Trust in Cohesive Communities*

Felipe Balmaceda
Diego Portales University†

Juan F. Escobar
University of Chile‡

June 4, 2013

Abstract

This paper studies which social networks maximize trust and cooperation when agreements are implicitly enforced. We study a repeated trust game in which trading opportunities arise exogenously and the social network determines the information transmission technology. We show that cohesive communities, modeled as social networks of complete components, emerge as the optimal community design. Cohesive communities generate some degree of common knowledge of transpired play that allows players to coordinate their punishments and, as a result, yield relatively high equilibrium payoffs. Our results provide an economic rationale for the commonly argued optimality of cohesive social networks.

JEL classification numbers: D85, C73, L14, D8

Keywords: repeated games, social networks, information transmission, coordination, cohesiveness

---

*This paper supersedes our previous working paper "Self Governance in Social Networks of Information Transmission" (2011).

†Economics Department, Diego Portales University, Avda. Santa Clara 797, Santiago, Chile, E-mail: felipe.balmace@udp.cl

‡Center for Applied Economics, Department of Industrial Engineering, University of Chile, Republica 701, Santiago, Chile, E-mail: jescobar@dii.uchile.cl

1 Introduction

The use of implicit mechanisms of misconduct deterrence has been widely recognized and documented by economists (Milgrom, North, and Weingast, 1990; Greif, 1993), political scientists (Ostrom, 1990; Fearon and Laitin, 1996), sociologists (Coleman, 1990; Raub and Weesie, 1990), and legal scholars (Bernstein, 1992). Crucial to their use is the way in which trading partners get informed about mischievous actions. As illustrated by Greif's (2006) study of contract enforcement between medieval Maghribi traders, a close-knit community can quickly disseminate information about its members' behaviors and, as a result, can align their members' incentives by employing implicit, community-based sanctions that punish behaviors deemed unacceptable or opportunistic.

The existence of implicit agreements in cohesive communities is also apparent in modern societies.[1] In the automobile industry, for example, firms usually outsource large amounts of work and suppliers are routinely called upon to make specific investments. Hold-up problems are overcome by the threat of future business losses. As McMillan (1995) documents, one of the keys to deter opportunistic behavior in vertical relationships is the existence of cohesive business associations, such as Japanese keiretsus or Korean chaebols, that facilitate information exchange about parties' previous performances.[2] More generally, several authors have highlighted a community cohesion as a pivotal ingredient to attain cooperation in trust-based transactions (Coleman, 1990; Greif, 1993; Rauch, 2001; Dixit, 2006). What are precisely the advantages, if any, a fully cohesive community has over partially cohesive social networks when agreements are implicit is not well understood.[3]

This paper provides a game theoretical rationale for the social optimality of fully cohesive communities when the enforcement of agreements is implicit and social networks determine how information is transmitted. Cohesive communities generate some degree of common knowledge of transpired play within each component and, as a result, players can employ relatively harsh

[1] By cohesion or cohesiveness, we mean the level of clustering or cliquishness of a social network, as in Section 2.2.3 in Jackson (2008).

[2] McMillan (1995) observes that "the institutionalization of links among firms that is provided by the keiretsu system arguably serve as ... an information-provision device." He also notes that "by providing a mechanism for keeping track of any opportunistic behavior ..., the keiretsu provides a disincentive to such behavior."

[3] As the literature on repeated games has shown (Abreu et al., 1990; Kandori, 2002, 1992; Mailath and Samuelson, 2006), efficient play can be enforced under a variety of assumptions when players are patient enough. One could therefore presume that when players are patient enough, the cohesiveness of the social network becomes irrelevant.

coordinated punishments. Since the availability of harsher punishments allows players to enforce higher equilibrium payoffs, cohesive communities are optimal.

Section 2 introduces a repeated trust game played by N investors and one agent. At each round $t \geq 1$ one out of the N investors is randomly and uniformly selected to play a trust game with the agent. More specifically, the investor decides whether or not to participate and, if he participates, he also picks an action or investment level; then the agent chooses whether to cooperate (or share the investment's return with the agent) or to defect (or appropriate the return in full). The equilibrium of the stage game is inefficient as the agent will defect after an investment is already made and, anticipating this behavior, the investor will not participate. The agent's temptation to defect may be curtailed by the existence of community sanctions governed by a social network of investors G. We assume that if the agent misbehaves when facing investor i, then i and all his connections in G (i.e., all those who are directly linked to i) become aware of that, and may act upon by changing their play in the continuation game. We focus on perfect Bayesian equilibria that sustain cooperation on the path of play.

Section 3 presents our main results. In Theorem 1, we establish the Pareto optimality of a social network of equally sized complete components (i.e., networks in which all players have the same number of connections and if a player is connected to two other players, then all three players are connected). Namely, a social network of equally sized complete components G^* can yield a higher expected payoff to each community member than any other network G in which each investor has at most the same number of connections as in G^*. Hence, even when each investor has the same number of connections in G^* and G, equilibrium payoffs sustained by appropriately designed trigger strategies in G^* Pareto dominate any equilibrium in G. As illustrated in Figure 1, this implies that when links are scarce and the number of connections per player is exogenously given, the formation of a fully cohesive social network is really the best that all game players can hope for.

The mechanism behind this result can be understood as follows. In a network in which two investors are not directly connected but have a common connection, a defection by the agent against one of them decreases the leverage the other investor has in the continuation game. After such defection, the second investor is not aware of the first defection, and defecting against him is profitable for the agent. A social network having an incomplete component allows the agent to defect twice by incurring a rather small loss in continuation value. In contrast, in a network of complete components, play within a component is common knowledge and punishments can

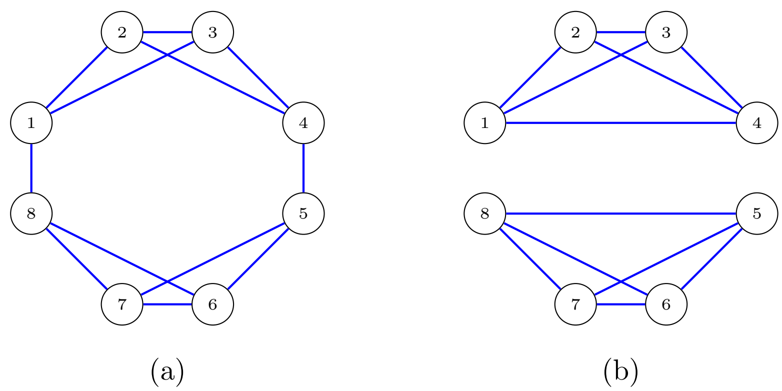

Figure 1: Illustration of Theorem 1. Even when each investor has the same number of connections in each network, network (b) yields higher payoffs than all equilibria of network (a).

be immediately implemented leaving no room for further off-path defections. Consequently, networks of complete components result in higher equilibrium payoffs.

Some simple but important corollaries can be deduced. In particular, Corollary 2 provides conditions under which networks of complete components maximize total welfare given a fixed number of links to be distributed among the investors. This result complements Theorem 1 by showing that cohesive communities maximize the total value that a community may create even when links can be freely allocated among investors. We also prove in Corollary 3 that given a subset V of investors that form a complete subnetwork (or clique) in G^*, it is possible to construct equilibria in the game with monitoring technology G^* such that no alternative network in which players in V have (weakly) less connections can result in higher equilibrium payoffs for investors in V. This result shows how a subgroup of players may favor the formation of a fully cohesive subnetwork, even when the whole social network ends up being barely cohesive.

