Cooperation and Network Formation

Felipe Balmaceda^{*} Centro de Economía Aplicada Industrial Engineering Department University of Chile[†]

Abstract

The present paper proposes a simple model for studying the interplay between self-enforcing cooperation and network formation. In particular, the model provides an answer to the ancient question of how cooperative behavior emerges in different communities and how the possibility of behaving cooperatively shapes the social structure of a community. In a sense, I provide an explanation of how trust, by which I mean the existence of self-sustainable cooperation, can emerge in a society and how the society is shaped by third-party enforcement.

^{*}I would like to thank participants in the regular seminar at the Economics Department of University of Chile, and the Latin American Meeting of the Econometric Society in Santiago-Chile. Any error however remains my own responsibility.

 $^{^{\}dagger}$ República 701, Santiago, 6521122, Chile fbalmace@dii.uchile.cl, http://www.dii.uchile.cl/economia/profe.htm

1 Introduction

It is by now widely known that many beneficial economic transactions are enforced by informal means. Social networks and reputation are frequently used as informal enforcement mechanisms to substitute for formal means like courts and explicit contracts and to enforce trust and cooperative behavior. Social scientists like Coleman (1988) and Putman (1993) have argued that social interactions and network closure -dense connections between network participants- are key determinants in fostering trust and cooperative relationships. Coleman illustrates this with Jewish diamond merchants in New York City. Jewish merchants lend bags of jewels worth thousands of dollars for examination overnight without signing any contract. By doing so, they save a great deal in lawyers' fees. Coleman argues that what prevents associates from not shirking their obligations is that anyone found guilty of malfeasance can kiss good-bye any future chance of being part of this profitable business, which is possible because merchants in this market belong to the same tight social circle and information flows well in that circle. Economists like Greif (1997) have documented that social norms were used to enhance exchange in Europe as early as the medieval period, and that informal enforcement mechanisms based on social networks have been utilized in economies with a relatively well developed legal system like the USA. In addition, he argues that social norms enhance efficiency relative to exchange based entirely on legally enforceable bilateral contracts.

The present paper examines the interplay between informal enforcement mechanisms and social interactions or network structure. In particular, the paper provides an answer to the question of how self-enforcing cooperative behavior emerges in different communities and how the possibility of behaving cooperatively shapes the social structure of a community. In a sense, an explanation of how trust emerges is provided (i.e. self-enforcing cooperative behavior) in a society and how the society is shaped by this possibility.

In so doing, the paper studies a framework where individuals in the early stage of their lifes invest in costly social links that have a long-term nature¹. Once links are formed, individuals engage in economic transactions captured by a repeated idiosyncratic Prisoner's Dilemma game with changing partners a-la-Kandori (1992), with the difference being that any given player can be matched with positive probability to more than one linked or unlinked player in each period. A crucial aspect of the model concerns the relationship between information transmission and social structure. In particular, it is assumed that when a player, say i, cheats another player, say j, then player j communicates that event, with positive probability, to each of his links. Yet, those links do not communicate that information to their links –that is, information flows one step and then stops. Thus, information concerning past actions flows only between linked players in a way that is consistent with the intuition that first-hand information carries more weight than second or third-hand information². In this setting, a strategy profile specifies the links that individuals have to form during the network formation phase and the actions they have to take during the interaction phase (hereinafter referred to as the continuation game). An equilibrium of the whole game is then a strategy profile such that the proposed network is pairwise stable –that is no pair of unlinked individuals has an incentive to deviate and form a link– and the strategy profile is a sequential equilibrium of the whole game.

¹For instance, this is the case when a supplier-customer relationship is formed or a joint-venture between two competing firms takes place.

²The structure of the model encompasses the two most emblematic repeated game models. When the probability of becoming informed is equal to one for each link and everyone is linked to everyone else, the model corresponds to the one in which everyone observes everyone else play in each period and everyone has the same chance to be matched to each other, while when that probability is zero, the model corresponds to the one in which no one else observes what happens between a given pair of individuals and everyone has the same chance to be matched to each other.

In what follows a network that is sustained by a equilibrium strategy profile is called a pair-wise equilibrium network (PWSE hereinafter).

When a trusting strategy is considered; that is, a strategy profile in which deviations occurring during the network formation phase result in cooperation becoming a sequential equilibrium during the continuation game are ignored³, the main results are as follows. When either cooperation is self-sustainable irrespective of the social structure –that is bilateral sanctions impose enough deterrence power to have selfenforcing cooperation– or there is no network structure under which cooperation is self-sustainable –that is the most severe community sanctions do not have enough deterrence power to stop players from defecting– no links are formed. Thus, when social structure has no bearing on whether cooperation is self-sustainable or not, the empty network is the only PWSE network since information has no benefits and forming links is costly.

When bilateral sanctions are not enough to induce cooperative behavior but community enforcement is capable of deterring defection under certain network structures the results are as follows. When the net gain from cooperation with a given player is higher than the cost of forming a new link, the symmetric network in which each individual forms the minimum number of links needed for cooperation to be selfsustainable is the only PWSE network. Because cooperation is enforced by all those players who become informed about a player's defection, self-sustainable cooperation requires a minimum number of players willing to punish a deviator. Once players reach that minimum span, they have no incentive to form more links since each extra link has a positive cost but no benefits. The reason for that is that the positive externality created by forming links –that is the possibility of behaving cooperatively in each encounter– is exhausted when each player reaches the minimum span⁴. In

³This strategy selects the Pareto efficient equilibrium.

⁴Because linked and unlinked players interact with each other with the same probability than

contrast, when the net gain from cooperation is lower than the cost of forming a new link and the minimum number of links needed for cooperation to be self-sustainable is lower than a threshold, the symmetric network in which each individual has the minimum span for self-enforcing cooperation is the only PWSE network, otherwise the empty network is the only PWSE network. The reason is that when the minimum number of links needed for self-enforcing cooperation is too large, the positive externalities of forming links do not outweigh the total cost of forming those links.

The uniqueness of PWSE networks is the result of using a trusting strategy. If a strategy profile in which any deviation occurring during the network formation phase triggers a punishment in the continuation game, there are multiple PWSE networks. Nevertheless, in any PWSE network other than the empty network, each player has at least the minimum number of links needed for self-sustainable cooperation and at most the maximum number of links under which cooperation remains self-sustainable⁵.

This paper relates two strands of the literature. First, our paper is related to the repeated game literature in which identical individuals, who may have different histories of play, belonging to a community are randomly matched to play a Prisoner's dilemma game (Ellison 1994; Kandori, 1992; Okuno-Fujiwara and Postlewaite, 1995). The main result of this literature is that cooperation can be made self-sustainable by mean of strategies that either punish deviators or reward conformers or use a mix of both. This result hinges on either the existence of information flows from each pair-wise match concerning past actions to the rest of the society or a contagious strategy; that is, players react to a deviation by punishing subsequent partners irretwo linked players do, when cooperation is self-sustainable between two linked players it also self-sustainable between a linked and unlinked player as long as the unlinked player has the minimum span.

⁵A maximum number of links exists when the speed at which the information flows decreases sufficiently fast with the number of links.

spective from their history of play. The main difference with our approach is that information flows are endogenously determined by the links or social network formed while their papers abstract from the effect of social structure by assuming a fixed and global pattern of play. This allows us to investigate network effects on the possibility of self-enforcing cooperation through social norms (third-party sanctions) and vice-versa. Still in the area of repeated games, an additional related paper is Haag and Lagunoff (2000). They consider a complete network where ex-ante heterogeneous players with respect to their discount factor play a prisoners' dilemma game with the whole population. They force players to play the same action–defect or cooperate– against each of their neighbors. Hence, the strategic implications of the exogenously given social network are the result of imposing players to play the same action rather than being the result of information transmission among an endogenously determined set of connected players.

Second, the paper borrows from the strategic network formation literature where individuals choose whether or not to form costly links (Jackson and Wolinsky 1996; Bala and Goyal, 2000; Jackson and Watts, 1999; and, Goyal and Joshi, 2003). None of these papers, however, deal with the issue of how social norms that induce selfenforcing cooperation affect the network formation process, and how then that affects individuals' cooperative behavior. In fact, our paper is one of the few that examines the interplay between network formation and cooperation. As far as we know Vega-Redondo (2002) is the only other paper that combines the network formation issue with the issue of self-enforcing cooperation. It considers a situation in which agents play idiosyncratic repeated prisoner's dilemma games with each of their neighbors. As in this paper, networks have a bearing on cooperation since players can communicate gradually via their links behavioral information about their links. Players can delete and form links as time goes on and only stable links are maintained; that is, links in which cooperation is self-sustainable. He shows that under certain conditions the network converges in the long-run to a stable and dense network in which cooperation is self-enforcing. Vega-Redondo's paper is more dynamic in nature than this one, with a focus on a different equilibrium concept and a different methodology. Yet, both of them arrive at similar conclusions, which is that self-enforcing cooperation requires a minimum network density, and thus they can be viewed as complementary.

