generalized model of inequity-aversion which parsimoniously explains interior outcomes in the *dictator game* and dynamics of outcomes in other games. This is done through introducing the equity-bias in the Fehr and Schmidt model. The paper postulates that a player's idea of equitable distribution is state-dependent, where the state is determined by psychological and structural parameters. The state could be fair, superior or inferior. Individuals in a fair state have zero equity-bias and split the pie evenly. Those in a superior (inferior) state have positive (negative) equity-bias and value more (less) than fair distribution as equitable.

***JEL Classification:*** C69; C70; C91; D03; D63; D64

***Keywords:*** *Experimental Economics; Social Preferences; Other-regarding behavior; Inequity-aversion; Altruism.*

I am thankful to Ian McDonald from the University of Melbourne for his valuable comments on an earlier draft of the paper.

## 1. Introduction

Economic agents are typically modeled as self-regarding selfish beings whose welfare is unaffected by welfare of others in the society. There have been long concerns about validity of the self-regarding assumption and calls to take *other-regarding preferences* seriously in economic theorizing. This assumption has heavily been under attack in recent years and efforts to take *other-regarding preferences* seriously escalated due to laboratory experiments in experimental and behavioral economics. These experiments suggest that individuals value fairness and some, if not all, behave altruistically and when allowed to do so, some sacrifice their self-interest to supposedly punish unfairness. This, naturally, led theoretical economists to look for theories that could explain the sort of behavior exhibited by subjects in these experiments. The main focus of this area of research has been to identify other-regarding individual-specific utility or social welfare functions that could explain patterns of behavior, such as fairness and reciprocity, observed in data from laboratory experiments. Literature in this area is expanding rapidly and so is variety of the proposed models. These models include both linear (such as [Fehr and Schmidt 1999](#); [Charness and Rabin 2002](#); [Levine 1998](#); [Rotemberg 2004](#); and [Erlei 2004](#)), and nonlinear models (such as [Bolton and Ockenfels 2000](#); [Cox et al. 2004](#); [Ottone and Ponzano 2005](#); and [Cox et al. 2007](#)).

Noting that altering the utility function allows one to explain just anything, [Camerer \(2003, p 101\)](#) writes

*“The goal is not to explain every different finding by adjusting the utility function just so; the goal is to find parsimonious utility functions, supported by psychological intuition, that are general enough to explain many phenomena in one fell swoop, and also make new predictions”*

Models proposed in this area of research however haven't been subjected to any formal tests to analyze their ability to explain many phenomenon in one fell swoop and ensure their conformity to psychological intuition. This paper formalizes Camerer's idea in the form of two theoretical tests that can be used to judge models of other-regarding preferences. The two tests are (i) the Variety test (V-Test) and (ii) the Psychological test (P-Test). This paper applies these tests to representative models of the literature and demonstrates that none of the models proposed to date passes these tests simultaneously. They fail to support psychological intuition when put to such a theoretical test, and most are not general enough to explain many phenomena in one fell swoop and do well in experiments of specific designs only. The paper then proposes a generalized model of *inequity-aversion* that passes the two tests simultaneously. This is done through introducing *equity-bias* in the [Fehr and Schmidt \(1999\)](#) model.

Individuals in the Fehr and Schmidt's model are inequality averse. [Fehr and Schmidt \(1999, p.822\)](#) justifies inequality-aversion by arguing that subjects enter the laboratory as equals who are randomly assigned different roles and do not know anything about each other. This paper highlights the difference between inequity and inequality and argues that individuals are inequity-averse than inequality-averse. This is because subjects cease to be equal once assigned into

different roles and given unequal property rights through whatever procedure. A dictator, for example, with all the power to give something or nothing to another player is not equal to the passive recipient who has no claim over the amount given to him/her. This bias in roles leads to bias in equity which is referred to as equity-bias. It argues that, in general, an individual's idea of equitable distribution is state-dependent where state is determined by psychological and structural parameters. The state could be fair, superior or inferior. Individuals in a fair state have zero *equity-bias* and split the pie evenly. Those in a superior (inferior) state have positive (negative) *equity-bias* and value more (less) than fair distribution as equitable distribution. *i.e.* bias in state leads to bias, not necessarily with the negative connotation, in equity. Given psychological tendencies of an individual, every experimental design or real world scenario assigns one of the three states to the player which leads to individual-specific valuation of equity and equitable outcomes.

Before going into details of the paper, let me point out that this paper will limit its discussion to the two most frequently used standard games, namely, the *dictator game* (DG) and the *ultimatum game* (UG). In the DG, a player called the *dictator* is given a certain amount of money with the option to share it with another player, the *recipient*, in any proportion, 0 to 100 percent, inclusive. In the UG, the dictator is lowered to the status of a *proposer* and the *recipient* is alleviated to the status of a *responder* who is allowed to either accept or reject a distribution proposed by the proposer. When accepted, each player keeps his/her share and when rejected both players get zero.

The model is developed in a two-player environment but results are applicable to games with multiple players competing or otherwise.

The rest of the paper is organized as follows. Section 2 introduces the two theoretical tests that can be used to judge models of other-regarding preferences. Section 3 provides a review of selected models representative of previous literature and evaluates their performance against the V-test and the P-test. Section 4 provides a discussion on equity vs equality and spells out the concept of *equity-bias*. Section 5 formally introduces the *state-dependent/equity-bias* model of inequity aversion. This section details dynamics of the model in the dictator and *ultimatum games*. Section 6 gives a brief comparison of the generalized model with those reviewed in Section 4. Section 7 discusses application of the model to different versions of the dictator and *ultimatum games* experiments conducted in the literature. Finally, Section 8 concludes the paper.

## **2. Theoretical Tests for Judging Models of Other-regarding Preferences**

A number of theoretical models have been proposed to explain the patterns of behavior observed during laboratory experiments. The number has continually been increasing and so has variety of the models proposed. This section identifies two theoretical tests that can be used to decide whether or not a mathematical model can be accepted as a generalized model of other-regarding preferences. These tests are discussed below.

## 2.1. The Variety Test (V-Test)

Variety of the laboratory experiments conducted in this area has grown richer and different versions of the games have been experimented with. A reasonable model should be general enough to perform consistently across (i) different/similar designs of the same game (e.g. across different versions of the DG and UG) and (ii) across different games of the same nature (e.g. across the dictator, ultimatum and impunity games). This test has been applied, not under the name V-test though, in the literature, as we will note in our discussion later.

## 2.2. The Psychological Test (P-Test)

As mentioned earlier, the main focus of the models in this area of research has been to explain patterns of the behavior observed during laboratory experiments. Whether or not a mathematical model that is able to explain outcomes in laboratory experiments conforms to psychological intuition is something that is often ignored and not tested for. The question however to ask is, is there any such test that can be used, together with some other test(s), to give a mathematical model the status of an other-regarding welfare function? Fortunately, there is at least one, Sen's Weak Equity Axiom. Think of two individuals, one normal and the other with a disability. The disabled person is less efficient in converting a dollar into utility relative to the normal person. Psychological intuition tells us that when distributing a given sum, one should be more altruistic towards the disabled person than towards the normal person. This is Sen's Weak Equity Axiom. The disability can be interpreted in general as poor socio-economic status. An even weaker version of the axiom is that the disabled person should be given at least as much as the normal person. Conformity to Sen's WEA can be checked through the sign of partial derivative of equilibrium pay-offs with respect to dollar-to-utility conversion efficiency of a player.

Formally, let  $V_i = V_i(u(x_i), u(x_j/\chi))$  be the social utility of an individual  $i$  where  $x_i$  is the pay-off of an individual  $i$ ,  $u(x_i)$  the selfish utility of  $i$  from  $x_i$ ,  $x_j$  is the pay-off of individual  $j$  and  $u(x_j/\chi)$  the selfish utility of individual  $j$  from  $x_j$ .  $\chi$  quantifies the inefficiency of individual  $j$  in converting  $x_j$  into utility relative to  $i$ . The larger the value of  $\chi$  the less efficient  $j$  is in converting a dollar into utility. The player's objective is to maximize  $V_i$  subject to a constraint (such as  $x_i + x_j = N$  where  $N$  is size of the pie). Let  $x_i^*$  and  $x_j^*$  be the equilibrium pay-offs after maximization. Sen's WEA requires  $d(x_j^*)/d\chi > 0$ .

In fact the P-test can be thought of as a special case of the V-test. It is a V-test that invokes variations in  $\chi$  as the change in structural parameters. Experiments show that Sen's WEA is satisfied. [Eckel and Grossman, \(1996\)](#) found a significant increase in donations in a *dictator game* when they increased the extent to which a donation goes to a "deserving" recipient. [Brañas-Garza \(2006\)](#) performed three different *dictator games*; the standard DG and two versions of the DG with poverty where the dictator is informed that their recipients were poor. Garza found that giving in the poverty game was significantly higher than in the standard DG.

Notice that social utility of the individual  $i$ ,  $V_i$ , is related to pay-offs through idiosyncratic selfish utility ( $u_i$  and  $u_j$ ) and is not expressed as a direct function of the pay-offs alone. This distinction is important from point of view of the application of the P-test.

### 3. Literature Review and Application of the Two Tests

This section reviews selected models of other-regarding preferences. These models include the FS model (Fehr and Schmidt, 1999), the BO model (Bolton and Ockenfels, 2000) and the CFG model (Cox, and Friedman and Gjerstad, 2007). The following justifies selection of the models in this section.

The model in this paper is primarily an extension of the FS model. The paper therefore starts the review with the Fehr and Schmidt (1999) model which is representative of piecewise linear models of inequality aversion. Results of the analysis therefore apply to any model with inequity modeled as inequality, such as Charness and Rabin (2002) and Erlei (2004). Appendix A demonstrates as an example that the FS model, or a restricted version of it, can be derived as a monotonic transformation of Charness and Rabin (2002). The same can be done for other piecewise linear models of the same sort such as Kohler (2003).

As well known, the FS model cannot explain interior solutions in the DG. Fehr and Schmidt attribute this flaw to the piecewise linearity of preferences in advantageous inequity. They claim that modifying their social welfare function to introduce non-linearity in the advantageous inequity could resolve the issue (Fehr and Schmidt, 1999 p. 823). Whereas this is true, this paper shows that nonlinearity is not necessarily needed to explain interior outcomes. The BO model can be shown to be the FS model with the proposed non-linearity. This is the reason why this model is reviewed as well and will be analyzed in the next section. Although I do not review other nonlinear models, the analysis in this paper is also applicable to other nonlinear models such as Ottone and Ponzano (2005).

Finally, most recently, Cox *et al.* (2007) proposed a parametric model of other-regarding preferences which depends on status, reciprocity, and perceived property rights. This allowed them to make distribution state-dependent, unlike the unconditional distributional preferences in the FS type models, which makes it similar to the model proposed in this paper. The CFG model is shown to be a special case of the generalized model when some additional restrictions are imposed and is therefore reviewed and analyzed in this section.

