Psycho-Social Equilibria: Theory and Applications*

PATRICIO DALTON[†] and SAYANTAN GHOSAL[‡]

University of Warwick, Economics Department
14 May 2007 - Preliminary version -

Abstract

This paper introduces a general class of simultaneous move games in which the payoff of each player depends not only on players' strategy profile, but also on their preference parameters which are, in turn, endogenously determined in equilibrium. The interaction among players generates a feedback effect on each player's preference parameters, which is not internalized. We label this class of games "psycho-social games". Psycho-social games offer a synthesis of the most influential theoretical literature on behavioural economics in a strategic setting. We show the existence of a psycho-social equilibrium in pure strategies assuming incomplete, non-convex preferences and strategic complementarity between actions and psycho-social states. We relax the assumption of strategic complementarity and show existence of a psycho-social equilibrium in mixed strategies. By defining an appropriate notion of embedding, we associate a set of psycho-social games to each normal form game where agents have only material payoffs. We show that, typically, the set of Nash equilibria and the set of psycho-social equilibria of an associated psycho-social game are distinct from each other. Our theoretical findings are consistent with the vast experimental findings in simultaneous move games. We also study how psycho-social games provide a theoretical framework to analyse issues of development in which psycho-social concerns play an important role, such as chronic poverty, aspirations, intrinsic motivation and empowerment. We present an application on aspirations formation.

JEL classification numbers: C72, C79, D01, Z10.

Keywords: Psycho-social Games; Psychological Games, Reference-dependent Preferences, Endogenous Preferences; Myopia, Incomplete Preferences; Aspiration Formation.

---

*Preliminary manuscript. Please do not cite or quote without the author's permission. We are grateful to John Geanakoplos for helpful discussions on an earlier version, and to Peter Hammond, Marcus Miller, Andres Carvajal, Amrita Dhillon, Javier Fronti, Tommaso Gabrieli and participants on the Warwick Economics Theory Workshop. Patricio Dalton gratefully acknowledges financial support from Marie Curie Host Fellowship for Early Stage Research Training and Warwick Economics Department. The usual disclaimer applies.

[†]P.S.Dalton@Warwick.ac.uk

[‡]S.Ghosal@Warwick.ac.uk

1 Introduction

This paper introduces a general class of simultaneous-move games in which the payoff of each player depends not only on her strategy profile, but also on her preference parameters. The preference parameters are, in turn, endogenously determined in equilibrium and they are not internalized by the player. We name this class of games "psycho-social games". In brief, a psycho-social game is a simultaneous move game in which preferences are endogenous and players don't fully internalize the consequences of their own actions on their preferences. The motivations we have in mind to study this class of games are two-fold.

First, psycho-social games provide a general framework to incorporate general insights from Social Psychology[1] into Economics, and thus gain a better understanding of some specific concerns of development, such as chronic poverty, aspirations failures, social conflict and empowerment that cannot be fully understood with the existing economics models[2].

Second, psycho-social games offer a synthesis of a seemingly different influential theoretical literature on behavioural economics in a strategic setting. In this sense this paper is an extension of Ghosal's (2006a) "self-influence" individual decision model for n-players case. The synthesis of this literature allows us to get a better understanding of the observed behaviour reported in a wide class of simultaneous-move games experiments by means of a very simple model. For example, our framework encompasses games that incorporate endogenous beliefs into player's payoffs such us Geanakopolos' et.al. (1989), or generalizes models that account for endogenous social preferences such us Rabin (1993), Charness and Rabin (2002), Charness and Dufwenberg (2006) and games that consider endogenous reference-dependent preferences with loss aversion such us Shalev (2000).

In the next section, we introduce a simple version of a psycho-social game that we use to illustrate the main conceptual features of the game. We also study the way in which psycho-social games provide a synthesis of apparently disconnected papers, and we make explicit the common structure that underlies most of them.

Then, in section 3, we introduce the $n-$person general model with a formal definition of a psycho-social game and its equilibrium solution concept. We provide two new existence results of a psycho-social equilibrium. Theorem 1 shows the existence of a psycho-social equilibrium in pure strategies assuming incomplete, non-convex pref-

[1] See for example Bandura (2001) or Forgas et. al (2000)

[2] For an exposition on the theoretical challenges faced by development economics, see Banerjee (2005) or Mullainathan (2006).

erences and strategic complementarity between actions and psycho-social states. Theorem 2 relaxes the assumption of strategic complementarity and shows existence of a psycho-social equilibrium in mixed strategies.

After showing equilibrium existence, we study the general characteristics of a psycho-social equilibrium, in section 4. By defining an appropriate way to associate - or embed - a psycho-social game with a normal form game, we compare the set of Nash Equilibria of an arbitrary standard normal form game with the set of psycho-social equilibria of the psycho-social games that are associated with this arbitrary normal form game. We show that, typically, the two sets are distinct from each other. Our theoretical findings are consistent with the vast experimental findings in simultaneous move games. Our theory suggests that in general, people play psycho-social games instead of normal-form games. Thus, it shouldn't be surprising that most experimental evidence in normal-form games reports that people don't play Nash's predictions.

In section 5 we further discuss two key issues of a psycho-social game. First, we explore what we mean by players not internalizing the feedback effect of their actions into their preferences. Second, we analyze why we need this assumption and we explore what would happen if the assumption was relaxed and we assumed instead that players are sophisticated. We find different results depending on the game we explore. For instance, in Rabin's (1993) one could get the same outcomes assuming sophisticated players. The same is true in Shalev's (2000) model. However, in other games (e.g. with guilt or commitment) the assumption of myopia introduces additional welfare ranked equilibria which can make the myopic player be worse-off or better-off than a sophisticated player, depending on the game..

In section 6 we link our framework with previous literature on Framing and we introduce an application of a psycho-social game in a model of aspiration failures.

Finally, in the last section we conclude and suggest further extensions, applications and future research.

2 A simple model with examples

Consider a simultaneous move game with two players $i = 1, 2$ whose payoff relevant variables are:

$i)$ the action set $A = A_1 \times A_2$ where $A_i = [\underline{a}_i, \bar{a}_i]$ is an interval of $\Re$.

$ii)$ the set of utility parameters (psycho-social states) $P = P_1 \times P_2$, where P_i is a subset of some metric space. For the moment, we will keep it simple and assume that $P_i \subset \Re$.

The preferences of each player i are represented by a utility function $v_i : A \times P_i \rightarrow \Re$. Given p_i, player i solves the following maximization problem:

$$Max_{\{a_i \in A_i\}} v_i(a_i, a_{-i}, p_i), \quad \text{for all } \; a_{-i} \in A_{-i} \; \text{ and } p_i \in P_i$$

In addition, suppose there is a feedback effect for each player i from the vector (a_i, a_{-i}, p_{-i}) to p_i represented by the map $\pi_i : A \times P_{-i} \rightarrow P_i$. At this stage, having endogenized p_i, we can take two different possible routes. On the one hand, we can assume that each player i is sophisticated to an extent that when she considers a deviation, she anticipates an appropriate change in her psycho-social state that is consistent with her deviation. In this scenario, each player i would solve the following optimization problem:

$$Max_{\{a_i \in A_i\}} u_i(a_i, a_{-i}), \text{ for all } a_{-i} \in A_{-i}$$

where $u_i(a_i, a_{-i}) = v_i(a_i, \pi_i((a_i, a_{-i}, p_{-i}))$. Observe that if we take this approach, we have a simultaneous move game with endogenous preferences and sophisticated players who perfectly take into account the consequences on their own and other's psycho-social states, and then react accordingly. The solution concept of such a game is analogous to a Nash Equilibrium.

We are not interested in this class of games though, since they would not lead us to different outcomes - at least in a simultaneous move framework - from those we already know from traditional game theory. In this paper, we are interested in taking a second possible route. In our definition of a psycho-social game, we shall assume that players do not take into account the consequences of their actions on their own and other's psycho-social states. When evaluating a deviation, the player does not change her psycho-social state but **takes as given** the feedback effect generated in the interaction. This simple assumption will allow us to capture in a simple model, the common but often ignore feature that underlies in apparently disconnected models in behavioural economics. From those that incoporate social preferences or intrinsic motivation to those that assume reference-dependent preferences, all of them implicitly assume the existence of an **endogenous** feedback effect which is **not internalized** by the player at the moment of computing her best response. Moreover, psycho-social games not only offer a synthesis of existing literature, but they also provide a single theoretical framework to analyze concerns that otherwise could not be studied, such us endogenous identity, endogenous commitment, aspirations formation, empowerment, etc.

In the remaining of the section we will analyze some of the literature that is encompassed within our framework. The key point is that each existing model specifies a different functional form for the map π_i.

2.1 Psychological Games

The class of games that are closest related with psycho-social games are Geanakoplos, Pearce and Stachetti (1989)[3] *normal form psychological games*[4]. In GPS's psychological games, the payoff to each player depends on players' action profiles and on their endogenous beliefs. When players compute their best responses, they fix the actions of the others and their own beliefs.

Formally, let $\Delta(A_i)$ and $\Delta(A_{-i})$ denote the set of probability distributions over A_i and A_{-i}, respectively, where $A_{-i} = \prod_{j \neq i} A_j$. For simplicity, let's define believes up to a second order. The set of first order beliefs of player i is given by $B_i^1 := \Delta(A_{-i})$, with $b_j \in B_i^1$ representing player $i's$ beliefs about what player j will do. The set of player $i's$ second order beliefs is $B_i^2 := \Delta\left(A_{-i} \times \Delta(A_i)\right)$, with $(b_j, c_i) \in B_i^2$ representing a profile of believes that player i holds, which include her beliefs about what player i will do, b_j, and also her beliefs about what player j believes player i will do, c_i. In a psychological game, each player is assumed to maximize her expected utility over the set of actions for each set of beliefs $p_i \in \bar{B}_i$, where $\bar{B}_i$ is her set of collectively coherent beliefs[5]. In equilibrium, all beliefs are assumed to conform to some commonly held view of reality. So, if $a^* = (a_i^*, a_j^*)$ is the equilibrium action profile, then the equilibrium set of beliefs $\left(p_i^*, p_j^*\right) = \left((b_j^*, c_i^*), (b_i^*, c_j^*)\right)$ with $a_i^* = b_i^* = c_i^*$ and $a_j^* = b_j^* = c_j^*$.

If we let $\bar{B}_i = P_i$ be a compact Polish Space and we let the map $\pi_i(a)$ be simply a probability distribution over $A_{-i} \times \Delta(A_i)$, then the normal form version of GPS's psychological game is a special case of a psycho-social game.

When comparing psychological with psycho-social games, we find some key distinctive features:

(i) both games consider a preference parameter that is endogenously determined within the interaction, and in both games the parameter is taken as given at the moment to compute the best response.

(ii) In GPS' framework the preference parameter is a coherent set of beliefs, whereas

[3] Henceforth GPS.

[4] GPS is also related in a less extent to Gilboa and Schmeidler (1988), Battigalli and Dufwenberg (2006) and Caplin and Leahy (2000). However, these three papers focus on dynamic psychological games, while we deal only with simultaneous move games.

[5] A coherent belief of player i is a belief that satisfies a particular marginal restriction (for details see GPS, 1989, p. 64). A set of collectively coherent beliefs of player i is her set of beliefs in which she is sure that it is common knowledge that beliefs are coherent.

in our framework, the preference parameters do not necessarily have to be beliefs. This feature allows us to study issues such us endogenous identity formation, endogenous aspirations or utility levels as reference points, which cannot be studied with GPS's psychological games.

(*iii*) If we wanted preference parameters to be just beliefs as in GPS framework, besides the 1^{st} and $k+1$-order beliefs considered in their framework, our framework allows also for 0-order beliefs (i.e. self-beliefs). These beliefs can be interpreted either as moral values (i.e. player's own beliefs about what is right and wrong to play) or as player's self-confidence (i.e. player's own beliefs about her own actions)[6].

2.2 Social Preferences

There exists two possible approaches to incorporate social preferences into strategic models: the "distributional" and the "reciprocity-guilt" approach. The first extends individual preferences and assumes that people do not only care about their own material welfare, but they also have exogenous social concerns[7]. The second approach, however, assumes that social preferences are determined endogenously within the same strategic interaction. Rabin (1993) provides the first contribution to this approach incorporating reciprocity motives into players' preferences. Charness and Rabin (2002) developed a general model that integrates endogenous reciprocity and exogenous inequity aversion. Recently, Charness and Dufwenberg (2006) introduces *guilt* as another motivation for endogenous social preference.

Psycho-social games encompass the three models, and further offer more degree of freedoms to capture other class of social preferences that cannot be captured in the existing literature model. We shall take Charness and Rabin (2002)[8] general model with $n-$players and show that it is a special case of a psycho-social game[9].

Let A_i be Player $i's$ pure strategies and $A_{-i} = \prod_{j \neq i} A_j$ the set of pure strategies of all players but player $i \in N$. The material payoffs are determined by the action profile $a \equiv (a_i, a_{-i})$ where $u_i(a)$ represents Player $i's$ payoffs given action profile $a \in A$. Let $p \equiv (p_i, p_{-i})$ be a demerit profile, where $p_i \in [0, 1]$ is a measure of all players but i beliefs about how much player i deserves[10] and p_{-i} is a vector $(p_1, ..., p_j, ..., p_n)$ for all $j \neq i$ and represents player i's disposition towards the other players. The higher the value of p_i the less player i deserves. Given profiles a, p_i and a set of parameters

[6] See Bandura (1997) or Bandura (2000) for references about how self-confidence affects subjective wellbeing.

[7] See for example Fehr and Schmidt (1999).

[8] Henceforth C&R

[9] We will work in pure strategies for expositional purposes only.

[10] p_i (resp. p_{-i}) here is analogous to d_i (resp. d_{-i}) in Charness and Rabin's (2002) paper.

$\kappa = (\lambda, \delta, b, k, f)$,player $i's$ preferences are defined as follows:

$$v_i(a, p_{-i}) \equiv (1-\lambda_i)u_i(a) + \lambda_i \left[\begin{array}{c} \delta \left[\min \{ u_i(a), \min_{j \neq i} \{ u_j(a) + bp_j \} \} \right] + \\ (1-\delta) \left[u_i(a) + \sum_{j \neq i} \max \{ 1 - kp_j, 0 \} u_j(a) \right] - \\ f \sum_{j \neq i} p_j u_j(a) \end{array} \right]$$

where $\lambda_i \in [0,1]$ measures how much player i *cares about pursuing the social welfare* versus his own self-interest; $\delta_i \in (0,1)$ measures the degree of concern for helping the worst-off person versus maximizing the total surplus and b, k and f are nonnegative parameters.