The remainder of the paper proceeds as follows. In the next section, Section 2, the model is presented. Next, in Section 3, the analysis is undertaken. This is split in three sub-sections. The first concerns the repeated game equilibrium given a social network. The next, Sub-section 3.2, discusses the equilibrium of the whole game. That is, the PWSE networks are characterized in terms of a set of architectures, and the main characteristics of the equilibrium architectures are discussed. And, in the third, the main assumptions are discussed in some detail. The last section, Section 4, discusses the importance of network density, network size and network cohesiveness, and presents concluding remarks.

2 The Matching Game

2.1 The Static Game

The society consists of $N + 1 \ge 3$ infinitely lived players who may interact through a collection of infinitely repeated games. At t = 0, before repeated interactions start, the N + 1 individuals form an undirected graph (network formation stage). A graph is composed of a set of nodes and a set of links; each node represents a player while each link indicates a bilateral relationship between two players. This process will be explained in more detail below. At each period $t \ge 1$, any two given individuals will meet at most once. However, since each meeting is probabilistic, two individuals can go several periods without meeting with one another. The probability that any

two individuals i and j meet in any given period is time independent and given by $p_{ij}(g): G \longrightarrow [0,1]$, where g is the social network to which player i and j belong. For each pair of players who actually interact, $i, j \in g$, the stage game they play is a symmetric prisoner's dilemma (PD, hereafter) with a payoff matrix given by:

	D	С
D	d_{ij}, d_{ij}	$b_{ij}, 0$
С	$0, b_{ij}$	c_{ij}, c_{ij}

where, as is customary, D stands for defection and C for cooperation, and payoffs satisfy the following restrictions: $b_{ij} > c_{ij} > d_{ij} > 0$ and $2c_{ij} > b_{ij}$ for all ij pairs⁶.

It is assumed that each of the PD games is choice independent, in the sense that players' decisions in the past do not restrict the feasible choices in the future. They need not be, however, strategically independent since the behavior in the future may be made contingent on information gained in the past.

A key aspect of the approach adopted here is that information on how players have behaved in the past diffuses through the social network gradually. In particular, it is assumed that if player i and j are matched and player i cheats on j, then player j informs that with probability $\pi_{jk}(g): G \longrightarrow [0,1]$ to each $k \in N_j(g)$, where $N_j(g)$ is the set of j's links or social contacts. This probability is assumed to be constant over time. Thus, information concerning players' history of play flows only between linked players. The main implication of the information structure assumed is that, in general, the architecture of the social network has bearing on the extent of cooperation that the network can support in a self-sustainable fashion. It is also assumed that player k does not communicate the information he learns about i cheating j to his contacts. That is, information flows one step and then stops. This is consistent

⁶This last assumption ensures that cooperation is the efficient outcome in the one shot game.

with the intuition that first-hand information carries more weight than second or third-hand information.

Notice that when $\pi_{jk}(g) = 1$ for all $jk \in g$ and everyone is linked to every one else, the model corresponds to the one in which everyone observes everyone else play in each period, while when $\pi_{jk}(g) = 0$ for all $jk \in g$, the model corresponds to the one in which no one else, but the current partners, observe what actions have been taken in a given period⁷. So the model encompasses the two most emblematic repeated game models. Thus, information can range from atomized (people know only what has happened in their own interactions) to comprehensive (people know what has happened in all previous interactions), and everything about the game is common knowledge.

Finally, each individual discounts the future with a discount factor equal to $\delta > 0$ and the payoff from being unmatched in any given period is normalized to 0.

2.2 Social Networks

The social network or network of connections among individuals is described by a graph $g \in G \equiv \{g \mid g \subseteq g^{N+1}\}$ which is an N + 1 * N + 1 matrix, where G is the set of graphs of N + 1. Each element of g is denoted by g_{ij} and g is a symmetric matrix, *i.e.*, $g_{ij} = g_{ji}$. The element $g_{ij} = 1$ means that i and j are linked to each other and $g_{ij} = 0$ means that they are not linked to each other. The set of i's direct

⁷It is worthwhile to keep in mind that when $\pi_{ij}(g) = 0$ for all $ij \in g$ and for all g, the relationships are not atomized in the way many sociologists like M. Granovetter suggest-that is anonymous interactions in perfectly competitive markets. It is true that third parties do not affect the actions taken by any given player, but individuals may interact over and over again with the same group of individuals and have specific information on their behavior on past encounters. This provides interactions with a non-anonymous dimension that opens the possibility of developing the reputation of being a cooperator.

contacts is $N_i(g) \equiv \{j \neq i : g_{ij} = 1\}$ which is of size $n_i(g)$. Thus, the size of g is $n(g) = \sum_{i \in N} n_i(g)/2$ and if $n_i(g) = v$ for all $i \in N + 1$, then g is a symmetric network of degree v, denoted by g^v , while if $n_i(g) \neq n_j(g)$ for some $i, j \in N + 1$, then g is an asymmetric network. Among the possible symmetric networks, two will be of crucial importance. These are the complete network g^N , i.e., $n_i(g) = N$ for all $i \in N + 1$, and the empty network g^0 , i.e., $n_i(g) = 0$ for all $i \in N + 1$. In addition, g + ij (resp. g - ij) denotes the network obtained by adding (resp. subtracting) the link formed by player i and j to (resp. from) g and the set of all links is defined as $N(g) = \bigcup_i N_i(g)$.

A network is connected if there exists a path between any pair $ij \in N + 1$. A network $g' \subset g$ is a component of g if for all $i, j \in g', i \neq j$, there exists a path in g' connecting i and j, and for all $i \in g'$ and $k \in g$, $g_{ik} = 1$ implies that $k \in g'$. A component is complete if $g_{ij} = 1$ for all $i, j \in g'$.

Let $N_1(g)$, $N_2(g)$,..., $N_m(g)$ be a partition of players corresponding to the number of links that players have, i.e., $i, j \in N_k(g)$, k = 1, 2, ..., m if and only if $n_i(g) = n_j(g)$. Note that k refers to the order in the partition and not the exact number of links that players have. An inter-linked star architecture has at least two members in the above partition, and the maximally and minimally linked groups, respectively, satisfy the following two conditions: (i) $n_i(g) = N - 1$ for $i \in N_m(g)$ and (ii) $N_i(g) = N_m(g)$ for $i \in N_1(g)$. The star network is a special case of such an architecture with $|N_m(g)| = 1$ and $|N_1(g)| = N - 1$, where $|\bullet|$ is the cardinality of the set. An exclusive group architecture is characterized by m+1 groups, a group of isolated players $A_1(g)$ and m distinct groups of completely connected players, $A_2(g), ..., A_{m+1}(g)$. Thus, $n_i(g) = 0$, for $i \in A_1(g)$, while $n_j(g) = |A_x(g)| - 1$, for $j \in A_x(g)$ for $x \in \{2, ..., m+1\}$. A special case of this architecture is the dominant group network in which there is only one complete component and the rest of the players are isolated-that is m = 1.

2.3 Link Formation

Before the matching game starts each individual announces the links he wants to form. For all $i, j \in N + 1$, $s_{ij} = 1$ if i wants to form a link with j, and $s_{ij} = 0$ otherwise. By convention $s_{ii} = 0$. A link is created if and only if $s_{ij} * s_{ji} = 1$. Thus, $g_{ij} = 1$ if and only if $s_{ij} * s_{ji} = 1$. Links are thus created by mutual consent and are costly. Individual i's cost of forming a link with individual j is r. Thus, the cost of each link is assumed to be independent of the network structure. In addition, it is assumed that before the matching phase starts, everyone observes the number of links that everyone else has formed.

2.4 The Equilibrium Concept

The equilibrium concept used is pairwise-equilibrium. That is, the strategy profile $(s_i, a_i) \equiv (s_{i1}, s_{i2}, ..., s_{iN}, a_{ij}^0, a_{ij}^1, ..., a_{ij}^t)$, where a_{ij}^t is player *i*'s action when playing the stage game with player *j* in period *t*, if (s_i, a_i) is a sequential equilibrium (SE, hereafter) in the whole game and $(s_{i1}, s_{i2}, ..., s_{iN})$ is pairwise stable –that is no pair of players gains by altering the current configuration of links by either adding a link or by eliminating any subset of existing links. Thus, *g* is a pairwise-equilibrium network (PWSE, here thereafter) if and only if there is a sequential equilibrium strategy profile (s_i, a_i) which supports *g* in a pair-wise stable fashion⁸, and (a_i) is a sequential equilibrium in the continuation game.

Let $g \in G$. Then, individual *i*'s expected payoff is given by:

$$U_{i}(g) \equiv (1-\delta) \sum_{t=0}^{\infty} \delta^{t} \sum_{j \in N+1/\{i\}} p_{ij}(g) u\left(a_{ij}^{t}, a_{ji}^{t}\right) - rn_{i}(g),$$

⁸The concept of SE by itself is too weak a concept. In fact the empty network is always a SE. More generally, for any pair i and j, it is always mutual best response for the players to offer to form no link.

where the strategy profile (s_i, a_i) has been omitted to save on notation.