Our analysis and insights can be adapted to study alternative models of repeated interactions in networks. Section 4 studies a model in which players are randomly matched to play a (two-sided) prisoners' dilemma. Before actually deciding whether to cooperate or to defect, one of the matched players decides the size of the match. Our model is similar to Kandori's (1992) community enforcement model, with the added twist that we introduce a social network of information transmission. As the community enforcement literature has forcefully argued, a desirable property of a social norm is its robustness to players' mistakes or trembles. An equilibrium with high on-path payoffs but utterly fragile does not seem like an attractive way of organizing a community.

Motivated by these considerations, we introduce a family of equilibria, called *contagion-free equilibria*, that keep punishments local. The idea here is to rule out fragile arrangements in which any (unmodeled) mistake may cause a community breakdown. Theorem 2 shows that networks of equally sized complete components are optimal, but the logic is slightly different from that of Theorem 1. In the two-sided model of Section 4, the restriction to contagion-free equilibria in networks of incomplete information implies a trade-off between high on-path investments and off-path incentives. In networks of complete components that trade-off is absent: once on-path incentives are satisfied, off-path incentives that keep punishments local are obtained for free. Networks of complete components are therefore optimal.

In sum, this paper shows how networks of complete components minimize the costs of off-path gaming in a family of repeated games where the monitoring technology is modeled as a network of information transmission. Such off-path gaming arises as, following an off-path defection, a player may find attractive to keep defecting when facing other players. These *contagion* effects have received considerable attention in related previous works and play an important role in our analysis. In the model studied in Sections 2 and 3, contagion is detrimental for incentives and therefore the optimal network design should avoid it. In the two-sided model of Section 4, an equilibrium involving contagion is normatively ruled out by our robustness criterion. In all these cases, this paper presents tools to prove that the formation of fully cohesive communities responds to the optimal network design.

We now contextualize our work. Cohesive networks play an important role in collective action games (Chwe, 1999; Morris, 2000). While some degree of common knowledge is also important to attain efficient coordination in those static games, our insights emphasize off-path coordination as a vehicle to attain high equilibrium payoffs.

Important antecedents for this paper come from work on community enforcement by Kandori (1992), Ellison (1994), Harrington (1995), and Okuno-Fujiwara and Postlewaite (1995). These authors study repeated games in which players interact infrequently, but the sources of information transmission are not captured by a social network. Our main application in Section 4 introduces a social network of information transmission in Kandori's (1992) model and can be seen as a natural extension of the community enforcement literature.

There is a growing literature studying repeated games on networks, with a particular focus on the social network as a determinant of transaction opportunities among players.[4] Some authors emphasize how heterogeneity in payoffs may make attractive the use of third party sanctions (Bendor and Mookherjee, 1990), favoring the formation of cohesive networks (Haag and Lagunoff, 2006; Lippert and Spagnolo, 2011). When players only observe their own interactions, a social network of trading opportunities has also a key role spreading information about distant interactions. In particular, as Mihm et al. (2009), Lippert and Spagnolo (2011) and Ali and Miller (2013) illustrate, cohesive networks fasten information dissemination and quicken punishments.[5] Jackson et al. (2012) study a repeated favor exchange model and show that equilibrium networks satisfying renegotiation-proofness and contagion-freeness (which they term "robustness to social contagion") must consist of cliques joined in a tree-like form. In their perfect information game, off-path behaviors take a very different form and therefore the forces behind our model and theirs are different.[6]

Ahn and Suominen (2001) study a model similar to ours, in which the network of information transmission is drawn at the beginning of each round. They do not study the problem of optimal network design, nor do they explore how networks of complete components minimize the costs of contagion. Wolitzky (2013) studies a repeated public provision game in which the social network determines the monitoring technology. Wolitzky (2013) does not focus on the optimality of networks of complete components, so our research question is different. As he does, we use some lattice theory techniques to prove our results, but in our model once a player defects whether or not defections keep occurring depends on the actions being implemented on the path of play. Since such off-path behavior is arbitrary in our model, on-path actions are neither complements nor substitutes in general networks.[7] To prove our optimality results, we obtain estimates of continuation values and derive relaxed incentive constraints.

[4]Other contributions to this topic include Raub and Weesie (1990), Bloch et al. (2008), Karlan et al. (2009), Fainmesser (2012) and Nava and Piccione (2013).

[5]The precise meaning of cohesiveness in some of these papers is rather different from ours. Mihm et al. (2009) show that enforcing cooperation requires triadic closure (Coleman, 1990), while Lippert and Spagnolo (2011) highlight the existence of cycles to spread punishments and attain cooperation. The main result in Ali and Miller (2013) shows that in networks of complete components, a contagion-equilibrium Pareto dominates all stationary equilibria of networks having incomplete components.

[6]Double-defections do not arise in Jackson et al.'s (2012) model. As we show in Lemma 1, incentive constraints deterring double-defections are key in networks of incomplete components.

[7]It is easy to construct examples of our model in which a player's own maximal on-path action increases in the neighbors' actions and decreases as a function of the actions of more distant players.

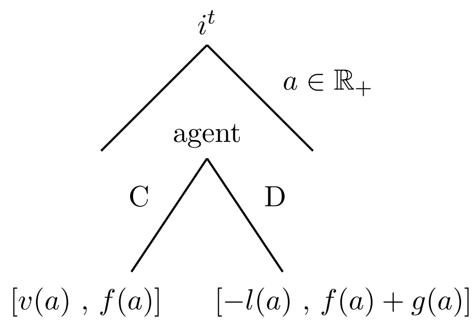

Figure 2: The game between investor i^t and the agent when the investor participates.

2 Set Up

2.1 The Environment

We consider a repeated game model in discrete time between $N+1$ players, with $N \geq 2$. At each round $t = 1, 2, \ldots$, an agent, hereinafter player 0, faces a randomly selected investor $i^t \in \{1, \ldots, N\}$ and they play a trust game. A social network G of investors determines the information transmission technology. Below we describe the details of the game.

2.1.1 Payoffs and Matching Technology

At the beginning of each round $t \geq 1$, an investor i^t is randomly and uniformly selected from the set $\{1, \ldots, N\}$. Investor i^t and the agent may produce some surplus in round t by playing a trust game as in Kreps (1996). First, the investor decides whether to participate or not (P or NP). If he chooses not to participate, then per-period payoffs are equal 0 and the game moves on to round $t+1$. If investor i^t participates, he chooses an action or investment $a \in \mathbb{R}_+$. This action can be thought of as the investor choosing how much to trust the agent. After observing the investor's decision, the agent decides whether to cooperate (C) or to defect (D). Figure 2 illustrates the game when investor i^t participates.

Investors who are not selected get a per-period payoff equal to 0. Given a stream of per-period payoffs $(u_i^t)_{t\geq 1}$ for player $i \in \{0, \ldots, N\}$, his utility function equals $\sum_{t\geq 1} \delta^{t-1} u_i^t$, where $\delta \in]0,1[$ is the common discount factor.

The following restrictions on payoffs are maintained throughout the paper.