Given a_{-i}, p_{-i} and κ, the set of player $i's$ actions that maximize her utility is:

$$A_i^*(a_{-i}, p_{-i}; \kappa) \equiv \{ a_i^* \in A_i | \, a_i^* \in \arg\max v_i(a_i, a_{-i}, p_{-i}; \kappa) \}$$

So far the only endogenous variable is a_i for all $i \in N$ and then, this is just a model with an extended utility that incorporates exogenous distributional concerns and exogenous concerns for reciprocity. If $p_i = 0$ for all $i \in N$, then this model becomes a simple model without psycho-social states. However, C&R endogenizes p_i and λ_i. It is assumed that the way player i cares about others' welfare (i.e. λ_i) depends on the action profile and on how much player i thinks that the others deserve. Formally, $\lambda_i(a, p_{-i})$ is assumed to be an upper hemi-continuous and convex-valued correspondence from (a, p_{-i}) into the set $[0,1]$ such that $\lambda_i(a, p_{-i}) \approx \{ \lambda | \, a_i^* \in A_i^*(a_{-i}, p_{-i}; \lambda) \}$. In turn, the function $\lambda_i(a, p_{-i})$ is a measure of *how appropriately other players feel that player i is behaving* when they determine how to reciprocate.

Then, C&R derive demerit profiles from these functions and assume that other players compare each $\lambda_i(a, p_i)$ with some exogenous selflessness standard $\hat{\lambda}$ - the weight they feel a decent person should put one social welfare. They define the equilibrium as follows:

The strategy profile a^* is a *reciprocal-fairness equilibrium* if for a given parameter profile $\kappa = (\hat{\lambda}, \delta, b, k, f)$, the following conditions hold for all $i \in N$:

i) given a_{-i}^*, p_i^* and p_{-i}^*, $a_i^* \in \arg\max v_i(a_i, a_{-i}^*, p_i^*, p_{-i}^*)$,

ii) given λ_i^*, $p_i^* \in \arg\max[\hat{\lambda} - \lambda_i^*, 0]$

iii) given a_{-i}^*, a_i^* and p_{-i}^*, , $\lambda_i^* \in \lambda_i(a_i^*, a_{-i}^*, p_{-i}^*)$

A reciprocal-fairness game is a special case of a psycho-social game in which the general map $\pi_i : A \times P \to \Re$ is a particular composite correspondence $(\lambda_i \circ p_i)(\kappa)$ that assigns a value from $\max[\hat{\lambda} - \lambda_i(a, p_{-i}), 0]$ to the interval $[0,1]$. It is important to point out that when player i chooses an optimal a_i, besides taking as given a_{-i}

and p_{-i}, she **takes as given** the endogenous $p_i(\lambda_i(a, p_{-i}))$ - i.e. how much other players think she deserves -, which in turn depends on her optimal action via the correspondence λ_i. Again, there is a feedback effect of player i action on her preference parameter that it is not internalized by the player when she chooses a best response. Thus a reciprocal-fairness game is a special case of a psycho-social game and therefore *all reciprocal-fairness equilibria* are psycho-social equilibria. By transitivity, all fairness equilibria (Rabin, 1993) are psycho-social equilibria too. In particular in Rabin's (1993) model $p_i = (b_j, c_i)$.

Finally, *guilt* motivations can be simply represented by means of second order beliefs $p_i = c_i$, i.e. player i beliefs about what player j believes player i will do. A guilt averse person will choose her best response trying to minimize the *guilt* caused by not conforming other's expectations. If for example, a person playing a public good game is guilt averse, and she beliefs that the others thinks she will cooperate, then, given these beliefs, she will end up cooperating.

It is true that *reciprocal-fairness* equilibria and *guilt* equilibria are also GPS's psychological equilibria. However, psycho-social games provide a richer framework than GPS's, allowing us to develop models where social preferences emerge due to different motivations. For instance, if we wanted to introduce in an strategic decision setting social concerns that are guided by individual's *commitment* to other players's welfare, psycho-social games would provide the most adequate theoretical framework to use. We shall devote the subsection below to extend more on this point.

2.2.1 Commitment and Sympathy

In "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory," Amartya Sen (1977) discusses the view held in traditional economics that "every agent is actuated only by self-interest." In his paper, Sen distinguishes two separate concepts: (i) *sympathy* and (ii) *commitment*. He argues that sympathy "corresponds to the case in which the concern for others directly affects one's own welfare. If the knowledge of torture of others makes you sick, it is a case of sympathy; if it does not make you feel personally worse off, but you think it is wrong and you are ready to do something to stop it, it is a case of commitment (p. 326)." Behaviour based on sympathy is in an important sense egoistic whereas the action based on commitment is non-egoistic. We claim that the most influential existing models on social preferences do not capture this distinction. Moreover, if they wanted to incorporate *individual commitment*, then they would need to move from a psychological game to a psycho-social game setting. We introduces *moral commitment* into a game as *0-order beliefs* affecting payoffs. For some reason, players have their *own beliefs* about what is right and what is wrong, and they enjoy when they choose an action that is consistent with those beliefs. Let's consider a discrete version of the dictator game. There are two players. Player i is the dictator who chooses either to keep (K) or give (G) to player j his experimental endowment (say £1). Player j, the recipient, has no real choice—she has to simply accept the dictator's decision. Now assume that the dictator has a moral commitment with player j, and his utility depends purely on conforming to her (endogenous) commitments. Let $\tilde{p}_i \in [0,1]$ be player $i's$ beliefs about her own actions and p_i be the probability attach to playing G.

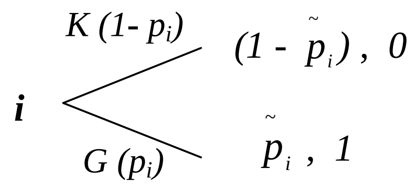


Figure 1: Commitment

This game has three psycho-social equilibrium: $p_i = \tilde{p}_i = 0$, $p_i = \tilde{p}_i = 1$ and $p_i = \tilde{p}_i = \frac{1}{2}$, with final payoffs being $1, 1$ and $\frac{1}{2}$ respectively.

When players have different motivations for social preferences, there is a possibility of miss-perception of players' kindness. A player might thing that she is being kind, when in fact, the others interpret she is being unkind. How player i interprets others' intentions will depend on player $i's$ moral frame. We shall illustrate this point with the following example proposed by Sen, about two boys who find two apples, one large and one small.

> Boy A tells boy B, "You choose." B immediately picks the larger apple. A is upset and permits himself the remark that this was grossly unfair. "Why?" asks B. "Which one would you have chosen, if you were to choose rather than me?" "The smaller one, of course," A replies. B is now triumphant: "Then what are you complaining about? That's the one you've got!" B certainly wins this round of the argument, but in fact A would have lost nothing from B's choice had his own hypothetical choice of the smaller apple been based on sympathy as opposed to commitment. A's anger indicates that this was probably not the case (p. 328-9)."

This story can be rationalized as a two-person game with only one active player chosen at random by nature with the following payoffs:

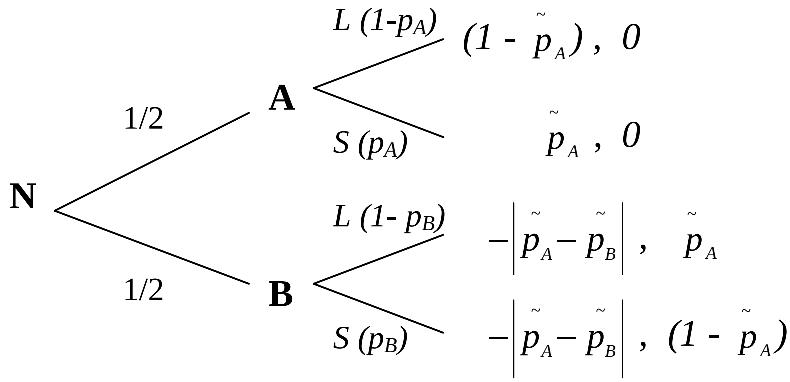


Figure 2: Commitment and sympathy

p_A (respectively p_B) represents the probability with which player A (respectively B) chooses to keep the small apple S. For expositional purposes, we shall assume that the material payoff derived from the large apple ($u_i(L) = 0$) is the same than that derived from the small one ($u_i(S) = 0$). In the standard (trivial) version of this game with just material payoffs, any possible strategy profile is an equilibrium profile.

Suppose now that Player A has *moral commitments* whereas player B just *sympathizes* with A and she cares about being fair with her.

For some reason, player A has her *own beliefs* about what is right and what is wrong, and she enjoys when she chooses an action that is consistent with her own beliefs. Moreover, she also enjoys when B chooses an action that matches her (player A) own beliefs.

Player B, on the other hand, likes to be fair with A in a particular way. We introduce fairness through *1st-order beliefs*. Player B enjoys giving player A what player A would choose to keep for her if she was to choose rather than player B[11].

Let's now focus on the payoffs of this game. If player A believes that keeping for her the small apple (i.e. playing S) is the right thing to do, then $p_A = 1$. Thus, if she had to make a choice, given her *values*, she would choose S. If player B is the one called to choose, $A's$ payoff is higher the closer is $B's$ action to $A's$ values. In other words, A prefers B choosing $p_B = 1$ rather than $p_B < 1$, since $p_B = p_A$ means that both players have the same values and player A enjoys that. Now looking at player $B's$ payoffs, he does not feel either negative or positive emotion regardless what player A chooses, since he just care about him being fair with player A. If he had to choose, he would choose to keep the large apple, L, since it gives the fairest outcome from his

[11] Note that this interpretation of fairness differes from the one taken by Rabin (1993).

point of view. However, player A would not consider this outcome to be fair when she compares $p_B = 1$ with $p_A = 0$, since from her point of view, choosing to keep L is something wrong to do.

Observe that we have here two *interconnected 1-active-player* games. Denote G-I (respectively G-II) to the game in which only player A (respectively B) is active. We say that both games are interconnected because the payoff of the players in each game depend on actions and beliefs held on the other game. Let $b_i^{k,I}$ (respectively $b_i^{k,II}$) be the k-order beliefs held by player i. Then, players' payoffs are as follows:

- In G-I: $v_A(a_A; b_A^0(a_A))$; $v_B(a_A)$
- In G-II: $v_A(a_B; b_A^0(a_A); b_A^1(a_B))$; $v_B(a_B; b_B^1(a_A))$

Both G-I and G-II are psycho-social games and G-II is also a GPS's psychological game.

Each active player take as given her beliefs and compute her best response as in any standard psychological or psycho-social game. In G-I, player $A's$ best response is $a_A^* \in \arg\max v_A(a_A; b_A^0)$ for a given b_A^0. In G-II, player $B's$ best response is $a_B^* \in \arg\max v_B(a_B; b_B^1)$ for a given b_B^1. In the equilibrium of the combined game (G-I and G-II), $b_A^{*0}(a_A^*) = a_A^*$ and $b_B^{*1}(a_A^*) = b_A^{*0}$.

This example has three psycho-social equilibria:

Type I: Player A believes that the right thing to do is to give player B the large apple ($p_A = 1$), so if she is the active player, she will choose to do so. When player B is the active player, given his beliefs, he will choose to keep the large apple for him ($p_B = 0$). Formally, Type I equilibrium is defined by the following quadruple ($a_A^* = S_A, a_B^* = L_B; b_A^{*0}(a_A^*) = S_A, b_B^{*1}(a_A^*) = b_A^{*0} = S_A$). The equilibrium payoffs are:

- If G-I is played: $(1, 0)$
- If G-II is played: $(-1, 1)$

Type II: Player A believes that the right thing to do is to give player B the small apple ($p_A = 0$), so if she is the active player, she will choose to do so. When player B is the active player, given his beliefs, he will choose to keep the small apple for him ($p' = 1$). Formally, Type II equilibrium is defined by the following quadruple ($a_A^* = L_A, a_B^* = S_B; b_A^{*0}(a_A^*) = L_A, b_B^{*1}(a_A^*) = b_A^{*0} = L_A$).The equilibrium payoffs are:

- If G-I is played: $(1, 0)$

- If G-II is played: $(-1, 1)$

Type III: Player A is "morally indifferent" between keeping the small or the large apple ($p_A = \frac{1}{2}$), so if she is active, she will randomize between her two options. When player B is the active player, he will also randomize ($p_B = \frac{1}{2}$). Formally, Type III equilibrium is defined as follows: ($p_A = \frac{1}{2}, p_B = \frac{1}{2}; b_A^{*0}(a_A^*) = \frac{1}{2}, b_B^{*1}(a_A^*) = b_A^{*0} = \frac{1}{2}$).The expected payoffs in equilibrium are:

- If G-I is played: $(\frac{1}{4}, 0)$
- If G-II is played: $(0, \frac{1}{4})$

Type I equilibrium describes the situation of Sen's example. Moreover, it can be inferred from his example that both players have a different notion of fairness. Player A whose behaviour is based on commitment, believes that player B is being fair with her when she observes ex-post that $|p_B - p_A| = 0$ or unfair otherwise. On the contrary, player B who sympathizes with A, thinks that he is being fair with A when $|p_B - p_A| \neq 0$ and unfair otherwise.

Now suppose that both players base their choices on sympathy. The payoffs of the game are:

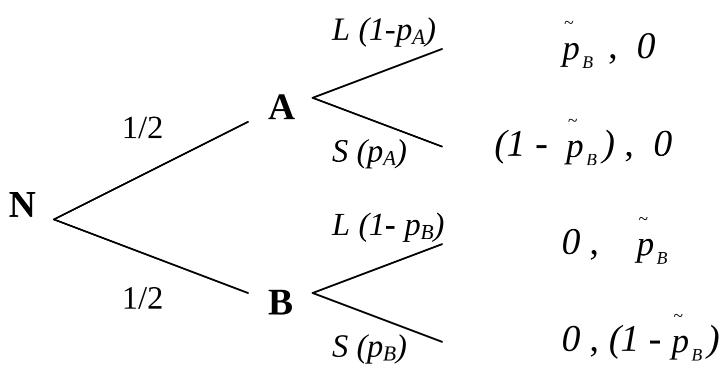


Figure 3: Both players base their choices on sympathy

Let's see what happen in the equilibrium described by Sen. If player $A's$ choice of the small apple ($p_A = 1$) was based on sympathy as opposed to commitment, then B choosing to keep the large apple ($p_B = 0$) would not be assessed by A as unfair since $|p_B - p_A| \neq 0$.