Apart from the distributional preference models above, as mentioned by Cox *et al.* (2007), there are alternative models of intention-based reciprocity (such as Rabin 1993; Dufwenberg and Kirchsteiger 2004; and Falk and Fischbacher 2001). These alternative models are complex and have many equilibria which makes them intractable in most applications (Cox. *et al.* 2007). Although I do not review intention-based models, the model proposed in this paper captures dynamics generated by intentions and encompasses this class of models as well. The analysis on the FS model in this section also applies to Levine (1998) who developed a simple tractable model with intentions where utilities are linear in one's own and the other's pay-offs.

#### 3.1. The FS (Fehr and Schmidt 1999) model

Fehr and Schmidt (1999) used a simple linear model to explain results of laboratory experiments including the *dictator* and *ultimatum games*. They modeled fairness as self-centered inequity- in fact inequality- aversion, whereby people are willing to sacrifice part of their material pay-offs to move in the direction of equality. Assuming two players in the game, their social welfare

function of an individual  $i$  is linearly increasing in self-pay-off,  $x_i$ , decreasing in advantageous and disadvantageous inequality,  $i.e.$

$$V_i = x_i - \beta_i \max(x_i - x_j, 0) - \alpha_i \max(x_j - x_i, 0) \quad 0 \leq \beta_i \leq 1 \text{ and } \alpha_i \geq \beta_i \quad (1)$$

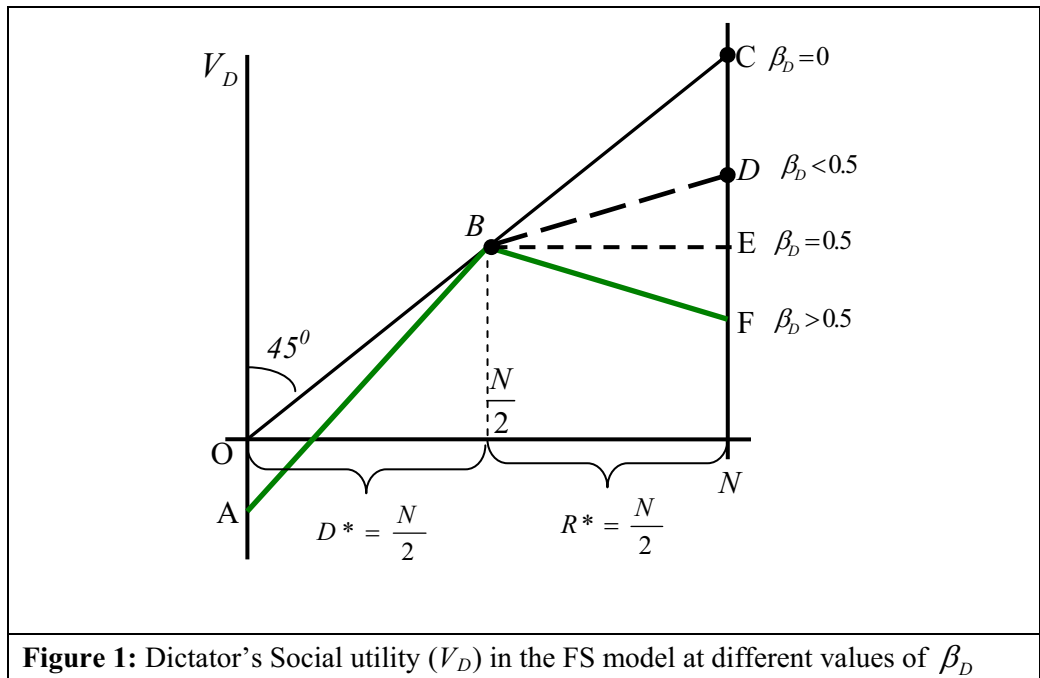
$\alpha_i$  is the social marginal utility when disadvantageous inequality ( $x_i - x_j < 0$ ) reduces by 1 unit. Similarly,  $\beta_i$  is the social marginal utility when advantageous inequality ( $x_i - x_j > 0$ ) decreases by 1 unit.  $\alpha_i > \beta_i$  implies that a unit increase in disadvantageous inequality hurts  $i$  more than a unit increase in advantageous inequality would.

Consider application of the model to the DG. Assuming  $D$  and  $R$  to be pay-offs of the dictator and recipient respectively, we can write social utility of the dictator, based on equation (1), as:

$$V_D = D - \beta_D \max(D - R, 0) - \alpha_D \max(R - D, 0)$$

The dictator's objective is to maximize  $V_D$  subject to  $D + R = N$ . Figure 1 plots  $V_D$  at different values of  $\beta_D$ . The following holds

$$V_D = \begin{cases} ABC & \text{when } \beta_D = 0 \\ ABD & \text{when } 0 < \beta_D < 0.5 \quad C < D < E \\ ABE & \text{when } \beta_D = 0.5 \\ ABF & \text{when } \beta_D > 0.5 \quad F < E \end{cases}$$



**Figure 1:** Dictator's Social utility ( $V_D$ ) in the FS model at different values of  $\beta_D$

Let  $D^*$  and  $R^*$  denote the values of  $D$  and  $R$  where  $V_D$  is maximum. The solution is easy to work out from the graph.

$$\left(\frac{D^*}{N}, \frac{R^*}{N}\right) = \begin{cases} (1, 0) & \text{when } \beta_D = 0 & \text{(point C)} \\ (1, 0) & \text{when } 0 < \beta_D < 0.5 & \text{(point D)} \\ (\in [0.5, 1], \in [0.5, 0]) & \text{when } \beta_D = 0.5 & \text{(line BE)} \\ (0.5, 0.5) & \text{when } \beta_D > 0.5 & \text{(point B)} \end{cases} \quad (2)$$

The solution is either  $D^*=N$  (with  $\beta_D < 0.5$ ),  $D^*=N/2$  (with  $\beta_D > 0.5$ ) or the range  $\frac{N}{2} \leq D^* \leq N$  (with  $\beta_D = 0.5$ ) with indifference. The range is not an equilibrium in a strict sense. All it says is that any value within the range is equally good and welfare-maximizing. The FS model can therefore only explain corner solutions ( $D=N$  and  $N/2$ ) in the DG, leaving the interior unexplained.

Now consider application of the model to the UG. Assuming  $P$  and  $R$  to be the pay-offs of the proposer and responder respectively, we can write social utility of the proposer and responder, based on equation (1), as:

$$V_P = P - \beta_P \max(P - R, 0) - \alpha_P \max(R - P, 0) \quad (3)$$

$$V_R = R - \beta_R \max(R - P, 0) - \alpha_R \max(P - R, 0) \quad (4)$$

The proposer's objective is to maximize  $V_P$  subject to  $P+R=N$  and make an offer to the responder good enough to make the responder's social utility non-negative, i.e. the proposer's objective is therefore to maximize  $V_R$  subject to  $P+R=N$  and  $V_R = 0$ . They are happy so long as their social utility is non-negative.  $V_R = 0$

requires  $R = r_{o,i}N$  where  $r_{o,i} = \frac{\alpha_R}{(1+2\alpha_R)}$  is the minimum acceptable offer by

individual  $i$ . Since we may have individuals with different  $\alpha_R$  ( $0 \leq \alpha_{RH} \leq \alpha_R \leq \alpha_{RH} < \infty$ ), corresponding to  $0 \leq r_{o,L} \leq r_{o,i} \leq r_{o,H} < 0.5$ , any offer within the lower and upper bound will be accepted with certain probability. Offers below  $r_{o,L}$  will be rejected and those equal to or greater than  $r_{o,H}$  will be accepted with certainty. This is depicted in Figure 2.  $p_{o,L} = 1 - r_{o,L}$  and  $p_{H,L} = 1 - r_{H,L}$ . The

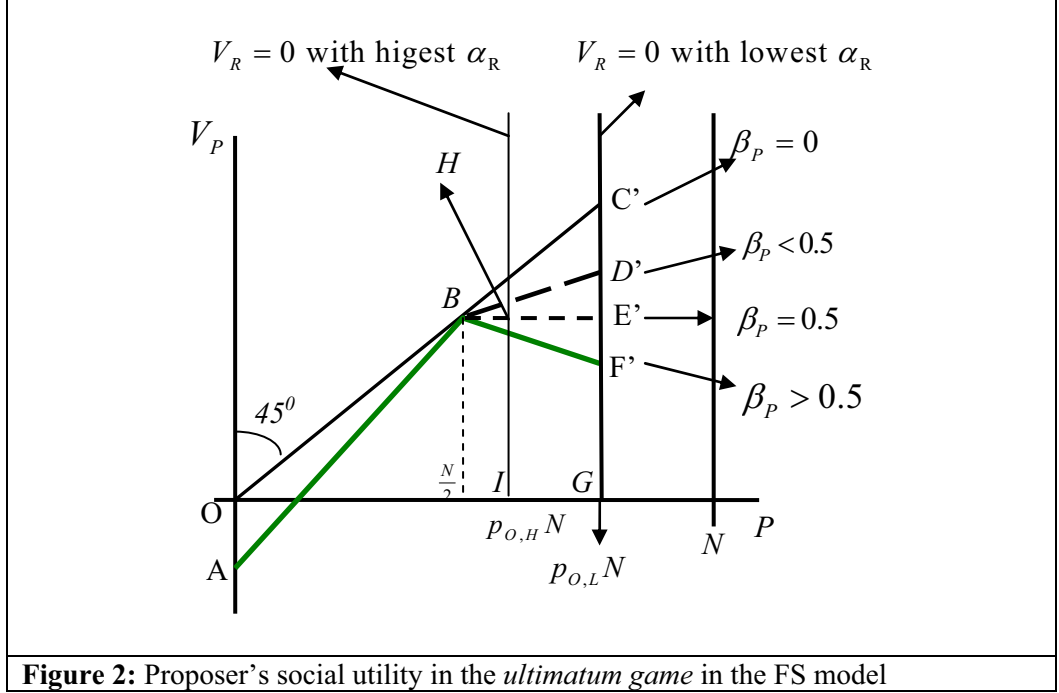
following holds under different values of  $\beta_P$  when constraints are implemented:

$$V_P = \begin{cases} ABC' & \text{when } \beta_P = 0 \\ ABD' & \text{when } 0 < \beta_P < 0.5 & C' < D' < E' \\ ABHE' & \text{when } \beta_P = 0.5 \\ ABF' & \text{when } \beta_P > 0.5 & F' < E' \end{cases}$$

The solution lies on the additional constraint  $V_R=0$  when  $\beta_P < 0.5$ , which implies an equilibrium offer  $R^0 = r_{o,i}N$ , on the line BH with indifference when  $\beta_P = 0.5$ , and at point B when  $\beta_P > 0.5$ . The reason why the solution of  $\beta_P = 0.5$  excludes points on HE' is that any offer on HE' has some probability of rejection, as a result the individual with  $\beta_P = 0.5$  will not be indifferent between BH and HE'. BH will be preferred to HE'.



$$\left(\frac{P^o}{N}, \frac{R^o}{N}\right) = \begin{cases} (\in [p_{o,L}, p_{o,H}], \in [r_{o,L}, r_{o,H}]) & \text{when } \beta_p \leq 0.5 \\ (\in [0.5, p_{o,H}], \in [0.5, r_{o,H}]) & \text{when } 0 < \beta_p = 0.5 \text{ (line BH with indifference)} \\ (0.5, 0.5) & \text{when } \beta_p > 0.5 \text{ (point B)} \end{cases} \quad (5)$$



**Figure 2:** Proposer's social utility in the *ultimatum game* in the FS model

The solution in UG is similar to the one in the DG (compare equation (5) with (2) and Figure 2 with Figure 1). The only difference is that the corners are now determined by the additional constraint  $R=r_{o,i}N$  rather than  $R=0$ . Risk-averse proposers will choose to offer closer to  $r_{o,H}$  and relatively risk lovers would offer close to  $r_{o,L}$ .