Pairwise stability implies that for all $g_{ij} = 0$, if $U_i(g + ij) > U_i(g)$ then $U_j(g + ij) < U_j(g)$.

3 The Analysis

3.1 The Repeated Game Phase

In this section I focus on the continuation game starting after the network formation phase has ended.

Notice that the model is stationary in nature and therefore the focus is on stationary equilibria in which, conditional on network q being formed in period 0, players rely on a "grim" trigger strategy that supports cooperation whenever it is self-sustainable. That is I abstract from the more complex issue of how individuals choose and arrive at that equilibrium instead of playing any other possible sub-game with perfect equilibria. The grim-trigger strategy considered in the continuation game, denoted by σ , is as follows. Player i when playing j, plays C if he has no information that j ever played D before, be it against i or against some third party h. On the other hand, if player i gets information that j has defected (against i or h), then i chooses defection himself against j in all interactions with j after receiving the information on j's defection. Mutual defections do not evoke sanctions, when they are not part of the prescribed pattern of behavior. That is, if k defects against i in order to punish the latter for cheating j, then k is not cheating but rather carrying out a prescribed punishment, so others observing the defection would not punish k in turn. It is also assumed that if k observes i cheating j then this is common knowledge between k and i.

Let \tilde{g} be the network formed in period 0. Then player i's expected payoff in the

continuation game when the grim-trigger strategy σ is followed by each player and $\tilde{g} = g^*$ is given by:

$$V_{it}\left(\sigma \mid g^{*}\right) \equiv (1-\delta) \sum_{j \in M_{t}} c_{ij} + \delta \sum_{j \in N+1/\{i\}} p_{ij}\left(g^{*}\right) c_{ij},$$

where M_t is the set of players with whom *i* is actually matched in period *t*.

The first term is the payoff from mutual cooperation in the current period and the second one is player i's long-run payoff from mutal cooperation in each possible encounter.

Player *i*'s expected payoff from deviating during the simultaneous move game against player k, who has a non-defection record, and then conforming to the grimtrigger strategy σ forever thereafter when everyone else is following σ is given by:

$$V_{ikt} (D, \sigma_{-i} \mid g^*) \equiv (1 - \delta) \left(b_{ik} + \sum_{j \in M_t / \{k\}} c_{ij} \right) + \delta \left(p_{ik} (g^*) d_{ik} + \sum_{j \in N + 1 / [N_k(g^*) \cup \{i,k\}]} p_{ij} (g^*) c_{ij} \right) + \delta \sum_{j \in N_k(g^*) / \{i\}} p_{ij} (g^*) \left[\pi_{kj} (g^*) d_{ij} + (1 - \pi_{kj} (g^*)) V_{ij} (\sigma \mid D, g^*) \right],$$

where $V_{ij}(\sigma \mid D, g^*) = (1 - \delta) c_{ij} + \delta [\pi_{kj}(g^*) d_{ij} + (1 - \pi_{kj}(g^*)) V_{ij}(\sigma \mid D, g^*)].$

The first term is self-explanatory. The second one is the probability of being matched with player k in the future and being punished by him forever thereafter. The third is the long-run payoff from being matched with someone different from k who is not linked to k and, therefore, cannot be informed that i defected against k in the past. The fourth term is the long-run payoff from being matched with someone different from k who is linked to k, and it is composed of two terms. The first corresponds to the long-run payoff when player j learns through k that i defected against player k during the last period, which is equal to mutual defection forever thereafter, and the other is the long-run payoff when player j does not learn right away that player i defected against player k during the last period.

Player *i* cooperates in each encounter with player *k* when everyone else including *k* follows σ if and only if $V_{it}(\sigma \mid g^*) \geq V_{ikt}(D, \sigma_{-i} \mid g^*)$. That is, cooperation is player *i*'s best response to σ_{-i} if and only if

$$\delta \left[\underbrace{\underbrace{p_{ik}\left(g^{*}\right)\left(c_{ik}-d_{ik}\right)}_{\text{gain from avoiding bilateral sanction}} + \underbrace{\sum_{j \in N_{k}\left(g^{*}\right)/\{i\}} \Pi_{kj}\left(g^{*}\right)p_{ij}\left(g^{*}\right)\left(c_{ij}-d_{ij}\right)}_{\text{gain from avoiding third-party sanctions}} \right] \geq \underbrace{\left(1-\delta\right)\left(b_{ik}-c_{ik}\right)}_{\text{short-run gain from deviation}},$$

$$(1)$$

where $\Pi_{kj}(g^*) \equiv \frac{\pi_{kj}(g^*)}{1 - \delta(1 - \pi_{kj}(g^*))}$.

This condition implies the following. First, i is more likely to cooperate with k when they are more likely to be matched with each other. When they meet each other more often, i suffers k's punishment more often. In other words, bilateral sanctions are more powerful when people meet each other more often. Second, the more links player k has, the more likely that i is willing to cooperate with k. The reason is that more people become informed of i's behavior through k and hence player i is more likely to be punished by someone other than k. In other words, third-party sanctions are more powerful, when the cheated player has more links. This implies that the network structure has bearing on self-enforcing cooperation. In fact, links have a positive spillover or externality since by adding links cooperation is more likely to be self-sustainable with any given player. Third, the larger the discount factor, the more likely that cooperation is self-sustainable.

After rearranging terms in equation 1, the following result is obtained.

Proposition 1 The strategy σ is a sequential equilibrium in the continuation game if and only if $p_{ik}(g^*) \ge P_{ik}(g^*)$ for all $i, k \in N+1$, where $P_{ik}(g^*) \equiv \max{\{\bar{p}_{ik}(g^*), \bar{p}_{ki}(g^*)\}}$ and

$$\bar{p}_{ik}(g^*) \equiv \frac{(1-\delta)(b_{ik}-c_{ik})}{\delta(c_{ik}-d_{ik})} - \sum_{j \in N_k(g^*)/\{i\}} \prod_{kj} (g^*) p_{ij}(g^*) \frac{(c_{ij}-d_{ij})}{(c_{ik}-d_{ik})}.$$
 (2)

The intuition is standard for a repeated game with changing partners. That is, cooperation between i and k is self-sustainable when they interact sufficiently often since they suffer each other punishment frequently if either of them were to defect. Yet, it immediately follows from 1 that the threshold $P_{ik}(g^*)$ is lower than that arising when only bilateral sanctions are allowed since the presence of information transmission to linked players allows for third-party sanctions. That is, a deviator is not only punished by the player cheated by him, but also by any other player who becomes aware of such behavior through the information provided by the cheated player. In fact, if $\pi_{kj}(g^*) = 0$ for all kj, third-party sanctions are not possible. Hence, the probability of interaction threshold depends on information flow, which in turn depends on the network architecture. For instance, if g^* is the empty network, only bilateral sanctions are possible while if g^* is the complete network and information flows perfectly, i.e., $\pi_{kj}(g^*) = 1$ for all $kj \in g^*$, then third-party sanctions are possible and have maximum deterrence power.

3.2 The Network Formation Phase

3.2.1 Preliminaries

Since the emphasis is on the role of information transmission, hereinafter the following simplifying assumptions are made⁹:

• Assumption 1: $p_{ij}(g) = p$ for all $i, j \in N$.

This assumption implies that the probability that any two players are matched is independent of the network architecture, yet the probability that any given player is matched with a linked player increases with the number of own links, while the probability of being matched with an unlinked player decreases with it.

 $^{^9\}mathrm{I}$ discuss the importance of these assumptions in more detail latter on.

• Assumption 2: $\pi_{kj}(g) = \pi(n_k(g)) g_{kj}, \pi(1) \leq 1, \lim_{n_k(g) \longrightarrow \overline{N}} \pi(n_k(g)) = 0,$ and $\pi(n_k(g)) \geq \pi(n_k(g) + 1)$ for all $n_k(g)$.

That is, the probability that k informs j of what was the behavior of his last period partner depends only on whether k and j are linked, and on the number of links that k has. In addition, when k has more links, he is less likely to communicate his experience to any given link.

• Assumption 3: $c_{ij} = c + \theta_{ij}$, $d_{ij} = d + \theta_{ij}$ and $b_{ij} = b + \theta_{ij}$ for all $i, j \in N + 1$.

It follows from equation (1) and assumptions 1, 2 and 3 that player i is willing to cooperate with player k when everyonelse including k follows σ if and only if the following holds:

$$p \ge \bar{p}_{ik}(n_k(g)) \equiv \frac{\bar{p}}{1 + (n_k(g) - g_{ik}) \prod (n_k(g))}.$$
(3)

where $\bar{p} \equiv \frac{(1-\delta)}{\delta} \frac{(b-c)}{(c-d)}$ and $\Pi(n_k(g)) = \frac{\pi(n_k(g))}{1-\delta(1-\pi(n_k(g)))}$.