Condition 1

1. $f, v, l, g > 0$.

2. *g is increasing, f is nondecreasing and v is increasing.*

3. $\frac{\delta}{1-\delta}[\frac{1}{N}f(0), \max f] \subseteq \text{range}(g)$, *where* range$(g)$ *is the range of g.*

Restriction 1 says that for any investment level, the agent's static best response is to defect. As a result, the investor decides not to participate and both players' payoffs equal 0. This outcome is Pareto dominated by the outcome in which the investor participates and the agent cooperates. Restriction 2 implies that when the agent cooperates, both players benefit from a higher investment. Yet, the higher the investment, the larger the agent's temptation to defect. The third restriction is mainly technical and allows us to bound equilibrium investments in the repeated game. It holds when f is bounded, g is continuous and unbounded, and $g(0) \leq \frac{\delta}{1-\delta}\frac{1}{N}f(0)$.

2.1.2 Social Networks, Information Flows and Histories

There is a social network of investors that determines the monitoring technology. More formally, a social network of investors is a symmetric matrix $G \in \{0,1\}^{n \times n}$ such that $G_{ij} = 1$ if and only if i and j are linked.[8] We also write $ij \in G$ whenever $G_{ij} = 1$. We assume that $G_{ii} = 0$. Denote by $N(i, G) = \{j \mid ij \in G\}$ the set of i's neighbors in G and define the closed neighborhood of i as $\bar{N}(i, G) = N(i, G) \cup \{i\}$. When there is no risk of confusion, we will simply write $N(i)$ and $\bar{N}(i)$ instead of $N(i, G)$ and $\bar{N}(i, G)$.

Information flows among investors are as follows. Let $i^t \in \{1, \ldots, N\}$ be the investor chosen at round t. If i^t plays NP, then the signal all investors receive is empty. If i^t chooses P, then all investors $j \in N(i^t)$ become aware of that and observe whether the agent cooperated or defected. More formally, if investor i^t participates and the agent played $x^t \in \{C, D\}$, then player $j \in N(i^t)$ receives a signal $s_j^t = (i^t, x^t)$, while if $j \notin \bar{N}(i^t)$ then j receives a signal $s_j^t = \emptyset$. If i^t did not participate, then all players $j \neq i$ receive signal $s_j^t = \emptyset$. Player i^t perfectly observes play during round t. Players receive signals only about current interactions; in particular, we assume that information does not travel any further. A history h_i^t for investor i at the beginning of round

[8]By considering a symmetric matrix G, we assume that the network is undirected

t will consist of all the signals $(s_i^1, \ldots, s_i^{t-1})$ he has received during past play. We assume the agent has perfect information and therefore his information sets are all singletons.

These monitoring assumptions capture the idea that the network determines the transactions each investor can observe. One could also interpret these assumptions as saying that the victim (and only the victim) of a mischievous action let all his connections know that the agent is a miscreant. All these interpretations seem appropriate in applications, and we postpone further discussion to Section 5.

2.1.3 Examples

Our model can be interpreted broadly and captures key aspects of several repeated interactions. The agent can be seen as a firm that may or may not hold-up the specific investments made by suppliers (investors), as in Williamson (1979). The social network represents all the business ties among the different suppliers (Greif, 1993; McMillan, 1995; Uzzi, 1996). The game can also be seen as a model in which consumers (investors) decide whether or not to buy experience goods from a monopolist (agent) who may be tempted to sell low-quality goods. The social network represents all sources of information on the monopolist performance (e.g. online feedback systems, word-of-mouth). The basic strategic structure of our repeated trust game also appears in models of relational contracting, in which the employer (agent) may renege payments to workers (investors).

2.2 Strategies That Sustain Cooperation

A pure strategy for investor i is a family of functions $\sigma_i = (\sigma_i^t)_{t \geq 1}$ such that each σ_i^t maps private histories $h_i^t = (s_i^1, \ldots, s_i^{t-1})$ to stage game actions in $\{NP\} \cup (P \times \mathbb{R}_+)$. A strategy σ_0 for the agent maps game histories, including the decision of investor i^t to participate in period t, to $\{C, D\}$. For any strategy profile $\sigma = (\sigma_0, \sigma_1, \ldots, \sigma_N)$, let $\mathbb{P}_\sigma$ be the probability measure induced over the set of histories by σ and let $\tilde{H} = \tilde{H}_\sigma$ be the set of on-path histories, or histories that have positive probability according to $\mathbb{P}_\sigma$.

We focus on equilibrium strategies in which the investors trust the agent and the agent cooperates. We say that a strategy profile σ *sustains cooperation* if σ is a perfect Bayesian equilibrium of the game and on the path of play, in all encounters, the selected investor participates and the agent cooperates. Let $\Sigma(G)$ be the set of all strategies σ that sustain cooperation. For any $\sigma \in \Sigma(G)$ and given any history of selected investors $(i^1, \ldots, i^t)$, we can define $\alpha_\sigma(i^1, \ldots, i^t) \in \mathbb{R}_+$ as the investment that player i^t chooses in round t on the path of play.

3 The Optimality of Networks of Complete Components

This section states and discusses our main results.

3.1 Main Results

This subsection presents results showing the optimality of fully cohesive networks. We first show and discuss the Pareto optimality of fully cohesive networks in Theorem 1. We then derive some corollaries and refinements of this result.

A network G has *complete components* if $ik \in G$ whenever $ij, jk \in G$ for some j. Analogously, a network G *has some incomplete component* if there exists i, j, k, all different, such that $ij, jk \in G$ but $ik \notin G$. We will say that G is *κ-regular*, for some $\kappa \in \mathbb{N}$, if $N(i, G) = \kappa$ for all i.

Given G and a strategy profile σ, let $u_i(\sigma, G)$ be the expected sum of discounted payoffs for player i when play evolves according to σ and the social network is G. We will say that G^* *Pareto dominates* G if there exists $\sigma^* \in \Sigma(G^*)$ such that for all $\sigma \in \Sigma(G)$, $u_i(\sigma^*, G^*) \geq u_i(\sigma, G)$ for all $i = 0, 1, \ldots, N$ with at least some strict inequality.

Theorem 1 *Let G^* be a κ-regular network of complete components. Let G be any other network having some incomplete component such that $|N(i, G)| \leq |N(i, G^*)|$ for all $i = 1, \ldots, N$. Then, G^* Pareto dominates G.*

By showing that all game players prefer a network of equally sized complete components over any other network in which no investor has more connections, this theorem provides a concrete answer to the question of what makes fully cohesive communities attractive. At an intuitive level, an incomplete network allows the agent to defect twice incurring a rather small loss in continuation values. The deterrence of those defections imposes bounds on the on-path investment levels. In contrast, in a social network of complete components, there is common knowledge of play within a given component and therefore connected investors can coordinate their punishment leaving no room for off-path gaming.