Let us call  $\beta_p = 0$ , selfish-regarding (SR),  $0 < \beta_p < 0.5$  not so SR or weakly other-regarding (WOR),  $\beta_p = 0.5$  equal regarding (ER) and  $\beta_p > 0.5$  strongly other-regarding (SOR). The following conclusions inevitably follow from the FS model;

- If there was no fear of rejection at  $R=0$ , the solution according to Fehr and Schmidt's formulation will be exactly the same as the DG, the corner solution. This leaves the interior solution unexplained.
- When there is fear of rejection the right corner solution on  $R=r_{o,i}N$  corresponds to  $0 \leq \beta_p < 0.5$  and it is hard to distinguish between SR and WOR individuals. Thus this solution wouldn't tell us whether preferences are weakly other-regarding or selfish.
- Knowing that an offer  $r=0.5$  will never be rejected, the SOR  $\beta_p > 0.5$  solution ( $r=0.5$ ) can only be interpreted as an altruistic solution, which is again a corner solution similar to the SOR dictator.
- Coming to the ER individual, the solution is again a range  $[0.5, r_{o,H}]$  with indifference which is not a unique interior solution.

What points (a) to (d) tell us is that, according to the FS model, a solution  $r=0.5$  can be interpreted as a solution with altruism if  $r_{O,H} < 0.5$ , may possibly be interpreted as a solution with fear of rejection (reciprocity) if for some individuals  $r_{O,H} = 0.5$ . In the first case preferences can be termed as SOR whereas in the second case they may or may not be. Given that in the standard dictator experiment an  $r \geq 0.5$  is not always observed,  $\beta_p > 0.5$  is not dominantly true. This means that the preferences of the proposer are either SR or WOR. i.e.  $0 \leq \beta_p < 0.5$ . Since SR and WOR have the same solution, the use of other-regarding preferences for the proposer is irrelevant. The problem with the ER solution is obvious. Indifference means any choice is equally good, which is not an explanation. Thus interior solutions in the UG in the FS model can be interpreted as belonging to a selfish (or WOR) proposer with inequality-averse responders. The UG solution of the FS model explains interior offers by proposers as a result of fear of rejection, leaving the altruistic part unexplained.

Coming to application of the V-test and the P-test, the FS model does not withstand the V-test. Firstly, it fails to explain interior solutions in the DG. Secondly, it fails to explain outcomes across different designs of the UG. If you modify the UG such that if a responder rejects an offer, 2 units are subtracted from all pay-offs, the model predicts that responders should never reject unequal offers (Camerer 2003, p. 112). This is because rejection would reduce pay-off of the responder without any effect on inequality. When individuals are inequality-averse in pay-offs the model automatically fails the P-test. However when we replace *inequality-aversion* in pay-offs by *inequality-aversion* in utility, the P-test is qualified for a special case of equal distribution of utility. This can be ascertained by replacing  $x_j$  in the FS model with  $x_j/\chi_i$  in equation (1), which becomes a special case of the model ( $e_i=1$ , implying equality in utility) in this paper and passes the P-test, as we shall see later.

### 3.2. The BO (Bolton and Ockenfels 2000) Model

The social welfare function of an individual  $i$  in the BO model is given (ignoring the  $x_i+x_j=0$  case) by

$$V_i = a_i x_i - \frac{b_i}{2} \left( \frac{x_i}{x_i + x_j} - \frac{1}{2} \right)^2 \quad a_i \geq 0 \quad \text{and} \quad b_i > 0$$

As  $x_i + x_j = N$ , it can be reduced to

$$V_i = a_i x_i - \frac{b_i}{8N^2} (x_i - x_j)^2 \quad (6)$$

Notice that  $V_i$  is linearly increasing in  $x_i$  and non-linearly decreasing in inequality. Recall that Fehr and Schmidt (1999) blamed linearity of the social utility function in advantageous inequity for the inability of their model to explain interior outcomes in the DG. (6) is nonlinear in inequality. The BO model is therefore the FS model with nonlinear inequality aversion.

Consider application of the model to the DG. The dictator's objective is to maximize (6) subject to  $x_i + x_j = N$ , which gives  $x_i^* = \frac{1}{2}N + \frac{a_i}{b_i}N^2$  and  $x_j^* = \frac{1}{2}N - \frac{a_i}{b_i}N^2$ .

The first problem with the solution is obvious. Pay-offs are non-linearly related to the size of the pie. Whereas this may be seen desirable at very high values of  $N$ , experiments with relatively higher  $N$  show that outcomes are not too sensitive to the value of  $N$  (Camerer 2003).

This model can explain interior solutions in the standard DG, it cannot, however adequately explain variety of the *dictator games* played (the V-test). For example consider the DG where individual  $i$  and  $j$  are given 10 dollars each. The dictator is given the option of either sharing his/her 10 dollars with the recipient or taking some of the recipient's money, as in Bardsley (2005). The BO model will give the same solution as the standard DG. These games do not give the same results in practice. Similarly, Camerer (2003, p. 111) notes that if you modify the UG such that when the responder rejects a proposal, the monetary payoffs are 10 per cent of the original offer, relative shares will be the same no matter whether the responder accepts or rejects. Thus the responder will always accept any offer; no matter how unequal it is (since the utility from rejection will be lower than the utility with acceptance). This shows that the model fails to pass the V-test.

It is a bit hard to implement the  $P$ -test here as the model is defined over pay-offs, not utility. One possibility is to think of  $x_i$  and  $x_j$  as utilities and replace  $x_j$  with  $x_j/\chi$ . The equilibrium value of  $x_j^*$  is ambiguously related to  $\chi$  but increasing with  $\chi$  at relatively high  $a_i/b_i$  and decreasing at relatively lower values of  $a_i/b_i$ .

### 3.3. The CFG (Cox, Friedman and Gjerstad 2007) Model

The model is given by

$$V_i = \frac{1}{\sigma} [x_i^\sigma + \theta_i x_j^\sigma] \quad \text{where } \sigma \leq 1$$

$$\theta_i = \theta_0 + \theta(r, s) + \varepsilon_i, \quad \theta(0, 0) = 0$$

where  $r$  is the reciprocity and  $s$  is the state variable. When  $r=s=0$ ,  $\theta_i = \theta_0 + \varepsilon_i$  which could be positive, negative or zero, implying a benevolent, malevolent or selfish player. Consider application of the game to the DG with status.  $\theta_i = \theta_0 - as + \varepsilon_i$ .  $s=1$  when individuals earn property rights to the sum  $N$  and  $0$  if the sum was *mana* from experimental heaven. The dictator's objective is to

maximize  $V_i = \frac{1}{\sigma} [(x_i)^\sigma + \theta_i (N - x_i)^\sigma]$  which gives  $x_i^* = \left[ \frac{1}{1 + \theta_i^{1-\sigma}} \right] N$  and

$$x_i^* = \left[ \frac{\theta_i^{\frac{1}{1-\sigma}}}{1 + \theta_i^{\frac{1}{1-\sigma}}} \right] N.$$

Notice that, depending upon the value of  $\theta_i$  and  $\sigma$ , the model explains the corner as well as interior solution observed in laboratory experiments.

Empirical estimates of the model in Cox *et al.* (2007) yield  $\sigma < 1$  and  $a > 0$ . This implies that  $\frac{dx_i}{ds} > 0$ . Intuitively the dictator decreases his/her weight  $\theta_i$  on utility of the recipient when  $s=1$ . This implies that the model passes the V-test.

In its application to the DG and UG, Cox *et al.* (2007, p. 25 and 32) restrict  $\sigma$  to be  $\geq 0$  (by mistake<sup>1</sup>). Cox *et al.* (2007) empirically estimate the value of  $\sigma$  and find that it is  $>0$  and  $<1$  in all empirical estimations<sup>2</sup>. To see if the model passes the P-test at the restricted values of  $\sigma$ , let us consider an alternative scenario where the recipient is less efficient in deriving satisfaction from a unit of pay-off than the dictator. The model can be reproduced as

$$V_i = \frac{1}{\sigma} \left[ (x_i)^\sigma + \theta_i \left( \frac{N - x_i}{\chi} \right)^\sigma \right]. \text{ Maximization yields } x_i^* = \left[ \frac{\chi^{\frac{\sigma}{\sigma-1}}}{\chi^{\frac{\sigma}{1-\sigma}} + (\theta_i)^{\frac{1}{1-\sigma}}} \right] N \text{ and}$$

$$x_j^* = \left[ \frac{(\theta_i)^{\frac{1}{1-\sigma}}}{\chi^{\frac{\sigma}{1-\sigma}} + (\theta_i)^{\frac{1}{1-\sigma}}} \right] N.$$

$$\frac{dx_j^*}{d\chi} < 0 \text{ when } 0 < \sigma < 1, \text{ implying that the dictator will be less altruistic}$$

towards the less efficient individuals, which is the opposite of what psychological intuition would predict<sup>3</sup>. Thus the model fails the P-test at the estimated value of  $\sigma$ . Qualification of the P-test requires  $\sigma < 0$ .

It will become clear later on that the CFG model with  $\sigma < 0$  is a special case of the model proposed in this paper.

## 4. Equity-bias: Equity vs Equality

Most of the *inequity-aversion* models, with rare exceptions such as the CFG, are models of *inequality-aversion* than inequity aversion. Fehr and Schmidt (1999) note:

*“The determination of the relevant reference group and the relevant reference outcome for a given class of individuals is ultimately an empirical question. The social context, the saliency of particular agents, and the social proximity among individuals are all likely to influence reference groups and outcomes. Because in the following we restrict attention to individual behavior in economic experiments, we have to make assumptions about reference groups and outcomes that are likely to prevail in this context. In the laboratory it is usually much simpler to define what is perceived as an equitable allocation by the subjects. The subjects enter the laboratory as equals, they do not know anything about each other, and they are allocated to different roles in the experiment at random. Thus, it is natural to assume that the reference group is simply the set of subjects*

<sup>1</sup> Their argument is that for  $\sigma < 0$ , utility is  $-\infty$  at  $x_j=0$  (which is observed in the data) so we must have  $\sigma \geq 0$ . This condition is not necessary as the solution can be explained by a  $\theta_i=0$  for any  $\sigma$  or by  $\theta_i < 1$  with  $\sigma=1$ .