Notice that the condition for cooperation being *i*'s best response to cooperation by player *k* is the same irrespective of whether *i* and *k* are linked or not since thirdparty sanctions are not affected by this. That is, the probability that player *k*'s links learn *i*'s behavior is the same whether *i* and *k* are linked, and the number of players linked to *k* that are willing to punish *i* is the same whether *i* and *k* are linked. In addition, assumption (2) implies that for each $ik \in N + 1$, there exists $\tilde{n} \in [2, \infty)$ such that $\bar{p}(n_k(g) + 1) < \bar{p}(n_k(g))$ for all $n_k(g) \leq \tilde{n}$ and $\bar{p}(n_k(g) + 1) > \bar{p}(n_k(g))$ for all $n_k(g) > \tilde{n}^{10}$. It readily follows from this that if $\bar{p} > p \geq \bar{p}(\tilde{n})$, then there is a minimum number of links, denoted by \underline{v} , such that $p = \bar{p}(\underline{v})$ and a maximum number of links, denoted by \bar{v} , such that $p = \bar{p}(\bar{v})^{11}$. This guarantees that there is an interval

 $^{^{10}\}mathrm{Here}$ it is assumed that \tilde{n} is the lowest integer that satisfies this condition.

¹¹If $\tilde{n} = \infty$, then $\overline{v} = \infty$.

for the number of links given by $[\underline{v}, \overline{v}]$ for which cooperation is *i*'s best response to cooperation by player *k*.

Lastly, but not least, it is easy to show that if the speed at which the information flows among linked individuals increases –that is $\pi(n_k(g))$ becomes higher for any $n_k(g)$ – then \underline{v} decreases and \overline{v} increases.

3.2.2 A Trusting Strategy

The strategy for the whole game considered here, called the trusting strategy, prescribes players forming a set of links that generate a network of social contacts g^{*12} . In the matching phase or continuation game starting in period 1 each player defects forever thereafter either when a network $\tilde{g} \notin g^* \cup G(\sigma)$ is observed or $G(\sigma) = \phi$ and follows σ from period 1 onwards otherwise, where $G(\sigma)$ is defined as follows:

 $G(\sigma) \equiv \{g \in G \mid \sigma \text{ is a SE in the continuation game}\}.$

It is worthwhile to remark on two things concerning this strategy. First, deviations from the prescribed network g^* that sustain σ as a SE in the continuation game starting in period 1 are ignored by all players. And, second, when the observed network \tilde{g} does not belong to $g^* \cup G(\sigma)$, the strategy prescribes to defect forever thereafter, which is a SE in the continuation game¹³. Therefore, only individuals' incentive to cooperate on-the-equilibrium path must be considered.

Given this strategy, assumptions 1, 2 and 3, and network \tilde{g} , player *i*'s long-run ¹²The name trusting strategy is meant to highlight that people trust each other to certain extent. That is, people believe that some of the deviations that may occur do not have negative consequences for other members of the society.

¹³In the next section a strategy in which any deviation from the prescribed network triggers a punishment in the continuation game.

payoff in the whole game is given by:

$$U_{i}\left(\tilde{g}\right) = I\left(\tilde{g}\right)\sum_{j\in N} p\left(c+\theta_{ij}\right) + \left[1-I\left(\tilde{g}\right)\right]\sum_{j\in N} p\left(d+\theta_{ij}\right) - rn_{i}\left(\tilde{g}\right),$$

where $I(\tilde{g})$ is an indicator function that takes the value 1 when $\tilde{g} \in g^* \cup G(\sigma)$ and $G(\sigma) \neq \phi$, and takes the value 0 otherwise.

The analysis is split into two cases: (i) either σ can be a SE in the continuation game starting in period 1 irrespective of the network architecture or $G(\sigma) = \phi$; and (ii) σ can be a SE in the continuation game starting in period 1 only under certain network architectures.

In the next proposition the case in which the action taken is the same irrespective from the network architecture is considered. The following result proves the existence of a unique PWSE network and characterizes it.

Proposition 2 (i) Suppose that N and \bar{p} are such $p < \bar{p} (\min \{\tilde{n}, N\})$, then the empty network is the unique PWSE network; and (ii) suppose that \bar{p} is such that $p \ge \bar{p}$, then the empty network is the unique PWSE network.

The intuition is as follows. When $p < \bar{p} (\min \{\tilde{n}, N\})$, there is no network architecture that induces self-enforcing cooperation between any pair of players and, therefore, the long-run net gain from forming a new link is $p(d + \theta_{ij}) - r - p(d + \theta_{ij}) < 0$. Thus, it is never worthwhile to form a link independent of what other players do. Whereas when $p \ge \bar{p}$, cooperation is self-sustainable by bilateral sanctions irrespective from the network architecture and, therefore, the long run net gain from a link is $p(c + \theta_{ij}) - r - p(c + \theta_{ij}) < 0$. Again, it is never worthwhile to form a link independent of what other players do. Notice this result is partially driven by the fact that no deviation from the prescribed network g^* triggers a punishment by other players since cooperation is never destroyed by such deviations. Thus, when the network architecture has no bearing on whether σ is self-sustainable, the empty network is the unique PWSE network.

The following proposition is concerned with the case in which cooperation is selfsustainable only under certain architectures and proves existence of a unique PWSE network and provides a characterization of it.

Proposition 3 Suppose that N and \bar{p} are such that $\bar{p} > p \ge \bar{p} (\min \{\tilde{n}, N\})$. If either pc < pd + r and $\underline{v} \le \frac{Np(c-d)}{r}$ or $pc \ge pd + r$, then the symmetric network $g^{\underline{v}}$ is the unique PWSE network, otherwise the empty network is the unique PWSE network.

The intuition is straightforward. When cooperation is induced by a social norm in which a deviating player is punished not only by the cheated player, but also by all those players who become informed about a player's defection, inducing cooperation requires a minimum number of players willing to punish a deviator. Since information about a player's last period action flows only to the cheated player's links, self-sustainable cooperation between two players demands that each of them forms the minimum number of links, \underline{v} . Notice that whether those players are linked themselves does not matter, what matters it is that the cheated player communicates his information to enough people who have a positive probability of being matched with the deviator¹⁴.

Once players have the minimum span required for self-enforcing cooperation they have no incentive to form more links since each extra link has a positive cost and no benefit. The reason is that the externality created by forming links, which is to ensure that cooperation is self-sustainable, is exhausted when the minimum span is reached. If a player has more links than the minimum span, then he has an incentive to delete all those links that are beyond that since that strategy saves a cost r per-link and *that*

¹⁴Behind this result is the fact that two unlinked players have a positive probability of being matched to each other in any given period.

kind of deviation is ignored by other players since they do not destroy self-enforcing cooperation. Thus, no one forms more than \underline{v} links.

Do players have an incentive to form less than \underline{v} links? The answer depends on whether $pc \leq pd + r$. When $pc \geq pd + r$, it is obvious that each player is willing to form \underline{v} links while when pc < pd + r individuals are willing to form \underline{v} links when $\sum_{j \in N} p(c + \theta_{ij}) - \underline{v}r \geq \sum_{j \in N} p(d + \theta_{ij})$ and no less despite that the cost of an extra link is larger than its direct benefit. The reason is that when the minimum span for self-enforcing cooperation is reached, cooperation becomes self-sustainable among linked and unlinked players and, therefore, the gains from cooperation with unlinked players $-(N - \underline{v})p(c - d) > 0$ - compensates for the loss from it with linked players $-\underline{v}(p(c - d) - r) < 0$. Whereas when $Npc - \underline{v}r < Npd$, the gains from cooperation with unlinked players do not compensate for the loss from cooperation with linked players and hence each player has an incentive to delete all his links.

There are a few worthwhile remarks concerning this result. Firstly, players choose to form a limited number of links despite the fact that as more links are formed the amount of information transmitted raises. The reason is that more information is costly and beyond certain amount that extra information has no value since cooperation is self-sustainable. Secondly, the faster the information travels the network, the smaller the minimum span. And, thirdly, $g^{\underline{v}}$ is efficient or total welfare maximizing¹⁵.

To see this notice that in a PWSE network g^* total welfare is given by:

$$W(g^{*}) \equiv \sum_{i \in N+1} \left(I(g^{*}) \sum_{j \in N} p(c + \theta_{ij}) + [1 - I(g^{*})] \sum_{j \in N} p(d + \theta_{ij}) - rn_{i}(\tilde{g}) \right), \quad (4)$$

where $I(g^*)$ is an indicator function that takes the value 1 when $g^* \in G(\sigma)$ and takes the value 0 otherwise.

It is easy to verify using equation 4 that the symmetric network of degree \underline{v} is

¹⁵This implies that the strategy proposed select the Pareto efficient equilibrium.

the unique efficient network when $\underline{v} < \frac{Np(c-d)}{r}$ and the empty network otherwise since each extra link beyond \underline{v} is costly but its net benefit is zero.

3.2.3 A Non-Trusting Strategy

The strategy considered here prescribes players forming a set of links that generate a network of social contacts g^* . In the matching phase or continuation game starting in period 1 each player defects whenever a network \tilde{g} , different from the prescribed network g^* , is observed, and follows σ from period 1 onwards otherwise.

