Scientific
Research
Publishing

Social Capital, Perceptions and Economic Performance

José A. Hernández[1], César Guerrero-Luchtenberg[2]

[1]Department of Applied Economic Analysis, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain
[2]Department of Economic Analysis, University of Zaragoza, Campus de Teruel, Teruel, Spain
Email: joseantonio.hernandez@ulpgc.es, cleonardo@unizar.es

Abstract

This paper describes how social capital emerges, relates to economic performance and evolves in the long run. Using the concept of psychological equilibrium, two types of individuals are generated in the population regarding their willingness to cooperate. We propose an evolutionary (learning) process over those types driven by the total payoffs of the psychological game, and provide a complete description of its dynamics. Macro-perceptions, defined as the individual perception of how cooperative the society is as a whole, are key to explain convergence to the full social capital state in the long run.

Keywords

1. Introduction

The aim of this paper is to explain how social capital is created, evolves and how it is related to economic performance and perceptions. Our main contributions are the following: Firstly, we introduce the concept of macro-perceptions as the individual and common perception of how likely cooperation is in a society. Social capital is framed through the likelihood of cooperation in a Prisoner's Dilemma game. Secondly, a standard replicator dynamics is adopted in this context. We show that if there are enough people favorable to social capital and if macro-perceptions are optimistic enough, then society tends to be populated only by those favorable to social capital in the long run. Also, the best economic outcome is achieved in this case. Finally, the paper characterizes long-run performance under various conditions and carries out several comparative statics in terms of macro-perceptions, the size of the economy or the material payoff weight in the utility function.

The term social capital refers in general to the degree of anonymous trust, the willingness to enforce contracts, to finance public goods or the social cohesion that exists in a society. It is a relevant variable to assess the economic efficiency and it is closely related to the ability of institutions to establish effective rules. Social capital mainly affects economic activity by reducing production costs in some cases (associated with litigation, security, surveillance, bureaucracy and complexity of contracts, for example) or in others, by easing transactions which would not take place otherwise. Intuitively, social capital may be seen as a catalyst for achieving efficiency improvements, or formally speaking, to reach Pareto optimal situations.

The proposed model studies social capital framed as the likelihood of cooperation in a Prisoner's Dilemma game by applying the concept of psychological equilibrium. Our main results can be summarized in three elements. First, we show that mutual cooperation can be achieved as an equilibrium if perceptions about cooperation are optimistic and agents are reciprocal enough. Second, we consider two types of agents, the so called *pro-social* and the *basic* one. Then, we introduce the concept of macro-perceptions as the probability that two pro-social players will in fact cooperate when matched. Third, by means of the replicator dynamics we endogenize the distribution of types in the population. The result of it is that when pro-social type agents are able to coordinate themselves on mutual cooperation through favorable macro-perceptions, they are better off than the *basic* agents and will be the only type of people in the population. In addition, the model characterizes long-run performances under different scenarios and carries out several comparative statics regarding the effects of education or institutional quality on economic performance.

Along with the main findings, this paper provides some original methodological contributions: 1) By means of the concept of psychological equilibrium we split the population into two types of individuals endogenously—and not as a preliminary assumption of the model. 2) Additionally, we suggest a way to extend such concept of psychological equilibrium to a dynamic game that through a natural selection process operates over the types previously generated by it. 3) The mechanics of this whole approach leads to the desired outcome of explaining as endogenous variables macro-perceptions and the distribution of types in the population among others.

This paper is organized as follows. Section 2 contains the related literature. Section 3 shows the basic static model structure and preliminary results. The following section describes the dynamic model. Section 5 discusses the main results and their main implications and Section 6 concludes the paper. The appendix contains proofs of the results.

2. Related Literature

There is an extensive and recent literature on the relationship between social capital and economic activity (see for an overview, [1] [2]). Also, there is a large empirical literature related to the measurement of social capital or the estimation of its relationship to economic activity (see, among others, [3]-[5]). In this section we discuss the literature close to our model.

Our paper static framework relies heavily on the kindness function and the corresponding psychological equilibrium found in [6], although we focus on the emergence of types resulting from this setup with the aim of endogenize perceptions, issues that are not considered in [6]. Regarding the dynamic analysis, the closest studies to our model are [7] and [8], where a similar concept to our macro-perceptions is introduced, although not endogenously as we do. Also, we share with those models that social capital is measured through the proportion of agents that cooperate in a one shot Prisoner's Dilemma. With a similar goal in mind the model found in [9] explains the distribution of interdependent altruist and spiteful preferences in the population. However, our model does not consider interdependent preferences, hence comparison of results is of little interest.

The existence of moral values is one of the arguments found in the literature of cooperative behavior. This precisely is the motivation found in [10]. This paper studies the evolution of moral values jointly with pure economic incentives when two types of individuals interact (*homo kantiensis vs. homo oeconomicus*).

Other approaches to study social capital are the following. In [11] [12] we find the basis of the institutional approach to social capital. The argument is that formal or informal institutions—conditioning behavioral rules—may determine collective behavior and individual disposition toward social cooperation. A different theoretical framework is to explain social capital by adopting the theory of networks. Some examples are found in [13]-[15].

Finally, social capital may be understood as trustworthiness which is the case of our model. This perspective

relies on the role of individual beliefs concerning partner's behavior, and it is generally formalized in a game theory environment, through preference specifications concerning a *type* parameter.

In [16] the author highlights the natural fact of considering perceptions as a relevant subject when we studying social capital. It is also pointed out the need to explain perceptions as an additional endogenous variable of the model. Some attempts are found in the literature regarding the former, but not the latter. In particular, in [17] we can find an example of incorporating perceptions on countries—somehow as a measure of global reputation—to explain international capital flows. The data indicate the existence of a positive correlation between trust-worthiness and investments flows into country. However, in that model, economic variables do not affect perceptions of trust inspired by a country, *i.e.*, perceptions are taken as given in the model. Following the argument of [16], in our model perceptions are endogenous.

Regarding the dynamic study of social capital there are mainly two alternatives, which are the natural selection and cultural selection approaches. The first one is based on an evolutionary rule that operates on the population that cooperates-usually in an interactive environment which, with appropriate interpretations, most of the time it resembles the Prisoner's Dilemma game. As a result, this evolutionary mechanism rewards the proportion of the population that cooperates through reciprocity. An evolutionary model categorized as institutional approach is found for instance in [20] (see references therein, mainly [12] and [21]). A key point in this line of literature is the existence of a mechanism punishing those who do not reciprocate. In our paper, we introduce a natural selection approach to study the dynamics of social capital, although we do not introduce any punishment and obtain converge to full social capital under general weak conditions.

In general, the natural selection approach can be summarized as follows. First, the static strategic interaction framework is settled through a sort of Prisoner's Dilemma. Second, a fitness function to establish whether a type grows or not, is defined-either in terms of the true preferences or in contrast, using some related values. Finally, the analysis of stability is carried out, usually assuming the replicator dynamics. There are at least two interpretations of this methodology. First, standard in biology, in which individuals are thought as automata following the precepts of a gene or pool of genes governing their behavior. Second, it is considered that people learn— mainly by imitation—those behaviors that are most successful from among the observable ones. (see among others [22] and [23]). We choose the latter interpretation for our set-up.[1]

Cultural selection is the second dynamic approach to social capital, as described in [24]. Its main idea is to consider the cooperative attitude of individuals as a cultural trait whose probability of being transmitted (from family or society) to descendants is subject to a dynamic rule. The object of study is the dynamics of the probability of a given trait to be present. In [25] this approach is applied to explain the relationship between social capital and economic development.