<sup>2</sup> The estimated value of  $\sigma$  changes in these estimations which needs justification.

<sup>3</sup> This points towards issues related to the estimation procedures adopted in Cox et al. 2007, which is worth investigating but is outside the scope of the paper.

*playing against each other and that the reference point, i.e., the equitable outcome, is given by the egalitarian outcome” (Fehr and Schmidt 1999, pp 821-22)*

Assuming individuals to be equal during experiments is an oversimplification which is not true in all contexts. Once an experimenter assigns subjects into different roles, through whatever procedure, and gives them unequal property rights, they cease to be equal. For example, a dictator with all the power to give something or nothing to another player is not equal to the passive recipient who has no claim over the amount given to him/her. There are biases that are hard-wired into them, such as self-serving bias, which interact with the subjects’ heterogeneous socio-economic background and real-life experiences (and their experience in the laboratory); this in turn affect their choices when given the power to do so. Bias in roles therefore leads to bias in ones idea of equitable distribution. It is therefore not right to assume that other-regarding individuals are universally inequality-averse irrespective of context. A reasonable postulation, in general, would be to think of one’s idea of equitable distribution as state-dependent, where the state is determined by psychological and structural parameters. Psychological parameters primarily include tendencies that are either hard-wired into human nature, such a self-serving bias, or a part of one’s culture; ones real-life experiences (including experience in the laboratory); one’s perception about the behavior of others (kind, selfish etc); and one’s perception about his/her socio-economic status relative to those of others. The “structural parameters” mainly relate to design of the experiment (such as how are different roles and property rights allocated, information about the socio-economic status of players, wording of the experiment, and role of the experimenter etc).

Let us define three different states, Fair (F), Superior (S), and inferior (INF). Given individual-specific psychological tendencies/parameters, every experimental design assigns one of the three states to players. When assigned a fair state, one’s idea of equity is a fair one (*i.e.* equal distribution); and when assigned a superior state (S) or inferior (INF) states, one’s idea of equity is a biased one. Let  $E_i = x_i/x_j$  be the measure of equity of an individual  $i$  over  $x_i$  (own pay-off) and  $x_j$  (other’s pay-off). When  $i$  is assigned a fair state, F, his/her idea of an equitable distribution is a fair one, *i.e.*  $E_i=1$ . However, when assigned a biased state, S or INF, he/she is emotionally locked into choosing a biased  $E_i=1+b_i$  where  $b_i$  quantifies *equity-bias* and assumes a non-zero value. In state S,  $b_i>0$  and  $i$  values more than fair distribution as equitable distribution. Similarly,  $b_i<0$  in state INF and  $i$  accepts/values less than fair distributions as equitable.

For example, in the DG, a dictator is assigned a state superior than the one assigned to the recipient. The dictator owns all the money and is assigned the right to use it as he/she pleases. The recipient is neither a party to the “production” of value nor legally entitled to receive any share of the money<sup>4</sup>. Thus, when behaving altruistically, he/she does not necessarily find it equitable to split the pie equally. Competition basically changes the relative location of players on the equity-bias chart, and hence their idea of equitable share. Similarly, a recipient in the *ultimatum game* is assigned an inferior state, and may accept less than fair offers as equitable. This plays an important role in explaining dynamics of the outcomes in variety of the two games, and other games of the same sort, as we shall see later.

---

<sup>4</sup> Even though it is possible for him/her to think of the sum as a result of the experiment to which the other player is also a party to some extent, which morally entitles him/her to some share.

## 5. The Model

With the above discussion in mind, let us generalize the FS model of inequality-aversion to a model of inequity-aversion.

Let

- $x_i$  = pay-off of individual  $i$
- $u_i(x_i)$  = idiosyncratic selfish utility function of individual  $i$  from own pay-off  $x_i$
- $u_{j,i}(x_j)$  = selfish utility of individual  $j$  as perceived by individual  $i$  from pay-off of individual  $j$   $x_j$ .
- $e_i$  = the equitable distribution of utility as perceived by individual  $i$  expressed as own-utility relative to utility of the other individual

The generalized FS model is given below

$$V_i = u_i(x_i) - \beta_i \max[u_i(x_i) - e_i u_{ji}(x_j), 0] - \alpha_i \max[e_i u_{ji}(x_j) - u_i(x_i), 0] \quad (7)$$

where  $u_i' > 0$ ,  $u_{ij}' > 0$ ,  $u_i'' \leq 0$ ,  $u_{ij}'' \leq 0$ ,  $\beta_i = \frac{\gamma_i}{1 + e_i \left( \frac{u_{ij}'}{u_i'} \right)}$  and  $\gamma_i > 1$  and  $\alpha_i \geq \beta_i$ .

The restrictions on the first and second derivatives of the idiosyncratic selfish utilities imply diminishing/constant marginal utility of money.  $\beta_i$  is the social marginal utility of a unit decrease in advantageous inequity referred to as the *social beta* of individual  $i$ . Similarly,  $\alpha_i$  is the social marginal utility of a unit decrease in disadvantageous inequity referred to as *social alpha*. The logic behind the value of  $\beta_i$  and the restriction on  $\gamma_i$  will be justified later in the paper. It might however be helpful to point out that this restriction is needed to explain interior outcomes and ensures that social utility is maximum at the equitable distribution of pay-offs. In the FS model this restriction was  $\beta_i > 1/2$ , which in our notation is equivalent to  $\beta_i = \frac{\gamma_i}{2}$  with  $\gamma_i > 1$ . The restriction on  $\beta_i$  could be

replaced with  $\beta_i > 1$  (as  $\beta_i > 1$  always implies  $\beta_i > \left[ 1 + e_i \left( \frac{u_{ij}'}{u_i'} \right) \right]^{-1}$ , see Appendix B

for proof). In order to understand the dynamics generated by the value of social beta and its interpretation, I will stick to the general restriction on *social beta* as above.

Without loss of generality, let us impose constant marginal utility of money and assume:

$$u_i(x_i) = f(x_i) = x_i \quad (8)$$

$$u_{ji}(x_j) = g(x_j) = f\left(\frac{x_j}{\chi_i}\right) = \left(\frac{x_j}{\chi_i}\right) \quad (9)$$

where  $\chi_i > 0$ ,  $f(x_i)$  is the preference technology of individual  $i$ , which converts  $x_i$  into utility and  $g_i(x_j)$  is the preference technology of individual  $j$ , as perceived by individual  $i$ , which converts  $x_j$  into the utility of  $j$ . The fact that the preference technology  $f$  may not necessarily be equal to  $g$  captures dynamics of differences in socio-economic status of players that leads to *equity-bias* in pay-offs. In the specification in equation (9),  $\chi_i$  embodies individual  $i$ 's belief about the socio-economic status of individual  $j$  relative to his/her own socio-economic status. For example, when individual  $i$  believes he/she is socially better off than individual  $j$  (say because of  $j$ 's disability or because  $j$  is poor relative to  $i$ ),  $\chi_i$  is greater than 1.  $\chi_i > 1$  implies that  $i$  believes that  $j$  is less efficient in deriving satisfaction from a dollar than  $i$  him/herself or a dollar given to  $j$  generates lesser utility than it does to  $i$  (vise-versa for  $\chi_i < 1$ ). Appendix C derives the utility function in equation (7) as a monotonic transformation of a social utility function that explicitly models differences in socio-economic status<sup>5</sup>.

Substituting (8) and (9) in (7) gives

$$V_i = x_i - \beta_i \max[x_i - E_i x_j, 0] - \alpha_i \max[E_i x_j - x_i, 0] \quad (10)$$

where  $\beta_i = \frac{\gamma_i}{1 + E_i}$ ,  $\gamma_i > 1$  and  $E_i = \left( \frac{e_i}{\chi_i} \right)$ .

$E_i$  is the equitable distribution of pay-offs (not utility) as perceived by individual  $i$  and  $e_i$  is an equivalent measure in terms of utility. The value of this is determined by *equity-bias* as  $E_i = 1 + b_i$  where  $b_i$  is equity-bias as discussed earlier. When an individual is assigned a fair state, the value of  $E_i$  is equal to 1 ( $b_i = 0$ ), when assigned a superior state  $E_i > 1$  ( $b_i > 0$ ) and when assigned an inferior state  $E_i < 1$  ( $b_i < 0$ ).

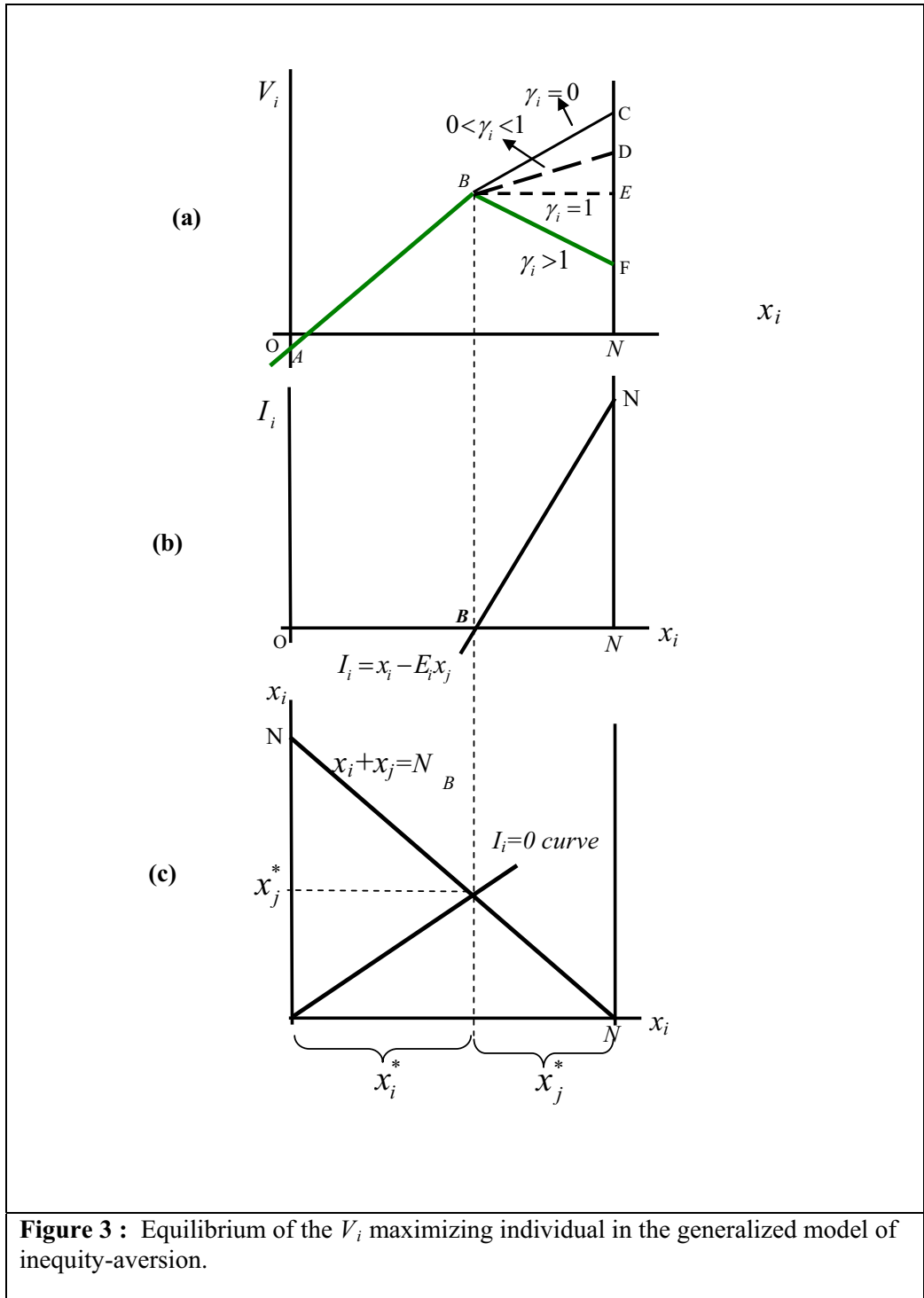
Notice that I mainly define states and bias in terms of pay-offs for convenience. It is possible for these states to be different in terms of utility. For example, consider a fair state in utility (i.e.  $e_i = 1$ ), which implies  $E_i = 1/\chi_i$ .  $E_i$  could be less than, equal to or greater than one (depending upon whether  $\chi_i$  is greater, equal to, or less than one respectively) which corresponds to positive, zero, and negative equity bias respectively. Thus one can think of an individual belonging to a superior state in utility and possibly belonging to any of the three states in pay-offs and vise versa.