The proof of Theorem 1 can be sketched as follows.[9] We first derive an upper bound for on-path investments and then show that such upper bound can be attained if and only if all network components are complete. To derive the upper bound, consider first the network of complete components G^*. By using trigger strategies –in which a defection against an investor triggers no further participation by all other investors in the same component– a stationary vector of investments $a \in \mathbb{R}^N$ can be implemented provided

$$g(a_i) \leq \frac{\delta}{1-\delta} \frac{1}{N} \sum_{j \in \bar{N}(i,G^*)} f(a_j) \tag{3.1}$$

for all $i = 1, \ldots, N$. It is relatively simple to show that this system of inequalities has a largest solution $\bar{a}^{G^*} \in \mathbb{R}^N$, which can be used to construct an equilibrium that Pareto dominates any other equilibrium in G^*. At such a largest solution $\bar{a}^{G^*}$, equation (3.1) binds for all $i = 1, \ldots, N$.

An important observation is that if $\sigma \in \Sigma(G)$, on-path investments satisfy an equation similar to (3.1), even when G has incomplete components. To see this, assume that σ implements stationary actions $a \in \mathbb{R}^N$ on the path of play. Since it is in the agent's interest to cooperate in each encounter, defecting against i and cooperating in subsequent encounters with players outside $\bar{N}(i, G)$ cannot be optimal. This yields an inequality similar to that in (3.1) for the on-path action vector a, with G taking the role of G^* in the sum. As a result, there exists a vector $\bar{a}^G$, with $\bar{a}^G \leq \bar{a}^{G^*}$, such that for any $\sigma \in \Sigma(G)$ on-path investments are bounded above by $\bar{a}^G$.

The crucial step in the proof, Lemma 1 stated and discussed below, is to prove that $\bar{a}^G$ cannot be implemented in network G. Theorem 1 then follows since $\bar{a}^{G^*} \geq \bar{a}^G$ and $\bar{a}^{G^*}$ can be implemented in G^*.

Lemma 1 *If G has some incomplete component, then for any $\sigma \in \Sigma(G)$,*

$$\mathbb{P}[\alpha_\sigma(i^1, \ldots, i^t) < \bar{a}^G_{i^t} \text{ for some } t \geq 1] = 1.$$

The idea behind this lemma is the following. Suppose that some equilibrium σ implements

[9] All proofs appear in the Appendix.

$\bar{a}^G$. Let $ij, jk \in G$ but $ik \notin G$. Following a defection against i, there is some positive probability of an encounter with player k in the following round. Since j must punish any defection against either i or k –otherwise the upper bounds $\bar{a}_i^G$ and $\bar{a}_k^G$ could not be attained– investor k loses leverage in the continuation game following the defection against i. Following a defection against i, the agent strictly prefers to defect once more when facing k. In particular, following a defection against i, the agent gets a continuation payoff strictly greater than the payoff he would obtain by cooperating in remaining encounters. The agent therefore prefers to defect when facing i and no equilibrium can attain the upper bound $\bar{a}^G$.[10]

Introducing some additional restrictions on payoffs, we can derive some simple but important corollaries that complement and sharpen Theorem 1.

Assumption 1 *f attains its maximum, g is unbounded.*

While this assumption is mainly technical, it is important for the results we will derive. Under this assumption, we can define

$$\bar{\delta} = \min\left\{\delta \mid \frac{\delta}{1-\delta}\frac{1}{N}\max f \geq g\Big(\min\big(\arg\max f\big)\Big)\right\}$$

which belongs to $]0,1[$. An important consequence of this assumption is that when $\delta > \bar{\delta}$, the i-th component of the upper bound on equilibrium investments, $\bar{a}_i^G$, depends on the network G only through the degree $|\bar{N}(i,G)|$. As shown in Lemma 4 in the Appendix, when $\delta > \bar{\delta}$, the i-th component of the upper bound $\bar{a}^G$ is large enough so that it lies in the flat portion of f and therefore it is entirely determined by the number of connections investor i has.

Corollary 1 *Under Assumption 1, for any $\delta > \bar{\delta}$, any network of complete components G^* and any network G having some incomplete component with $|N(i,G)| \leq |N(i,G^*)|$ for all $i = 1,\ldots,N$, G^* Pareto dominates G.*

In contrast to Theorem 1, this corollary allows us to compare networks that are not necessarily regular, at least for the class of models in which Assumption 1 holds. It follows by noting

[10] Finding a Pareto optimal equilibrium given a network can be thought of as a vector maximization problem subject to equilibrium constraints. Theorem 1 proceeds by relaxing some incentive constraints. Lemma 1 shows that such relaxation is nontrivial in networks of incomplete components; in particular, Lemma 1 shows a family of incentive constraints that are likely to bind in incomplete networks.

that $\bar{a}^G \leq \bar{a}^{G^*}$ and that, as Lemma 1 shows, the upper bound can be attained only in network G^*.

We will say that G^* *gives strictly more total welfare* than G if there exists $\sigma^* \in \Sigma(G^*)$ such that for all $\sigma \in \Sigma(G)$, $\sum_{i=0}^{N} u_i(\sigma^*, G^*) > \sum_{i=0}^{N} u_i(\sigma, G)$.

Corollary 2 *Under Assumption 1, for any $\delta > \bar{\delta}$, any network of complete components G^* and any network G having some incomplete component with $\sum_{i=1}^{N} |N(i,G)| \leq \sum_{i=1}^{N} |N(i,G^*)|$, the following hold:*

i. If $v \circ g^{-1}$ is linear, then G^* gives strictly more total welfare than G

ii. If $v \circ g^{-1}$ is strictly concave and G^* is a network of equally sized (complete) components, then G^* gives strictly more total welfare than G

This result shows that networks of complete components maximize total utilitarian welfare given an exogenously given number of links to be freely allocated among different investors. The main novelty here is the restriction on $v \circ g^{-1}$.[11] To understand its role in the result, note first that under Assumption 1 and for $\delta > \bar{\delta}$, $\bar{a}_i^G$ can be written as $g^{-1}\big(\varphi |\bar{N}(i, G^*)|\big)$ where φ is a parameter determined by f, N and δ. The map from the degree of each investor to his equilibrium utility when the upper bound $\bar{a}^G$ is implemented in the game is given (up to a constant) by $v \circ g^{-1}$. The optimal assignment of links among the different investors depends then on the shape of $v \circ g^{-1}$. When $v \circ g^{-1}$ is linear, networks of complete components are optimal as the upper bound for on-path investments can be attained in those networks. The strict concavity of $v \circ g^{-1}$ is a force towards equally sized neighborhoods which, as shown in Lemma 1, must be complete to attain the upper bound on equilibrium investments.[12]

By identifying complete induced subnetworks, we can actually rank arbitrary graphs at least from the perspective of the players belonging to the subnetwork. Recall that an *induced subnetwork* $G\,|_V$, where $V \subseteq \{1, \ldots, N\}$, is a network among members of V with $ij \in G\,|_V$ iff $ij \in G$ for all $i, j \in V$. We will say that investors in $V \subseteq \{1, \ldots, N\}$ *strictly prefer* G^* over G if

[11] Observe that $v \circ g^{-1}$ is concave provided v is concave and g is convex.

[12] In Appendix C we show that when $v \circ g^{-1}$ is "sufficiently convex", the welfare maximizing network is a star. While the upper bound on equilibrium investments cannot be implemented in a star, we show that the benefit of having an extremely uneven distribution of links outweighs the implementation losses. In particular, Appendix C shows conditions under which network (a) gives strictly more total welfare than networks (b) and (c) in Figure 3.