2.3 Reference-dependent Preferences in a strategic context

Since Kahneman and Tversky's (1979) seminal paper, a large number of papers in both theoretical and experimental literature have supported the hypothesis that preferences depend on a reference point. Psycho-social games allow to incorporate this type of reference-dependent preferences endogenously in a strategic interaction environment. A crucial concern that we need to address to do so is how to define and derive the reference point. This seems to be a quite understudied issue and so far there is not a clear agreement in the literature. In general, the reference point is associated with individual's status quo, being it an initial endowment of an object (e.g. mugs, money, chocolates) or a current state of the world (e.g. current action, situation, norms, etc.). Nevertheless, there are a couple of exceptions to that view. On the one hand, Shalev (2000) specifies the reference point as being equal to individual's *reference-dependent (expected) utility.* On the other hand, Kozsegi and Rabin (2006) equate the reference point with individual's *expectations,* particularly with the probabilistic *beliefs* that a person held in the recent past about outcomes.

We shall show that the first two approaches are special cases of a psycho-social game. It is not the case for the third approach, since it requires a dynamic model. Again, for the sake of exposition, we just deal with pure strategies, but it can be easily extended to mixed strategies.

If we required the reference point to be equal to *individual's status quo measured as a current state* (approach 1), we would just need to take a general psycho-social game and assume that for each $i \in N$, $\pi_i(a_i, a_{-i})$ is the identity map and $P_i \equiv A_i$.Thus, the *consistent* reference point for player i is simply an endogenous action profile $p_i = (a_i, a_{-i})$, which is taken as given when she computes her best response. If we wanted to model reference-dependent preference using Shalev's (2000) approach, then $P_i \equiv U_i$ and $p_i = u_i(a_i, a_{-i})$.

In addition, Shalev's introduces loss aversion and then he models players utility function satisfying the properties of Kahneman and Tversky's (1979) value function.

$$(2)\ v_i(a_i, a_{-i}, p_i) = \begin{cases} u_i(a_i, a_{-i}) & \text{if } u_i(a_i, a_{-i}) \geq p_i \\ u_i(a_i, a_{-i}) - \lambda_i \left[p_i - u_i(a_i, a_{-i})\right] & \text{if } u_i(a_i, a_{-i}) < p_i \end{cases}$$

where $\lambda_i \in \Re_+$ specifies player's i degree of loss aversion. Higher values of λ_i represent greater loss aversion. When $\lambda_i = 0$ player i is not loss averse[12].

In this second section, we have shown that three apparently disconnected branches of the literature on behavioural economics are all special cases of a psycho-social game.

[12] Note that this formulation captures loss-aversion for $\lambda_i > 0$ but we have not made specific assumptions in order to capture diminishing sensitivity (i.e. the value function is convex on the domain of losses and concave on the domain of gains).

Besides, we can make use of our theoretical framework to rationalize the way in which other psycho-social concerns such us identity (Akerlof and Kranton, 2000), intrinsic motivation (Benabou and Tirole, 2003), aspirations (Ray 2006, Heifetz and Minelly, 2006) and self-confidence (Benabou and Tirole, 2002) affect people's choices in a strategic environment. We shall present one of these applications in the last sections of the paper and we leave the rest for future applications.

3 Psycho-social equilibrium

3.1 The general model

Formally, the game is structured as follows. There is a finite set $N = \{1, ..., n\}$ of players (indexed by i) and for each player a finite set of (pure) actions A_i. In addition, each player is characterized by a preference parameter (or psycho-social state) $p_i \in P_i$, where P_i is a corresponding subset of some metric space. Let $A = \prod_{i \in N} A_i$, $A_{-i} = \prod_{j \in N \setminus \{i\}} A_j$ and $P = \prod_{i \in N} P_i$, $P_{-i} = \prod_{j \in N \setminus \{i\}} P_j$. We use the following notation to account for the cardinality of the sets:$m_i := |A_i|$, $\quad m := \sum\limits_{i \in N} m_i$, $\quad m^* := \prod\limits_{i \in N} m_i$ and $\rho_i := |P_i|$, $\quad \rho := \sum\limits_{i \in N} \rho_i$, $\quad \rho^* := \prod\limits_{i \in N} \rho_i$. A generic element of A_i (resp. A) is denoted by a_i (resp. a) and a generic element of P_i (resp. P) is denoted by p_i (resp. p). It is assumed that $A \subset \Re^{nm_i}$ and $P \subset \Re^{n\rho_i}$ are subsets of a finite dimensional Euclidian space[13]. For the purposes of this paper we shall consider only pure strategies.

Definition. A *consistent (pure) psycho-social state* for player i is a $\tilde{p}_i \in P_i$ such that $p_i = \pi_i(a_i, a_{-i}, p_{-i})$. The set of (pure) consistent psycho-social states is $\tilde{P}_i = \{\tilde{p}_i |\, p_i = \pi_i(a_i, a_{-i}, p_{-i}),$ for all $i \in N,\ a \in A$ and $p_{-i} \in P_{-i}\}$.

For the purpose of equilibrium existence, we will require $\pi_i(a_i, a_{-i}, p_{-i})$ to be non empty and close relative to P for each a_{-i} and p_{-i}. We do not make any other assumption on $\pi_i(\cdot)$ as we want to keep the model as general as possible.

Player *i's utility function* $v_i : A \times P_i \to \Re$ depends on the outcomes (as in the standard literature) and *also* her and other players preference parameters (or psycho-social states)[14]. We assume that player i seeks to maximize v_i given p and a_{-i}.

Definition. A *normal form psycho-social game* $\psi = (A_i, P_i; v_i, \pi_i)$ consists on an action set A_i,a set of utility parameters P_i, an utility function $v_i : A \times P_i \to \Re$, and a

[13]Note that if we wanted to include GPS psychological games as a particular case of a psycho-social game, we should consider the set P_i to be a subset of a Polish space. However, this would invalidate our proof for existence in pure strategies and so, we rather assume that P_i is a subset of a finite dimensional Euclidean Space.

[14]For expositional purposes, we assume in this section that preferences can be represented by a utility function. However, for our existence proof, we will not require preferences to be able to be represented by a utility function, i.e. we do not require preferences to be rational (complete and transitive).

map $\pi_i : A \times P_{-i} \to P_i$ for each player i.

Definition. A *(pure) psycho-social equilibrium* of a normal form psycho-social game is defined by a pair $(a^*, p^*) \in A \times P$ if, for each i,

(i) given p_i^*, p_{-i}^*, a_{-i}^*, $a_i^* \in \arg\max v_i(a_i, a_{-i}^*, \tilde{p}_i^*)$

(ii) given a_i^*, p_{-i}^* and a_{-i}^*, $\tilde{p}_i^* \in \pi_i(a_i^*, a_{-i}^*, p_{-i}^*)$

In the next section, we provide a new existence proof of a psycho-social equilibrium in both, pure and mixed strategies.

3.2 Existence in pure strategies

In this section, we examine the conditions under which a psycho-social equilibrium in pure strategies exists. To be coherent with the model we have in mind, we allow for non-convex and incomplete set of preferences due to two main reasons. Firstly, since player's preferences depend on their psycho-social states which is a reference parameter, their preferences sets may not be convex. Secondly, we allow for incomplete preferences since status quo is one of the equilibria that the game may have and Mandler (2004) and (2005) show that status quo maintenance can be rationalized by means of allowing preferences to be incomplete. Therefore, we examine the conditions for equilibrium existence in the case we allow for incomplete and non-convex preferences. The presence of non-convexities does not allow us to apply Kakutani's fix-point theorem. We apply Tarski's fix point theorem instead. Ghosal (2006b) introduces a general existence result for games with incomplete and non-convex preferences. We take Ghosal's general proof and we adapt it to our framework. We will introduce some additional notation.

Lets define an individual psycho-social state as $\sigma_i = (p, a_{-i})$ where $p = (p_1, ..., p_i, ..., p_n) \in P$ and $a_{-i} \in A_{-i}$. The primitives of the model are two maps, $\succ_{i,\sigma_i}: P \times A_{-i} \to A_i \times A_i$ and $\pi_i : A \times P_{-i} \to P_i$. The first map is a preference relation over A_i. The expression $(a_i, a_i') \in \succ_{i,\sigma_i}$is written as $a_i \succ_{i,\sigma_i} a_i'$ and is to be read as "a_i is preferred to a_i' when the psycho-social state is p and the actions chosen by other players are a_{-i}." Note that in the general model, we assume that preferences of player i are not only determined by her actions and psycho-social states, but also by the psycho-states of the other players. Define the sets $\succ_{i,\sigma_i} (a_i) = \{a_i' \in A_i : a_i' \succ_{i,\sigma_i} a_i\}$ (the upper section of $\succ_{i,\sigma_i}$) and $\succ_{i,\sigma_i}^{-1} (a_i) = \{a_i' \in A_i : a_i \succ_{i,\sigma_i} a_i'\}$ (the lower section of $\succ_{i,\sigma_i}$). We write $a_i' \notin \succ_{i,\sigma_i} (a_i)$ as $a_i \not\succ_{i,\sigma_i} a_i'$ and $a_i' \in \succ_{i,\sigma_i} (a_i)$ as $a_i' \succ_{i,\sigma_i} a_i$.

The second map specifies the set of psycho-social states consistent with the actions chosen by each individual and the psycho-social states of the others. Throughout this section, it is assumed that $\pi_{i,a,p_{-i}}$ is non-empty and closed relative to P for each a and p_{-i}.

As stated above, a (pure) psycho-social equilibrium is a pair (a^*, p^*) such that for each $i \in N$, (i) given p^* and a^*_{-i}, $\succ_{i,\sigma^*_i}(a^*_i) \cap A_i = \varnothing$ where $\sigma^*_i = (p^*, a^*_{-i})$ and (ii) given a^* and p^*_{-i}, $p^*_i \in \pi_{i,a^*,p^*_{-i}}$.

Consider the following additional set of assumptions:

Assumption 1 (AS-1): For each $i \in N$, both A_i and P are compact lattices and for each p_{-i} and a, $\pi_{i,a,p_{-i}}$ is a compact sublattice of P.

Assumption 2 (AS-2): For each $i \in N$, p and a_{-i}, $\succ^{-1}_{i,\sigma_i}(a_i)$ is open relative to A_i, i.e. $\succ^{-1}_{i,\sigma_i}$ has an open lower section.

Assumption 3 (AS-3): For each $i \in N$, p and a_{-i}, $\succ^{i}_{\sigma^i}$ is acyclic, i.e. there is no finite set $\{a_i^1, ..., a_i^n\}$ such that $a_i^k \succ^{i}_{\sigma^i} a_i^{k-1}$, $k = 2, ..., n$ and $a_i^1 \succ^{i}_{\sigma^i} a_i^n$.

Assumption 4 (AS-4): For each $i \in N$, p_{-i}, a_{-i} and $a_i \geq a'_i$, if $p_i \in \pi_{i,a,p_{-i}}$ and $p'_i \in \pi_{i,a'_i,a_{-i},p_{-i}}$, then $\sup(p_i, p'_i) \in \pi_{i,a,p_{-i}}$ and $\inf(p_i, p'_i) \in \pi_{i,a',p_{-i}}$. (psycho-social states of each player are increasing in her actions)

Assumption 5 (AS-5): For each $i \in N$, p, a_{-i} and each pair of actions $a_i, a'_i \in A_i$, (i) if $\inf(a_i, a'_i) \not\succ_{i,\sigma_i} a_i$, then $a'_i \not\succ_{i,\sigma_i} \sup(a_i, a'_i)$, and (ii) if $\sup(a_i, a'_i) \not\succ_{i,\sigma_i} a_i$, then $a'_i \not\succ_{i,\sigma_i} \inf(a_i, a'_i)$. (quasi-supermodularity)

Assumption 6 (AS-6): For each $i \in N$, $a_i \geq a'_i$, $a_{-i} \geq a'_{-i}$, $p_i \geq p'_i$ and $p_{-i} \geq p'_{-i}$, (i) if $a'_i \not\succ_{i,a_i,a'_{-i},p} a_i$, then $a \not\succ_{i,\sigma_i} a_i$, (ii) if $a_i \not\succ_{i,\sigma_i} a'_i$, then $a_i \not\succ_{i,a_i,a'_{-i},p} a'_i$, (iii) if $a'_i \not\succ_{i,a,p'_i,p_{-i}} a_i$, then $a'_i \not\succ_{i,\sigma_i} a_i$, (iv) if $a_i \not\succ_{i,\sigma_i} a'_i$, then $a_i \not\succ_{i,a,p'_i,p_{-i}} a'_i$, (v) if $a'_i \not\succ_{i,a,p_i,p'_{-i}} a_i$, then $a'_i \not\succ_{i,\sigma_i} a_i$, (vi) if $a_i \not\succ_{i,\sigma_i} a'_i$, then $a_i \not\succ_{i,a,p_i,p'_{-i}} a'_i$ (single-crossing property in actions and psycho-social states[15])

Assumption 7 (AS-7): For each $i \in N$, p, a_{-i} and $a_i \geq a'_i$, (i) if $\succ_{i,\sigma^i}(a'_i) \cap A_i = \varnothing$ and $a'_i \not\succ_{i,\sigma_i} a_i$, then $\succ_{i,\sigma^i}(a_i) \cap A_i = \varnothing$, and (ii) if $\succ_{i,\sigma^i}(a_i) \cap A_i = \varnothing$ and $a'_i \not\succ_{i,\sigma_i} a'_i$, then $\succ_{i,\sigma^i}(a'_i) \cap A_i = \varnothing$ (monotone closure)

Theorem 1: Under assumptions 1-7, there exists a psycho-social equilibrium.

Proof.

Step 1. Define a map $\Psi : A \times P \to A \times P$ as follows:

$\Psi(a, p) = \left(\Psi^1(a, p), ..., \Psi^i(a, p), ..., \Psi^n(a, p)\right)$, where

$\Psi^i(a, p) = \left(\Psi^i_1(a_{-i}, p), \Psi^i_2(a, p_{-i})\right)$ and for each i, a and p,

$\Psi^i_1(a_{-i}, p) = \left\{a'_i \in A_i : \succ_{i,\sigma^i}(a'_i) \cap A_i = \varnothing\right\}$, and $\Psi^i_2(a, p_{-i}) = \left\{p'_i \in P : p'_i \in \pi_{i,a,p_{-i}}\right\}$.

Step 2. We want to show that $\Psi^i_1(a_{-i}, p)$ is a closed and compact sublattice of A.

Since for each i, p and a_{-i} $\succ_{i,\sigma^i}$ is acyclic (*AS-3*), $\succ^{-1}_{i,\sigma_i}(a_i)$ is open relative to A_i (*AS-2*) and A_i is compact (*AS-1*), then by Bergstrom (1975), $\Psi^1_i(a_{-i}, p)$ is not empty. Note that the complement of the set $\Psi^i_1(a_{-i}, p)$ in A_i is the set $\Psi^{i,c}_1(a_{-i}, p) =$

[15]Single crossing property is the ordinal analog of increasing differences.