The players' objective is to maximize (10) subject to  $x_i + x_j = N$ . Figure 3 panel (a) plots  $V_i$  at different values of  $\gamma_i$ . The following holds:

$$V_i = \begin{cases} ABC & \text{when } \gamma_i = 0 \\ ABD & \text{when when } 0 < \gamma_i < 1 & C < D < E \\ ABE & \text{when } \gamma_i = 1 \\ ABF & \text{when } \gamma_i > 1 & F < E \end{cases}$$

---

<sup>5</sup> Note that laboratory experiments are conducted in a controlled environment; it is possible for the impact of these problems to be minimized, if not completely avoided. When experiments are blind or double blind, self-serving bias may still be at play.



The solution can be summarized thus:

$$\left( x_i^*, x_j^* \right) = \begin{cases} (1, 0) & \text{when } \gamma_i = 0 \quad (\text{point C}) \\ (1, 0) & \text{when } 0 < \gamma_i < 1 \quad (\text{point D}) \\ \left( \in [x_i^*, 1], \in [N - x_i^*, 0] \right) & \text{when } \gamma_i = 1 \quad (\text{on line BE with indifference}) \\ (x_i^*, N - x_i^*) & \text{when } \gamma_i > 1 \quad (\text{on point F}) \end{cases}$$



where  $x_i^* = \left[ \frac{E_D}{1+E_D} \right] N$  as discusses below.

With  $\gamma_i > 1$ , the problem reduces to minimization of the deviation of the inequity gap from zero,  $I_i = x_i - E_i x_j$ , which occurs at  $I_i = 0 \Rightarrow x_i = E_i x_j$  (see Figure 3 panel (b)) Thus equilibrium is the solution to the following two equations:

$$I_i = 0 \text{ equation} \quad x_i = E_i x_j \quad (11)$$

$$\text{and the budget constraint} \quad x_i + x_j = N \quad (12)$$

$$\text{Solving (11) and (12) gives} \quad x_i^* = \left[ \frac{E_i}{1+E_i} \right] N \text{ and } x_j^* = N - x_i^* = \left[ \frac{1}{1+E_j} \right] N.$$

The solution is simply the intersection of  $I_i = 0$  curve and the budget constraint, as shown in panel (c) of Figure 3.

This equilibrium solution can also be arrived at using a CES (Constant Elasticity of Substitution) social utility function under conditions required to ensure conformity to the  $P$ -test. Section 6 derives and discusses the issue in detail.

As a passing note, sometimes it might be easier to work with another version of the model given below:

$$V_i = x_i - \tilde{\beta}_i \max[x_i - x_j - \tilde{b}_i N, 0] - \tilde{\alpha}_i \max[\tilde{b}_i N + x_j - x_i, 0]$$

where  $\tilde{\beta}_i > 0.5$  and  $\tilde{\alpha}_i \geq \tilde{\beta}_i$ . Maximization subject to  $x_i + x_j = N$  gives  $x_i = 0.5(1 + \tilde{b}_i)N$  and  $x_j = 0.5(1 - \tilde{b}_i)N$ , where  $\tilde{b}_i$  captures *equity-bias*.  $\tilde{b}_i = 0$  means no *equity-bias*,  $\tilde{b}_i > 0$  means positive *equity-bias* and  $\tilde{b}_i < 0$  means negative *equity-bias*. The paper will mainly stick to the specification in (10).

### 5.1. Application to the Dictator Game

Let  $D$  be pay-off of the dictator, and  $R$  be pay-off of the recipient. Based on equation (10), social utility of the dictator can be written as:

$$V_D = D - \beta_D \max[D - E_D R, 0] - \alpha_D \max[E_D R - D, 0] \quad (13)$$

where  $\beta_D = \frac{\gamma_D}{1+E_D}$  and  $\gamma_D > 1$ .

Maximization of equation (13) subject to  $D+R=N$  gives  $d^* = \frac{D^*}{N} = \left[ \frac{E_D}{1+E_D} \right]$ .

Theoretically  $\gamma_D$  can assume any value. We restricted it to greater than 1 ( $\gamma_D > 1$ ) in the model because it ensures unique interior solutions observed in the laboratory experiments. This is why we assumed this value of *social beta* in the model and restricted  $\gamma_D$  to be greater than unity. This is proved in proposition 2 below.

**Proposition 1:** *The equilibrium offer,  $1-d^*$ , made by the dictator varies with state-dependent equity-bias and can be anywhere between 0 and 1 inclusive.*

**Proof:** As  $1-d^* = \frac{1}{1+E_D} = \frac{1}{2+b_D}$ . Theoretically, the *equity-bias* can take any positive or negative value. Thus, when the dictator's valuation of equity is infinitely biased, the offer will be equal to zero (i.e.  $b_D = \infty \Rightarrow (1-d^*) = 0$ ). When the dictator's valuation is negatively biased, say  $b_D = -1$ , the offer is equal to 1;  $b_D = 0$  means the dictator is in a fair state and splits the sum fairly, i.e. half-half. The less than infinitely *equity-biased* dictator in a superior state (i.e.  $0 < b_D < \infty$ ) will give offers in the interior  $0.5 > (1-d^*) > 0$  and the dictator in an inferior state with  $-1 < b_D < 0$  will lead to the interior solution in  $0 > (1-d^*) > 0.5$ .

**Proposition 2:** *Any  $0 \leq d^* \leq 1$  is a unique other-regarding equilibrium with  $\beta_D > (1-d^*)$*

**Proof:** For  $d^*$  to be a unique equilibrium  $\gamma_D$  has to be greater than 1. Recall that  $\beta_D = \frac{\gamma_D}{1+E_D}$ .  $\gamma_D > 1 \Rightarrow \beta_D > \frac{1}{1+E_D} \Rightarrow \beta_D > (1-d^*)$ . The interpretation of this condition is straight forward. The dictator will have to sacrifice  $\frac{1}{1+E_D} = (1-d^*)$  dollars to reduce advantageous inequity by 1 unit. The one unit reduction in inequity will increase social utility by  $\beta_D$ . Thus for the dictator to find it optimal to sacrifice an extra dollar, it must increase his/her social-utility more than the sacrifice in terms of selfish utility.

The model clearly explains the corner as well as interior outcomes observed during experiments. This model acknowledges that individuals are heterogeneous and differ in their valuation of equity; hence their social utility is maximized at different offers. This leads them to optimally offer different amounts. In fact the corner solution, which is the selfish Nash equilibrium, can also be interpreted as an equitable equilibrium with  $E_D = \infty$ .

Recall that the restriction on  $\beta_i$  ensures that social utility has a maximum, at the state-dependent equitable distribution. The *social beta* is inversely related to  $E_i$ , implying that for a relatively larger *equity-bias*, a relatively lower *social beta* is required for the social utility to have a maximum.

## 5.2. Application of the Model to the *Ultimatum Game*

Now consider application of the model to the standard *ultimatum game* experiment. Let  $P$  be pay-off of the proposer and  $R$  be pay off of the responder.

The utility function of the proposer and the responder can be written (using equation (10)) as:

$$V_P = P - \beta_P \max(P - E_P R, 0) - \alpha_P \max(E_P R - P, 0) \quad (14)$$

$$V_R = R - \beta_R \max(R - E_R P, 0) - \alpha_R \max(E_R P - R, 0) \quad (15)$$

The proposer's objective is to maximize (14) subject to  $P+R=N$ . The proposer's offer will be equal to  $R^O = \left(\frac{1}{1+E_P}\right)N$ , leaving the responder an

amount  $P^O = \left[\frac{E_P}{1+E_P}\right]N$ . If this offer is accepted by the responder, both will keep

the positive sum; if rejected both will end up getting zero. The responder will accept the offer if it is greater than or equal to his/her minimum acceptable offer level. Following Fehr and Schmidt, assume minimum acceptable offers are offers which makes the responders just happy i.e. where  $V_R=0$ <sup>6</sup>. Assuming ( $E_R P - R > 0$ ), this implies  $R - \alpha_R (E_R P - R) = 0$ , which gives a minimum acceptable offer level of  $R^A = \frac{\alpha_R E_R N}{1 + \alpha_R (1 + E_R)}$ .<sup>7</sup> The superscript  $A$  is added to

denote the minimum acceptable offer level. Notice that when  $E_R=1$ , this expression reduces to the one derived by Fehr and Schmidt 1999. Also notices that when  $\alpha_R$  approaches  $\infty$ ,  $R^A/N$  approaches  $\frac{E_R}{1+E_R}$  which can be any value between 0 and 1, inclusive, depending upon the value of  $E_R$ . As  $E_R$  goes to zero (infinity),  $R^A$  goes to 0 (1). This means that even when  $\alpha_R > 0$ ,  $R^A$  could be any value between 0 and N, inclusive, depending upon  $E_R$ .