[16] highlights that perceptions are relevant for a deep understanding of social capital and its relationship with economic activity. This author also points out the need to explain perceptions as an additional endogenous variable of the model. Some attempts are found in the literature. In particular, the model found in [17] considers perceptions on countries—somehow as a measure of their reputation—to explain international capital flows. The data indicate the existence of a positive correlation between trustworthiness and investments flows to a country. However, in that model, economic variables do not affect perceptions of trust inspired by a country. That is, perceptions are taken as given in that paper whereas in our model these are endogenous.

We can summarize that in the previous work of social capital literature more closely related to our research, in general types of individuals are given, perceptions in general are ignored or are not endogenous. Our research allows for a rich interpretation of the aspects that may affect perception and hence social capital as education, information or institutional quality for instance.

3. Theoretical Framework

In this section we introduce basic concepts and micro-foundations of the model in a static setup. Also, we show preliminary results, which provide ground to extract useful implications for the full specification of the model.

We first propose the Prisoner's Dilemma (PD) game to formalize how social capital is created, economic efficiency and its conflict with individual incentives-as usual in the related literature. The payoff matrix of the PD game is shown next in **Table 1**.

---

[1]Possible Darwinian causes for cooperation has been observed in biology (see [26]), among which it is worth to mention: 1) Genetic kinship, as for example, suicide bees. See [27]-[29]); 2) Reciprocation, as found in symbiosis (or in vampire bats, for example). See [30] and [31]; and 3) The benefit of conspicuous generosity, as a way to signal the true type, as the case of the Arabian babblers (see [32]-[35]).

Table 1. PD game payoff matrix.

	C	D
C	rX, rX	$0, dX$
D	$dX, 0$	X, X

where $d > r > 1$, C and D are for *Cooperate* and *Defect* actions respectively and $X > 0$ is the scale or size of the game. Introducing this variable will lead us to extract some implications of interest. Similar specifications of this game to study the subject of social capital are easily found in the literature—see [18] and [6] for instance.

The utility function that we propose is a convex combination of two elements: a material and a *psychological* payoff—a specification close to the utility functions found in [19] and [6]. The *psychological payoff* captures agents' preference for reciprocity, what is formalized through the interaction of individual *kindness* functions—to be introduced. In order to define these functions first we need to define some elements and sets.

We define player i's set of actions as $S_i = S = \{C, D\}$; we denote by $a_i \in S_i$ player i's action and by $b_j \in S_j$ player i's belief about player j's action—first order beliefs. That is, b_j is what player i believes about player j's action. Also, denote by $c_i \in S_i$ player i's belief about player j's beliefs about the action chosen by i—second order beliefs. That is, c_i is what player i thinks that player j thinks that he (i) is choosing.

Given the previous elements we define vectors of actions and beliefs. In particular, let $a = (a_1, a_2) \in S \times S$ be the vector of actions and $\beta_1 = (c_1, b_2)$ and $\beta_2 = (b_1, c_2)$, the vector of beliefs of player 1 and player 2 respectively—where $\beta_i \in S \times S$ for $i = 1, 2$. Next, we define the vector of micro-perceptions β as $\beta = (\beta_1, \beta_2) = ((c_1, b_2), (b_1, c_2)) \in (S \times S)^2$. Finally, from the previous actions and micro-perceptions vectors we define the Actions-Beliefs (AB) array as $e = (a, \beta) = (a_1, a_2, ((c_1, b_2), (b_1, c_2))) \in (S \times S)^3$.

From actions and beliefs we now define the *player 1's kindness to player 2* function. We first consider the set of all feasible payoffs to player 2 for a given first order belief b_2 depending on which action player 1 chooses, as $\Pi_2(b_2) = \{\pi_2(a, b_2) | a \in S\}$. Let $\pi_2^h(b_2) = \max \Pi_2(b_2)$ and $\pi_2^{\min}(b_2) = \min \Pi_2(b_2)$. Let the equitable payoff be $\pi_2^e(b_2) = \dfrac{\pi_2^h(b_2) + \pi_2^{\min}(b_2)}{2}$. This element provides a reference payoff level against which measuring how kind is player 1 to player 2 when choosing a_1. Taking into account these elements, given b_2, player 1's kindness to player 2 function is given by

$$f_1(a_1, b_2) = \frac{\pi_2(a_1, b_2) - \pi_2^e(b_2)}{\pi_2^h(b_2) - \pi_2^{\min}(b_2)}.$$

Following a similar procedure we construct the *player 1's expected player 2's kindness to player 1* function. First we consider a given player 1's second order belief about his own action, c_1. Thus, for a given c_1 we consider the set of all feasible payoffs to player 1 that are associated to a particular player 2 expected action b. Formally, this is the set $\Pi_1(c_1) = \{\pi_1(c_1, b) | b \in S\}$, in which we identify the highest, lowest and equitable values, denoted by $\pi_1^h(c_1)$, $\pi_1^{\min}(c_1)$ and $\pi_1^e(c_1)$ respectively. Taking into account these values, player 1's expected player 2 kindness to him is given by

$$\tilde{f}_2(c_1, b_2) = \frac{\pi_1(c_1, b_2) - \pi_1^e(c_1)}{\pi_1^h(c_1) - \pi_1^{\min}(c_1)}. \tag{1}$$

By operating in a similar manner we define the *player 2's kindness to player 1* function, denoted by $f_2(b_1, a_2)$, and the *player 2 expected player 1 kindness to player 2* function, denoted by $\tilde{f}_1(b_1, c_2)$.

These kindness functions are used next to define players' utility functions. We adopt the general framework of the utility function proposed by [19], although we propose some innovations as shown in the next definition. Agents interaction takes place when each player chooses an action to maximize his utility, which involves taking into account both a material and a psychological payoff. This last element is formalized by each player kindness to the other player when choosing a_i, given the kindness he expected from the other player. In short, the psychological payoff contains kindness functions interaction showing players preference for reciprocity—that is, choosing to be friendly (hostile) when he expects a friendly (hostile) action from the other player.

Definition 1. *Players* 1 *and* 2 *utility functions are given by the following equations respectively*:

$$U_1(a_1, a_2, \beta) = \alpha \pi_1(a_1, a_2) + (1-\alpha)G(X)\left[\tilde{f}_2(c_1, b_2)\right]\left[1 + f_1(a_1, b_2)\right] \qquad (2)$$

and

$$U_2(a_1, a_2, \beta) = \alpha \pi_2(a_1, a_2) + (1-\alpha)G(X)\left[\tilde{f}_1(b_1, c_2)\right]\left[1 + f_2(b_1, a_2)\right] \qquad (3)$$

where $\alpha \in (0,1)$ serves as a weighting parameter for the material payoff, $G(X)$—the psychological size effect, such that $\dfrac{G(X)}{X} \equiv \tilde{G}(X)$ is a decreasing function and $\lim_{X \to \infty} \tilde{G}(X) = 0$. The motivation for $G(X)$ is to balance the role of X in both components of the utility function. Although assumptions on $G(X)$ are mainly technical, the way we introduce X in the utility function is consistent with related literature suggesting that the size of the game increases the probability of choosing D in the PD or similar games—see [18] or [6].

Next we define the equilibrium concept linked to the utility function just considered.

The Action-Beliefs vector (a_1^*, a_2^*, β^*), where $\beta^* = ((c_1^*, b_2^*), (b_1^*, c_2^*))$ is a Psychological Equilibrium (PE) if

1) $U_1(a_1^*, a_2^*, \beta^*) \geq U_1(a_1, a_2^*, \beta^*)$ for all $a_1 \in S$.
2) $U_2(a_1^*, a_2^*, \beta^*) \geq U_2(a_1^*, a_2, \beta^*)$ for all $a_2 \in S$.
3) $a_1^* = b_1^* = c_1^*$ and $a_2^* = b_2^* = c_2^*$.