$\{a_i' \in A_i : \succ_{i,\sigma^i} (a_i') \cap A_i \neq \varnothing\}$. If $\Psi_1^{i,c}(a_{-i},p) = \varnothing$, then $\Psi_1^{i,c}(a_{-i},p) = A_i$ is necessarily compact. So suppose $\Psi_1^{i,c}(a_{-i},p) \neq \varnothing$. For each $a_i' \in \Psi_1^{i,c}(a_{-i},p)$, there is a $a_i'' \in A_i$ such that $a_i'' \succ_{i,\sigma^i} a_i'$, i.e. $a_i' \succ_{i,\sigma_i}^{-1} (a_i'')$. By *(AS-2)* $\succ_{i,\sigma_i}^{-1} (a_i'')$ is open relative to A_i. By definition of $\Psi^1(a,p)$, $\succ_{i,\sigma_i}^{-1} (a_i'') \subset \Psi_1^{i,c}(a_{-i},p)$. Therefore, $\succ_{i,\sigma_i}^{-1} (a_i'')$ is a non-empty neighborhood of $a_i' \in \Psi_1^{i,c}(a_{-i},p)$. Now, consider a sequence $\{a_i^k : k \geq 1\}$ such that for each $k \geq 1$, $a_i^k \in \Psi_1^i(a_{-i},p)$ but $\lim_{k\to\infty} a_i^k = \hat{a}_i \in \Psi_1^{i,c}(a_{-i},p)$. Now, by assumption, for each $a_i' \in \Psi_1^{i,c}(a_{-i},p)$, $\succ_{i,\sigma^i} (a_i') \cap A_i$ is open relative to A_i and therefore, there exists a neighborhood of $\hat{a}_i$, $N(\hat{a}_i) \subset \Psi_1^{i,c}(a_{-i},p)$, a contradiction as there exists a $K > 1$ such that for each $k > K$, $a_i^k \in N(\hat{a}_i)$. It follows that $\hat{a}_i \in \Psi_1^i(a_{-i},p)$ and therefore, $\Psi_1^i(a_{-i},p)$ is closed and since A is compact[16] *(AS-1)*, $\Psi_1^i(a_{-i},p)$ is also compact. Moreover, as $\succ_{i,\sigma^i}$is quasi-supermodular *(AS-5)*, $\Psi_1^i(a_{-i},p)$ is also ordered and therefore is a compact (and hence, complete) sublattice of A. Thus, $\Psi_1^i(a_{-i},p)$ has a maximal and a minimal element denoted by $\bar{a}_i(a_{-i},p)$ and $\underline{a}_i(a_{-i},p)$ respectively.

Step 3. Fix p. For $a_{-i} \geq a_{-i}'$, let $a_i \in \Psi_1^i(a_{-i},\cdot)$ and $a_i' \in \Psi_1^i(a_{-i}',\cdot)$. We want to show that $\sup(a_i,a_i') \in \Psi_1^i(a_{-i},\cdot)$ while $\inf(a_i,a_i') \in \Psi_1^i(a_{-i}',\cdot)$.

First, note that since $a_i' \in \Psi_1^i(a_{-i}',\cdot)$, $\inf(a_i,a_i') \not\succ_{i,a_i,a_{-i}',p} a_i'$. By part (i) of quasi-supermodularity *(AS-5)*, it follows that $a_i \not\succ_{i,a_i,a_{-i}',p} \sup(a_i,a_i')$. By part (i) of single-crossing *(AS-6)*, it follows that $a \not\succ_{i,\sigma^i} \sup(a_i,a_i')$. Since $a_i \in \Psi_1^i(a_{-i},\cdot)$, $\succ_{i,\sigma^i} (a_i) \cap A_i = \varnothing$ and therefore, by part (i) of monotone closure *(AS-7)* since $a_i \not\succ_{i,\sigma^i} \sup(a_i,a_i')$, $\succ_{i,\sigma^i} (\sup(a_i,a_i')) \cap A_i = \varnothing$. It follows that $\sup(a_i,a_i') \in \Psi_1^i(a_{-i},\cdot)$.

Next, note that since $a_i \in \Psi_1^i(a_{-i},\cdot)$, $\sup(a_i,a_i') \not\succ_{i,\sigma^i} a_i$. By part (ii) of single-crossing property, it follows that $\sup(a_i,a_i') \not\succ_{i,a_i,a_{-i}',p} a_i$.By part (ii) of quasi-supermodularity, it follows that $a_i' \not\succ_{i,a_i,a_{-i}',p} \inf(a_i,a_i')$. Since $a_i' \in \Psi_1^i(a_{-i}',\cdot)$, $\succ_{i,a_i,a_{-i}',p} (a_i') \cap A_i = \varnothing$ and therefore, by part (ii) of monotone closure, since $a_i' \not\succ_{i,a_i,a_{-i}',p} \inf(a_i,a_i')$, $\not\succ_{i,a_i,a_{-i}',p} \inf(a_i,a_i') \cap A_i = \varnothing$. It follows that $\inf(a_i,a_i') \in \Psi_1^i(a_{-i}',\cdot)$.

Step 4. Fix a_{-i} and p_{-i}. For $p_i \geq p_i'$, let $a_i \in \Psi_1^i(a_{-i},p_i,p_{-i})$ and $a_i' \in \Psi_1^i(a_{-i},p_i',p_{-i})$. We want to show that $\sup(a_i,a_i') \in \Psi_1^i(p_i,\cdot)$ while $\inf(a_i,a_i') \in \Psi_1^i(p_i',\cdot)$.

First, note that since $a_i' \in \Psi_1^i(p_i',\cdot)$, $\inf(a_i,a_i') \not\succ_{i,a,p_i',p_{-i}} a_i'$. By part (i) of quasi-supermodularity, it follows that $a_i \not\succ_{i,a,p_i',p_{-i}} \sup(a_i,a_i')$. By part (iii) of single-crossing, it follows that $a_i \not\succ_{i,\sigma^i} \sup(a_i,a_i')$. Since $a_i \in \Psi_1^i(p_i,\cdot)$, $\succ_{i,\sigma^i} (a_i) \cap A_i = \varnothing$ and therefore, by part (i) of monotone closure, as $a_i \not\succ_{i,\sigma^i} \sup(a_i,a_i')$, $\succ_{i,\sigma^i} (\sup(a_i,a_i')) \cap A_i = \varnothing$. It follows that $\sup(a_i,a_i') \in \Psi_1^i(p_i,\cdot)$.

Next, note that since $a_i \in \Psi_1^i(p_i,\cdot)$, $\sup(a_i,a_i') \not\succ_{i,\sigma^i} a_i$. By part (iv) of single-crossing property, it follows that $\sup(a_i,a_i') \not\succ_{i,a,p_i',p_{-i}} a_i$. By part (ii) of quasi-supermodularity,

[16] If A_i is compact, $A = \prod_{i\in I} A_i$ is also compact.

it follows that $a_i' \not\succ_{i,a,p_i',p_{-i}} \inf(a_i, a_i')$. Since $a_i' \in \Psi_i^1(p_i', \cdot)$, $\succ_{i,a,p_i',p_{-i}} (a_i') \cap A_i = \varnothing$ and therefore, by part (ii) of monotone closure, as $a_i' \not\succ_{i,a,p_i',p_{-i}} \inf(a_i, a_i')$, $\not\succ_{i,a,p_i',p_{-i}} \inf(a_i, a_i') \cap A_i = \varnothing$. It follows that $\inf(a_i, a_i') \in \Psi_i^1(p_i', \cdot)$.

Step 5. Fix a_{-i} and p_i. For $p_{-i} \geq p_{-i}'$, let $a_i \in \Psi_1^i(a_{-i}, p_i, p_{-i})$ and $a_i' \in \Psi_1^i(a_{-i}, p_i, p_{-i}')$. We want to show that $\sup(a_i, a_i') \in \Psi_1^i(p_{-i}, \cdot)$ while $\inf(a_i, a_i') \in \Psi_1^i(p_{-i}', \cdot)$.

First, note that since $a_i' \in \Psi_1^i(p_{-i}', \cdot)$, $\inf(a_i, a_i') \not\succ_{i,a,p_i,p_{-i}'} a_i'$. By part (i) of quasi-supermodularity, it follows that $a_i \not\succ_{i,a,p_i,p_{-i}'} \sup(a_i, a_i')$. By part (v) of single-crossing, it follows that $a_i \not\succ_{i,\sigma^i} \sup(a_i, a_i')$. Since $a_i \in \Psi_1^i(p_{-i}, \cdot)$, $\succ_{i,\sigma^i} (a_i) \cap A_i = \varnothing$ and therefore, by part (i) of monotone closure, as $a_i \not\succ_{i,\sigma^i} \sup(a_i, a_i')$, $\succ_{i,\sigma^i} (\sup(a_i, a_i')) \cap A_i = \varnothing$. It follows that $(\sup(a_i, a_i')) \in \Psi_1^i(p_{-i}, \cdot)$.

Next, note that since $a_i \in \Psi_1^i(p_{-i}, \cdot)$, $\sup(a_i, a_i') \not\succ_{i,\sigma^i} a_i$.By part (vi) of single-crossing property, it follows that $\sup(a_i, a_i') \not\succ_{i,a,p_i,p_{-i}'} a_i$. By part (ii) of quasi-supermodularity, it follows that $a_i' \not\succ_{i,a,p_i,p_{-i}'} \inf(a_i, a_i')$. Since $a_i' \in \Psi_i^1(p_{-i}', \cdot)$, $\succ_{i,a,p_i,p_{-i}'} (a_i') \cap A_i = \varnothing$ and therefore, by part (ii) of monotone closure, as $a_i' \not\succ_{i,a,p_i,p_{-i}'} \inf(a_i, a_i')$, $\not\succ_{i,a,p_i,p_{-i}'} \inf(a_i, a_i') \cap A_i = \varnothing$. It follows that $\inf(a_i, a_i') \in \Psi_i^1(p_{-i}', \cdot)$.

Step 6. It follows that both $\bar{a}_i(a_{-i}, p)$ and $\underline{a}_i(a_{-i}, p)$ are increasing in p and in a_{-i}. Further, since for each p_{-i} and a, $\pi_{i,a,p_{-i}}$ is a compact (and therefore complete) sublattice of P *(AS-1)*, $\Psi_2^i(a, p_{-i})$ has a maximal and a minimal element (in the usual component wise vector ordering): denote these by $\bar{p}_i(a, p_{-i})$ and $\underline{p}_i(a, p_{-i})$ respectively. As $\pi_{i,a,p_{-i}}$ is increasing in a_i *(AS-4)*, both $\bar{p}_i(a, p_{-i})$ and $\underline{p}_i(a, p_{-i})$ are increasing in a_i as well. It follows that the $(\bar{a}_i(a_{-i}, p), \bar{p}_i(a, p_{-i}))$ is an increasing function from $A \times P$ to itself and since $A \times P$ is compact (and hence, complete) lattice, by applying Tarski's fix-point theorem (Tarski, A., 1955), it follows that $(\bar{a}, \bar{p}) = (\bar{a}_i(\bar{a}_{-i}, \bar{p}), \bar{p}_i(\bar{a}, \bar{p}_{-i}))$ is a fix-point of Ψ. By a symmetric argument $(\underline{a}_i(a_{-i}, p), \underline{p}_i(a, p_{-i}))$ is an increasing function from $A \times P$ to itself and therefore, $(\underline{a}, \underline{p}) = (\underline{a}_i(\underline{a}_{-i}, \underline{p}), \underline{p}_i(\underline{a}, \underline{p}_{-i}))$ is also a fix-point of Ψ. Moreover,

$$(\bar{a}, \bar{p}) = \sup\{(a, p) \in A \times P : (\bar{a}_i(\bar{a}_{-i}, \bar{p}), \bar{p}_i(\bar{a}, \bar{p}_{-i})) \geq (a, p)\}$$

and

$$(\underline{a}, \underline{p}) = \inf\left\{(a, p) \in A \times P : (\underline{a}_i(\underline{a}_{-i}, \underline{p}), \underline{p}_i(\underline{a}, \underline{p}_{-i})) \leq (a, p)\right\}$$

Therefore, $(\bar{a}, \bar{p})$ and $(\underline{a}, \underline{p})$ are, respectively, the largest and smallest (in the usual component wise vector ordering) fix-points of Ψ. ■

Remark. The seminal equilibrium existence proof in games with incomplete preferences is due to Shafer and Sonnenschein (1975). They show the existence of equilibria in generalized games with incomplete, non-ordered preferences under continuity and convexity assumptions on preferences and actions sets. Their continuity and convexity assumptions on preferences ensure that a fix-point exists via an application of Kakutani's fix-point theorem.

3.3 Existence of a mixed psycho-social equilibrium

When we proved the existence of a psycho-social equilibrium in pure strategies, we paid the cost of assuming that psycho-social states are increasing in player's actions (AS- 4). This assumption is appropriate for some applications such as motivation and aspiration failures, but it is not for others. Of course, by assuming that p_i is increasing in a_i, we gained in many other aspects. In particular, we did not have to make the standard assumptions such as A and P are convex sets, preferences are complete, transitive and have open graphs. In this section, we relax (AS-4) and we look at the existence of a mixed psycho-social equilibrium. As it was the case for the existence of an equilibrium in pure strategies, we will not require preferences to have an expected utility representation. Further, we will assume that preference has open and lower sections, which is a weaker continuity assumption than the standard assumption of continuity proposed in the literature but it is still stronger than (AS-2) made in Theorem 1.[17]

As before, for each $i \in N = \{1, ..., n\}$ there are two sets, a set A_i of pure actions, $A_i \subset \Re^{m_i}$ and a set P_i of psycho-social states, $P_i \subset \Re^{\rho_i}$, where $\Re^{m_i}$ and $\Re^{\rho_i}$ are finite dimensional Euclidian spaces. For each $i \in N$, there is a preference map $\succ_{i,\sigma_i}$: $P \times A_{-i} \rightarrow A_i \times A_i$, where for each p and a_{-i}, describes a preference relation over A_i Recall that $P = \prod_{i \in N} P_i$. Let the space of all Borel probability distributions over P_i (respectively A_i) be denoted by $\Delta(P_i)$ (respectively $\Delta(A_i)$),for each $i \in N$. As P_i (respectively A_i) is separable (in the usual topology) and $\Delta(P_i)$ (respectively $\Delta(A_i)$) is endowed with the topology of weak convergence, $\Delta(P_i)$ (respectively $\Delta(A_i)$) is separable and metrizable by the Lévy-Prokhorov metric[18].