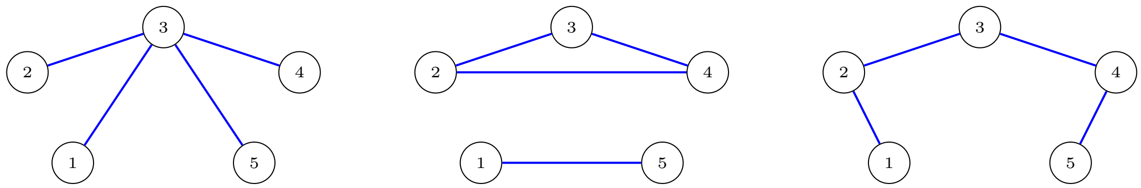

Figure 3: As Corollary 3 shows, among all networks that have 4 links, network (a) is the most preferred network by investor 3. Corollary 1 shows that network (b) Pareto dominates network (c). Corollary 2 shows that network (b) gives strictly more total welfare than networks (a) and (c) when $v \circ g^{-1}$ is linear.

there exists $\sigma^* \in \Sigma(G^*)$ such that for any $\sigma^* \in \Sigma(G^*)$, $u_i(\sigma^*, G^*) \geq u_i(\sigma, G)$ for all $i \in V$ with strict inequality for some $i \in V$.

Corollary 3 *Under Assumption 1, for any $\delta > \bar{\delta}$, any $V \subseteq \{1, \ldots, N\}$, any network G^* such that $G^* |_V$ is complete, and any network G such that $|N(i,G)| \leq |N(i,G^*)|$ for all $i \in V$ and $G |_V$ is incomplete, investors in V strictly prefer G^* over G.*

This result shows that it may be in the interest of a subgroup of investors to push for the formation of a fully cohesive subnetwork, even when the whole social network ends up being not cohesive at all. The crucial observation to derive this corollary is that the upper bound on investments can only be implemented in networks of complete components. Assumption 1 and the restriction to $\delta > \bar{\delta}$ are important, as the upper bound can be implemented in V regardless of whether the connections outside V result in complete components. An immediate implication of this Corollary is that given an exogenous number of links to be assigned among the different investors, the best that an investor can hope for is to be the center of a star. Figure 3 illustrates the corollaries.

3.2 Some Variations

In this subsection, we explore how our results can be extended to accommodate alternative modeling assumptions.

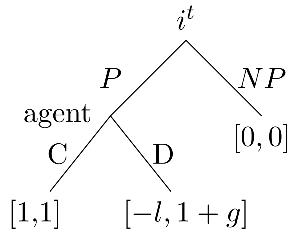

Figure 4: The game between investor i^t and the agent. We assume that $l, g > 0$.

3.2.1 On The Continuous Actions Assumption

We now explore a model in which the scale of the project cannot be chosen by the investor. The only decision that the investor i^t makes is whether or not to participate. If the investor participates, the agent decides whether to cooperate or to defect. Figure 4 illustrates the stage game. The monitoring technology is also determined by a social network G of investors such that if i^t decides to participate in round t, then all of i^t's neighbors become aware of that and observe the behavior of the agent in period t. No further information is transmitted.

We import the definition of strategies that sustain cooperation from Section 2. We say that a strategy profile σ *sustains cooperation* if it is a perfect Bayesian equilibrium and on the path of play all investors participate and the agent cooperates.

Proposition 1 *Let σ be any strategy profile sustaining cooperation in network G. Let G^* be any network of complete component such that $N(i, G) \leq N(i, G^*)$ for all i. Then, there exists a strategy profile σ^* sustaining cooperation in network G^*.*

This result shows that if a strategy profile can sustain cooperation in a network having an incomplete component, then it is possible to construct strategies –indeed trigger strategies– that sustain cooperation in a network of complete components. As in Theorem 1, Proposition 1 shows that forming a network of complete components is the best that game players can hope for. The logic behind both results is the same: in networks of complete components there is less room for off-path gaming. Yet, Proposition 1 is weaker as it is still possible that, given an exogenous number of links, networks G and G^* result in the same equilibrium payoffs.[13]

[13]This model is studied in Balmaceda and Escobar (2011), where conditions are shown under which networks of complete components are (strictly) optimal.

3.2.2 On The Random Matching Assumption

In our model, trading opportunities arise randomly and uniformly. In this subsection, we explore how our results can be extended to allow for alternative matching assumptions.

Suppose first that the $(i^t)_{t\geq 1}$ follows a Markov chain with transition matrix P defined on $\{1,\ldots,N\}$. (We obtain the model in Section 2 by setting $P(i,j) = \frac{1}{N}$.) We assume that $P(i,j) > 0$ for all i,j and assume that i^1 is drawn from the ergodic distribution of P. All the other details of the game remain unchanged. The key to extend the results in Section 3 is to keep the symmetry in the model. In the random matching model, investors are selected with the same frequency and, more subtly, the intensity with which each link is used does not depend on the identity of the linked players. Arbitrary transition matrices can introduce persistence among some subset of investors, implying that the formation of some connections can be more valuable than others as they can be used more intensely. Those asymmetries are ruled out by the following strong restriction.

Assumption 2 *There exists $\lambda > 0$, such that for all $i \neq j$, $P(i,i) = \lambda$ and $P(i,j) = \frac{1-\lambda}{N-1}$.*

This assumption implies that conditional on the investor in $t+1$ being different from the investor selected in $t+1$, the investor in $t+1$ is uniformly selected from $\{1,\ldots,N\}\setminus\{i^t\}$. Under Assumption 2, Theorem 1 applies without modifications.

We can also consider models in which the social network G not only determines information flows but also how trading opportunities arise. For example, suppose that after player i has been selected in round t, each of his neighbors is equally likely to be selected while investors who are not connected to i have a lower probability of being chosen in round $t+1$. More formally, we assume that given G the transition matrix takes the form

$$P(i,j;G) = \begin{cases} p(|\bar{N}(i,G)|) & \text{if } j \in \bar{N}(i,G) \\ q(|\bar{N}(i,G)|) & \text{if } j \notin \bar{N}(i,G) \end{cases}$$

for all i,j, where $p > q > 0$ and for all $n \geq 1$, $np(n) + (N-n)q(n) = 1$. This model has as a special case a quisi-random walk on a network, in which given i^t, i^{t+1} is uniformly drawn from $\bar{N}(i^t,G)$ with probability λ, and uniformly drawn from $\{1,\ldots,N\}$ with probability $(1-\lambda)$. We can derive the transition matrix by setting $p(n) = \frac{1-\lambda}{N} + \frac{\lambda}{n}$ and $q(n) = \frac{1-\lambda}{N}$ for all $n \geq 1$.

In this model, we assume that $\frac{\delta}{1-\delta}[\min_{n=2,\ldots,N} p(n)f(0), \max f] \subseteq \text{range}(g)$ instead of Restriction 3 in Condition 1.

Proposition 2 *Let G^* and G be κ-regular networks. Assume that G^* has complete components, but G has some incomplete component. Then, G^* Pareto dominates G.*

The logic behind this result is similar to that of Theorem 1. The fact that in a social network of complete components it is more likely to keep trading within a neighborhood of a player reinforces the attractiveness of those networks as punishments are even more severe.