Thus, if

$$s = \begin{cases} 1 & \text{Accept offer when } R^O \geq R^A \\ 0 & \text{Reject offer when } R^O < R^A \end{cases}$$

the solution will be

$$\left(\frac{P^A}{N}, \frac{R^A}{N}\right) = \begin{cases} (p^*, 1 - p^*) & \text{when } R^O \geq R^A (s = 1) \\ 0 & \text{when } R^O < R^A (s = 0) \end{cases}$$

$$\text{where } 0 \leq p^* = \left[\frac{E_P}{1+E_P}\right] \leq 1.$$

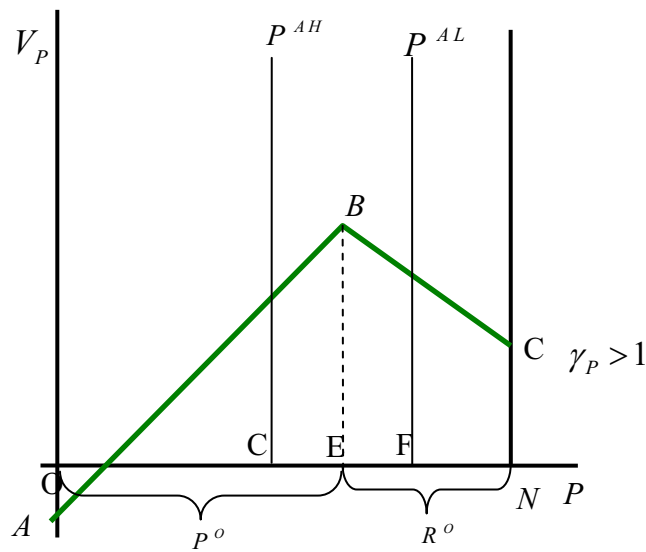
---

<sup>6</sup> An alternative approach would be to maximize social utility of the responder and workout the minimum acceptable threshold at minimum  $E_R$ . The social utility at this equilibrium may be equal to or greater than zero. This innovation however does not buy the model any interesting insight.

<sup>7</sup> It is possible for some players to find  $R^O > R^A$  offensive and reject them (this is when  $R^O > \beta_R E_R N / (\gamma_R - 1)$ ). I ignore such a possibility in discussions but the model does allow for such interpretations and solution.

Since, in practice, we may have agents with different *equity-biases*, let us capture this heterogeneity by expressing  $P^A$  of an individual of type  $n$  ( $=1, \dots, n$ ) by  $P^{An}$ , which is distributed with support  $P^{AH}$  and  $P^{AL}$ .  $P^{AH}$  belongs to a responder with the highest minimum acceptable offer ( $R^{AH}$ ) and  $P^{AL}$  belongs to a responder with the lowest minimum acceptable offer ( $R^{AL}$ ). Thus the outcome of the experiment will depend on whom the proposer is playing with. Consider Figure 4, where the proposer's utility is maximum at point B,  $P^O = OE$  and  $R^O = NE$ . The equilibrium is:

$$\left( \frac{P^*}{N}, \frac{R^*}{N} \right) = \begin{cases} (p^*, 1-p^*) & \text{when } R^A \in [NF, NE] \\ (0,0) & \text{when } R^A \in [NE, NC] \text{ excluding point E} \end{cases}$$



**Figure 4:** Equilibrium in the *ultimatum game* for the generalized model of inequity-aversion.

Of course, when  $E=C$  the proposer's equitable offer will always be accepted irrespective of type of the responder. The  $(0,0)$  equilibrium implies that the proposer does not offer more than  $NE$  because he/she thinks it is inequitable and should be accepted by any responder; if not, he/she is happy to face the consequences. Here the proposer can clearly be classified as other-regarding, unlike the FS model. The responder, on the other hand, rejects the offer, believing that it is less than what would make him/her just happy. The main reason for this rejection is that the offer is less than the minimum acceptable threshold. This rejection could be motivated by reciprocity, intentions, social punishment, self-assertiveness or any other reason. When rejections inflict monetary losses to proposers (as in the *ultimatum game*), all of these reasons could be in play. *Impunity games* narrow down the list to social punishment and self-assertiveness.

When  $\gamma_P \leq 1$ , the solution is similar to the FS model as discussed earlier.

## 6. Comparison with Representative Models of the Literature:

As obvious, the FS model is a special case of the generalized model in this paper. The generalized model reduces to the FS model when  $E_i=1$  which is relevant only when players are assigned into a fair state. The generalized model allows  $E_i$  to vary with state which is related to changes in *equity-bias*. Recall that in the FS model, the other-regarding behavior of the proposer was irrelevant for interior solution in the UG, whereas it is not irrelevant in the generalized model as the solution is not necessarily determined by the  $V_R=0$  constraint (see Figure 4). As indicated earlier, the FS model cannot explain interior solutions in the DG. Fehr and Schmidt argue that this is because of the piecewise linearity of preferences in advantageous inequity and that introducing non-linearity in the advantageous inequity could resolve the issue. As shown earlier, the BO model is the FS model with the proposed non-linearity. This specification however, like the FS model, fails the V-test (and, at some values of the preference parameters, the P-test). This paper shows that the nonlinearity is not necessarily needed to explain interior outcomes. Unlike the FS and BO models, the generalized model simultaneously passes the two proposed tests.

When  $\alpha_i = -\beta_i$ , the generalized model reduces to:

$$V_i = u_i(x_i) - \beta_i [u_i(x_i) - e_i u_{ji}(x_j)] = (1 - \beta_i) u_i(x_i) + \beta_i e_i u_{ji}(x_j) \quad (16)$$

Assuming  $\beta_i < 1$ , this is equivalent to the CFG model with  $\theta_i = \left( \frac{\beta_i}{1 - \beta_i} \right) e_i$ . As

obvious, the CFG model is decreasing in advantageous inequity (which is consistent with the FS model and the model in this paper) and increasing in disadvantageous inequity which is the opposite of what is postulated in the FS model and its extension here<sup>8</sup>. As shown earlier, the CFG model fails the P-test when  $\sigma \geq 0$ . Imposing  $\alpha_i = -\beta_i$  therefore comes at the cost of the P-test unless  $\sigma$  is restricted to non-positive values. No such restriction is needed in the generalized model when  $\alpha_i > -\beta_i/\gamma$  which includes any positive  $\alpha_i$ .

Similarly, when  $\sigma=1$ , indifference curves are linear and would give corner solution only ( $x_i=1$  or  $0$ ). The CFG model therefore requires diminishing marginal utility of money to explain interior outcomes, the model in this paper doesn't as it can explain any solution at any given  $\sigma$ .

## 7. Application of the Model

As mentioned earlier, there are a variety of dictator and *ultimatum games* experimental economists have tested in the laboratory. Even though solutions of the model in this paper were derived for the standard dictator and *ultimatum games*, they can easily be applied to different designs, contexts and multiple player versions of these games. All we need to do is work out what impact would a particular innovation to an experimental design have on the relative state of players and for that matter their *equity-bias* and predictions about the solution

---

<sup>8</sup> Fehr and Schmidt (1999) cites empirical evidence to support  $\alpha_i \geq \beta_i > 0$ .

follow intuitively. Understandably, it may not be possible in some cases to predict what impact a particular innovation to an experimental design would have on the relative state of a player beforehand; experiments could be used to understand the dynamics of such innovations on states. Similarly, laboratory experiments could be used to quantify the impact of alternative policy tools on *equity-bias* which will help in choosing efficient policies. The notion of *equity-bias* is general in nature and embodies all information related to socio-economic status, intentions, reciprocity, social distance, design of the experiment role of the experimenter, etc. The model therefore provides a unified framework to understand outcomes of the research in a broader context.

In the standard DG, the state assigned to a dictator is superior than the one assigned to the receiver. The recipient is neither a party to the creation of value, nor legally entitled to any share in the sum. Thus it seems natural for the dictator to have a notion of equity that is biased towards his/her welfare and make offers in the interior. The *ultimatum game* introduces two changes to the DG. Firstly, it lowers the status of the dictator to a relatively inferior position by assigning him/her the role of a proposer. Secondly, it alleviates the status of the recipient to that of a responder who becomes an active partner to the creation of value. This arrangement leads to a reduction in *equity-bias* of the dictator (now the proposer) and an improvement in *equity-bias* of the recipient (now responder). In a one-shot *ultimatum game* the proposer is in a relatively superior state than the responder (by just being the first mover), which allows for the possibility of less than fair splits as equitable offers. In a repeated *ultimatum game* their relative states converge to a fair state and the offers converge to an even split.

The wording of the experiment and instruction list (Bolton *et al.* 1998; Hoffman *et al.* 1999), identity of the experimenter, social distance, and design of the experiment seem to have an impact on outcomes (Hoffman *et al.* 1994; 1996). All these factors change the relative positions of players in the game, and hence their *equity-bias*, which show up in their offers and changes in the threshold for accepting or rejecting offers. Hoffman *et al.* (1994), for example, found reduction in offers by dictators when anonymity was increased, implying that anonymity increases *equity-bias*. This means that part of the altruism in the DG is to avoid being labeled as too selfish, to look good or not to look bad. Similarly, the experiment by Brañas-Garza (2006) cited earlier, which introduces poverty in the DG, concludes that informing the dictator about the poverty of the recipient leads to more altruism. Again, this is because information about the socio-economic status of players leads to changes in relative states and *equity-biases*, which leads to different outcomes.

Camerer and Thaler (1995 p. 216) mentioned that when instead of giving the dictator a sum as “manna from heaven”, if the dictator is made to feel as if he earned the right to the sum, then sharing shrinks. Making the dictator feel as if he/she earned the right to the sum basically changes his/her location to a relatively more superior (from superior to more superior) state, which leads to an increase in his emotional state of *equity-bias*. Similarly, Schotter *et al.* (1996) introduced property rights in two-stage-survival dictator and *ultimatum games*. In the first stage proposers were competing with each other in offering higher amounts to a single responder. They earn property rights to the sum when a proposer accepts the responder’s offer. They move to the second stage with property rights they earned in the first state. In the second stage they offered lesser amounts and responders rejected smaller amounts less often. The offers were still significantly higher than zero and considered to be fair by player 2. This is because earning

property rights in the second stage increased *equity-bias* of the proposers. The responders also made note of that and respected it by revising their lowest threshold (less biased).

Similarly, structure/design of the experiment also plays a role. Consider two types of *dictator games*. (i) First, assume that the dictator is given \$20 to share with another anonymous recipient. (ii) Second, consider the same two players now given 10 dollars each, with the dictator having the option of giving or taking away up to 10 dollars. Both experiments involve sharing 20 dollars but individuals are not located on the same position on the bias-chart. The dictator in (i) is in a more superior state than in (ii). Thus we will expect the share of the dictator in (ii) to be lower than, or at least as much as, in (i). Similarly, in the *ultimatum game* we would expect the responder in (ii) to reject offers higher than in (i); this is because he/she is not in as inferior state in (ii) as he/she is in (i). The model also gives insight into why the hypotheses in experiments such as [Bardsley \(2005\)](#) are erroneous and provide a context for explaining results of his experiment. Bardsley hypothesized that the standard DG would give the same solution if the sum was instead distributed and the dictator was given the option of (giving and) taking money from the recipient. Arguing that the dictator is facing a similar problem of allocating the same budget, optimal allocation should be the same. The key assumption in Bardsley's argument is that preferences of the dictator are the same in the two experiments. The model in this paper postulates that preferences are state-dependent and the two models belong to two different states; expecting it to give the same solution is simply wrong.