Let $\hat{\succ}_i : \Delta(P_i) \times \Delta(P_{-i}) \times \Delta(A_{-i}) \rightarrow \Delta(A_i) \times \Delta(A_i)$, be a map that describes preferences over probability distributions over P and A_{-i} where for each $\mu \in \Delta(P_i) \times \Delta(P_{-i})$ and $s_{-i} \in \Delta(A_{-i})$, $\hat{\succ}_{i,\mu,s_{-i}}$describes a preference relation over $\Delta(A_i)$. For s_i and s'_i

[17] In general it is assumed in the literature that preferences have open graphs (see Shafer and Sonnenschein, 1975).

[18] The Lévy-Prokhorov metric is a metric (i.e. a definition of distance) on the collection of probability measures on a given metric space. For more details see Billingsley (1999).

$\in \Delta(A_i)$,the expression $(s_i, s_i') \in \hat{\succ}_{i,\mu,s_{-i}}$is written as $s_i \ \hat{\succ}_{i,\mu,s_{-i}} s_i'$ and is to be read as "s_i is preferred to s_i' by player i when the distribution over the set of psycho-social states and the actions chosen by other players is μ_i and s_{-i} respectively." For any set X, note that the set of Dirac probability measures[19] over X is simply X itself. Let $\hat{\succ}_i^D : P \times A_{-i} \to A_i \times A_i$ denote the restriction of $\hat{\succ}_i$to Dirac probability measures. We assume the preferences of player i over probability distributions are consistent with her preferences over pure actions, i.e. $\hat{\succ}_i^D \equiv \succ_i$. As before, define the sets $\hat{\succ}_{i,\mu,s_{-i}}(s_i) = \left\{s_i' \in A_i : s_i' \hat{\succ}_{i,\mu,s_{-i}} s_i\right\}$ (the upper section of $\hat{\succ}_{i,\mu,s_{-i}}$) and $\hat{\succ}_{i,\mu,s_{-i}}^{-1}(s_i) = \left\{s_i' \in A_i : s_i \hat{\succ}_{i,\mu,s_{-i}} s_i'\right\}$ (the lower section of $\hat{\succ}_{i,\mu,s_{-i}}$). We write $s_i' \notin \hat{\succ}_{i,\mu,s_{-i}}(s_i)$ as $s_i \hat{\not\succ}_{i,\mu,s_{-i}} s_i'$ and $s_i' \in \hat{\succ}_{i,\mu,s_{-i}}(s_i)$ as $s_i' \hat{\succ}_{i,\mu,s_{-i}} s_i$.

As before, there is a map $\pi_i : A \times P_{-i} \to P_i$ that specifies the set of psycho-social states consistent with the actions chosen by each individual and the psycho-social states of the others. In this part of the paper, it is assumed that $\pi_{i,a,p_{-i}}$ is non-empty for each $a \in A$ and $p_{-i} \in P_{-i}$ and π_i is a continuous function on $A \times P_{-i}$.

A *(mixed) psycho-social equilibrium* is a pair (s^*, μ^*) such that for each $i \in N$, (i) given μ^* and s^*_{-i}, $\hat{\succ}_{i,\mu^*,s^*_{-i}}(s_i^*) \cap A_i = \varnothing$ and (ii) for each $i \in N$, and for each p such that $\mu^*(p) >> 0$, there is $a \in A$ such that $s^*(a) >> 0$ and $\mu_i^*(p_i) = \mu_i^*(\pi_i(a^*, p^*_{-i})) = s^*(a)$.

We make the following assumptions:

Assumption 1' (AS-1'): For each $i \in N$, both A_i and P are compact and hence, $\Delta(A_i)$ and $\Delta(P)$ are compact sets.

Assumption 2' (AS-2'): For each $i \in N$, μ and s_{-i}, both $\hat{\succ}_{i,\mu,s_{-i}}$and $\hat{\succ}^{-1}_{i,\mu,s_{-i}}$are open relative to $\Delta(A_i)$, i.e. $\hat{\succ}_{i,\mu,s_{-i}}(s_i)$ has both open upper and lower sections.

Assumption 3' (AS-3'): For each $i \in N$, μ and s_{-i}, $\hat{\succ}_{i,\mu,s_{-i}}$is acyclic, i.e. there is no finite set $\left\{s_i^1, ..., s_i^n\right\}$ such that $s_i^k \hat{\succ}_{i,\mu,s_{-i}} s_i^{k-1}$, $k = 2, ..., n$ and $s_i^1 \hat{\succ}_{i,\mu,s_{-i}} s_i^n$.

Assumption 4' (AS-4'): For each $i \in N$, μ, s_{-i}, s_i and s_i', if $\hat{\succ}_{i,\mu,s_{-i}}(s_i) \cap A_i \neq \varnothing$ and $\hat{\succ}_{i,\mu,s_{-i}}(s_i') \cap A_i \neq \varnothing$, then $\hat{\succ}_{i,\mu,s_{-i}}(\lambda s_i + (1-\lambda) s_i') \cap A_i \neq \varnothing$ for each $\lambda \in [0,1]$ (convexity)

Theorem 2: Under assumptions *(AS-1')* to *(AS-4')*, a mixed strategy psycho-social equilibrium exists.

Proof.

Step 1'. Define a map $\hat{\Psi} : \Delta(A) \times \Delta(P) \to \Delta(A) \times \Delta(P)$,
$\hat{\Psi}(s, \mu) = \left(\hat{\Psi}^1(s,\mu), ..., \hat{\Psi}^i(s,\mu), ..., \hat{\Psi}^n(s,\mu)\right)$, where
$\hat{\Psi}^i(s,\mu) = \left(\hat{\Psi}^i_1(s_{-i},\mu), \hat{\Psi}^i_2(s,\mu_{-i})\right)$ and for each $i \in N$, s and μ,

[19] The Dirac measure is a measure δx on a set X (with any sigma algebra of subsets of X) that gives the singleton set $\{x\}$ the measure 1, for a chosen element $x \in X$.

$$\hat{\Psi}^i_1(s_{-i},\mu) = \left\{ s'_i \in \Delta(A_i) : \hat{\succ}_{i,\mu,s_{-i}}(s'_i) \cap A_i = \varnothing \right\}$$

and

$$\hat{\Psi}^i_2(s,\mu_{-i}) = \left\{ \begin{array}{c} \mu_i \in \Delta(P_i) : \mu^*(p) >> 0, \text{ iff } \exists\, a \in A \\ \text{s.t. } s^*(a) >> 0 \text{ and } \mu^*_i(p_i) = \mu^*_i(\pi_i(a^*, p^*_{-i})) = s^*(a). \end{array} \right\}$$

Step 2'. Using a similar argument to the one used in Theorem 1, step 2, we know that $\hat{\Psi}^i_1(s_{-i},\mu)$ is non-empty and compact.

Step 3'. Now we want to show that $\hat{\Psi}^i_1(s_{-i},\mu)$ is upper semi-continuous. As the range of $\hat{\Psi}^i_1(s_{-i},\mu)$ is compact, $\hat{\Psi}^i_1(s_{-i},\mu)$ is upper semi-continuous if $\hat{\Psi}^i_1(s_{-i},\mu)$ has the closed graph property. Consider four convergent sequences $\left\{ s^\tau_i, s^\tau_{-i}, \mu^\tau_i, \mu^\tau_{-i} : \tau \geq 1 \right\}$ such that $\lim_{\tau\to\infty} s^\tau_i = \hat{s}_i$, $\lim_{\tau\to\infty} s^\tau_{-i} = \hat{s}_{-i}$, $\lim_{\tau\to\infty} \mu^\tau_i = \hat{\mu}_i$, $\lim_{\tau\to\infty} \mu^\tau_{-i} = \hat{\mu}_{-i}$ and for each $\tau \geq 1$, $s^\tau_i \in \hat{\Psi}^i_1(s^\tau_{-i},\mu^\tau)$, with $\mu^\tau = (\mu^\tau_i, \mu^\tau_{-i})$ but $\hat{s}_i \notin \hat{\Psi}^i_1(\hat{s}_{-i},\hat{\mu})$, with $\hat{\mu} = (\hat{\mu}_i, \hat{\mu}_{-i})$ i.e. $\hat{s}_i \in \Psi^{i,c}_1(\hat{s}_{-i},\hat{\mu}) = \left\{ s'_i \in \Delta(A_i) : \hat{\succ}_{i,\mu,s_{-i}}(s'_i) \cap A_i \neq \varnothing \right\}$. Again, by the assumption that $\hat{\succ}_{i,\mu,s_{-i}}(\hat{s}_i)$ has an open lower section (AS-2'), arguments similar to those used in Theorem 1, step 2 show that there exists a non-empty neighborhood of $\hat{s}_i$, $N(\hat{s}_i) \subset \Psi^{i,c}_1(\hat{s}_{-i},\hat{\mu})$, a contradiction as there exists a $\bar{\tau} > 1$ such that for each $\tau > \bar{\tau}$, $s^\tau_i \in N(\hat{s}_i)$. Now for each $s'_i \notin \hat{\Psi}^i_1(\hat{s}_{-i},\hat{\mu})$, there is a $s''_i \in \Delta(A_i)$ such that $s'_i \hat{\succ}_{i,\hat{\mu},\hat{s}_{-i}} s''_i$ i.e. $s'_i \in \hat{\succ}_{i,\hat{\mu},\hat{s}_{-i}}(s''_i)$. By assumption (AS-2'), $\hat{\succ}_{i,\hat{\mu},\hat{s}_{-i}}(s''_i)$ is open relative to $\Delta(A_i)$ and therefore there is a neighborhood $N(s'_i) \subset \hat{\succ}_{i,\hat{\mu},\hat{s}_{-i}}(s''_i)$. As $s^\tau_i \in \hat{\Psi}^i_1(s^\tau_{-i},\mu^\tau)$, there is some $s'_i \in \hat{\Psi}^i_1(\hat{s}_{-i},\hat{\mu})$ and $\breve{\tau} \geq 1$ such that for all $\tau > \breve{\tau}$, $s^\tau_i \in N(s'_i)$ and therefore, s'_i is a limit point of the sequence $\{s^\tau_i : \tau \geq 1\}$, a contradiction as all the subsequences of convergent sequence must have the same limit. It follows that $\hat{\Psi}^i_1(s_{-i},\mu)$ has the closed graph property. Moreover, by (AS-4') $\hat{\Psi}^i_1(s_{-i},\mu)$ is also convex.

Step 4'. By the continuity of the map $\pi_i(.)$, $\hat{\Psi}^i_2(s,\mu_{-i})$ is also a continuous function.

Step 5'. It follows that $\hat{\Psi}$ satisfies all the assumption of Fan-Glicksberg fix-point theorem and therefore has a fix-point (s^*, μ^*), which by construction, is a random psycho-social equilibrium. ■

4 Nash vs. Psycho-social equilibria

> "...In essence, we constantly ignore Kelley and Thibaut's (1978) long accepted observation that the people in our experiments transform the payoff matrices that we give them, and then they act in ways that maximize their transformed outcomes. Yet we constantly ignore this knowledge, probably so that we don't have to complicate our work too much." (in J. Murnighan and A. Roth, forthcoming, p.8)

What do people play when they play a normal-form game? Do they actually play the game that the experimenter thinks they play? This is an open question and the vast experimental evidence showing that players do not behave in a simultaneous move game as it is predicted by Nash is still puzzling. Some scholars argue that people may misunderstand the game and that is why they don't play Nash. However, experimenters are very careful by making the rules and the payoffs of the game simple and clear in order to minimize any possible misunderstanding of the game. In some cases such as the Dictator or the Ultimatum Game, the simplicity of the game is so clear that it is difficult to imagine that people don't understand the game. In this paper we argue that people may play psycho-social games instead of playing normal form games, and thus we provide a theoretical framework to understand a general behaviour in the experimental setting. Indeed, we generalize what Rabin (1993), Charness and Rabin (2002), Charness and Dufwenberg (2006) or Shalev (2000) and others have done by providing partial answers to our initial question by means of "implicitly" assuming particular functional forms for π_i. Rabin (1993), in particular, derives a "reciprocity" psychological game from some basic material games providing a link between both classes of games. However, Rabin's analysis is restrictive to the particular functional form he assumes for the map π_i. In principle, one could use other sensible functional forms to derive other psycho-social games from the same normal form game. One single normal-form game can be associated with many psycho-social games, not just a "reciprocity" one.

Example 1. Consider the following prisoner's dilemma game with material payoffs where $X > 0$:

		Player 2	
		C	D
Player 1	C	$4X, 4X$	$0, 6X$
	D	$6X, 0$	X, X

(Table 1)

One can think on deriving many sensible (mathematically infinite) psycho-social games from this standard prisoner's dilemma. Let's consider some examples. Let a_i be player i beliefs about what she chooses, b_j be player i beliefs of what player j chooses and c_i be player i beliefs about what player j believes player i chooses.

a) Fairness-Reciprocity Game: If $p_i = (b_j, c_i)$ and $v_i(a_i, p_i) = x_i(a_i, b_j) + \rho_i \left[f_i(a_i, b_j)\tilde{f}_j(b_j, c_i) \right]$ with $\rho_i \geq 0$ being a constant measuring $i's$ sensitivity to reciprocity, $f_i(a_i, b_j)$ being $i's$ kindness towards j and $\tilde{f}_j(b_j, c_i)$ being $j's$ kindness towards i, then on can derive a "fairness-reciprocity" game from the standard prisoner's dilemma game (Rabin, 1993).

b) Guilt Game: If $p_i = (c_i)$ and $v_i(a_i, p_i) = x_i(a_i, b_j) - \gamma_i \max\{0, x_j(c_i) - x_j(a_i)\}$, with $\gamma_i \geq 0$ being a constant measuring $i's$ guilt aversion, $x_j(c_i)$ and $x_j(a_i)$ being $j's$ material payoffs given c_i and a_i respectively, then one can derive a "guilt-aversion" game from the same standard prisoner's dilemma game.

c) Commitment Game: If $p_i = (a_i)$ and $v_i(a_i, p_i) = x_i(a_i, b_j) - \alpha_i \max\{0, x_j(a_i) - x_j(a_i)\}$, with $\alpha_i \geq 0$ being a constant measuring $i's$ sensitivity to moral commitment and $x_j(a_i)$ being $j's$ material payoffs given a_i respectively, then one can derive a "commitment" game from the same standard prisoner's dilemma game.