It is worth remarking on two things concerning this strategy. As opposed to a trusting strategy, deviations from the prescribed network g^* that do not destroy cooperation are no longer ignored and, second, when the observed network \tilde{g} is different from the prescribed network g^* , the rational strategy prescribes defecting forever thereafter which is a sequential equilibrium in the continuation game (out-of-theequilibrium path) starting in period 1. Therefore, as with the trusting strategy, only the incentives to cooperate on-the-equilibrium path must be considered.

Given this strategy, assumptions 1, 2 and 3, and network \tilde{g} , player *i*'s long-run payoff in the whole game is given by:

$$U_{i}\left(\tilde{g}\right) = I\left(\tilde{g}\right)\sum_{j\in N} p\left(c+\theta_{ij}\right) + \left[1-I\left(\tilde{g}\right)\right]\sum_{j\in N} p\left(d+\theta_{ij}\right) - rn_{i}\left(\tilde{g}\right),$$

where $I(\tilde{g})$ is an indicator function that takes the value 1 when $\tilde{g} = g^*$ and takes the value 0 otherwise.

The analysis is again split into two cases: (i) either σ can be a SE in the continuation game irrespective of the network architecture or $G(\sigma) = \phi$; and (ii) σ can be a SE in the continuation game only under certain network architectures.

The following result proves the existence of PWSE networks and characterize them whenever possible for case (i).

Proposition 4 (i) Suppose that N and \bar{p} are such that $p < \bar{p}(\min\{N,\tilde{n}\})$, then the empty network is the unique PWSE network; and (ii) suppose that \bar{p} is such that $p \ge \bar{p}$, then any network in which the maximally connected player has at most $\frac{Np(c-d)}{r}$ links is a PWSE network.

Notice that when bilateral sanctions are enough for self-enforcing cooperation, then many different architectures are PWSE networks among which one could find the star network with a central player having at most $\frac{Np(c-d)}{r}$ links, the exclusive group architecture and the dominant group architecture. However, every one but the empty network is inefficient since no links are needed to have self-enforcing cooperation between any pair of players. The reason for this is that any deviation in the network formation phase is severely punished by defection forever thereafter in each encounter.

In the next proposition the case in which σ can be a SE in the continuation game only under certain network architectures is considered. The following result proves existence of a unique PWSE network and characterize it.

Proposition 5 Suppose that N and \bar{p} are such that $\bar{p} > p \geq \bar{p}(\min\{\tilde{n}, N\})$. If pc < pd + r and $\underline{v} < \frac{Np(c-d)}{r}$ then any symmetric network g^v with $v \in [\underline{v}, \overline{v}]$ and $v \leq \frac{Np(c-d)}{r}$ is a PWSE network and any asymmetric network in which the maximally connected player has at most $n_m(g^*) \leq \min\left\{\frac{Np(c-d)}{r}, \overline{v}\right\}$ links and the minimally connected player has at least $n_1(g^*) \geq \underline{v}$ links is a PWSE network while if $\underline{v} \geq \frac{Np(c-d)}{r}$, the empty newtork is the unique PWSE network; and (ii) if $pc \geq pd+r$, then any symmetric network g^v with $v \in [\underline{v}, \overline{v}]$ is a PWSE network and any asymmetric network and the minimally connected player has at least $n_1(g^*) \geq \underline{v}$ links is a PWSE network and any asymmetric network in which the maximally connected player has at player has at least $n_p(\underline{v}, \overline{v})$ is a PWSE network and any asymmetric network in which the maximally connected player has at player has at least $n_1(g^*) \geq \underline{v}$ links is a PWSE network and any asymmetric network in which the maximally connected player has at most n_m(g^*) \leq \overline{v} links and the minimally connected player has at most n_m(g^*) \leq \overline{v} links and the minimally connected player has at player has at

Notice that in any PWSE network, no one can have less than 2 links and, therefore, neither the star nor the interlinked star, and neither the exclusive groups architecture nor the line architecture can be PWSE networks since in each of them there are players with less than 2 links. Furthermore, if $\overline{v} < N$, the entire network is not a PWSE network.

The multiplicity of equilibria is due to the fact that any deviation from the prescribed strategy triggers a punishment in the continuation game. When a player deletes any number of links other players defect forever thereafter, and thus when deciding whether to defect or not he has to take into account that cooperation is fully destroyed. When the number of links that a player has to form in any given network g^* is larger than $\frac{Np(c-d)}{r}$, he prefers to save the cost of forming links and defect in each encounter forever thereafter.

Finally, when $\underline{v} < \frac{Np(c-d)}{r}$ any PWSE network different from the symmetric network of degree \underline{v} is inefficient since the total cost of forming links is larger than that under \underline{v} while the total benefit from self-sustainable cooperation is the same across PWSE networks. When $\underline{v} > \frac{Np(c-d)}{r}$, the empty network is the unique efficient PWSE network.

3.3 Robustness Issues

These results have been derived under several assumptions. Firstly, a particular kind of payoff heterogeneity was assumed. Secondly, the probability of information transmission depends only on number of links that the transmitter of information has. And, thirdly, linked as well as unlinked players have a positive probability of interacting with each other. In this section, I will briefly discuss each of these in turn.

If a more general kind of heterogeneity were to be assumed, then it would be more difficult to characterize PWSE networks. The reason for that is the conditions for self-sustainable cooperation will differ across links. In particular, under assumptions 1 and 2, player i's best response to cooperation by player k is cooperation if and only if the following holds:

$$p > \bar{p}_{ik} \left(n_k \left(g^* \right) \right) \equiv \frac{\bar{p}_{ik}}{1 + \Pi \left(n_k \left(g^* \right) \right) \sum_{j \in N_k(g^*) / \{i\}} \frac{(c_{ij} - d_{ij})}{(c_{ik} - d_{ik})}},$$

where $\bar{p}_{ik} \equiv \frac{(1-\delta)(b_{ik}-c_{ik})}{\delta(c_{ik}-d_{ik})}$.

In order for cooperation to be sustained in a self-enforcing fashion, player *i* needs to rank each possible link by $\bar{p}_{ik}(n_k(g^*))$ and, then to be sure that he is willing to cooperate with the link with the largest $\bar{p}_{ik}(n_k(g^*))$. This suggests that players will form the most profitable links first, and that the minimum span for each player would be different. Thus, in general no symmetric network would be a PWSE network, except for the empty and the complete network.

Secondly, if the information transmission probability were to depend in a more general fashion on the network g, then under assumptions 1 and 3, player i's best response to cooperation by player k is cooperation if and only if the following holds

$$p > \bar{p}_{ik}(g^*) \equiv \frac{\bar{p}}{1 + \sum_{j \in N_k(g^*)/\{i\}} \prod_{jk} (g^*)}$$

Again, for a trusting strategy to be an equilibrium, each player *i* has to be willing to cooperate with the players with the highest probability threshold, $\bar{p}_{ik}(g^*)$. For instance, if $\Pi_{jk}(g^*) = \Pi(n_j(g^*), n_k(g^*))$ and this decreases with $n_j(g^*)$ and $n_k(g^*)$, then it is still true that each player will form the minimum span required for selfenforcing cooperation. In addition, the minimum span will be the same for each link.

In each of these two extensions, the basic economics of the model remain unchanged; that is, under a trusting strategy each player forms the minimum number of links so that the externality of forming links is materialized. In other words, players invest in links just up to the point where cooperation becomes self-sustainable in each encounter. Hence, only network span matters. Lastly, if the matching probability $p_{ij}(g)$ were to depend on the network structure in a more general fashion, then under assumptions 2 and 3, player *i*'s best response to cooperation by player k is cooperation if and only if the following holds:

$$p_{ik}(g^*) \ge \bar{p}_{ik}(g^*) \equiv \bar{p} - \Pi(n_k(g^*)) \sum_{j \in N_k(g^*)/\{i\}} p_{ij}(g^*).$$

How the results change depends a lot on the specification of $p_{ij}(g)$. For the sake of brevity, the discussion will focus on of the most interesting case which is $p_{ij}(g) = pg_{ij}$. That is, only linked individuals have a positive probability of being matched. In this case the condition for self-enforcing cooperation becomes:

$$p \ge \bar{p}_{ik}(g^*) \equiv \frac{\bar{p}}{1 + \Pi(n_k(g^*)) n_{ik}(g^*)},$$
(5)

where $n_{ik}(g^*)$ is the cardinality of the set $N_k(g^*) \cap N_i(g^*)$.

It is easy to show that if $p(d + \theta_{ij}) \ge r$ for all ij the complete network is the unique PWSE network while if $p(c + \theta_{ij}) < r$ for all ij, the empty network is the unique PWSE network. So, let us consider in more detail the case in which $p(c + \theta_{ij}) \ge r >$ $p(d + \theta_{ij})$ for some ijs, and for the sake of simplicity let us assume that $\Pi(n_k(g^*)) =$ Π for all $n_k(g^*)$. That is, the speed at which the information flows does not decrease with the number of links. Thus, there is a minimum number of mutual links that player i has to have in order for his best response to cooperation by k is cooperation. Let us denote that cutoff by \underline{v}_{ik}^2 and notice that $\underline{v}_{ik}^2 = \underline{v}^2$ is the same for all $ik \in N+1$.

In this case the following is proposition is formally shown in the appendix.