An increase in competition on the proposer's side is expected to reduce the proposer's *equity-bias*, leading to relatively larger offers and increasing the responder's *equity-bias*, leading to rejection of relatively larger offers. However, if proposers were to compete for lowest offers, the proposer's *equity-bias* is expected to increase, leading to smaller offers, and that of the responders is expected to move further away from fair offers, leading them to accept relatively lower offers. This holds only if the responders are aware of the nature of the competition. [Camerer and Thaler \(1995\)](#) writes;

*"A good general theory of fairness predicts that fair-minded players behave self-interestedly in some situations. Two experiments show that competition can push ultimatum offers closer to zero, in ways consistent with fairness. [Schotter et al. \(1994\)](#) created competition among Proposers. Eight Proposers made offers in a first stage. The four who earned the most in the first-stage game could then play a second-stage game (with a different player). Sensible fairness theories would say that Proposers now have an excuse for making low offers-they must compete for the right to play again-so low offers are not as unfair, and Responders will accept them more readily. That is what happened."*

Information asymmetry may also play its role in positioning players on the state/bias chart and so may intentions. About the role of intentions in rejections in *ultimatum games*, [Fehr and Schmidt \(2005\)](#) conclude *"Taken together, the evidence from [Blount \(1995\)](#), [Kagel et al. \(1996\)](#), [Offerman \(1999\)](#), [Brandts and Sola \(2001\)](#) and [Falk et al. \(2000a, 2000b, 2003\)](#) supports the view that subjects want to punish unfair intentions or unfair types. Although the evidence provided by the initial study of [Blount](#) was mixed, the subsequent studies indicate a clear role of these motives."*

Similarly, in experiments by [Kagel et al. \(1996\)](#) where subjects had to divide 100 chips in an UG, chips were convertible to monetary pay-offs at different prices across players. For example in one treatment responders' chips were valued 10 cents each and that of proposers 30 cents each. Players knew their own conversion rate but not necessarily that of others. When the proposer is aware that his/her chips are valued three times more than that of the responder, an equal monetary split would require the proposer to give 75 chips to the responder. When the information was available to responders, they rejected unequal money splits more frequently than when they were not aware of the difference in chips' money value. Thus unequal proposals were rejected at higher rates than unintentional unequal proposals ([Fehr and Schmidt, 2005](#)). Similarly, another important insight of the experiment was that proposers offered close to 50 percent when there was information asymmetry. This implies that proposers prefer to seem fair than be fair. [Camerer and Thaler \(1995\)](#) believes this is "an important reminder that self-interested behavior is alive and well, even in *ultimatum games*". This also points towards an important distinction between altruism/fairness as a natural instinct and altruism/fairness as a code or altruism as an instinct complemented by religious or moral affiliations.

This could intuitively be predicted by the model as well. Information asymmetry in this experiment positions proposers in a relatively superior state and they make offers closer to 50% than 75%. This is when they prefer to seem fair than act fair. What the model also predicts is that individuals with commitment to moral codes through say religious or cultural affiliations would offer closer to the 75%. Information asymmetry should as such not change their relative state. It will be interesting to investigate whether or not individuals with different religious affiliation would behave differently.

Similarly it can be shown that the framework in this paper explains the wide variety of data in laboratory experiments such as [Blount \(1995\)](#) and [Falk et al. \(2003\)](#). The framework also provides a rationale for the three-player games by [Güth and Damme \(1998\)](#), [Kagel and Wolf \(2001\)](#), and [Berby-Meyer and Nuedereke \(2005\)](#) by invoking player and design-specific *equity bias*.

## 8. Summary and Conclusions

This paper takes note of the variety and multiplicity of models of other-regarding preferences proposed in the literature on experimental and behavioral economics and points towards the need for certain criteria to judge these models. The paper identifies two tests, the Variety Test (ability of a model to explain outcomes under variety or alternative scenarios) and the Psychological Test (ability of a model to conform to psychological intuition), that can be used to judge any model of other-regarding preferences. It is argued that, for a mathematical model to qualify as a social welfare function, it must simultaneously pass the two tests. It is shown that none of the models proposed to date passes the two tests simultaneously with the exception of the [Cox et al. \(2007\)](#) model which simultaneously passes the two tests when some additional restrictions are imposed.

This paper extends the [Fehr and Schmidt \(1999\)](#) model of *inequality-aversion* to a generalized model of *inequity-aversion*. [Fehr and Schmidt \(1999\)](#) assume that subjects enter laboratory as equals and should therefore use egalitarian outcome as a reference outcome. This is why the Fehr and Schmidt model cannot explain interior outcomes in the dictator game. This paper argues that once assigned into different roles and given unequal property rights, subjects cease to be equal and



equal-distribution is not necessarily the equitable norm. The paper introduces the concept of *equity-bias* and postulates that one's idea of equitable distribution is state-dependent where state is determined by psychological and structural parameters. The state could be fair, superior or inferior. When assigned a fair state, one's valuation of equity is a fair one (even split), and when assigned a biased-state (superior or inferior) one's valuation of equity is a biased one. *i.e.* bias in state leads to bias in equity. Individuals in a fair state have zero *equity-bias* and split the pie evenly. Those in a superior (inferior) state have positive (negative) *equity-bias* and value more (less) than a fair distribution as an equitable distribution. Bias in state therefore leads to bias in equity. Given the psychological tendencies/state of an individual, every experimental design assigns one of the three states to the player, which leads to individual-specific valuation of equity. Predictions about outcomes in different experiments, or the same experiment with different designs, can be made through predicting its impact on *equity-bias*.

The notion of *equity-bias* is general in nature and embodies all information related to socio-economic status, intentions, reciprocity, social distance, design of the experiment, role of the experimenter, etc. The model therefore is all-encompassing and provides a unified framework to understand outcomes of research in a broader context. For example, in the standard dictator game, the state assigned to the dictator is superior than the one assigned to the receiver. This is because the recipient is neither a party to the creation of value, nor legally entitled to any share in the sum. Thus it seems natural for the dictator to have a notion of equity that is biased towards his/her welfare. The standard *ultimatum game* introduces two changes to the dictator game. Firstly, it assigns the dictator the role of the proposer, which reduces his/her superiority and secondly, it alleviates the status of the recipient to that of a responder who becomes an active partner to the creation of value. This arrangement leads to reduction in the *equity-bias* of the dictator (now the proposer) and increase in the *equity-bias* of the recipient (now the responder). Thus the fear of rejection, along with other factors, changes *equity-bias* of the players, hence equitable distribution. This is the main reason why offers in the dictator game are positive but lower than in the *ultimatum game*. Competition changes the relative position of the players and their valuation of equitable distribution. So does the design of experiments, such as the way property rights are assigned (e.g. earned or manna from heaven), the wording of the experiment/instructions, role of the experimenter etc.

The model is more general than its previous counterparts. It parsimoniously explains interior solutions in the dictator game and provides a framework to understand outcomes in other experiments. It provides a framework to understand why outcomes change with design of the experiment and across different experiments of the same nature.

The framework in this paper is simple and doesn't require individuals to process complex information. It rationalized all kind of choices, smart or otherwise, as state-dependent other-regarding utility maximizing outcomes. Policy makers can benefit from understanding the evolution of relative states and equity biases. The effectiveness of alternative policy reforms could be explored in the laboratory through its impact on equity-bias. Research therefore should be directed to unfold the dynamics and evolution of equitable states relevant to policy debates and positive analysis of issues of interest. The model provides a framework to understand the relevance and importance of equity-bias which is determined by psychological (natural tendencies as as self-serving bias, individual characteristics etc) and structural parameters (design of the experiment, the

manner in which property rights are assigned, wording of the experiment, real world scenarios etc). Laboratory experiments could be used to determine the evolution of these states.

The existing literature and experimental results obtained thus far can be re-organized in the light of their impact on equity-bias for organized conclusions. The framework can also be used to understand outcomes of other games, such as *public good* and *trust* games.

## References

- Bardsley, N. (2005). [Altruism or Artefact? A Note on Dictator Game Giving](#). [Discussion Papers](#) 2005-10, The Centre for Decision Research and Experimental Economics, School of Economics, University of Nottingham.
- Bereby-Meyer, Y., Niederle, M. (2005). Fairness in Bargaining. *Journal of Economic Behavior and Organization*, 56, 173-186.
- Blount, S. (1995). When Social Outcomes aren't Fair: The Effect of Causal Attributions on Preferences. *Organizational Behavior and Human Decision Processes* LXIII, 131-144.
- Bolton, G., Katok, E., Zwick, R. (1998). Dictator Game Giving: Rules of Fairness Versus Acts of Kindness. *International Journal of Game Theory*, 27 269-299.
- Bolton, G., E., Ockenfels, A. (2000). A Theory of Equity, Reciprocity and Competition. *American Economic Review*, 90(1), 166-193.
- Brañas-Garza, P. (2006). Poverty in Dictator Games: Awakening Solidarity. *Journal of Economic Behavior and Organization*, 60(3), 306-320
- Brandts, J., Sola, C. (2001). [Reference Points and Negative Reciprocity in Simple Sequential Games](#). [Games and Economic Behavior](#), Elsevier, 36(2), 138-157.
- Brandts, J., Charness, G. (2004). Gift-Exchange with Excess Supply and Excess Demand. mimeo, Pompeu Fabra, Barcelona.
- Camerer, C., Thaler, R., H. (1995). Anomalies: Ultimatums, Dictators and Manners. *The Journal of Economic Perspectives*, 9(2), 209-219
- Camerer, C., F. (2003). *Behavioral Game Theory, Experiments in Strategic Interaction*. Princeton University Press, Princeton.
- Charness, G., Rabin, M. (2002). Understanding Social Preferences with Simple Tests. *Quarterly Journal of Economics*, 117, 817-869.
- Cox, J. C., Friedman, D., Gjerstad, S. (2004). A Tractable Model of Reciprocity and Fairness. mimeo, University of Arizona.
- Cox, J. C., Friedman, D., Gjerstad, S. (2007). A Tractable Model of Reciprocity and Fairness. *Games and Economic Behavior*, Elsevier, 59(1), 17-45.
- Dufwenberg, M., Kirchsteiger, G. (2004). A Theory of Sequential Reciprocity. *Games and Economic Behavior*, 47, 268-98.
- Eckel, C., Grossman, P. (1996). Altruism in anonymous dictator games. *Games and Economic Behavior*, 16, 181-191.
- Erlei, M. (2004). Heterogeneous Social Preferences. mimeo, Clausthal University of Technology.