Interestingly, a cooperative equilibrium (C,C) exists in a any of this three psycho-social games associated with the prisoner's dilemma material game. In particular, (C,C) is a "fairness-reciprocity" equilibrium if $\rho_i \geq 4X$, it is a "guilt-aversion" equilibrium if $\gamma_i \geq \frac{1}{2}$, and it is a "commitment" equilibrium for any $\alpha_i \geq \frac{1}{2}$.[20]

In general, one could think in other type of psycho-social games that can be derived from an arbitrary material normal form game. In the remaining of the section, we introduce a general way to associate a normal form game with a psycho-social game. Then, we show that if we take any normal form game and we look at the psycho-social equilibria of those psycho-social games associated to it, then we shall find that they are generically different from the Nash equilibria of the original normal form game we took. Throughout this section we will work just in pure strategies to simplify the notation and to gain intuition, but the analysis shall be extended to mixed strategies in the near future.

4.1 Embeddedness

Let $\psi := \{A_i, P_i; v_i : A \times P_i \to \Re; \pi_i : A \to P_i\}$ denote any finite n-person **psycho-social game** as defined in section 3.3[21]. Let $\psi \in \Psi(A_1, ..., A_n)$ where $\Psi(A_1, ..., A_n)$ is the *set of all finite psycho-social* games with strategy spaces $A_1, ..., A_n$. Now, let $\lambda := \left\{\tilde{A}_i; u_i : \tilde{A} \to \Re\right\}$ denote any finite n-person **standard normal form game** where $\tilde{A}_i$ is a finite nonempty set of pure strategies. Let $\lambda \in \Lambda(\tilde{A}_1, ..., \tilde{A}_n)$ where $\Lambda(\tilde{A}_1, ..., \tilde{A}_n)$ is the *set of all finite standard normal form* games with strategy spaces $\tilde{A}_1, ..., \tilde{A}_n$.

Definition 1: Take any λ from the set Λ and any ψ from the set Ψ. We say that

[20] $(\mathbf{C}, \mathbf{C})$ is a "reciprocity-fairness" equilibrium iff for $i = 1, 2$: $x_i(\mathbf{C}, \mathbf{C}) + \rho_i \left[\tilde{f}_j(\mathbf{C}, \mathbf{C}) f_i(\mathbf{C}, \mathbf{C})\right] \geq x_i(\mathbf{D}, \mathbf{C}) + \rho_i \left[\tilde{f}_j(\mathbf{C}, \mathbf{C}) f_i(\mathbf{D}, \mathbf{C})\right] \Leftrightarrow 4X + \rho_i[\frac{1}{2}\frac{1}{2}] \geq 6X + \rho_i[\frac{1}{2}(-\frac{1}{2})] \Leftrightarrow \rho_i \geq 4X$

$(\mathbf{C}, \mathbf{C})$ is a "guilt-aversion" equilibrium iff for $i = 1, 2$:

$x_i(\mathbf{C}, \mathbf{C}) - \gamma_i \max\{0, x_j(\mathbf{C}) - x_j(\mathbf{C})\} \geq x_i(\mathbf{D}, \mathbf{C}) - \gamma_i \max\{0, x_j(\mathbf{C}) - x_j(\mathbf{D})\} \Leftrightarrow 4X \geq 6X - 4X\gamma_i \Leftrightarrow \gamma_i \geq \frac{1}{2}$

And it can be similarly shown that $(\mathbf{C}, \mathbf{C})$ is a "commitment" equilibrium.

[21] To simplify the notation, we shall assume that preferences do not depend on p_{-i} and we shall also rule out the possibility that p_{-i} affects p_i.

$\lambda \in \Lambda$ is **consistently embedded** into $\psi \in \Psi$ (with the following notation: $\lambda \overrightarrow{e} \psi$) if, for all $i \in N$,

(1) $A_i \equiv \tilde{A}_i$ and (2) $v_i(a, p_i) = u_i(a)$ for any $p_i = \pi_i(a)$ and all $a \in A$.[22]

There are two other - rather extreme - ways to define a normal form game embedded into a psycho-social game, although we decide to work with the "consistently embedded" definition for two main reasons. First, we could have said that λ is "weakly embedded" into ψ if, for all $i \in N$, condition (1) and (2) hold for any $p_i \in P_i$, which does not need to be a consistent p_i. We rule-out this restriction since it does not have much intuition behind. Further, we could have said that λ is "strongly embedded" into ψ if for all $i \in N$, condition (1) and (2) hold for some $p_i = p_i^* \in P_i$. If we consider this definition, we are imposing the overly-strong restriction that the ψ has at least one psycho-social equilibrium in pure strategies, and then we rule out games without pure strategy psycho-social equilibrium.

Now we have defined a general way to associate a normal form game to a psycho-social game, we shall compare the set of equilibria of both classes of games.

Conjecture. Let NE be the set of Nash equilibrium strategy profiles of $\lambda \in \Lambda$ and PSE be the set of psycho-social equilibrium strategy profiles of $\psi \in \Psi$. If $\lambda \overrightarrow{e} \psi$ then generically, under standard regularity assumptions, both sets of equilibria in pure strategies are distinct from each other.

Proof.

Let $\lambda \overrightarrow{e} \psi$, so that we work with $A_i \equiv \tilde{A}_i$ for all $i \in N$.

Assume that for all $i \in N$

i) $A_i = [\underline{a}, \bar{a}]_i$ is an interval of $\Re$,

ii) $v_i : A \times P_i \to \Re$ and $u_i : A \to \Re$ are smooth strictly concave functions in A_i for all $i \in N$.

Then

a) the strategy profile $a^* = (a_i^*, a_{-i}^*) \in ((\underline{a}_i, \bar{a}_i), (\underline{a}_{-i}, \bar{a}_{-i}))$ is a (pure strategy) Nash Equilibrium (interior solution) of any $\lambda \in \Lambda$ iff

$$\frac{\partial u_i}{\partial a_i}(a_i^*, a_{-i}^*) = 0, \qquad \text{for all } i \in N$$

b) the strategy profile $a^* = (a_i^*, a_{-i}^*) = (\underline{a}_i, \underline{a}_{-i})$ is a (pure strategy) Nash Equilibrium (corner solution) of any $\lambda \in \Lambda$ iff

[22] Note that there is no technical reason for which we have to imposse "payoff equivalece" between the normal form game and the projected psycho-social game. It would be sufficient to state that $a^* = \tilde{a}^*$, i.e. both are strategically equivalent. However, we prefer to work with "payoff equivalence" because we think that it simplifies the notation and it is innocuous.

$$\frac{\partial u_i}{\partial a_i}(a_i^*, a_{-i}^*) \leq 0, \qquad \text{for all } i \in N$$

c) the strategy profile $a^* = (a_i^*, a_{-i}^*) = (\bar{a}_i, \bar{a}_{-i})$ is a (pure strategy) Nash Equilibrium (corner solution) of any $\lambda \in \Lambda$ iff

$$\frac{\partial u_i}{\partial a_i}(a_i^*, a_{-i}^*) \geq 0, \qquad \text{for all } i \in N$$

Analogously,

d) the profile $(a^*, p^*) = (a_i^*, a_{-i}^*, p_i^*, p_{-i}^*) \in \big(((\underline{a}_i, \bar{a}_i), (\underline{a}_{-i}, \bar{a}_{-i})), p^*\big)$ is a (pure) Psycho-social Equilibrium (interior solution) of any $\psi \in \Psi$ iff for all $i \in N$

$$\frac{\partial v_i}{\partial a_i}(a_i^*, a_{-i}^*, p_i^*) = 0, \qquad \text{and} \quad p_i^* = \pi_i(a_i^*, a_{-i}^*, p_{-i}^*)$$

e) the profile $(a^*, p^*) = (a_i^*, a_{-i}^*, p_i^*, p_{-i}^*) = \big(\underline{a}_i, \underline{a}_{-i}, p^*\big)$ is a (pure) Psycho-social Equilibrium (corner solution) of any $\psi \in \Psi$ iff for all $i \in N$

$$\frac{\partial v_i}{\partial a_i}(a_i^*, a_{-i}^*, p_i^*) \leq 0, \qquad \text{and} \quad .p_i^* = \pi_i(a_i^*, a_{-i}^*, p_{-i}^*)$$

f) the profile $(a^*, p^*) = (a_i^*, a_{-i}^*, p_i^*, p_{-i}^*) = ((\bar{a}_i, \bar{a}_{-i}), p^*)$ is a (pure) Psycho-social Equilibrium (corner solution) of any $\psi \in \Psi$ iff for all $i \in N$

$$\frac{\partial v_i}{\partial a_i}(a_i^*, a_{-i}^*, p_i^*) \geq 0, \qquad \text{and} \quad .p_i^* = \pi_i(a_i^*, a_{-i}^*, p_{-i}^*)$$

Since $\lambda \overrightarrow{e} \ \psi$, then

$$u_i(a_i^*, a_{-i}^*) = v_i(a_i^*, a_{-i}^*, \pi_i(a_i^*, a_{-i}^*))$$

Thus, in a interior solution, the strategy profile $a^* = (a_i^*, a_{-i}^*)$ is a (pure strategy) Nash Equilibrium of $\lambda \overrightarrow{e} \ \psi$ iff

$$\frac{\partial u_i}{\partial a_i}(a_i^*, a_{-i}^*) = \frac{\partial v_i}{\partial a_i} + \frac{\partial v_i}{\partial \pi_i}\frac{\partial \pi_i}{\partial a_i} = 0 \qquad \text{for all } i \in N \tag{1}$$

whereas, as shown in d) the same equilibrium strategy profile is part of a (pure) psycho-social equilibrium of ψ in which λ is embedded into if:

$$\frac{\partial v_i}{\partial a_i}(a_i^*, a_{-i}^*, p_i^*) = 0 \text{.......for all } i \in N$$

An analogous analysis can be done in the two cases with corner solutions.

Therefore, the set of Nash equilibria of $\lambda\overrightarrow{e}\ \psi$ will coincide with the set of psycho-social equilibria of ψ iff the second term of equation (1) is equal cero. This condition holds if:

i) the trivial case: $p_i = 0$ or

ii) $\pi_i(a)$ is just a constant or

iii) $(v_i(., \pi_i))$ is just a constant.

iv) We are in a loss-aversion model

But these are just isolated cases. ■

Remarks:

1. We focus on the interior solutions just to simplify the analysis. Note that the essential point of proposition 2 is the second term of eq. (1). This term measures the externality that is in fact internalized by player i when she chooses an action to maximizes her utility given the actions of the others. If we assume that there is a feedback effect generated in the interaction (given by any consistent p_i), then a non myopic player will take this feedback effect into account, and choose a best response accounting for the consequences that her optimal choice will have on her psycho-social state and thus, on her utility. So, it is not puzzling to observe that in general people don't play as predicted by Nash.

2. Both, the definition of embeddedness and proposition 2 consider equilibria in pure strategies. If we wanted to allow for mixed strategies, then the first order conditions used for the proof would not be valid any more.

5 Myopia vs Sophistication

In this section, we explore the different meanings of a player i fixing p_i when computing her best response. We analyze why we need this assumption and we explore what would happen if the assumption was relaxed and instead we assumed that players do consider the change that their own actions produce on their preferences. We find different results depending on the game we explore. For instance, in Rabin's (1993), although players are assumed to be myopic, one could get the same results assuming sophisticated players. The same is true in Shalev's (2000) model and it is shown in his Proposition 2. However, in other games (e.g. with guilt or commitment) the assumption of myopia introduces additional welfare ranked equilibria which can make the myopic player be worse-off or better-off than a sophisticated player, depending on the game.

5.1 Interpretation

Let us consider two impressionistic descriptions of a psycho-social game to start the analysis about the meanings of p_i.

Example 2. Paul has two options: he can go to the school $\bar{a}_i$ or stay at home $\underline{a}_i$. Suppose that his self-confidence is higher when he goes to school than when he stays at home, and suppose that his utility is increasing in his self-confidence. If Paul did internalize the effect of his action on his *self-confidence*, he would choose to go to school, since $u_i(\bar{a}_i, \bar{p}_i) > u_i(\underline{a}_i, \underline{p}_i)$, with $\bar{p}_i > \underline{p}_i$. However, if he did not internalize the effect of his actions on p_i, not going to school could be a possible equilibrium. This will happen for instance, if his preferences are such that $u_i(\bar{a}_i, \underline{p}_i) < u_i(\underline{a}_i, \underline{p}_i)$. In words, if Paul does not believe on herself enough, he prefers staying at home than going to the school.

Example 3. Paul has to decide whether to contribute $\bar{a}_i$ or not $\underline{a}_i$ with some money to a common pool made by his mates to build a common football pitch in the neighbourhood. Paul does not even play football, but he is *guilt averse*. Let the guilt be represented by the psycho-social state $p_i = c_i$ (i.e. Paul's second order beliefs). Suppose Paul's preferences are such that given $\bar{p}_i = \bar{c}_i$, $u_i(\bar{a}_i, \bar{p}_i) > u_i(\underline{a}_i, \bar{p}_i)$ and given $\underline{p}_i = \underline{c}_i$, $u_i(\bar{a}_i, \underline{p}_i) < u_i(\underline{a}_i, \underline{p}_i)$. That is, regardless what the others will do and despite he will not use the football pitch, Paul will contribute given that his beliefs are $\bar{p}_i = \bar{c}_i$.

Observe that in both examples, Paul does not internalize the consequences of his actions on his preferences and ends up doing something that from the outsider's point of view is suboptimal. If we allowed Paul to internalize the effect on his different p_i, he would rather go to school and he would rather not contribute to the football pitch. But is Paul being irrational or boundedly rational? The answer is a matter of how we define bounded-rationality. What is clear here, is that in the first case, being sophisticated means being a determined person and in the second case it means being an unscrupulous person.

In general, a psycho-social game adds an "extra dimension" to a standard normal form game in the form of an endogenous but fixed preference parameter. It is indeed something that players do not control when they evaluate the gains from deviating[23]. But why it is the case that players do not control it?. There are three different answers to this question and which answer is more convincing will depend on the specific situation we want to model. First, it may be that they just **cannot do it** due to some

[23] In this sense, this paper is related to the literature of affective decision making (see Bracha, 2004 for an application to insurance markets)

psychological constraint (example 2) or even moral (example 3) or cultural constraints. In these cases, players' autonomy to choose (i.e. players' agency) is endogenously and internally restrained in the interaction. Other reason may be that they **do not know how to** control it. This may happen whenever they have to make a choice for the first time in their life or just because it is very costly to get the introspection need to be sophisticated in a psycho-social context. Finally, it may just be that they just **do not want to** control it, as in some situations, the myopic player may be better-off than a sophisticated player. We illustrate this point with the following example.