Conditions 1) and 2) establish that the pair of actions (a_1^*, a_2^*) is a Nash equilibrium given the exogenous beliefs vector β^*, while condition 3) implies that beliefs should be consistent with actions in the PE.

It is convenient to remark that since beliefs are exogenous within the definition of PE, it might be tempting to think that the concept of PE barely contribute to explain the relationship between beliefs, actions and economic performance. Nevertheless, this is a first step in the development of the model, and once it is fully specified, beliefs will become endogenous. Hence, within the dynamic aggregate setup, we will be able to describe how beliefs affect actions and economic performance in the long run and *vice versa*.

Preliminary Results

We are now in conditions to identify the PE actions-beliefs vectors in a psychological PD game as well as to show related implications. Since condition 3) is to be hold by a PE, only some AB vectors are to be considered. Let the following AB vectors $e_1 = (C, C, ((C,C), (C,C)))$, $e_2 = (D, D, ((D,D), (D,D)))$, $e_3 = (C, D, ((C,D), (C,D)))$ and $e_4 = (D, C, ((D,C), (D,C)))$ be the set of feasible PE candidates, as only for them, $a_i = b_j = c_i$ hold.

Before formally showing the existing PE, it might be convenient to show in an example the total payoffs matrix of the psychological game for a given micro-perceptions vector. This exercise may illustrate the concept and identification of the existing PE. Taking micro-perceptions as $\beta_c \equiv ((C,C), (C,C))$, since $b_i = c_i = C$ for all $i = 1, 2$, it is possible to compute player 1 utility level for all possible action pairs (a_1, a_2) from player 1's utility function. The resulting total payoffs matrix of a psychological PD given $\beta = \beta_c$ is shown in **Table 2**. Only player 1's payoffs are shown since the game is symmetric and we assume so far that parameters are the same for both players.

A psychological equilibrium is a Nash equilibrium of the psychological game generated by a given micro-perception vector. Hence, for our example, to identify a PE we just need to prove that a pair of actions is a Nash equilibrium using **Table 2** information. It is of course of interest to find out whether the pair of actions (C,C) could be a Nash equilibrium under certain conditions. Since variations of material and psychological payoffs have opposite sign when player 1 deviates from C, the pair of actions (C,C) would be in fact a Nash

Table 2. Psychological PD total payoffs for $\beta = \beta_c$.

	C	D
C	$\alpha r X + (1-\alpha)\dfrac{3}{4}G(X)$	$(1-\alpha)\dfrac{3}{4}G(X)$
D	$\alpha d X + (1-\alpha)\dfrac{1}{4}G(X)$	$\alpha X + (1-\alpha)\dfrac{1}{4}G(X)$

equilibrium depending on α and X.

We show next a full formalization of the existing PE generated by the psychological game, for the set of all possible beliefs which are consistent with the AB vectors of interest. For a clear understanding of the next result, we consider the possibility of different player specific values for the parameter α, where α_i denotes player i's parameter value, for $i = 1, 2$.

Proposition 1. *Consider the utility functions given in Definition* 1. *Then*:
1) Neither e_3 nor e_4 can be PE;
2) For any $\alpha_i \in [0,1]$ with $i = 1, 2$, e_2 is a PE;
3) Taking

$$\tilde{\alpha}(X) = \frac{\tilde{G}(X)}{2(d-r) + \tilde{G}(X)} \tag{4}$$

then: 1) If $\alpha_i \geq \tilde{\alpha}(X)$ with $i = 1$ or $i = 2$, e_2 is the unique PE; 2) If $\alpha_i < \tilde{\alpha}$ with $i = 1, 2$, e_1 (and e_2) is a PE.

4) $\tilde{\alpha}(X)$ is decreasing in X and $\lim_{X \to \infty} \tilde{\alpha}(X) = 0$.

The proof can be found in the Appendix.

Some interpretations and implications of interest come out from the proposition above. Items 1), 2) and 3) imply that there are only two PE: e_1 and e_2. Thus, adding a psychological term to the standard utility function allows to achieve the cooperative solution in a PD game under certain conditions. On the other hand it is worth saying that by construction, one or another PE is to be achieved depending on the micro-perception vector considered by each player, as well as on the value of the parameter α. In particular, e_1 is achieved if and only if $\beta_1^* = \beta_2^* = (C, C)$ and $\alpha_i < \tilde{\alpha}(X)$ for all $i = 1, 2$. If $\alpha_i > \tilde{\alpha}(X)$ for $i = 1$ or $i = 2$, then e_2 is the unique existing PE, and obviously it is achieved if $\beta_1^* = \beta_2^* = (D, D)$.

Although a relative low weight to the material payoff is needed in order to achieve the cooperative PE, item 4) shows that such weight has to decrease if the size of the game, X, increases. In other words, the higher the economic transaction in hand is, the lesser has to be α to achieve the cooperative PE.

So far this static framework relies closely on the [6] model so the previous result can be easily derived from it. However, it was needed to state it properly for the sake of the forthcoming development of the model. In particular, we draw one important implication from Proposition 1 which is the following. Given X, the critical value $\tilde{\alpha}(X)$ leads us to consider two types of agents within the population. On the one hand, those whose parameter is $\alpha_C < \tilde{\alpha}(X)$, and on the other hand, those characterized by the parameter value $\alpha_D \geq \tilde{\alpha}(X)$. These are called the *pro-social* type and the *basic* type respectively.

4. Aggregation and Dynamic Setup

In this section we describe the aggregated economy and its dynamic behavior. Aggregation is modeled using a structure close to that of the overlapping generation models, in an infinite periods horizon where time is a discrete variable. Thus, in each period a new generation is born, each individual living two periods. As previously mentioned, there exist two types of individuals, called *pro-social* and *basic* types, denoted as τ_C and τ_D respectively.

We propose the replicator dynamics approach for the determination and evolution of types.[2] Roughly, this general mechanism operates as follows: during the first period of life each individual acquires his type (either in a biological sense or in a learning interpretation of the process) for his second period of life. The acquired type by this evolutionary process is the one that belongs to the group with a higher level of expected utility than the average expected payoff of the population. Because of that, the group of individuals of this type increases within the population at expenses of the other groups.

For simplicity, we will assume that only PE, that is, e_1 or e_2, are achieved through strategic interaction and that there exist perfect information on types.[3] Once the type is acquired, players are randomly matched and interact in the framework of the psychological game described in the previous section, where macro-perceptions

---

[2]Although the proofs are done using the *sign preserving dynamics*, a more general form of dynamics.

[3]It would be relevant to consider also e_3 and e_4, either by allowing basic type players to cheat a pro-social one or simply as a result of an error in micro-perceptions. Nevertheless, the model becomes more complex and does not provide new insights on what we already have.

will play a crucial role to determine simultaneously micro-perceptions and actions. The process thus described continues period by period through generations, until the proportion of individuals of each type reaches a steady state under the conditions that we consider.

4.1. Dynamic Analysis

First, we introduce new notation in order to define the fitness function to carry out the dynamic analysis-recall that only PE are achieved through the game. Let $U(e_1, \alpha_C)$ represent the utility level of a type τ_C player, conditioned to beliefs $\beta_i = (C, C)$, for $i = 1, 2$. Similarly, $U(e_2, \alpha_C)$ is the utility of a player of the same type conditioned to beliefs $\beta_i = (D, D)$, for $i = 1, 2$. Finally, $U(e_2, \alpha_D)$ denotes the utility level of a type τ_D player.