- Falk, A., Fehr, E., Fischbacher, U. (2000a). Informal Sanctions. Institute for Empirical Research in Economics, University of Zurich, Working Paper No. 59.
- Falk, A., Fehr, E., Fischbacher, U. (2000b). Testing Theories of Fairness - Intentions Matter. Institute for Empirical Research in Economics, University of Zurich, Working Paper No. 63.
- Falk, A., Fischbacher, U. (2001). A Theory of Reciprocity. CEPR Discussion Paper no. 3014, University of Zurich.
- Falk, A., Fehr, E., Fischbacher, U. (2003). On the Nature of Fair Behavior. *Economic Inquiry*, 41, 20-26.
- Fehr, E., Schmidt, K., M. (1999). A Theory of Fairness, Competition and Co-operation. *Quarterly Journal of Economics*, 114, 817-868.
- Fehr, E., Schmidt, M. K. (2005). The Economics of Fairness, Reciprocity and Altruism - Experimental Evidence and New Theories. Discussion paper 2005-20, Department of Economics, University of Munich.
- Güth, W., Damme, V. E. (1998). Information, strategic behavior and fairness in ultimatum bargaining: An experimental study. *Journal of Mathematical Psychology*, 42, 227-47.
- Hoffman, E., McCabe, K., Shachat, K., Smith, V. (1994). Preferences, Property Rights and Anonymity in Bargaining Games. *Games and Economic Behavior*, 7, 346-80.
- Hoffman, E., McCabe, K., Smith, V. (1996). Social Distance and Other-Regarding Behavior in Dictator Games. *American Economic Review*, 86(3), 653-660.
- Hoffman, E., McCabe, K., Smith, V. (1999). Social Distance and Other-Regarding Behavior in Dictator Games: Reply. *The American Economic Review*, 89 (1), 340-341.
- Kagel, J. H., Kim, C., Moser, D. (1996). Fairness in Ultimatum Games with Asymmetric Information and Asymmetric Payoffs. *Games and Economic Behavior*, 13, 100-110.
- Kagel, J. H., Wolfe, K. (2001). Tests of Fairness Models Based on Equity Considerations in a three-person Ultimatum Game. *Experimental Economics*, 4, 203-219.
- Kohler S. (2003). Difference Aversion and Surplus Concern – An Integrated Approach. Mimeo.
- Levine, D. (1998). Modeling Altruism and Spitefulness in Experiments. *Review of Economic Dynamics*, 1, 593-622.
- Offerman, T. (1999). Hurting hurts more than helping helps: The Role of the self-serving Bias. mimeo, University of Amsterdam.
- Ottone, S., Ponzano, F. (2005). An Extension to the Model of Inequity Aversion by Fehr and Schmidt. P.O.L.I.S. department's Working Papers 51 Department of Public Policy and Public Choice – POLIS
- Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics. *American Economic Review*, 83 (5), 1281-1302.
- Rotemberg, J. (2004). Minimally Acceptable Altruism and the Ultimatum Game. mimeo, Harvard Business School.
- Schotter, A., Weiss, A., Zapater, I. (1994). Fairness and Survival in Ultimatum Games. Unpublished paper, New York University Department of Economics.
- Schotter, A., Weiss, A., Zapater, I. (1996). Fairness and Survival in Ultimatum and Dictatorship Games. *Journal of Economic Behavior & Organization*, 31 37-56.

## APPENDIX A: A Second Example of the Linear Model

Charness-Rabin (*quasi-maximin preferences*) can be written as under (in a two player environment with  $x_i + x_j = 1$ )

$$U_i = (1 - \gamma)x_i + \gamma W$$

$$W = \delta \min(x_1, x_2) + (1 - \delta)(x_1 + x_2)$$

This implies

$$U_i = \begin{cases} (1 - \gamma)x_i + \gamma\delta x_j + \gamma(1 - \delta) & \text{if } x_i > x_j \\ \lceil 1 - \gamma(1 - \delta) \rceil x_i + \gamma(1 - \delta) & \text{if } x_i \leq x_j \end{cases}$$

$$U_i - \gamma(1 - \delta) = \begin{cases} \lceil 1 - \gamma(1 - \delta) \rceil x_i - \gamma\delta(x_i - x_j) & \text{if } x_i > x_j \\ \lceil 1 - \gamma(1 - \delta) \rceil x_i & \text{if } x_i \leq x_j \end{cases}$$

$$\frac{U_i - \gamma(1 - \delta)}{\lceil 1 - \gamma(1 - \delta) \rceil} = \begin{cases} x_i - \frac{\gamma\delta}{\lceil 1 - \gamma(1 - \delta) \rceil}(x_i - x_j) & \text{if } x_i > x_j \\ x_i & \text{if } x_i \leq x_j \end{cases}$$

Assume  $V_i = \frac{U_i - \gamma(1 - \delta)}{\lceil 1 - \gamma(1 - \delta) \rceil}$  and  $\beta_i = \frac{\gamma\delta}{\lceil 1 - \gamma(1 - \delta) \rceil}$ . The model reduces to

Fehr and Schmidt's 1999 model with  $\alpha_i = 0$ . i.e.

$$V_i = x_i - \beta_i \max[(x_i - x_j), 0]$$

The fact that  $\alpha_i = 0$  makes the model inferior under the specification as it would give corner solution in both the *dictator game* as well as the *ultimatum game*.

## APPENDIX B: Derivation of the Restriction on Social Beta

$$\text{As } V_i = u_i(x_i) - \beta_i \max[u_i(x_i) - e_i u_{ij}(x_j), 0] - \alpha_i \max[e_i u_{ij}(x_j) - u_i(x_i), 0]$$

Differentiating w.r.t  $x_i$  gives (assume  $u_i' = \left(\frac{du_i}{dx_i}\right)$  and  $u_{ij}' = \left(\frac{du_{ij}}{dx_j}\right)$ )

$$\frac{d(V_i)}{dx_i} = \begin{cases} u_i' + \alpha_i [e_i u_{ij}' + u_i'] & \text{when } u_i < e_i u_{ij} \\ u_i' & \text{when } u_i = e_i u_{ij} \\ u_i' - \beta_i [u_i' + e_i u_{ij}'] & \text{when } u_i > e_i u_{ij} \end{cases}$$

This shows that the slope is positive when  $u_i < e_i u_{ij}$  and  $\alpha_i > -\frac{\beta}{\gamma}$ . For the point  $u_i = e_i u_{ij}$  to be the unique maximum, the social utility must decline when  $u_i > e_i u_{ij}$ . Thus  $\frac{d(V_i)}{dx_i}$  has to be  $< 0$  when  $u_i > e_i u_{ij}$  .i.e.  $u_i' - \beta_i (u_i' + e_i u_{ij}') < 0$ .

This implies  $\beta_i > \left[1 + e_i \left(\frac{u_{ij}'}{u_i'}\right)\right]^{-1}$ . Thus, if we assume  $\beta_i = \gamma_i \left[1 + e_i \left(\frac{u_{ij}'}{u_i'}\right)\right]^{-1}$ ,  $\gamma_i$

has to be greater than 1 for the condition to be satisfied.  $\gamma_i$  can also be equal to 1 when  $u_i'' < 0$  and  $u_{ij}'' < 0$  (i.e. diminishing marginal utility in pay-offs).  $\gamma_i$  must be greater than 1 in case of constant marginal utilities i.e.  $u_i'' = u_{ij}'' = 0$ .

The condition can further be simplified as well. Proposition 2 shows that this condition is equivalent to  $\beta_i > 1 - x_i^*$  which will always hold when  $\beta_i > 1$ . Thus

the model will still give the same solution if we replace  $\beta_i > \left[1 + e_i \left(\frac{u_{ij}'}{u_i'}\right)\right]^{-1}$

condition with  $\beta_i > 1$  for simplicity.

## APPENDIX C: Derivation of the Main Model

Assume the individual  $i$  has the following social utility function

$$\tilde{V}_i = u_i \left( \frac{x_i}{\delta_i w_i} \right) - \beta_i \max \left[ u_i \left( \frac{x_i}{\delta_i w_i} \right) - e_i u_{ij} \left( \frac{x_j}{\delta_{ij} w_{ij}} \right), 0 \right] - \alpha_i \max \left[ e_i u_{ij} \left( \frac{x_j}{\delta_{ij} w_{ij}} \right) - u_i \left( \frac{x_i}{\delta_i w_i} \right), 0 \right]$$

$$u_i' > 0, u_{ij}' > 0, u_i'' \leq 0, u_{ij}'' \leq 0, \beta_i = \gamma_i \left[ 1 + e_i \left( \frac{u_{ij}'}{u_i'} \right) \right]^{-1} \text{ and } \gamma_i > 1.$$

$w$  stands for wealth and  $\delta$  quantifies the efficiency/inefficiency with which individuals convert pay-offs into utility. The subscript  $ij$  on represents individuals  $i$  belief about individual  $j$  since  $i$  doesn't have perfect information about  $j$ 's utility, wealth and efficiency.

Notice that selfish utility is determined by pay-off relative to wealth which basically acknowledges the fact that a dollar received by a wealthier person generates lesser utility than when it is received by a poor person. The parameter  $\delta$  is there to capture heterogeneity in preference technology. When  $\delta_i > \delta_{ij}$  individual  $i$  is more efficient in converting a dollar into utility than individual  $j$  (because of some socio-economic feature other than wealth, say disability). This specification therefore acknowledges socio-economic status as one of the determinant of selfish utility generated by pay-offs.

Let us assume the selfish utility functions are homogeneous of degree  $k$  in its arguments. We can re-write the social utility function as under

$$\tilde{V}_i = \left( \frac{x_i}{\delta_i w_i} \right) - \beta_i \max \left[ \left( \frac{x_i}{\delta_i w_i} \right) - e_i \left( \frac{x_j}{\delta_{ij} w_{ij}} \right), 0 \right] - \alpha_i \max \left[ e_i \left( \frac{x_j}{\delta_{ij} w_{ij}} \right) - \left( \frac{x_i}{\delta_i w_i} \right), 0 \right]$$

Multiplying both sides by  $(\delta_i w_i)$  and substitute  $(\delta_i w_i) \tilde{V}_i = V_i$  and  $\chi_i = \left( \frac{\delta_j}{\delta_i} \right) \left( \frac{w_j}{w_i} \right)$ , we get

$$V_i = x_i - \beta_i \max \left[ x_i - e_i \left( \frac{x_j}{\chi_i} \right), 0 \right] - \alpha_i \max \left[ e_i \left( \frac{x_j}{\chi_i} \right) - x_i, 0 \right]$$

which is the model in the paper. As argued,  $\chi_i$  in the model captures individual  $i$ 's valuation of his/her socio-economic status relative to that of individual  $j$ .