Example 4. Paul has two options: he can either smoke, $\bar{a}_i$ or not smoke, $\underline{a}_i$. He feels calm when he smokes, $\underline{p}_i = \pi_i(\bar{a}_i)$ and anxious when he does not smoke, $\bar{p}_i = \pi_i(\underline{a}_i)$. When he is anxious he prefers smoking to abstaining, $u_i(\bar{a}_i, \bar{p}_i) = 1 > u_i(\underline{a}_i, \bar{p}_i) = 0$, while when he is calmed he prefers abstaining to smoking, $u_i(\underline{a}_i, \underline{p}_i) = 1 > u_i(\bar{a}_i, \underline{p}_i) = 0$. In this example, if the player was sophisticated, he would get a payoff of 0, $u_i(\underline{a}_i, \bar{p}_i) = u_i(\bar{a}_i, \underline{p}_i) = 0$. However, if he is myopic, there is a unique psycho-social equilibrium in mixed strategies $(\frac{1}{2}\underline{a}_i + \frac{1}{2}\bar{a}_i, \frac{1}{2}\underline{p}_i + \frac{1}{2}\bar{p}_i)$ with an expected payoff of $\frac{1}{2} > 0$.

5.2 Myopia in Social Preferences

How does the "fixed p_i" assumption works in the endogenous social preferences models? The existing models on endogenous social preferences assume myopia in players. However, it turns out that the particular specification of Rabin's (1993) model makes this assumption innocuous - i.e. the set of equilibria are the same under any of the two assumptions. It is not the case, however, with "guilt-aversion" and "commitment" models.

Consider the three motivations for social preferences discussed in section 4, example 1. If (a^*, p^*) is a psycho-social equilibrium profile of a "fairness with reciprocity" game, "guilt" game and "commitment" game, then given p^*, $v_i(a_i^*;p_i^*) \geq v_i(a_i';p_i^*)$ for each player i. In particular:

a) in a "fairness with reciprocity" game, player i does not have individual incentives to deviate from playing a_i iff $v_i(a_i;b_j, c_i) \geq v_i(a_i';b_j, c_i)$. When a myopic reciprocal player i computes the gains from deviating, she re-computes $f_i(a_i', b_j)$ but she leaves $\tilde{f}_j(b_j, c_i)$ unchanged, i.e. she does not fully internalize the consequence of her deviation on her set of beliefs because c_i remains fixed. However, if the reciprocal player was sophisticated, she would have incentives to deviate from playing a_i iff $v_i(a_i;b_j, c_i) \geq v_i(a_i';b_j, c_i')$ and $a_i' = c_i'$.[24] Then, given the new set of beliefs $p_i = (b_j, c_i')$, sophisticated-reciprocal player

[24] Note that first order beliefs b_2 always are kept fixed when evaluating a deviation, as in any standard game.

i will re-compute $f_i(a_i', b_j)$ **and** $\tilde{f}_j(b_j, c_i')$ and evaluate the gains from deviating. It turns out that given the particular specification of Rabin's (1993) model $\tilde{f}_j(b_j, c_i') = \tilde{f}_j(b_j, c_i)$, so the set of equilibria assuming either sophisticated or myopic players is the same. Intuitively, it seems more reasonable to think that $\tilde{f}_j(b_j, c_i') \neq \tilde{f}_j(b_j, c_i)$. Going back to the prisoner's dilemma example, one would expect that if player j was reciprocal, he would be less kind with player i if he believes player i will defect when j cooperates. However, Rabin's specification implies that player i interprets that when j cooperates, he is being equally generous regardless what player i does. This contra-intuitive result seems to highlight a failure in the specification of the model. Intuitively, one would expect $\tilde{f}_j(C, D)$ to be higher than $\tilde{f}_j(C, C)$.

b) In "guilt" and "commitment" games, it is clear that the assumption of myopia may change the set of equilibria (see example 3 for a "guilt" game). Consider the pure "commitment" game in figure 1, section 2. In that example, if the player were sophisticated, he would solve the following problem

$$\underset{p_i}{Max}(1-p_i)(1-\tilde{p}_i) + p_i\tilde{p}_i \quad st.p_i = \tilde{p}_i, p_i \in [0,1] \quad \Leftrightarrow$$

$$\underset{p_i}{Max}(1-p_i)^2 + p_i^2 \text{ s.t. } p_i \in [0,1]$$

which has two corner solutions or equilibria: $p_i = \tilde{p}_i = 0$ and $p_i = \tilde{p}_i = 1$ (both with payoff of 1). Note that there is no mixed strategy equilibrium in the sophisticated case, although a mixed strategy psycho-social equilibrium with payoff of $\frac{1}{2}$ does exist when we assume myopic players.

5.3 Theoretical Remarks

One concern that may come to the reader's mind is whether a psycho-social game can be understood just as any standard normal form game with $n+1$ players, in which the additional player chooses a p_i from "her action set" P_i, given the actions of the other n players. However, our game is not analogous to such a game for two main reasons. First, the map $\pi_i : A \to P_i$ should be required to be the best response correspondence of the additional player, which would be just a particular case of the infinite classes of maps π_i that we allow in our game. More important, if n players are randomizing over actions, then it must be the case that this particular $n+1$ player is also randomizing. Nevertheless, this cannot be possible if the best response in mixed strategies is the function π_i, because the best response in mixed strategies must make other players indifferent among the actions in the support of A_{-i}, so any way of randomizing over the set of pure actions can be a best response. So, there must be always more than one best response, and this cannot be possible if the best response is simply the function π_i.

6 Endogenous and Exogenous Frames

Each psycho-social state can be interpreted as an *endogenous* frame (beliefs, emotions, utility reference levels, motivations, aspirations, self-confidence, etc) that mediates players' decisions. However, these are not the type of frames regarded as "framing effects" in the behavioural economics literature (see for instance, Tversky and Kahneman, 1986). The latter are simply *exogenous frames* such us wording, colour, labeling, context, etc., that have been largely empirically proved to affect observed behaviour.

Both *exogenous* and *endogenous* frames may be interrelated with each other:

a) Endogenous (e.g. p_i) and exogenous (e.g. θ_i) frames can affect preferences separately, $v_i(a_i, a_{-i}, p_i(a_i, a_{-i}), \theta_i)$ or

b) Exogenous frames can affect preferences through affecting endogenous frames, $v_i(a_i, a_{-i}, p_i(a_i, a_{-i}, \theta_i)$. Dufwenberg et. al (2006) show that exogenous frames such us labelling and valence affect first and second order beliefs and players' contributions to a public good.

In the following subsection, we present a model for aspiration formations as an application of a psycho-social game in which both frames affect preferences separately.

6.1 Application: Aspirations formation and chronic poverty

Most of the theoretical models in the development literature[25], with some exceptions such us Ray (2003) and Heifetz and Minelli (2006), look for the causes of persistent poverty on people's external constraints: market or institution failure. Consequently, they disregard the endogenous psycho-social constraints that are inherent to the condition of chronic poverty. Inner problems such us lack of self-confidence and lack of aspiration, are particularly well documented in the literature of development, psychology, sociology and anthropology. Mookherjee (2006) argues that "long-run poverty is fundamentally self-perpetuating [and] the entrapment goes hand in hand with [...] lack of hope". Moreira (2003) adds that this lack of hope together with low self-esteem is also a common characteristic in the personality of the Brazilian North-eastern. "As the poor lose their values, they no longer believe in themselves. They go through a process of Nihilism [denial of hope]". Stern (2004) refers to this issue arguing that "an individual can be constrained by their aspirations and perceptions of their role, so that development depends on relaxing these constraints." Then he adds "to understand path

[25] See for example Azaridis and Stachurski (2004) or Azaradis (2004) for a literature review on Poverty Traps

out of poverty, we have to focus not only on the growth of opportunity but also on [...] internal constraint on aspirations and behaviour [...] that limit poor people's ability to participate." Disregarding these psycho-social endogenous constraints does not just imply the existence of a theoretical gap in the economic literature, but also a "real world" problem when it comes to develop anti-poverty policies. For example, creating jobs is not necessarily an effective policy to solve an aspiration failure. As Atkinson (1998) argues, ending social exclusion will depend on the nature of these new jobs. Do they restore a sense of control? Do they provide an acceptable relative status? Do they offer prospects for the future? Likewise, income support programs may be exclusionary, as they could make the recipients feel excluded by the state (Atkinson 1998). The psycho-social games we introduce in this paper provide an appropriate framework to account for this type of interdependence between individual's preferences, agency and her (relative) extrinsic circumstances.

6.1.1 A simple example

Based on Appadurai (2004), Ray (2003) argues that poverty and failure of aspirations may be reciprocally linked in a self-sustaining trap. Here we present a very simple example of a psycho-social game in which such a self-sustaining aspiration trap emerges as one of the multiple possible equilibria.

The game consists of two interdependent individuals who make a decision of either maintaining the existing status quo or changing it by means of, for example, undertaking higher education.

Consider two individuals $i = 1, 2$ whose payoff relevant variables are:

- the action set $A = A_1 \times A_2$ where $A_i = \{\underline{a}_i, \bar{a}_i\}$ representing, respectively, maintaining the existing status quo and changing the status quo.
- the set $\Theta = \Theta_1 \times \Theta_2$ of exogenous frame: relevant external environment (wealth, health, nutrition, housing, etc.), where $\Theta = \left[\underline{\theta}, \bar{\theta}\right]$.
- the set $P = P_1 \times P_2$ of psycho-social states (or endogenous frame), with P_i representing the set of $i's$ psycho-social considerations such us aspirations (or confidence, intrinsic motivation, etc.). $P \subset \Re$ and for simplicity we assume that $p = p_1 = p_2$.

In addition, consider a map $\pi : A \to P$ generating players aspirations for each configuration of both players' actions.

Finally, there is a map $s_i : A_i \times \Theta_i \to \Theta_i$ that generates a new external environment for the individual as a function of her actions and the initial external environment. Let $s_i = (a_i, \theta_i)$ increases with a_i and in θ_i.

Further, assume that the preferences of each player can be represented by the utility function $v_i(a_i, p, \theta_i) = b_i(\tilde{\theta}_i) - c_i(a_i, p)$ where $b_i(\tilde{\theta}_i)$ is the benefit that each individual obtains when the realized external circumstances is $\tilde{\theta}_i$ and $c_i(a_i, p)$ is the cost of effort, which we assume that is decreasing in p but increasing in a_i. Let $u_i(a_i, p, \theta_i) = v_i(a, p, s_i(a_i, \theta_i)) = b_i(s_i(a_i, \theta_i)) - c_i(a_i, p)$. For simplicity, assume that the individual's utility from preserving the status quo $u_i(\underline{a}_i, p, \theta_i)$ is normalized to zero for all values of p, θ_i and $u_i(\bar{a}_i, p, \theta_i)$ is the gain (or loss) to each individual in deviating from status quo. Further, assume that for each a_i, $u_i(a_i, p, \theta_i)$ is continuous in p and θ_i and that both individuals have the same preferences (i.e. $u_i = u_j = u$).

For each p and θ_i, let $\alpha(p, \theta_i)$ be the set of actions that maximize individual $i's$ payoffs.

A psycho-social equilibrium $(P - S.E)$ is a pair (a^*, p^*) such that for each θ_i, (i) given p^* and a_j^*, $a_i^* \in \alpha(p^*(a_j^*), \theta_i)$ and (ii) given a^*, $p^* \in \pi(a^*)$

Interpersonal complementarity

Let's assume that $i's$ psycho-social states is increasing in $j's$ actions (a_j).

Since $c_i(a_i, p)$ is decreasing in p, and p is increasing in actions, then the model implies that,

(i) for each θ_i and a_j, $u(\bar{a}_i, p, \theta_i) > u(\bar{a}_i, p', \theta_i)$, $p > p'$

(ii) for each θ_i and p, $u(\bar{a}_i, p, \theta_i) > u(\bar{a}_i, p(\bar{a}_i, a_j'), \theta_i)$, $a_j > a_j'$

In words, given θ_i, $u(a_i, p, \theta_i)$ has strictly increasing differences in (a, p)[26], i.e. player $i's$ marginal return of going to school is higher the higher her aspirations and if player j goes to school.

Moreover, since $s_i = (a_i, \theta_i)$ is increasing in θ_i,

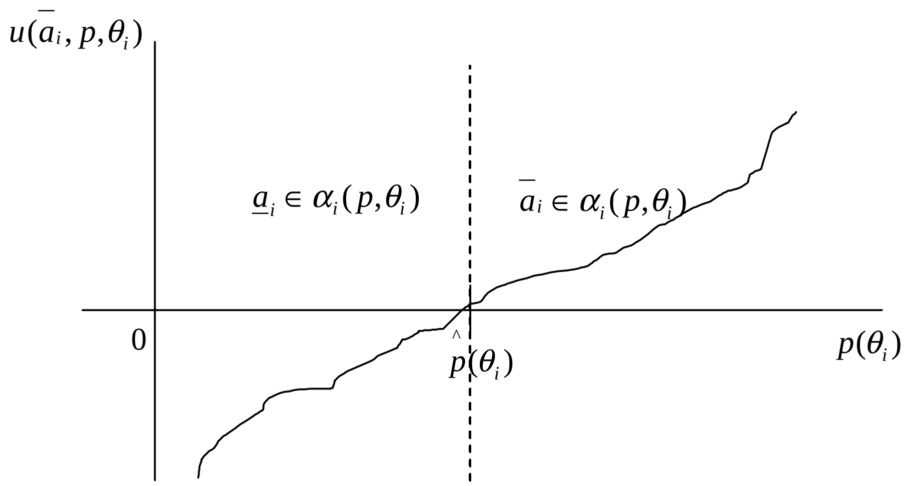
(iii) for each θ_i and p, $u(\bar{a}_i, p, \theta_i) > u(\bar{a}_i, p, \theta_i')$, $\theta_i > \theta_i'$

What are the psycho-social equilibria of this example and how do they relate to (θ_i, θ_j)?

Under the assumptions made so far, there exist a unique solution $\hat{p}(\theta_i)$ for the equation $u(\bar{a}_i, p, \theta_i) = 0$. That is, the level of psycho-social state that player i needs in order to be indifferent between going to school or not, is higher the poorer she is.