Let $\rho(t) \in [0, 1]$ be the proportion of type τ_C players in the population at period t. Let $\tilde{U}(\tau_C, t)$ and $\tilde{U}(\tau_D, t)$ be the expected utility for types τ_C and τ_D players respectively at period t. Finally, the average utility of the game is $\bar{U}(t) = \rho(t)\tilde{U}(\tau_C, t) + (1 - \rho(t))\tilde{U}(\tau_D, t)$.

The process governing the evolution of ρ is the Replicator Dynamics, given by the following equation:

$$\rho(t+1) - \rho(t) = \rho(t)\frac{\tilde{U}(\tau_C, t) - \bar{U}(t)}{\bar{U}(t)}. \tag{5}$$

The Replicator Dynamics generates the same steady states and stability as any other sign preserving dynamics (see, for instance, [36]), which we define as follows. A dynamic process $\rho(t+1) = F(\rho(t), \tilde{U}(\tau_C, t), \bar{U}(t))$ — where F is a continuous function in all its arguments—is a sign preserving dynamic if $sign[\rho(t+1) - \rho(t)] = sign[\tilde{U}(\tau_C, t) - \bar{U}(t)]$.[4] Since there are only two types of individuals within the population, when $\rho(t) \neq 1$, the previous condition is equivalent to

$$sign[\rho(t+1) - \rho(t)] = sign[\tilde{U}(\tau_C, t) - \tilde{U}(\tau_D, t)]. \tag{6}$$

For simplicity we assume that $\rho(t+1)$ follows a sign preserving dynamics hereinafter.

The purpose of our dynamic analysis is to identify the steady states for ρ and study their stability. A steady state is a sequence $(\rho(t))_{t=0}^{\infty}$ such that $\rho(t+1) - \rho(t) = 0$ for all $t \geq 0$. Considering equation (6) we only need to know the sign of $\tilde{U}(\tau_C) - \tilde{U}(\tau_D)$ in order to conduct the dynamic analysis. Thus, if $\tilde{U}(\tau_C) - \tilde{U}(\tau_D) > 0$ (< 0), ρ increases (decreases) over time. Finally, if $\tilde{U}(\tau_C) - \tilde{U}(\tau_D) = 0$ then ρ remains constant-which is the steady state condition.

Given the previous argument we need to study the condition $\tilde{U}(\tau_C) - \tilde{U}(\tau_D) = 0$ for identifying the steady states values. The computation of $\tilde{U}(\tau_D)$ is immediate, as type τ_D players only choose D, hence only the PE e_2 will arise regardless of the type of his opponent. Formally,

$$\tilde{U}(\tau_D) = \rho U(e_2, \alpha_D) + (1 - \rho)U(e_2, \alpha_D) = U(e_2, \alpha_D). \tag{7}$$

The computation of $\tilde{U}(\tau_C)$ is not so trivial, as type τ_C players can achieve any of the existing PE. A player of this type will meet a type τ_D player with probability $1 - \rho$ —hence the only achievable equilibrium is e_1 —and obtain a utility level $U(e_2, \alpha_C)$. On the other hand, a type τ_C player would meet a player of his same type with probability ρ. In this case he faces a dilemma generated by the existence of two feasible beliefs vectors—which conditioned his behavior and thus the achievement of e_1 or e_2. Thus, it then comes up as a natural question which beliefs vector players are going to choose.

4.2. Macro-Perceptions

Given the need to provide a satisfactory answer to the coordination problem found in a type τ_C players matching, we introduce the concept of *macro-perceptions*. The intuitive argument behind this concept is quite simple and a standard one in the economic literature. In short, we solve the mentioned problem by assuming that type τ_C players share a common perception about the probability of coordinating their actions. Next we

---

[4]It is easy to check that this equation implies that $[\rho(t+1) - \rho(t)] = 0$ if and only if $[\tilde{U}(\tau_C, t) - \bar{U}(t)] = 0$, except when $\rho(t) = 0$. If $\rho(t) = 0$, we impose $\rho(t+1) = 0$. That is, a type cannot emerge if it did not exist previously.

provide a formal definition of this concept.

Definition 3. *Given a population in which the proportion of pro-social people is ρ, we define macro-perceptions through the continuous function $\lambda(\rho):[0,1] \rightarrow [0,1]$, which for an entrant pro-social player measures the conditional probability of achieving e_1, given that he is matched to another pro-social player. The function $\lambda(\rho)$ satisfies the following assumptions: $\lambda(1)=1$ and $\rho\lambda(\rho)$ is increasing in ρ.*[5]

Some comments are in order. First, note by definition $\rho\lambda(\rho)$ measures the probability of a new entrant pro-social player of achieving e_1. That is, in a society where ρ is the proportion of pro-social people, micro-perceptions leading to e_1 occur with probability $\rho\lambda(\rho)$. Second, it would have been sufficient to just require continuity on $\lambda(\rho)$ and nothing else. Nevertheless we have imposed two more assumptions—$\rho\lambda(\rho)$ increasing and $\lambda(1)=1$—for simplicity.

From the definition of macro-perceptions we can now compute the expected utility of a type τ_C player, $\tilde{U}(\tau_C)$, which is,

$$\tilde{U}(\tau_C) = \rho\left[\lambda(\rho)U(e_1,\alpha_C) + (1-\lambda(\rho))U(e_2,\alpha_C)\right] + (1-\rho)U(e_2,\alpha_C).$$

By rearranging terms properly in the previous equation, we have,

$$\tilde{U}(\tau_C) = \gamma U(e_1,\alpha_C) + (1-\gamma)U(e_2,\alpha_C) \tag{8}$$

where $\gamma = \rho\lambda(\rho)$. Equation (8) is consistent with the definition of λ, since for a type τ_C player, micro-perceptions leading to *C* occur (and PE e_1 is achieved) with probability γ, and also those leading to *D* occur (and PE e_2 is achieved) with probability $(1-\gamma)$. Note that $1-\gamma > 1-\rho$; that is, for a type τ_C player, the probability of achieving the PE e_2 is larger than the probability of matching another τ_D player since not only this type (τ_D) leads to such PE.

Taking into account Equations (7) and (8), we are now in condition to compute the difference $\tilde{U}(\tau_C) - \tilde{U}(\tau_D)$. Thus, by substituting the considered equations, we have

$$\tilde{U}(\tau_C) - \tilde{U}(\tau_D) = \gamma U(e_1,\alpha_C) + (1-\gamma)U(e_2,\alpha_C) - U(e_2,\alpha_D). \tag{9}$$

The steady states for ρ are found by setting equation (9) equal to zero. In the next section we conduct the full dynamic analysis. It is important to remind for further interpretations the meaning of γ in equation (8), as this value is related to the probability of convergence to the full social capital situation.

5. Results

In this section we present the main results of the model: identify the steady states and discuss convergence to the optimal steady state under different set of assumptions. Overall, our main achievement is that we explain endogenous and simultaneously social capital, economic performance, micro and macro-perceptions.

Proposition 2. *Consider the dynamics given in* (6). *Then*:

(a) If $U(e_1,\alpha_C) > U(e_2,\alpha_D)$ then there exist only three steady states: $\rho_0^* = 0$, $\rho_1^* = 1$ and $\rho_I^* \in (0,1)$, such that $\rho_I^* = \dfrac{A^*}{\lambda(\rho_I^*)}$, where $A^* = \dfrac{U(e_2,\alpha_D) - U(e_2,\alpha_C)}{U(e_1,\alpha_C) - U(e_2,\alpha_C)}$. Also, $\rho_0^* = 0$ and $\rho_1^* = 1$ are asymptotically stable and ρ_I^* is unstable.