Proposition 6 Suppose that $p(d + \theta_{ij}) < r$ for all $ij \in N + 1$ and $p(c + \theta_{ij}) \ge r$ for some $ij \in N + 1$. Then under a trusting strategy in any PWSE network all profitable cooperative links are mutually linked and no player can have less than $\hat{n}_i(g^*)$ cooperative links. Furthermore, if $p(c + \theta_{ij}) \ge r$ for all $ij \in N + 1$, then the complete network is the unique PWSE network. This proposition shows that all profitable cooperative links –that is links in which $p(c + \theta_{ij}) \geq r$ – must be mutually linked, and that players may in fact form cooperative links that are not profitable. The reason is that self-enforcing cooperation requires that each pair of players has a minimum number of common links and thus it is optimal if there is not enough profitable links to make cooperation self-enforcing to form some unprofitable cooperative links. In other words, is still the case that a minimum span is required to exploit the externality created by information transmission through linked players. However, in contrast to the case in which the probability of interaction is positive across all pairs, here not only the minimum span or network density matters, but also network cohesiveness. Not only is it necessary that any given individual has to know enough people, but also those he knows must know each other to certain extent. In fact when all cooperative links are profitable, the complete network is the unique PWSE network and thus everyone is mutually linked or know each other directly. This implies full cohesiveness.

4 Discussion and Conclusions

The results in propositions 3 and 4 highlight the importance of network density, which formally means the ratio of the total number of links to the total number of links in complete network. That is, $\frac{v}{N}^{16}$. On the one hand, they show that when bilateral sanctions are not powerful enough to enforce self-enforcing cooperation, but the most severe third-party sanctions are, a minimum span per individual is needed for selfenforcing cooperation. On the other hand, they show that there is maximum span under which cooperation is self-enforcing. The reason is that, while as a network becomes denser the number of potential players that can punish a deviator increases,

¹⁶This follows from that in a network with N+1 individuals there are N(N+1) links in a complete network and in the symmetric network $g^{\underline{v}}$ each of the N+1 individuals has \underline{v} .

the probability of being punished by any one of them decreases. This implies that the expected total punishment increases with network span, when this is sufficiently low, and decreases when that is sufficiently high. Thus, a network yields self-enforcing cooperation when it is neither too spare nor too dense. Furthermore, it was shown that there is an efficient or welfare maximizing density level which is given by the minimum span needed for self-enforcing cooperation \underline{v} .

The results also show that the role of the size of the population on self-enforcing cooperation is ambiguous. On one hand, when network size N is such that $N \ll \tilde{n}$ and $p < \bar{p}(N)$, cooperation is never self-sustainable and hence no player has incentive to form links. On the other hand, when $N \gg \overline{v}$ and $p < \overline{p}(\tilde{n})$, there is no network architecture that induces players to cooperate in each possible encounter and thus no player has an incentive to form links. Thus, when the population size is either too small or too large and the probability of interaction is small between any pair of individuals, cooperation cannot be sustained in a self-enforcing fashion since the deterrence power of third-party sanctions is small. In contrast, when the population size is neither too small nor too large, changes in the population size do not have an effect on the level of cooperation that can be sustained in a self-enforcing fashion by any PWSE network, but they can have an effect on total welfare. When a non-trusting strategy is used in any PWSE network each player must have no less than \underline{v} and no more than \overline{v} links. However, the most efficient network is the symmetric $g^{\underline{v}}$. Thus, for any $N \in [\underline{v}, \overline{v}]$ as N decreases, total welfare increases whenever a PWSE network requires a player to form more than \underline{v} links. In other words, a smaller population size puts a limit on network density, which in turn decreases the total cost of link formation.

This provides a theoretical foundation for the empirical importance of network density as emphasized by sociologists like J. Coleman and M. Granovetter. In particular, they argue that network closure or network density make enforcement of group cooperative behavior more effective since denser networks increase the quality and reliability of third-party enforcement through an improvement of information transmission among individuals. For instance, studies of the rapidly growing economies of East Asia almost always emphasize the importance of dense social networks. These networks, often based on the extended family or close-knit ethnic communities like the overseas Chinese, are seen as responsible for fostering trust, lowering transaction costs, and facilitating information transmission. In fact, studies show that China's extraordinary economic growth over the last decade has depended less on formal institutions than on personal connections to allocate resources efficiently and make contracts self-enforceable. Network density has also been important in the development of advanced Western economies. Mark Granovetter has pointed out that economic transactions like contracting or job searches are more efficient when they are embedded in social networks. Studies of highly efficient, highly flexible industrial districts emphasize networks of collaboration among workers and small entrepreneurs; e.g; Silicon Valley. In the OECD countries, the most comprehensive finding is that controlling for other key variables, well-connected people are more likely to be housed, healthy, hired and happy.

The result in proposition 6 emphasizes a different dimension of social networks, which is cohesiveness. In other words, the mutuality of links. Coleman has emphasized not only the importance of network density, but also network closure or network cohesivneess. In fact, he defines the term Social Capital in terms of network cohesivness. He argues that social capital is, often, underprovided since the strong externalities associated with it are typically not internalized by individuals' own link investment decisions. While the results here emphasize the importance of network closure when only linked individuals can interact with each other as Coleman (1990) does, the results also show, in contrast to Coleman's view, that individuals do have an incentive to invest in links up to the point where the externality associated with them is fully internalized. This is clear from the fact that individuals are even willing to invest in non-profitable cooperative links when the number of profitable links available is not enough for cooperation to be sustained in a self-enforcing fashion.

This paper has contributed to understanding the emergence of self-enforcing cooperative behavior in different communities and how the possibility of self-sustainable cooperative behavior shapes the social structure of a community. In a sense, it is an explanation of how trust emerges in a society is provided (i.e. self-enforcing cooperative behavior) and how the society is shaped by that possibility. In particular, it was shown that individuals are willing to invest resources to form links in order to take advantage of third-party or community enforcement at least up to the point where cooperation can be made self-sustainable. This leads to communities or social networks that are neither too spare nor too dense. That is, individuals form, at least, the minimal number of links needed for cooperation to be sustained in a self-enforcing fashion and never form more links than those under which cooperation is self-sustainable. In short, social networks are formed so that linked as well as unlinked individuals trust each other.

Furthermore, in equilibrium not only do social networks emerge in which individuals take advantage of network externalities created by information transmission between linked players, but also the network structures that emerge are, under a trusting strategy at least, Pareto efficient. Thus, individuals are by themselves able to solve the difficult problem of devising an efficient social network structure or they internalize the strong externalities associated to link formation at a personal cost.

I will end by briefly discussing two applications which I believe are worth pursuing. The first is the role of formal organizations –technological progress– in shaping social networks and how that effects cooperative behavior. This will help us better understand the consequences which tremendous progress in communication technology has had on network density and network closure, which is sometimes referred to as Social Capital, and shed some light on the importance of it. It will also help in understanding why informal enforcement mechanisms based on social networks have been utilized in economies with relatively well developed legal systems like the USA. In fact, the model suggests that informal enforcement mechanisms like third-party enforcement enhance efficiency relative to exchange based entirely on legally enforceable bilateral contracts¹⁷.

The second application concerns the study of social norms based on ascriptive features. Given the prevalence throughout history of social norms in which individuals condition their behavior on group identity or personal traits like ethnicity and religion, the study of these types of social norms is of a great empirical importance. For instance, Greif (1993) presents an interesting example of social norms based on personal traits used by Maghribi traders in the Mediterranean during the 11th century. He documents that the lack of formal institutions able to enforce overseas trading, induced the Maghribi traders to adopt the following social norm: no Maghribi trader would trade with another Maghribi trader who had cheated a Maghribi trader before. The study of ascriptive social norms will help us to understand another feature of social networks, which is segregation by ascriptive features like race and religion even when those are payoff-irrelevant. In fact, work done by the author in Balmaceda (2005b) suggests that total welfare of adopting a within-group social norm is larger than that when a between-group social norm is adopted, but a within-group norm gives rise to a more segregated society than a between-group norm. Thus, the prevalence of within group social norms and segregation by ascriptive features may be rationalized in terms of efficiency by following the interplay between network formation and cooperation proposed in this paper.

 $^{^{17}}$ See, Balmaceda (2005a).

A Appendix

Proof. of proposition 2.

(i) Suppose that $p < \bar{p} (\min \{\tilde{n}, N\})$, then $G(\sigma) = \phi$ since cooperation is not selfsustainable. Then, for any network g that players can form, the strategy prescribes in the continuation game to play defection and therefore the long-run payoff from that strategy is $p(d + \theta_{ij})$. Each player's long-run payoff from a link is $p(d + \theta_{ij}) - r$, take any two players, say i and j, that are not linked and suppose that they form a link, then each gets a long-run benefit equal to $p(d + \theta_{ij})$ and each faces a cost equal to $r + p(d + \theta_{ij})$. Because $r \ge 0$, they do not have an incentive to deviate and form a link and each player has an incentive to delete all his links. Thus, the empty network is the unique PWSE network.