4 Application: Contagion-Free Community Enforcement

We now consider a community enforcement model similar to that studied by Kandori (1992). In our model, players are randomly matched to play a prisoners' dilemma game, and a social network determines the monitoring technology. By restricting attention to *contagion-free equilibria*, in which mistakes have local but not global consequences on continuation play, we show how the tools and insights introduced in previous sections can be applied to other models of repeated interaction. In particular, we argue that social networks of complete components are optimal.

4.1 The Game

There are $N \geq 2$ players. At each round $t \geq 1$, two players are randomly chosen to play a game. More precisely, a pair (i^t, j^t) is randomly and uniformly chosen from $\{(i, j) \mid j \neq i\}$ and both i^t and j^t simultaneously decide whether to participate (P) or not (NP). If one of them does not participate, they obtain a per-period payoff equal to 0 and the game moves to round $t + 1$. If both participate, player i^t chooses an action $a \geq 0$ and, after such action is observed by j^t, they play the following stage game:

	C	D
C	$v(a)$	$l(a)$
D	$v(a) + g(a)$	$f(a)$

The game is symmetric and the bimatrix shows the payoffs for the row player. We assume that $0 < l(a) < f(a) < v(a) < v(a) + g(a)$ for all $a \in \mathbb{R}_+$ and therefore the stage game is a prisoners dilemma. As in Section 2, we restrict payoffs so that v and g are increasing, v is bounded, g is unbounded and continuous, and $g(0) \leq \frac{\delta}{1-\delta}\frac{2}{N(N-1)}v(0)$.

There is a social network G of players that determines the monitoring technology. If the match in round t is (i^t, j^t) and one of them does not participate, then all players $k \notin \{i^t, j^t\}$ receive a signal $s_k^t = \emptyset$. If both players participate, the outcome of the game is $o^t = (i^t, j^t, a^t_{i^t,j^t}, x^t_{i^t}, x^t_{j^t}) \in \{1, \ldots, N\}^2 \times \mathbb{R}_+ \times \{C, D\}^2$, then players $l \in N(k)$, for each $k \in \{i^t, j^t\}$, receive a signal $s_l^t = (i^t, j^t, x^t_{-k})$, where $-k \in \{i^t, j^t\} \setminus \{k\}$, indicating the matched players and the action chosen by player k's rival; players $k \notin \bar{N}(i) \cup \bar{N}(j)$ receive a signal $s_k^t = \emptyset$. Players (i^t, j^t) observe play during round t perfectly.

Under this monitoring technology, if players i and j play the prisoners dilemma in round t, but player i defects, then all of j's neighbors will observe that defection. Intuitively, when i defects, he can hide his mischievous action from his connections, players in $N(i)$. We note that none of our qualitative results would be altered if we assumed that a defection by i would be observed by players in $N(i) \cup \bar{N}(j)$.

4.2 Strategies and Contagion-Freeness

A strategy σ_i for player i is a function mapping private histories into stage game strategies. Given a strategy profile $\sigma = (\sigma_i)_{i=1}^N$, we denote by $\tilde{H}_\sigma$ the set of all on-path histories. For an arbitrary history h, we write $h' \succsim_\sigma h$ if h' follows history h when play evolves according to the strategy profile σ. We will say that σ *sustains cooperation* if σ is a perfect Bayesian equilibrium and on the path of play, all players participate and cooperate. Let $\Sigma = \Sigma(G)$ be the set of perfect Bayesian equilibria σ that sustain cooperation.

As it is clear from the definition, if an equilibrium $\sigma \in \Sigma$ sustains cooperation it is entirely possible that after an off-path defection, players optimally defect when facing players in the continuation game. We illustrate such contagion in the following example.

Example 1 (Cooperation and Contagion) *We assume that for all i, $N(i, G) = \emptyset$. The model then becomes closely related to Kandori's (1992) community enforcement and his construction of equilibrium strategies can be adapted to our set up as follows. Fix $\bar{a} \in \mathbb{R}_+$. We* *construct a strategy σ_i^* for each player i by introducing a state variable $\omega_i^t \in \{G, B\}$ evolving according to $\omega_i^1 = G$ and $\omega_i^t = G$ if and only if $\omega_i^{t-1} = G$ and player i's private history shows that play during round $t-1$ had both players participating and cooperating given an action $\bar{a}$. If $\omega_i^t = G$, player i participates, chooses $\bar{a}$ and cooperates (if put in that subgame), while if $\omega_i^t = B$ player i participates, chooses $\bar{a}$ and defects (regardless of what action was actually implemented).*

On the path of play, player i's incentives to cooperate arise as any defection is eventually spread to all the members of the community. That contagion implies that the community breaks down, continuation payoffs are low and thus defections on the path of play can be deterred. Kandori (1992) shows conditions under which the strategy profile $\sigma^* = (\sigma_i^*)_{i=1}^N$ constitutes a sequential equilibrium.[14]

As this example shows, even when players cannot monitor how other players interact and any two players rarely meet each other, a large community can still align players' incentives to attain widespread cooperation as an equilibrium outcome. A problematic aspect of these equilibrium strategies is that they are very fragile as any mistake causes a complete community break down. As Kandori (1992) forcefully argues, when evaluating the merits of a social norm one should also consider its robustness to mistakes. An arrangement in which small trembles have incommensurate impact on continuation play seems inappropriate. Our focus therefore will be on equilibria that are immune to mistakes.

We present two preliminary definitions. Let h' and $\bar{h}'$ be two histories of the same length T. We say that h' and $\bar{h}'$ *coincidentally select players* if for all $\tau \in \{1, \ldots, T\}$ and all $i, j \in \{1, \ldots, N\}$, $(i^\tau, j^\tau) = (i, j)$ in h' if and only if $(i^\tau, j^\tau) = (i, j)$ in $\bar{h}$. We will say that h' and $\bar{h}'$ *have equivalent outcomes at* (i, j), with $i \neq j$, if for all $\tau \in \{1, \ldots, T\}$, for all $a \geq 0$ and all $x, y \in \{C, D\}$, $o^\tau = (i, j, a, x, y)$ in h' iff $o^\tau = (i, j, a, x, y)$ in $\bar{h}'$.

A strategy profile $\sigma \in \Sigma$ is *contagion-free* if for all $h \in \tilde{H}_\sigma$, with $h = (i^1, j^1 a^1, C, C, \ldots, i^T, j^T a^T, C, C)$, any $\bar{h} = (i^1, j^1 a^1, C, C, \ldots, i^{T-1}, j^{T-1}, a^{T-1}, C, C, i^T, j^T a^T, x_{i^T}^T, x_{j^T}^T)$, with $(x_{i^T}^T, x_{j^T}^T) \in \{(C, D), (D, C)\}$ and $d^T \neq c^T \in \{i^T, j^T\}$ such that $x_{d^T}^T = D$, and all continuation histories $h' \succsim_\sigma h$ and $\bar{h}' \succsim_\sigma \bar{h}$, if h' and $\bar{h}'$ coincidentally select investors, then h' and $\bar{h}'$ have equivalent outcomes at $(i, j) \notin \left(\{d^T\} \times \bar{N}(c^T)\right) \cup \left(\bar{N}(c^T) \times \{d^T\}\right)$. Let $\Sigma^{cf} = \Sigma^{cf}(G)$ be the set of all contagion-free strategy profiles.