(b) If $U(e_1,\alpha_C) \leq U(e_2,\alpha_D)$, then only two steady states exist: $\rho_0^* = 0$ and $\rho_1^* = 1$, and for all $\rho(0) \neq 1$, $\rho(t)$ converges to ρ_0^*.

The proof can be found in the Appendix.

Item (a) shows the most favorable conditions for convergence to full social capital. The interpretation of the assumption stated in (a) is simple. First, if $U(e_1,\alpha_C) > U(e_2,\alpha_D)$—say the *cooperative condition* henceforth—does not hold, it would be difficult to observe convergence to $\rho_1^* = 1$, since each one of the terms in that inequality is associated to the expected utility of the corresponding type.

We require $\rho\lambda(\rho)$ to be an increasing function on ρ. See that in this case there exists a unique root for equation $\rho\lambda(\rho) = A^*$ in $(0,1)$—which is a less restrictive condition. Nevertheless, we impose increasing $\rho\lambda(\rho)$ for simplicity. In any case, assuming that only one root exist in $(0,1)$ is an extension of the focal

---

[5]There is no need to impose any condition on $\lambda(0)$, except being a well defined number.

point argument used in coordination games (See [37]). Indeed, the cooperative equilibrium works as a focal point, thus the more pro-social people there are in the society, the more attractive the cooperative equilibrium is, and the more likely it is to coordinate in that equilibrium.

If the cooperative condition do not hold, convergence to full social capital fails, as shown in (b). In intuitive terms, full social capital is not reached in the long run if the probability of cooperation between pro-social players is too low or if the return of cooperation is not high enough.

The following graph ($\rho(t+1) - \rho(t)$ vs. $\rho(t)$, generated by the Replicator Dynamics) shows the dynamics of variable ρ for the cases (a-i) and (b) of Proposition 2. The case (a-i) corresponds to the existence of an interior steady state ρ_I^*. If for the initial condition $\rho(0) > \rho_I^*$, the economy converges to $\rho_1^* = 1$ in the long run. In case (b), there is no interior steady state, thus the economy converges to $\rho_0^* = 0$ (see **Figure 1**).

To conclude, recall that $\gamma = \rho\lambda(\rho)$ is the probability of occurrence C, so the item (a-i) of Proposition 2 allows us to interpret A^* as the lower bound of the probability of achieving C so that the economy converges to $\rho_1^* = 1$ in the long run. The intuition of the above argument is that people turn out to be individually cooperative (*i.e.* $\rho\lambda(\rho)$ is high enough) when they perceive themselves in such a way that as a result of their behavior, the proportion of pro-social players increases. In other words, an individual will cooperate when by choosing "cooperate" it is likely to achieve cooperation, or in other words, when he perceives that himself belongs to a sufficiently large group. So, because of this perception of himself, and thus that cooperation was achieved, the proportion of pro-social players increases and the economy converges to $\rho_1^* = 1$. This argument leads us to interpret informally $\rho\lambda(\rho)$ as a measure of the degree to which people perceive themselves as *addable* or member of a group in which their contribution is effective.

Before presenting some implications, it is worth to establish the precise conditions under which the *co-operative condition* (CC) holds.

Proposition 3. *Consider the two types of agents given by the previous proposition, characterized by* α_C *and*

α_D. *Then* $U(e_1, \alpha_C) > U(e_2, \alpha_D)$—*the CC*—*is satisfied iff* $\dfrac{\alpha_D}{\alpha_C} < \dfrac{r + \tilde{G}(X)\left(\dfrac{1}{\alpha_C} - \dfrac{3}{4}\right)}{1 + \dfrac{\tilde{G}(X)}{4}} \equiv n(X, \alpha_C).$[6]

Proof follows from algebraic manipulations and hence omitted.

The CC says that pro-social playing e_1 are better off than basic type people if and only if basic people are not too materialist—in relation to pro-social people—, that is, if and only if $\alpha_D < \alpha_C n(X, \alpha_C)$.
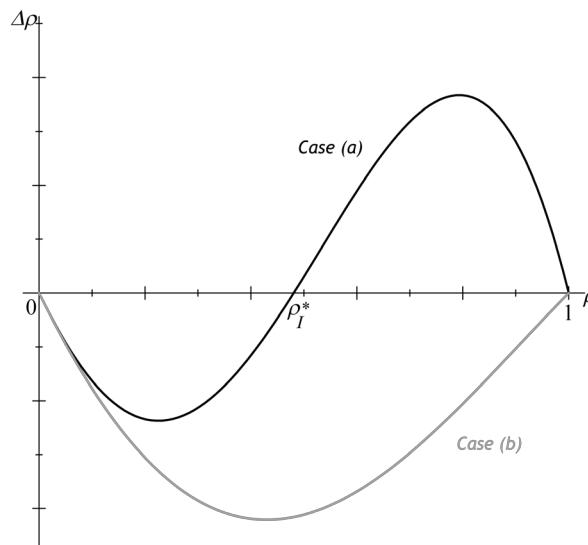

Figure 1. Replicator dynamics for the rate of pro-social players.

[6]The inequality $\dfrac{\alpha_D}{\alpha_C} < n(X, \alpha_C)$ only is possible if $\alpha_C \neq 0$. If $\alpha_C = 0$, the Cooperative Condition is not satisfied.

We can easily prove that $\dfrac{\mathrm{d}n(X,\alpha_C)}{\mathrm{d}X}>0$ for $\dfrac{4}{r+3}<\alpha_C<\tilde{\alpha}(X)$, and that $n(X,\alpha_C)$ is always larger than one. Therefore, the CC and the requirements $\alpha_C<\tilde{\alpha}(X)<\alpha_D$ are not inconsistent if, for instance $\dfrac{1}{n(X,\alpha_C)}<\tilde{\alpha}(X)$, which holds if and only if $\tilde{G}(X)$ is large enough. Observe, furthermore, that if $\dfrac{1}{n(X,\alpha_C)}<\alpha_C<\tilde{\alpha}(X)$, then α_D can be arbitrarily close to one, without violating any condition. In particular, if α_C is sufficiently small but *not too small,* the larger is X, the larger is $n(X,\alpha_C)$, and therefore the larger can be α_D, without violating the CC.

Comparative Statics

Previous results allow to discuss the effects of exogenous changes in α_D, X or λ on the probability of convergence to steady state $\rho_1^*=1$ (value that may be estimated by $1-\rho_I^*$). These implications could be useful to explain some empirical regularities concerning corruption or institutional quality and its relationship to economic performance. The proofs are given in the appendix.

First we highlight some previous conclusions drawn from Proposition 2 when the CC holds. In this case, we show that there exist convergence to $\rho_1^*=1$ if and only if the initial condition $\rho(0)$ is such that $\rho(0)\lambda(\rho(0))>A^*$. Thus, if $\rho(0)\lambda(\rho(0))<A^*$, there would be no convergence to $\rho_1^*=1$ but to $\rho_0^*=0$ and because of that, the larger is A^* the lower the probability of achieving convergence to $\rho_1^*=1$. Finally, note that under the conditions stated, there exist a positive relationship between ρ_I^* and A^*. Thus, considering two interior steady states $\rho_I^*(1)$ and $\rho_I^*(2)$ such that $\rho_I^*(1)\lambda(\rho_I^*(1))=A^*(1)$ and $\rho_I^*(2)\lambda(\rho_I^*(2))=A^*(2)$, if $A^*(2)>A^*(1)$, then $\rho_I^*(2)>\rho_I^*(1)$ since $\rho\lambda(\rho)$ is increasing.