(ii) Suppose next that $p \geq \bar{p}$, then $G(\sigma) = G$ since bilatareal sanctions are enough to induce cooperation between any pair of players. Then players are willing to cooperate in each encounter irrespective from the number of links that each player has. Then the empty network is the unique PWSE network since each player has an incentive to unilaterally delete all his links and no one has an incentive to add a link. To see this take any network $g^* \in G(\sigma)$ and take any two players, say *i* and *j*, that are not linked and suppose that they form a link, then each gets a long-run benefit equal to $p(c + \theta_{ij})$ and each faces a cost equal to $r + p(c + \theta_{ij})$ since a deviation from the prescribed network g^* by adding a link does not trigger a punishement in the continuation game. Because r > 0, they do not have an incentive to deviate and form that link. Now consider any player *i*, then his payoff in any network g^* is $\sum_{j\in N} p(c + \theta_{ij}) - n_i(g^*)r$. Then, deleting any positive number of links *x* is optimal since by doing so player *i*'s payoff is $\sum_{j\in N} p(c + \theta_{ij}) - (n_i(g^*) - x)r$ which is larger than $\sum_{j\in N} p(c + \theta_{ij}) - n_i(g^*)r$ since r > 0. Thus, adding links is not profitable and deleting links is and, therefore, the empty network is the unique PWSE network. **Proof.** of proposition 3.

Suppose that $\bar{p} > p \geq \bar{p}_i (\min{\{\tilde{n}, N\}})$ and let $N_1(g^*), N_2(g^*), ..., N_m(g^*)$ be a partition of players corresponding to the number of links that each player has in the PWSE network g^* ; i.e., $i, j \in N_k(g^*), k = 1, 2, .., m$ if and only if $n_i(g^*) = n_j(g^*) = k$, where k refers to the order in the partition but not to the exact number of links that players in partition k have.

Notice that $G(\sigma)$ is non-empty and that $G(\sigma) \neq G$. If $g^* \notin G(\sigma)$, then σ is not an equilibrium in the continuation game starting in period 1 and therefore (g^*, σ) is not an equilibrium of the whole game. Suppose now that $g^* \in G(\sigma)$; that is, cooperation is self-sustainable in each encounter. Player *i*'s payoff from g^* is $\sum_{j \in N} p(c + \theta_{ij}) - \theta_{ij}$ $rn_{i}(g^{*})$. Deviating by adding links is never worthwhile since either $g^{*} + ij \in G(\sigma)$ in which case adding a link only increases the cost or $g^{*} + ij \notin G\left(\sigma\right)$ in which case adding a link increases the cost and decreases the payoff from each link. Consider now player *i*'s incentive to delete links. Suppose player *i* deletes a subset $X_i(g^*)$ of links, with $x_i \equiv |X_i(g^*)|$. Then his payoff is $\sum_{j \in N} p(c + \theta_{ij}) - (n_i(g^*) - x_i) r$ if x_i is such that $g^* - X_i(g^*) \in G(\sigma)$ while his payoff is $\sum_{j \in N} p(d + \theta_{ij}) - (n_i(g^*) - x_i)r$ if x_i is such that $g^* - X_i(g^*) \notin G(\sigma)$. Notice that if the subset of links to be deleted leads to the second case, then player i's best deviation is to delete all his links. Thus, player *i* is not willing to delete all his links if and only if $n_i(g^*) \leq \frac{Np(c-d)}{r}$. Thus, in any PWSE network g^* , $n_m(g^*) \leq \frac{Np(c-d)}{r}$. Notice also that if $X_i(g^*)$ is such that $g^* - X_i(g^*) \in G(\sigma)$, player *i* has an incentive to delete any link that is not strictly needed to maintain self-sustainable cooperation in the network. Thus, in any PWSE network g^* , players form the minimum number of links needed to achieve self-enforcing cooperation.

Recall that \underline{v} is defined as the minimum the degree such that $p = \overline{p}_{ik}(g^*) \equiv \frac{\overline{p}}{1+(v-g_{ik})\Pi(v)}$. Then it readily follows from this that no player wants to have more than \underline{v} links. Thus, if $\underline{v} \leq \frac{Np(c-d)}{r}$, the symmetric network g^* with degree \underline{v} is a

PWSE network.

Notice that if $pc \ge r + pd$, then $\frac{Np(c-d)}{r} > N$ and therefore $\underline{v} \le \frac{Np(c-d)}{r}$ holds, while this is not necessarily the case when pc < r + pd.

Suppose now that g^* is an asymmetric PWSE network. Then, the maximally connected player has at most $n_m(g^*) \leq \frac{Np(c-d)}{r}$ links and the minimally connected player has at least $n_1(g^*) \geq \underline{v}$ links, otherwise cooperation is not self-sustainable.

Now consider any player $i \in N_k(g^*)$, then it was shown already that player i has an incentive to delete all the links that are not needed for cooperation to be self-sustainable and therefore in any PWSE network g^* each player will form \underline{v} links. Thus, if $\underline{v} < \frac{Np(c-d)}{r}$, then the symmetric network of degree \underline{v} is the unique PWSE network.

Proof. of proposition 4.

(i) The proof is the same as the one for part (i) in proposition 2.

(ii) Suppose next that if $p \geq \bar{p}$. Since bilatareal sanctions are enough to induce cooperation between any pair of players then for any prescribed network, σ is an equilibrium in the continuation game. Then, player *i*'s payoff in any network g^* is $\sum_{j \in N} p(c + \theta_{ij}) - rn_i(g^*)$. Adding links is never optimal because it increases the cost and triggers a punishment. Deleting any positive number of links *x* yields the following payoff is $\sum_{j \in N} p(d + \theta_{ij}) - (n_i(g^*) - x)r$ and, therefore, the best deviation is to delete all links. In this case player *i*'s payoff from deleting all his links is smaller than $\sum_{j \in N} p(c + \theta_{ij}) - n_i(g^*)r$ whenever $n_i(g^*) \leq \frac{Np(c-d)}{r}$. Thus, in any PWSE network g^* , $n_m(g^*) \leq \frac{Np(c-d)}{r}$ and $n_1(g^*) \geq \underline{v}$.

Proof. of proposition 5.

Suppose that $\bar{p} > p \geq \bar{p}_i (\min{\{\tilde{n}, N\}})$ and let $N_1(g^*), N_2(g^*), ..., N_m(g^*)$ be a partition of players corresponding to the number of links that each player has in the PWSE network g^* ; i.e., $i, j \in N_k(g^*), k = 1, 2, ..., m$ if and only if $n_i(g^*) = n_j(g^*) = k$, where k refers to the order in the partition but not to the exact number of links that

players in partition k have.

If $g^* \notin G(\sigma)$, then σ is not an equilibrium in the continuation game starting in period 1 and therefore (g^*, σ) is not an equilibrium of the whole game. Suppose now that $g^* \in G(\sigma)$; that is, cooperation is self-sustainable in each encounter. Furthermore, if $g^* \in G(\sigma)$, no player can have less than $\underline{v} \geq 2$ links and no more than \overline{v} . Player *i*'s payoff from g^* is $\sum_{j \in N} p(c + \theta_{ij}) - n_i(g^*) r$.

Take any two players, say $i \in N_k(g^*)$ and $j \in N_{k'}(g^*)$, with $g_{ij} = 0$. Then if i and j deviate and form a link each obtains $\sum_{j \in N} p(d + \theta_{ij}) - (n_i(g^*) + 1)r$ for any $n_i(g^*)$ since any deviation from g^* triggers a punish in the continuation game. This is lower than $\sum_{j \in N} p(c + \theta_{ij}) - n_i(g^*)r$ and therefore no pair of players have an incentive to add a link. Consider now player $i \in N_k(g^*)$ incentive to delete links. Suppose player i deletes a subset $X_i(g^*)$ of links, with $x_i \equiv |X_i(g^*)|$. Then his payoff is $\sum_{j \in N} p(d + \theta_{ij}) - (n_i(g^*) - x_i)r$ and therefore his best deviation is to delete all his links, in which case his payoff is $\sum_{j \in N} p(d + \theta_{ij})$. Thus, player i is not willing to delete all his links if and only if $n_i(g^*) \leq \frac{Np(c-d)}{r}$. Thus, in any PWSE network g^* , $n_m(g^*) \leq \min\left\{\frac{Np(c-d)}{r}, \overline{v}\right\}$ and $n_1(g^*) \geq \underline{v}$.

Notice that if $pc \ge r + pd$, then $\frac{pN(c-d)}{r} > N$ and therefore $\overline{v} \le \frac{Np(c-d)}{r}$ holds, while this is not necessarily the case when pc < r + pd.

Suppose now that g^* is a symmetric network of degree $v(g^*)$. Then, if $v(g^*) \in [\overline{v}, \underline{v}]$ and $v(g^*) \leq \frac{Np(c-d)}{r}$, g^* is a PWSE network.

Then, the maximally connected player has at most $n_m(g^*) \leq \frac{Np(c-d)}{r}$ links and the minimally connected player has at least $n_1(g^*) \geq \underline{v}$ links, otherwise cooperation is not self-sustainable.