[14] Off-path incentives are subtle as once a defection occurs a player may want to cooperate in order to slow down contagion. Ellison (1994) shows that introducing a public randomization device, Kandori's (1992) sufficient conditions can be significantly improved.

Contagion-free strategies restrict off-path play following a defection to remain unchanged in matches involving players unaware of the defection. This normative prescription is motivated by the observation that (unmodeled) mistakes may occur, but those mistakes should have a moderate impact on the whole community.[15] While after a single defection play among unaware players remain unchanged, a second defection may unchain a sequence of defections and cause a complete community breakdown. Alternative robustness requirements could be explored too. For example, one may actually want to impose equilibrium robustness to a larger number of mistakes. In particular, while players could make an arbitrary number of mistakes, contagion-free equilibria are robust only to a single mistake. As our Theorem 2 shows, optimal networks are robust to an arbitrary number of mistakes.[16]

4.3 The Optimality of Networks of Complete Components

We say that G^* Pareto dominates G if there exists $\sigma^* \in \Sigma^{cf}(G^*)$ such that for all $\sigma \in \Sigma^{cf}(G)$, $u_i(\sigma^*, G^*) \geq u_i(\sigma, G)$ for all $i = 1, \ldots, N$ with at least some inequality strict.

The following result shows that κ-regular networks Pareto dominate any other network in which players have weakly less connection.

Theorem 2 *Let G^* be a κ-regular network of complete components. Let G be any other network such that $|N(i,G)| \leq |N(i,G^*)|$ for all $i = 1, \ldots, N$. Then, G^* Pareto dominates G.*

To intuitively understand this result, it is perhaps instructive to compare the networks in Figure 5. Suppose that in network (a), players 1 and 2 coordinate their punishment by refusing trade in all rounds following a defection against one of them. The advantage of having players 1 and 2 coordinating their punishments is that the deterrence power of link 12 is used to the maximum and therefore relatively high investments can be implemented. On the other hand, the costs of using link 12 to the maximum is that when player 2 is punishing, say because 4 defected against 1, player 3 looses leverage in the continuation game. In a contagion-free equilibrium, when players 3 and 4 are matched following the defection by 4 against 1, they must be playing

[15] In a contagion-free equilibrium, a player's incentives remain unchanged if he believes there was a single defection. Yet, this is different from belief-free equilibria (Ely and Välimäki, 2002; Ely et al., 2005) as in a contagion-free equilibrium if a player believes two or more defections occurred his behavior could be modified.

[16] In general, the contagion-free restriction has costs in terms of equilibrium payoffs. In Example 1, contagion-freeness implies that cooperation can rely only on the punishment available in the match. Under contagion-strategies, in contrast, punishments are more severe as the whole community breaks down.

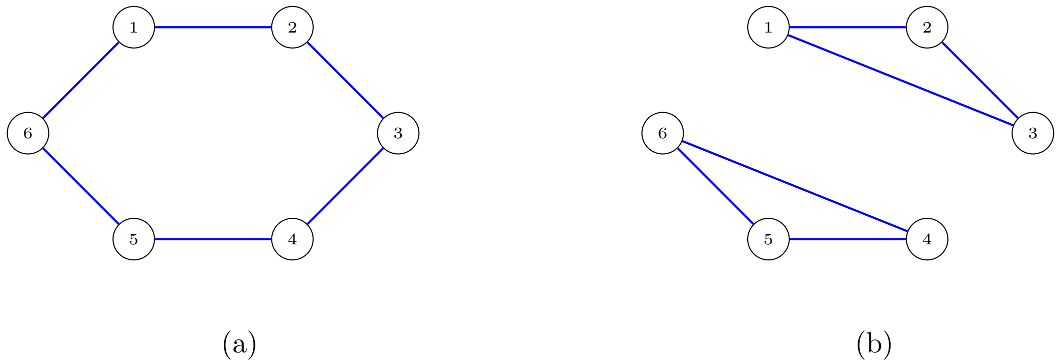

Figure 5: The network on the left is Pareto dominated by the network on the right.

as if no defection against 1 had occurred. In particular, on the path of play, player 2 cannot be called upon to punish a defection by player 4 against 3. In network (a), the cost of having players 1 and 2 coordinating their punishments is that player 2 cannot be called upon to punish a defection against 3. In contrast, in network (b), when link 12 is used to the maximum, player 2 can also be called upon to punish a defection against 3. This is so because it is socially desirable to include player 3 in the punishment following a defection against 1. Therefore, each link in the subnetwork consisting of players 1, 2, and 3 can be used to its maximum extent and on-path investments in network (b) are larger than those that can be implemented in network (a).

As the above discussion suggests, under the optimal equilibrium of network G^* (or network (b) in Figure 5), when one group observes one of its members was cheated, the whole group reacts by refusing trade against the miscreant in all subsequent rounds (even if the miscreant belongs to the same group). Consistent with our focus on contagion-free equilibrium, all other interactions are kept as if no defection had occurred.

Theorem 2 allows us to rationalize the existence of ethnic groups engaging in cooperative relationships. Fearon and Laitin (1996) provide evidence of cooperation between individuals from different ethnicities. They construct a random matching repeated prisoners dilemma in which players belong to cohesive ethnic groups and show conditions for cooperation. Theorem 2 suggests that such cohesive ethnic groups play a pivotal role in maximizing the gains from trade when individuals are randomly matched. The existence of different groups identified by ethnicity can be seen as an optimal societal response to enhance cooperative attitudes in the population.

The proof of Theorem 2 is similar to that of Theorem 1. We derive a family of incentive compatibility inequalities and show that any solution to that system of inequalities is bounded above by the largest solution of the system. The argument is completed by showing that no contagion-free equilibria can implement the largest solution in a network G having some incomplete component, while the largest solution can be implemented in network G^*.

5 Concluding Remarks

This paper studies the emergence of cooperation and trust in a repeated game model in which information flows are governed by the community architecture. We argue that networks of complete components allow players to coordinate their moves, reduce the scope for off-path gaming, and therefore result in high equilibrium payoffs. In contrast, social networks in which some components are not complete leave room for off-path behaviors that are detrimental for incentives and should be avoided.

These basic principles seem new to the literature, show up in a variety of set-ups (one- and two-sided incentive problems, discrete and continuous actions), and provide justification for the commonly discussed optimality of cohesive networks. In our theory, ethnic groups (Fearon and Laitin, 1996) and cohesive business associations (Greif, 1993; McMillan, 1995; Uzzi, 1996) can be seen as providing the required common knowledge to attain high cooperation in trust-based transactions.

It is possible to consider more general monitoring assumptions in our baseline model of Section 2. For example, news could propagate through the network, perhaps with noise.[17] Social networks of complete components need not be optimal in such richer set-up. However, we can obtain Theorem 1 and all the corollaries of Section 3 by restricting attention to contagion-free equilibria, as in Section 4. An interesting question for further research is to make explicit the trade-off between on-path investments and robustness in repeated games of imperfect information on networks.[18]

[17] Alternatively, investors could observe neighbors' investments, and those investments could be used to provide information about more distant interactions.

[18] This question can be addressed either in the model of Section 2 under the assumption that news propagate through the network, or in the community enforcement model of Section 4.