Given θ_i, the best responses for player i (and by symmetry for player j) are:

[26]Let $S \subset \Re$ and $A \subset \Re$ be the parameter and action sets respectively. A function $F : S \times A \to \Re$ has (strictly) increasing differences in (s, a) if $F(s', a') - F(s, a')(>) \geq F(s', a) - F(s, a), \forall a' > a, s' > s$



(i) if $p > \hat{p}\ (\theta_i)$, then $u(\bar{a}_i, p, \theta_i) > 0$, so $\bar{a}_i = \alpha(p, \theta_i)$

(ii) if $p < \hat{p}(\theta_i)$, then $u(\bar{a}_i, p, \theta_i) < 0$, so $\underline{a}_i = \alpha(p, \theta_i)$

(iii) if $p = \hat{p}(\theta_i)$, then $u(\bar{a}_i, p, \theta_i) = 0$, so $\{\underline{a}_i, \bar{a}_i\} = \alpha(p, \theta_i)$

Since p is generated by a_i and a_j, let $\pi(\underline{a}_i, \underline{a}_j) = p_A$, $\pi(\bar{a}_i, \underline{a}_j) = p_B$, $\pi(\underline{a}_i, \bar{a}_j) = p_C$ and $\pi(\bar{a}_i, \bar{a}_j) = p_D$.

In this context, there exist four types of equilibrium:

- **Type A equilibrium**: if $p_A \leq \hat{p}(\theta_i)$ and $p_A \leq \hat{p}(\theta_j)$ then for each θ_i and θ_j, there exist a unique equilibrium (a^*, p^*) where $a^* = (\underline{a}_i, \underline{a}_j)$ and $p^* = p_A$.

Proof. By (ii) we know that if $p < \hat{p}(\theta_i)$ and $p < \hat{p}(\theta_j)$, then $\underline{a}_i = \alpha(p, \theta_i)$ and $\underline{a}_j = \alpha(p, \theta_j)$. Since $p = p_A = \pi(\underline{a}_i, \underline{a}_j)$, then $p^* \in \pi(a^*)$ and the second condition for an equilibrium holds.

Now, if $p_A = \hat{p}(\theta_i) = \hat{p}(\theta_j)$, by (iii) the best responses are $\alpha(p, \theta_i) = \{\underline{a}_i, \bar{a}_i\}$ and $\alpha(p, \theta_j) = \left\{\underline{a}_j, \bar{a}_j\right\}$. But since $p = p_A = \pi(\underline{a}_i, \underline{a}_j)$, then $(\bar{a}_i, \bar{a}_j)$ is not consistent with this equilibrium. Therefore, if $p_A = \hat{p}(\theta_i) = \hat{p}(\theta_j)$ the equilibrium is $a^* = (\underline{a}_i, \underline{a}_j)$ and $p^* = p_A$.

Finally, if $p_A > \hat{p}\ (\theta_i)$ and $p_A > \hat{p}\ (\theta_j)$, then the best responses are $\alpha(p, \theta_i) = \bar{a}_i$ and $\alpha(p, \theta_i) = \bar{a}_j$. but this is not consistent with $p = p_A = \pi(\underline{a}_i, \underline{a}_j)$. Therefore if $p_A > \hat{p}$ (θ_i) and $p_A > \hat{p}\ (\theta_j)$, type A equilibrium does not exist.

- **Type B and C equilibria**: if $p_B \leq \hat{p}(\theta_i)$ and $p_B > \hat{p}(\theta_j)$ then for each θ_i and θ_j, there exist a unique equilibrium (a^*, p^*) where $a^* = (\underline{a}_i, \bar{a}_j)$ and $p^* = p_B$

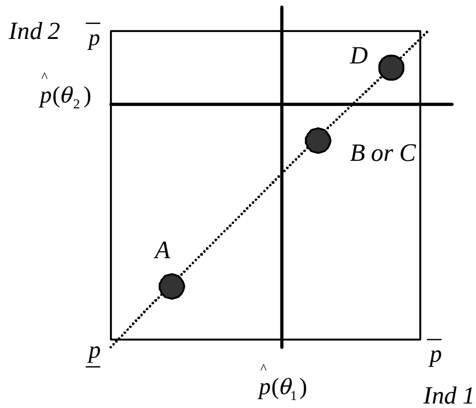


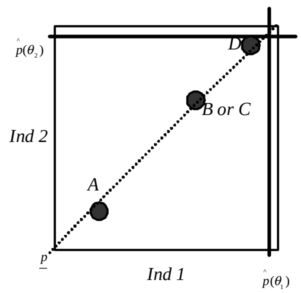
Figure 1: Multiple equilibria

(Type B). Likewise, if $p_C > \hat{p}(\theta_i)$ and $p_C \leq \hat{p}(\theta_j)$ the equilibrium is (a^*, p^*) where $a^* = (\bar{a}_i, \underline{a}_j)$ and $p^* = p_2$ (Type II").

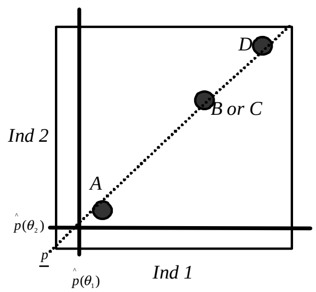
Proof. We know that if $p \leq \hat{p}(\theta_i)$ and if $p > \hat{p}(\theta_j)$ then $\underline{a}_i = \alpha(p, \theta_i)$ and $\bar{a}_j = \alpha(p, \theta_i)$. Since $p = p_B = \pi(\underline{a}_i, \bar{a}_j)$, then $a^* = (\underline{a}_i, \bar{a}_j)$ with $p^* = p_B$ is an equilibrium. Analogously with type C equilibrium.

- **Type D equilibrium**: if $p_D > \hat{p}(\theta_i)$ and $p_D > \hat{p}(\theta_j)$ then for each θ_i and θ_j, there exist a unique equilibrium (a^*, p^*) where $a^* = (\bar{a}_i, \bar{a}_j)$ and $p^* = p_D$. The proof is analogous to the others.

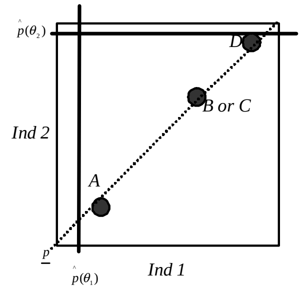
Note that by *(iii)* $\hat{p}(\theta_i)$ and $\hat{p}(\theta_j)$ are decreasing in θ_i and θ_j, respectively, so, given the preferences of each individual and (θ_i, θ_j), the model may present up to four different equilibria.



Both very poor: Type A is the unique equilibrium



Both very rich: Type D is the unique equilibrium



1 rich and 2 poor: Type C is the unique equilibrium

7 Conclusion

In this paper, we have introduced a general class of simultaneous move games in which the payoff of each player depends not only on her strategy profile, but also on her preference parameters. The preference parameters are, in turn, endogenously determined in equilibrium. We named this class of games "psycho-social games".

We have shown existence of a (pure strategy) psycho-social equilibrium under incomplete and acyclic preferences without concavity assumptions. We also provided an existence proof on mixed strategies without an explicit expected utility representation.

Further, we have studied how psycho-social games provide an appropriate theoretical framework to analyze issues of development in which psycho-social concerns play an important role, such as chronic poverty, aspirations, intrinsic motivation, and empowerment. These issues have not only theoretical relevance but also important policy implications.

We have also outlined the way in which our framework also generalizes some existing theoretical models in the behavioural economics literature. Our paper provides a general picture to rationalize most of the existing experimental results on simultaneous move games. Once we have a psycho-social game in mind, there is not reason to be puzzled anymore when we observe that people don't play simultaneous move games as Nash predicted. We have shown that, typically, the set of Nash equilibria and the set of psycho-social equilibria of an associated psycho-social game are distinct from each other.

There are several possible extensions of the work reported here. On the theoretical side, we are currently working on the links between Quantal Response Equilibrium and Psycho-social Equilibrium. Further, we shall analyze the eventual evolution of psycho-social states in repeated and dynamic games settings. On the applications side, we shall explore how the presence of endogenous and non-observed psycho-social states explain the emergence of prejudices, social conflict and anti-social behaviour.

References

[1] Atkinson, A.B. (1998) "Social Exclusion, Poverty and Unemployment," in *Exclusion, Employment and Opportunity*, A. B. Atkinson and J. Hills (Eds.) Centre for Analysis of Social Exclusion, January.

[2] Akerlof G. and R. Kranton (2000) "Economics and Identity," *Quarterly Journal of Economics*, CXV(3), August, 715-733.

[3] Azariadis C. and J. Stachurski (2004) "Poverty Traps," Forthcoming in the *Handbook of Economic Growth*, Philippe Aghion and Steven Durlauf, eds.

[4] Azariadis, C. (2004): "The theory of poverty traps: what have we learned?" in *Poverty Traps*, S. Bowles, S. Durlauf and K. Ho, eds.

[5] Appadurai (2004) "The Capacity to aspire," in V. Rao and M. Walton (eds.), *Culture and Public Action*, The International Bank for Reconstruction and Development, The World Bank, Washington, DC

[6] Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman

[7] Bandura A. (2000) "Self-efficacy: The foundation of agency," in W. J. Perrig (2000) (Ed.), *Control of human behavior, mental processes, and consciousness*, p. 17-33. Erlbaum

[8] Bandura, A. (2001) "Social Cognitive Theory: An Agentic Perspective," *Annual Review of Psychology*, 2001, 52, pp. 1-26.

[9] Banerjee, A. V. (2005) "New Development Economics and the Challenge to Theory," *Economical and Political Weekly*, October 1.

[10] Battigalli, P. and M. Dufwenberg (2005) "Dynamic Psychological Games," mimeo.

[11] Benabou, R. and Tirole, J. (2002) "Self-Confidence and Personal Motivation," *Quarterly Journal of Economics*, 117, 3, 871-915.

[12] Benabou R. and J. Tirole (2003) "Intrinsic and Extrinsic Motivation," *Review of Economic Studies*, 70, 489-520

[13] Bergstrom, T.C. (1975) "Maximal elements of acyclic relations on compact sets," *Journal of Economic Theory*, 10 (3), 403-404.

[14] Billingsley, Patrick (1999). *Convergence of Probability Measures*. John Wiley & Sons, Inc., New York.

[15] Bracha, Annat (2004) "Affective Decision Making in Insurance Markets," Yale ICF

Working Paper No. 04-03

[16] Camerer, C. (2003) *Behavioral Game Theory. Experiments in Strategic Interaction*, Russell Sage Foundation, Princeton University Press, New Jersey.

[17] Caplin, A. and Leahy (2001) "Psychological Expected Utility Theory and Anticipatory Feelings," *Quarterly Journal of Economics*, 116, 55-79.

[18] Charness G. and M. Rabin (2002) "Understanding Social Preferences with a Simple Tests," *Quarterly Journal of Economics,* 117, 817-869.

[19] Charness G. and M. Dufwenberg (2006) "Promises and Partnership," *Econometrica*, 74, 6, 1579-1601.

[20] Dufwenberg, M., Gachter S and H. Hennig-Schmidt (2006) "The Framing of Games and the Psychology of Strategic Choice," *CeDEx Discussion Paper No 2006-20.*

[21] Fehr, E. and K. M. Schmidt (1999) "A Theory Of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, 114, 3, 17-868

[22] Forgas, J., K. D. Williams and L. Wheeler (2000) *The Social Mind. Cognitive and Motivational Aspects of Interpersonal Behavior*, Cambridge University Press, UK, 460 pages.

[23] Geanakoplos, J., D. Pearce and E. Stacchetti (1989) "Psychological games and sequential rationality," *Games and Economics Behavio*r, 1,1, 60-79.

[24] Ghosal, S (2006a) "Self-influence," mimeo, University of Warwick.

[25] Ghosal, S (2006b) "Equilibrium with Incomplete Preferences: Non-convexity and Strategic Complementarity," mimeo, University of Warwick.

[26] Gilboa, I and D. Schmeidler (1988) "Information Dependent Games: Can Common Sense Be Common Knowledge?," *Economics Letters*, 27, 215-221.

[27] Heifetz, A and E. Minelli (2006) "Aspiration Traps," mimeo.

[28] Kahneman D. and A. Tversky (1979) "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, XLVII, 263-291.

[29] Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations. A theory of interdependence* New York: Wiley-Interscience

[30] Mandler, M. (2004) "Status Quo Maintenance Reconsidered: Changing or Incomplete Preferences," *Economic Journal*, 114, F518-F535.

[31] Mandler, M. (2005) "Incomplete Preferences and Rational Intransitivity of Choice," *Games and Economic Behavior*, 50, 255-277.

[32] Mookherjee D. (2006) "Poverty Persistence and Design of Anti-Poverty Policies," in Banerjee, R. Benabou and D. Mookherjee (eds), *Understanding Poverty*, Oxford University Press.

[33] Moreira V. (2003) "Poverty and Psychopathology," in Carr S. C. and Sloan T. S. (eds.) *Poverty and Psychology. From Global Perspective to Local Practice*, Kluwer Academic/Plenun Publishers New York.

[34] Mullainathan S. (2006) "Better Choices to Reduce Poverty," in Banerjee, R. Benabou and D. Mookherjee (eds), *Understanding Poverty*, Oxford University Press.

[35] Murnighan, J.K. and A. E. Roth, "Some of the Ancient History of Experimental Economics and Social Psychology: Reminiscences and Analysis of a Fruitful Collaboration," *Social Psychology and Economics*, D. De Cremer, M. Zeelenberg, and J.K. Murnighan, editors, Lawrence Erlbaum Associates, Inc.: Mahwah, NJ. forthcoming.

[36] Rabin, M. (1993) "Incorporating Fairness into Game Theory and Economics," *American Economic Review*, 83, 5, 1281-1302.

[37] Ray, D. (2006) "Aspirations, Poverty and Economic Change," in A. Banerjee, R. Benabou and D. Mookherjee (eds) *Understanding Poverty*, Oxford University Press

[38] Rao, V. and M. Walton (2004) "Culture and Public Action: Relationality, Equality of Agency, and Development," in V. Rao and M. Walton (eds.), *Culture and Public Action*, The International Bank for Reconstruction and Development, The World Bank, Washington, DC.

[39] Sen Amartya (1977) "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory," *Philosophy and Public Affairs*, 6, 4, 317-344.

[40] Shafer, W. and H. Sonnenschein (1975) "Equilibrium in abstract economies without ordered preferences," *Journal of mathematical Economics*, 2, 345-348.

[41] Shalev J. (2000) "Loss Aversion Equilibrium," *International Journal of Game Theory*, 29, 269-287.

[42] Stern N., J-J. Dethier and F. H. Rogers (2005) *Growth and Empowerment. Making Development Happen*, Munich Lectures in Economics, The MIT Press, Cambridge, Massachusetts, and London, England

[43] Tarski, A. (1955): "A lattice theoretical fixed-point theorem and its applications," *Pacific J. of Math*, 5, 285–309

[44] Tversky A. and D. Kahneman (1986) "Rational Choice and the Framing of Decisions," *Journal of Business*, 59, 4, 5251-78.