Suppose now that g^* is an asymmetric PWSE network. Then, the maximally connected player has at most $n_m(g^*) \leq \min\left\{\frac{Np(c-d)}{r}, \overline{v}\right\}$ links and the minimally connected player has at least $n_1(g^*) \geq \underline{v}$ links, otherwise cooperation is not selfsustainable. Thus, if $\underline{v} > \frac{Np(c-d)}{r}$ then the empty network is the unique PWSE network.

Proof. of proposition 6.

Suppose that $\bar{p} > p \ge \bar{p}_{ij} (\min \{\underline{v}^2, N\})$ for all $ij \in N + 1$ and let $N_1(g^*), N_2(g^*), ..., N_m(g^*)$ be a partition of players corresponding to the number of links that each player has in the PWSE network g^* ; i.e., $i, j \in N_k(g^*), k = 1, 2, .., m$

if and only if $n_i(g^*) = n_j(g^*) = k$, where k refers to the order in the partition but not to the exact number of links that players in partition k have.

Notice that $G(\sigma)$ is non-empty and that $G(\sigma) \neq G$. If $g^* \notin G(\sigma)$, then σ is not an equilibrium in the continuation game starting in period 1 and therefore (g^*, σ) is not an equilibrium of the whole game. Suppose now that $g^* \in G(\sigma)$; that is, cooperation is self-sustainable in each encounter. Player i's payoff from g^* is $\sum_{j \in N_i(g^*)} p(c + \theta_{ij}) - rn_i(g^*)$. Take now any two players *i* and *j* with $g_{ij} = 0$ such that $p \geq \bar{p}_{ij} (g^* + ij)$. The the benefit from deviating and forming a link is $p(c + \theta_{ij})$ while the cost is r. Thus, as long as $p(c + \theta_{ij}) \ge r$, i and j have an incentive to deviate and form the *ij* link. Notice that a deviation does not trigger a punishment since that deviation does not destroy cooperation with anyone else. This implies that each pair of players that can form a sufficiently profitable cooperative link must be mutually linked. Consider now any two players i and j with $g_{ij} = 1$ such that $p \ge \bar{p}_{ij}(g^*)$ but $p(c + \theta_{ij}) < r$. Suppose that player *i* deviates by unilaterally deleting the *ij* link. Then he saves r and looses $p(c + \theta_{ij}) + A_i(g^*)$, where $A_i(g^*)$ is the set of all players with whom self-enforcing cooperation is detroyed by deleting the ij link. Notice that if $A_i(g^*) = \phi$, then player *i* has an incentive to delete the *ij* link since $r > p(c + \theta_{ij})$ since that triggers no punishment in the continuation game because self-enforcing cooperation is not destroyed. Suppose now that $A_i(g^*) \neq \phi$, then by deleting the *ij* link player is punished by all his links in the continuation game, given this if player i deletes the ij link unilaterally, he is better-off deleting all his links. Thus, player iis not willing to delete all his links if and only if $\sum_{j \in N_i(g^*)} p(c + \theta_{ij}) - rn_i(g^*) \ge 0$.

This condition boilds down to $n_i(g^*) \ge \hat{n}_i(g^*) \equiv \frac{n_i(g^*)r - \sum_{j \in N_i(g^*)} p\theta_{ij}}{pc}$.

Thus, in any PWSE network all cooperative links are mutually linked and no player can have less than $\hat{c}_i(g^*)$ cooperative links. Furthermore, if $p(c + \theta_{ij}) \ge r$ for all $ij \in N + 1$, then the complete network is the unique PWSE network.

References

- [1] Axelrod R. The Evolution of Cooperation, New York, 1984 Basic Books.
- Balmaceda, F., Cooperation: Social networs versus Formal Institutions, Mimeo, CEA, University of Chile, 2005a.
- [3] Balmaceda, F., Why Do We live in Segregated Societies?, Mimeo, CEA, University of Chile, 2005b.
- [4] Bendor, J. and Mookherjee, D., Regulating Intergroup Conflict: Ascriptive versus Universalistic Norms, Working Paper, Boston University, 1998.
- [5] Degenne, A. and Forse, M Introducing Social Networks, Sage Publication, 1999.
- [6] Durlauf, S.N (2002) On the Empirics of Social Capital, Economic Journal, vol 112 (November) pag. 459-478.
- [7] Dutta, S. Building Trust, Discussion Paper, No TE/96/305, LSE, March 1996.
- [8] Dutta, B., and M.O. Jackson "The Stability and Efficiency of Directed Communication Networks," Review of Economic Design, 5, 251-272, 2000.
- [9] Dutta, B., and Mutuswami, S. "Stable Networks, Journal of Economic Theory, 76, 322-344, 1997.
- [10] Ellison, G., Cooperation in the Prisoners Dilemma with Anonymous Random Matching, Rev. Econ. Stud., 61: 567-588, 1994.
- [11] Fukuyama, F., Trust: The Social Virtues and the Creation of Prosperity, The Free Press Paperbacks, New York, 1995.
- [12] Goyal, S., Self-Organization in CommunicationNetworks," Econometrica, 68, pp 1181-1230, 2000.

- [13] Granovetter, M., "The Strength of Weak Ties", American Journal of Sociology, 78, 1360-1380, 1973.
- [14] Granovetter, M. S., Getting a Job: A Study of Contacts and Careers, University of Chicago, Press, Chicago, 1995. 2nd edition.
- [15] Grannovetter, M. S., Economic Action and Social Structure: The Problem of Embeddedness, American Journal of Sociology, 481-510, 1985.
- [16] Greif, A., Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders Coalition, Amer. Econ. Rev., 83(3): 525-548, June 1993.
- [17] Greif, A., Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies, J. Polit. Econ., 102(5): 912-950, 1994.
- [18] Jackson, M.O. (2001) "The Stability and Efficiency of Economic and Social Networks," forthcoming: Advances of Economic Design, edited by Murat Sertel, Springer-Verlag.
- [19] Jackson, M.O. and A. Watts (1998) "The Evolution of Social and Economic Networks," forthcoming: Journal of Economic Theory.
- [20] Jackson, M.O. and A. Wolinsky (1996) "A Strategic Model of Social and Economic Networks," Journal of Economic Theory, 71, 44-74.
- [21] Kali, R., Endogenous Business Networks, Journal of Law and Economic Organization, 15(3): 615-36, October 1999.
- [22] Kandori, M., Social Norms and Community Enforcement, Review of Economic Studies, 59: 63-80, 1992.

- [23] Kranton., R. E. The Formation of Cooperative Relationships, Journal of Law, Economics and Organization, v12, No1, 214-233.
- [24] Kranton, R. E., Reciprocal Exchange: A Self-Sustaining System, Amer. Econ. Rev., 86(4): 830-851, September 1996.
- [25] Kranton, R. and D. Minehart (1998) "A Theory of Buyer-Seller Networks," forthcoming: American Economic Review.
- [26] McMillan, J. and Woodruff, C., Interfirm Relationships and Informal Credit in Vietnam, Quarterly Journal of Economics, 114(4): 1285-1320, November 1999.
- [27] Milgram, S. The Small World Problem, Pyschology Today, 1: 61-7, 1967.
- [28] Milgrom, P. R., North, D. C., and Weingast, B., The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs, Economics and Politics, 2(19): 1-23, 1991.
- [29] Montgomery, J. D., Social Networks and Labor-Market Outcomes: Toward an Economic Analysis, American Economic Review, 81(5): 1408-1418, December 1991.
- [30] North, D. C., Institutions, Institutional Change, and Economic Performance, Cambridge University Press.
- [31] Portes, A. and Landolt, P. Unsolved Mysteries: The Tocqueville Files II. The American Prospect no.26, Spring 1996.
- [32] Putman, R. D., The Prosperous Community: Social Capital and Public Life. The American Prospect no.13, Spring 1993.
- [33] Putman, R. D.(1995), Bowling Alone: America's Declining Social Capital. Journal of Democracy, 6, n~1, 65-78.

- [34] _____(2000), Bowling Alone: The Collapse and Revival of American Community, New York: Simon & Schuster.
- [35] Raub, W. and Weesie, J., Reputation and Efficiency in Social Interactions: An Example of Network Effects, Amer. J. Sociology, 96(3): 626-54, November 1990.
- [36] S. H. Strogatz, Exploring complex networks. Nature 410, 268-276, 2001.
- [37] Vega-Redondo, F. (2002) Building Social Capital in a Changing World: Repeated Interactions in an Evolving Social Network, Working Paper N 2002-26, Instituto Valenciano de Investigaciones Economicas.
- [38] Wasserman, S. and K. Faust (1994) Social Network Analysis: Methods and Applications, Cambridge University Press.
- [39] Watts, D. J., Netwroks, Dynamics and the Small-World Phenomenon, American Journal of Sociology, v. 105, n. 2, 1999, 493-597
- [40] Watts, D. J. and Strogatz, S. H., Collective dynamics of 'small-world' networks. Nature 393, 440-442, 1998.
- [41] Weber, M, The Theory of Social and Economic Organizations, The Free Press, 1947.