Corollary 1 (α_D increases). *If $\alpha_D<\alpha_C n(X,\alpha_C)$, $\alpha_C<\tilde{\alpha}(X)$, $\lambda(1)>A^*$ and $\rho\lambda(\rho)$ is increasing, then there exist a negative relationship between α_D and the probability of convergence to $\rho_1^*=1$.*

To highlight the idea of this corollary, it may be useful to assume that $\dfrac{1}{n(X,\alpha_C)}<\alpha_C<\tilde{\alpha}(X)$, as in this case α_D can be arbitrarily close to one. Indeed, in this last case, the graphical representation of the result in the space ($\rho(t+1)-\rho(t)$ vs. $\rho(t)$) would show that the inner root ρ_I^* shifts to the right as α_D increases, up to some limiting value smaller than one. Since the probability of convergence to the optimal solution is $1-\rho_I^*$, we conclude that there is a negative relationship between such probability and this and α_D.

In intuitive terms, this corollary may explain the paradox of individuals weighting to high the material payoff, and by doing so penalize individually and collectively the options to achieve the best material results.

Corollary 2 (X increases). *If CC hold, $\lambda(1)>A^*$, $\rho\lambda(\rho)$ is increasing and α_C is large enough* ($\alpha_C>\dfrac{4}{3+r}$), *then, if X increases, as ρ_I^* decreases, the probability of convergence to the steady state $\rho_1^*=1$ increases.*

We show in Section 3, that as X increases, the interaction leads eventually to the non-cooperative equilibrium, which in dynamic terms means that there is convergence to the steady state $\rho_0^*=0$. However, it is possible to find convergence to $\rho_1^*=1$ if parameter α_C is not too small or r is large enough. The intuition is that provided r is large enough, then, the negative effect of large X not only disappear, but also become positive. That is, if the reward for achieving cooperation is high enough, then it overtakes the negative effect of a large X, so that the larger is X, the larger the probability of convergence to full social capital state.

Corollary 3 (λ shifts). *If CC hold, $\lambda(1)>A^*$ and $\rho\lambda(\rho)$ increasing, then societies with a higher level of optimism about anonymous cooperation, converge with higher probability to the optimum steady state $\rho_1^*=1$. Similarly, societies with lower optimism converge with higher probability to the inefficient steady state $\rho_0^*=0$.*

We can consider that the level of optimism about the anonymous cooperation could be measured by macro-perceptions function, $\lambda(\rho)$. Thus, an increase in the level of optimism would shift the function λ upwards for a given ρ, which imply that ρ_I^* decreases and therefore that the probability of convergence to $\rho_1^*=1$ increases.

Some variables that could affect that level of optimism about anonymous cooperation may in general be related to institutional quality, as corruption, democratic participation, transparency in the financial sector and the legal system guarantees, or mass media political power, for instance. These variables may be formalized through macro-perceptions in our model. So, if there is a shock related to those variables, it would be understood as a shift in the level of optimism about the anonymous cooperation, which in our model is formalized through λ, hence affecting the probability of convergence to $\rho_1^* = 1$.

6. Conclussions

In this paper we have proven that provided that the CC is satisfied two variables are crucial in order to obtain full social capital in the long run, the macro-perceptions and the proportion of pro-social people in the society. That is, if two societies only differ in macro-perceptions, the will not reach the same distribution of types and economic outcome in the long run. In fact, that with optimistic macro-perceptions will achieve better economic performance. These results beg the natural question of how to change macro-perceptions.

This is a complex issue indeed although some considerations are in order. Formal and informal education is one of the main sources of perception on trust. Specially informal education is relevant as people tend to imitate or learn successful behavior, which in our model is captured by the fact that the dynamic process is understood as a learning process. People tend to imitate, then to learn, successful behavior, so it is important to observe which in particular is the successful behavior. On the other hand, we cast doubt on how strongly formal education may help to increase social capital. Moreover, we conjecture that the effect of formal education over social capital is much lower than the effect of the imitation process, an issue that is left for future research. Finally, it may be relevant for institutions to establish effective rules to control the influence of lobbies on mass media and also on political actors.

References

[1]   Fukuyama, F. (1995) Trust: The Social Virtues and the Creation of Prosperity. The Free Press, New York.

[2]   Durlauf, S. and Fafchamps, M. (2005) Social Capital. In: Aghion, P. and Durlauf, S., Eds., *Hand-Book of Economic Growth*, Vol. 1, Ch. 26, Elsevier, Amsterdam, 1639-1699. http://dx.doi.org/10.1016/s1574-0684(05)01026-9

[3]   Knack, S. and Keefer, P. (1997) Does Social Capital Have an Economic Payoff? A Cross-Country Investigation. *The Quarterly Journal of Economics*, **112**, 1251-1288. http://dx.doi.org/10.1162/003355300555475

[4]   La Porta, R., de-Silanes, F.L., Shleifer, A. and Vishny, R.W. (1998) Law and Finance. *Journal of Political Economy*, **106**, 1113-1155. http://dx.doi.org/10.1086/250042

[5]   Temple, J. and Johnson, P. (1998) Social Capability and Economic Growth. *Quarterly Journal of Economics*, **113**, 965-990. http://dx.doi.org/10.1162/003355398555711

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

[7]   Poulsen, A. and Poulsen, O. (2010) Prisoner's Dilemma Payoffs and the Evolution of Cooperative Preferences. *The Journal of Socio-Economics*, **39**, 158-162. http://dx.doi.org/10.1016/j.socec.2009.10.002

[8]   Poulsen, A. and Svendsen, G. (2005) Social Capital and Endogenous Preferences. *Public Choice*, **123**, 171-196. http://dx.doi.org/10.1007/s11127-005-0266-0

[9]   Sethi, R. and Somanathan, E. (1996) The Evolution of Social Norms in Common Property Resource Use. *American Economic Review*, **86**, 766-788.

[10] Alger, I. and Weibull, J. (2013) Homo Moralis-Preference Evolution under Incomplete Information and Assortative Matching. *Econometrica*, **6**, 2229-2302.

[11] North, D.C. (1990) Institutions, Institutional Change and Economic Performance. Political Economy of Institutions and Decisions. Cambridge University Press, Cambridge. http://dx.doi.org/10.1017/CBO9780511808678

[12] Güth, W. and Ockenfels, A. (2005) The Coevolution of Morality and Legal Institutions: An Indirect Evolutionary Approach. *Journal of Institutional Economics*, **1**, 155-174. http://dx.doi.org/10.1017/S1744137405000159

[13] Kandori, M. (1992) Social Norms and Community Enforcement. *The Review of Economic Studies*, **59**, 63-80. http://dx.doi.org/10.2307/2297925

[14] Putnam, R.D., Leonardi, R. and Nanetti, R. (1993) Making Democracy Work: Civic Traditions in Modern Italy. Princeton University Press, Princeton.

[15] Annen, K. (2003) Social Capital, Inclusive Network and Economic Performance. *Journal of Economic Behavior and Organization*, **50**, 449-463. http://dx.doi.org/10.1016/S0167-2681(02)00035-5

[16] Dasgupta, P. (2003) Social Capital and Economic Performance: Analytics. In: Ostrom, E. and Ahn, T., Eds., *Foundations of Social Capital*, Edward Elgar, Cheltenham, 309-339.

[17] Guiso, L., Sapienza, P. and Zingales, L. (2009) Cultural Biases in Economic Exchange? *Quarterly Journal of Economics*, **124**, 1095-1131. http://dx.doi.org/10.1162/qjec.2009.124.3.1095

[18] Dawes, R.M. and Thaler, R.H. (1988) Anomalies: Cooperation. *Journal of Economic Perspectives*, **2**, 187-198. http://dx.doi.org/10.1257/jep.2.3.187

[19] Geanakoplos, J., Pearce, D. and Stacchetti, E. (1989) Psychological Games and Sequential Rationality. *Games and Economic Behavior*, **1**, 60-79. http://dx.doi.org/10.1016/0899-8256(89)90005-5

[20] Jackson, K. (2013) Contract Enforceability and the Evolution of Social Capital. *The Journal of Law*, *Economics*, *& Organization*, **29**, 60-77. http://dx.doi.org/10.1093/jleo/ewr016

[21] Tabellini, G. (2008) The Scope of Cooperation: Values and Incentives. *Quarterly Journal of Economics*, **123**, 905-950. http://dx.doi.org/10.1162/qjec.2008.123.3.905

[22] Izquierdo, L.R., Izquierdo, S.S. and Vega-Redondo, F. (2012) Learning and Evolutionary Game Theory. In: Seel, N., Ed., *Encyclopedia of the Sciences of Learning*, Springer, New York, 1782-1788.

[23] Vega-Redondo, F. (1996) Evolution, Games, and Economic Behaviour. Oxford University Press, Oxford. http://dx.doi.org/10.1093/0198774729.001.0001

[24] Cavalli-Sforza, L.L. and Feldman, M.W. (1981) Cultural Transmission and Evolution: A Quantitative Approach. Vol. 16 of Monographs in Population Biology. Princeton University Press, Princeton.

[25] Francois, P. and Zavojnik, J. (2005) Trust, Social Capital and Economic Development. *Journal of the European Economics Association*, **3**, 51-94. http://dx.doi.org/10.1162/1542476053295304

[26] Dawkins, R. (2006) The God Delusion. Houghton Mifflin Co., Boston.

[27] Hamilton, W.D. (1964) The Genetical Evolution of Social Behaviour I. *Journal of Theoretical Biology*, **7**, 1-16. http://dx.doi.org/10.1016/0022-5193(64)90038-4

[28] Hamilton, W.D. (1967) Extraordinary Sex Ratios. *Science*, **156**, 477-488. http://dx.doi.org/10.1126/science.156.3774.477

[29] Hamilton, W.D. and Zuk, M. (1982) Heritable True Fitness and Bright Birds: A Role for Parasites? *Science*, **218**, 384-387. http://dx.doi.org/10.1126/science.7123238

[30] Trivers, R. (1971) The Evolution of Reciprocal Altruism. *Quarterly Review of Biology*, **46**, 35-57. http://dx.doi.org/10.1086/406755

[31] Axelrod, R. and Hamilton, W.D. (1981) The Evolution of Cooperation. *Science*, **211**, 1390-1396. http://dx.doi.org/10.1126/science.7466396

[32] Grafen, A. (1990) Biological Signals as Handicaps. *Journal of Theoretical Biology*, **144**, 517-546. http://dx.doi.org/10.1016/S0022-5193(05)80088-8

[33] Zahavi, A. (1975) Mate Selection—A Selection for a Handicap. *Journal of Theoretical Biology*, **53**, 205-214. http://dx.doi.org/10.1016/0022-5193(75)90111-3

[34] Zahavi, A. (1977) The Cost of Honesty. *Journal of Theoretical Biology*, **67**, 603-605. http://dx.doi.org/10.1016/0022-5193(77)90061-3

[35] Zahavi, A. (1997) The Handicap Principle: A Missing Piece of Darwin's Puzzle. Oxford University Press, Oxford.

[36] Ritzberger, K. and Weibull, J.W. (1995) Evolutionary Selection in Normal-Form Games. *Econometrica*, **63**, 1371-1399. http://dx.doi.org/10.2307/2171774

[37] Schelling, T.C. (1960) The Strategy of Conflict. Harvard University Press, Cambridge, MA.

Appendix

A.1. Proof of Proposition 1

First we prove (a). Consider $e_3 = \left(C, D, \left((C,D), (C,D) \right) \right)$ and the player 1. We will show that a deviation from playing C is advantageous to player 1. Imagine player 2 chooses D, and player 1 chooses C. Recalling that $U_1(a_1, a_2, \beta) = \alpha \pi_1(a_1, a_2) + (1-\alpha) G(X) \left[\tilde{f}_2(c_1, b_2) \right] \left[1 + f_1(a_1, b_2) \right]$, and noting that

$\tilde{f}_2(c_1, b_2) = \tilde{f}_2(C,D) < 0$, $f_1(C,D) > 0$ and $f_1(D,D) < 0$ we have that both terms of $U_1(a_1, a_2, \beta)$, the psychological term $\left((1-\alpha) G(X) \left[\tilde{f}_2(c_1, b_2) \right] \left[1 + f_1(a_1, b_2) \right] \right)$ and the material term $\left(\alpha \pi_1(a_1, a_2) \right)$, increase if it chooses D. Indeed, $\pi_1(C,D) < \pi_1(D,D)$, and

$(1-\alpha) G(X) \left[\tilde{f}_2(C,D) \right] \left[1 + f_1(C,D) \right] < (1-\alpha) G(X) \left[\tilde{f}_2(C,D) \right] \left[1 + f_1(D,D) \right]$, due to $\tilde{f}_2(C,D) < 0$,

$f_1(C,D) > 0$ and $f_1(D,D) < 0$. An identical reasoning can be made for e_4 and player 2. Therefore the item (a) is done.

The proof of (b) is analogous to (a) and hence omitted (both parts of the payoff decrease by playing C against D).

In order to prove (c) we consider player 1, rename for practical purposes $U_1(e_1)$ as $U_1(e_1, \alpha)$ (recall that $e_1 = \left(C, C, \left((C,C), (C,C) \right) \right)$), define $U_1(D, \alpha) = U_1 \left((D, C, (C,C,C,C)), \alpha \right)$, and finally define the function $h(\alpha)$ as

$$h(\alpha) = U_1(D, \alpha) - U_1(e_1, \alpha).$$

Thus, we have that e_1 is not a PE if and only if $h(\alpha) > 0$, and that e_1 is a PE if and only if $h(\alpha) \leq 0$ (if we assume that when $h(\alpha) = 0$, the cooperative equilibrium holds). Now notice that $h(\alpha)$ is a linear function of α and thus it is enough to show that $h(1) > 0$ and $h(0) < 0$ and take $\tilde{\alpha}$ such that $h(\tilde{\alpha}) = 0$. We have that $U_1(D, C, (C,C,C,C), 1) = d > r = U_1(e_1, 1)$ so that $h(1) > 0$; on the other hand,

$U_1(D,0) = \tilde{G}(X) \left[\tilde{f}_2(C,C) \right] \left[1 + f_1(D,C) \right]$ and $U_1(e_1, 0) = \tilde{G}(X) \left[\tilde{f}_2(C,C) \right] \left[1 + f_1(C,C) \right]$, thus $h(0) = \tilde{G}(X) \left[\tilde{f}_2(C,C) \right] \left[f_1(D,C) - f_1(C,C) \right]$. Now, we have that

$$\tilde{f}_2(C,C) = \frac{\pi_1(C,C) - \pi_1^e(C)}{\pi_1^h(C) - \pi_1^{\min}(C)} = \frac{r - \dfrac{r+0}{2}}{r - 0} = \frac{1}{2}, \quad f_1(C,C) = \frac{\pi_2(C,C) - \pi_2^e(C)}{\pi_2^h(C) - \pi_2^{\min}(C)} = \frac{r - \dfrac{r+0}{2}}{r - 0} = \frac{1}{2} \quad \text{and}$$

$$f_1(D,C) = \frac{\pi_2(D,C) - \pi_2^e(C)}{\pi_2^h(C) - \pi_2^{\min}(C)} = \frac{0 - \dfrac{r+0}{2}}{r - 0} = -\frac{1}{2}, \quad \text{thus} \quad h(0) = -\tilde{G}(X)\frac{1}{2} < 0.$$

Thus,

$$\tilde{\alpha} = \frac{\tilde{G}(X)}{2[d-r] + \tilde{G}(X)} \tag{10}$$

and hence, if $\alpha_1 < \tilde{\alpha}$, and beliefs are (C,C,C,C), then C is a best replay to C. The same reasoning applies for the player 2, and thus the item (c) is done.

The proof of (d) is a direct consequence of $\tilde{G}(X)$ being decreasing and $\lim_{X \to \infty} \overline{G}(X) = 0$, since any function of the form $f(x) = \dfrac{x}{a+x}$, with $a > 0$ is increasing.

A.2. Proof of Proposition 2

First, we prove (a). To see $\rho_0^* = 0$ and $\rho_1^* = 1$ are steady states, notice that if $\rho_t = 0$, then $\rho_{t+1} = 0$ by definition of the dynamics, and the same reasoning applies for $1 - \rho_t = 0$ (the other type).

For the rest of the items, as now we are looking for interior steady states, we notice that $0 < \rho^* < 1$ is an interior steady state if and only if $\tilde{h}(\rho^*) = 0$, where $\tilde{h} : [0,1] \to \mathbb{R}$ given by

$$\tilde{h}(\rho) = \tilde{U}(\tau_C, \rho) - \tilde{U}(\tau_D, \rho) \tag{11}$$

where $\tilde{U}(\tau_D, \rho) = \rho U(e_2, \alpha_D) + (1-\rho)U(e_2, \alpha_D) = U(e_2, \alpha_D)$ and
$\tilde{U}(\tau_C, \rho) = \rho\lambda(\rho)U(e_1, \alpha_C) + (1-\rho\lambda(\rho))U(e_2, \alpha_C)$. Thus,
$\tilde{h}(\rho) = \rho\lambda(\rho)U(e_1, \alpha_C) + [1-\rho\lambda(\rho)]U(e_2, \alpha_C) - U(e_2, \alpha_D)$. Then, to see that there is only one non-trivial
steady state, it is enough to show that $\tilde{h}(\rho)$ has only one root in the set $(0,1)$. Indeed, as $\tilde{h}(\rho)$ is an
increasing linear function in $\rho\lambda(\rho)$, we have that ρ^* satisfies $\tilde{h}(\rho^*) = 0$ if and only if $\rho^*\lambda(\rho^*) = A^*$
(with $A^* < 1$, since we must have $U(e_1, \alpha_C) > U(e_2, \alpha_D)$), and as $\rho^*\lambda(\rho^*) = A^*$ has only one solution by

hypothesis, then $\tilde{h}(\rho^*) = 0$ has only one solution, which is denoted by ρ_I^* and satisfies $\rho_I^* = \dfrac{A^*}{\lambda(\rho_I^*)}$.

Finally, to see that $\rho_0^* = 0$ and $\rho_1^* = 1$ are asymptotically stable, due to that $\tilde{h}(\rho)$ is continuos, it is enough
to show $\tilde{h}(0) < 0$, and $\tilde{h}(1) > 0$, and that is the case as $\tilde{h}(0) = U(e_2, \alpha_C) - U(e_2, \alpha_D) < 0$ and
$\tilde{h}(1) = U(e_1, \alpha_C) - U(e_2, \alpha_D) > 0$. Indeed, we have that $\tilde{h}(\rho) < 0$ for all $\rho < \rho_I^*$, since $\tilde{h}(\rho)$ is increasing,
due to that $\lambda(\rho)$ is increasing; thus if $\rho_0 < \rho_I^*$ (ρ_0, the initial condition of the system), $\rho(t)$ is strictly
decreasing and then tends to some $\underline{\rho}$; now, again due to that $\tilde{h}(\rho)$ is increasing and continues, we must have
$\underline{\rho} = 0$; indeed, recalling that we defined in Section 4 the general dynamic rule as
$\rho(t+1) = F(\rho(t), \tilde{U}(\tau_C, t), \bar{U}(t))$, where F is continuous in all its arguments, we must have

$\underline{\rho} = F(\underline{\rho}, \lim_{t\to\infty}\tilde{U}(\tau_C, t), \lim_{t\to\infty}\bar{U}(t))$, and thus $\underline{\rho}$ must be a steady state, what is impossible if $\underline{\rho} > 0$, due
to that $\tilde{h}(\underline{\rho}) < 0$ and $\tilde{h}(\rho)$ is increasing. Similarly, we have that $\tilde{h}(\underline{\rho}) > 0$ for all for all $\rho > \rho_I^*$, and
hence if $\rho_0 > \rho_I^*$, we have that $\rho(t)$ tends to one as t tends to infinite. Consequently, ρ_I^* is unstable. The
item (i) is then proven.

Now we come to the item (b). Notice that if $U(e_1, \alpha_C) \le U(e_2, \alpha_D) < 0$, then $A^* \ge 1$, and therefore
$\tilde{h}(\rho) < 0$ for $\rho < 1$ (as $\rho\lambda(\rho) < 1$, for $\rho < 1$), and thus there are no interior steady states, and for all
$\rho_0 < 1$, $\rho(t)$ tends to zero as t tends to infinite. Also, if $\rho\lambda(\rho) = A^*$ does not have roots in $(0,1)$, then
$\tilde{h}(\rho) < 0$ for $\rho < 1$ (recall that $\tilde{h}(0) < 0$), and therefore the same reasoning as before applies.

A.3. Proof of Corollary 1

Taking into account that $A^* = \dfrac{(\alpha_D - \alpha_C)}{\alpha_C}\left[\dfrac{\left(1 + \dfrac{1}{4}\tilde{G}(X)\right)}{(r-1) + \dfrac{1-\alpha_C}{\alpha_C}\tilde{G}(X)}\right]$, it can be seen that A^* is a linear function

of α_D, with positive slope. Then, if α_D increases, A^* increases too, and as it was shown in Section 5 were
we presented the results, $1 - \rho_I^*$ decreases.

A.4. Proof of Corollary 2

First, it can be easily checked that $\dfrac{\partial A^*}{\partial\tilde{G}(X)} > 0 \leftrightarrow \alpha_C > \dfrac{4}{3+r}$. With this in mind, the proof of this corollary is

similar to that of Corollary 1. Simply notice that that if $\alpha_C > \dfrac{4}{3+r}$, then as X increases, $\tilde{G}(X)$ and A^*

decreases and thus $1 - \rho_I^*$ increases.

A.5. Proof of Corollary 3

Let Q represent institutional quality for instance and consider that $\lambda(\rho, Q_0) < \lambda(\rho, Q_1)$ for all $\rho \ne 0$. Thus,
$\rho\lambda(\rho, Q_0) < \rho\lambda(\rho, Q_1)$, and then $\rho_I^*(1)\lambda(\rho_I^*(1), Q_0) < \rho_I^*(1)\lambda(\rho_I^*(1), Q_1) = A^*$. Since $\rho\lambda(\rho, Q_0)$ is increasing, we have that $\rho_I^*(1) < \rho_I^*(0)$, what proves the first part of the corollary. The second part is totally
analogous to the first one, and hence omitted.