Appendix

This Appendix consists of three parts. Appendix A presents proofs for Section 3. Appendix B presents proofs for Section 4. Finally, Appendix C presents a result that complements Corollary 2. In particular, some conditions are provided under which a star maximizes total welfare.

A Proofs for Section 3

Proof of Theorem 1. The proof is organized as follows. In Lemma 2, we derive necessary conditions for implementability and then, in Lemma 3, we show that those necessary conditions have a larger solution. Lemma 1, showing that in networks having some incomplete component the upper bound cannot be attained, is then proven. We then complete the proof of the Theorem.

Lemma 2 *Let $\sigma \in \Sigma(G)$, $(i^1, \ldots, i^T) \in \{1, \ldots, N\}^T$ be any selection of randomly chosen investors, and a^T be the action chosen by i^T on the path of play. Then,*

$$g(a^T) \leq \sum_{t=T+1}^{\infty} \delta^{t-T} \sum_{j \in \bar{N}(i^T)} \frac{1}{N} \mathbb{E}_\sigma\big[f(a^t_{i^t}) \mid i^1, \ldots, i^T,\ i^t = j\big]. \tag{A.1}$$

To prove this lemma, denote by $V(C)$ (resp. $V(D)$) the continuation value accruing to the agent by cooperating (resp. defecting) against investor i^T at the on path history selecting $(i^1, \ldots, i^T)$. Then,

$$g(a^T) \leq V(C) - V(D).$$

Now, since σ sustains cooperation

$$V(C) = \sum_{t=T+1}^{\infty} \delta^{t-T} \mathbb{E}_\sigma[f(a^t_{i^t}) \mid i^1, \ldots, i^T] = \sum_{t=T+1}^{\infty} \delta^{t-T} \sum_{j=1}^{N} \frac{1}{N} \mathbb{E}_\sigma[f(a^t_{i^t}) \mid i^1, \ldots, i^T, i^t = j].$$

Now, after a defection against i^T, it is still feasible for the agent to cooperate when facing players $i \notin \bar{N}(i^T)$. Thus, following a defection against i^T, the agent can secure a total discounted payoff

$$V(D) \geq \sum_{t=T+1}^{\infty} \delta^{t-T} \sum_{j \notin \bar{N}(i^T)} \frac{1}{N} \mathbb{E}_\sigma[f(a^t_{i^t}) \mid i^1, \ldots, i^T, i^t = j].$$

It follows that

$$
\begin{aligned}
g(a^T) &\le \sum_{t=T+1}^{\infty} \delta^{t-T} \mathbb{E}_\sigma[f(a^t_{i^t}) \mid i^1, \ldots, i^T] - \sum_{t=T+1}^{\infty} \delta^{t-T} \sum_{j \notin \bar{N}(i^T)} \frac{1}{N} \mathbb{E}_\sigma[f(a^t_{i^t}) \mid i^1, \ldots, i^T, i^t = j] \\
&= \sum_{t=T+1}^{\infty} \delta^{t-T} \sum_{j \in \bar{N}(i)} \frac{1}{N} \mathbb{E}_\sigma\big[f(a^t_{i^t}) \mid i^1, \ldots, i^T,\ i^t = j\big].
\end{aligned}
$$

This completes the proof of the lemma.

Consider now the set

$$
\mathcal{A} = \Big\{\alpha \mid \alpha\colon \bigcup_{t=1}^{\infty} \{1, \ldots, N\}^t \to \mathbb{R}_+\Big\}.
$$

For any $\sigma \in \Sigma(G)$, we can generate the same distribution over on path actions by using a particular element $\bar{\alpha} = \bar{\alpha}_\sigma \in \mathcal{A}$. Inspired by this observation and the incentive constraint (A.1), we define $\mathcal{T}^G\colon \mathcal{A} \to \mathcal{A}$ by

$$
\Big(\mathcal{T}^G(\alpha)\Big)(i^1, \ldots, i^T) = g^{-1}\Big(\sum_{t=T+1}^{\infty} \delta^{t-T} \sum_{j \in \bar{N}(i^T)} \frac{1}{N} \mathbb{E}\big[f(\alpha(i^1, \ldots, i^T, i^{T+1}, \ldots, i^t)) \mid i^1, \ldots, i^T,\ i^t = j\big]\Big).
$$

We observe that this quantity is well defined as

$$
\begin{aligned}
\Big(\sum_{t=T+1}^{\infty} \delta^{t-T} \sum_{j \in \bar{N}(i^T)} \frac{1}{N} \mathbb{E}\big[f(\alpha(i^1, \ldots, i^T, i^{T+1}, \ldots, i^t)) \mid i^1, \ldots, i^T,\ i^t = j\big]\Big) &\in \frac{\delta}{1-\delta}[\frac{1}{N}\min f, \max f] \\
&\subseteq \mathrm{range}(g).
\end{aligned}
$$

Lemma 3 *For any network G, there exists $\bar{a}^G \in \mathbb{R}^N_+$ which is the largest solution to the equation $a = \Phi(a, G)$, with*

$$
\Phi_i(a, G) = g^{-1}\Big(\frac{\delta}{1-\delta} \frac{1}{N} \sum_{j \in \bar{N}(i,G)} f(a_j)\Big).
$$

Moreover, any solution $\alpha \in \mathcal{A}$ to the system $\alpha \le \mathcal{T}^G(\alpha)$ satisfies

$$
\alpha(i^1, \ldots, i^T) \le \bar{a}^G_{i^T}
$$

for all $(i^t)_{t=1}^T \in \{1, \ldots, N\}^T$ and for all T. Letting G and G^* be as in the statement of Theorem 1, $\bar{a}^G \le \bar{a}^{G^*}$.

To prove this lemma, take $\bar{A} = \max_{a \geq 0} g^{-1}(\frac{\delta}{1-\delta} f(a))$ and note that for any α, $\mathcal{T}^G(\alpha)(i^1, \ldots, i^T) \leq \bar{A}$ for all $(i^1, \ldots, i^T) \in \{1, \ldots, N\}^T$ and all T. Define $\mathcal{A}^* = \mathcal{A} \cap \{\alpha \mid \alpha(i^1, \ldots, i^T) \leq \bar{A}\}$. In order to find solutions to the system $\alpha \leq \mathcal{T}^G(\alpha)$, it is enough to restrict the domain and range of $\mathcal{T}^G$ to $\mathcal{A}^*$ and, abusing notation, we write $\mathcal{T}^G \colon \mathcal{A}^* \to \mathcal{A}^*$. Let $\alpha^* \in \mathcal{A}^*$ be the largest element in $\mathcal{A}^*$ defined as $\alpha^*(i^1, \ldots, i^T) = \bar{A}$ for all $(i^1, \ldots i^T)$. Since $\mathcal{T}^G$ is increasing, Tarski's fixed point theorem implies the existence of a largest fixed point $\bar{\alpha} \in \mathcal{A}$, which can be computed as the limit point of the sequence $((\mathcal{T}^G)^n(\alpha^*))_{n \geq 1}$. Observing that $(\mathcal{T}^G)^n(\alpha^*)(i^1, \ldots, i^T)$ is only a function of i^T (and not of the vector $(i^1, \ldots, i^{T-1})$), we can actually represent the sequence $((\mathcal{T}^G)^n(\alpha^*)) \subseteq \mathcal{A}^*$ as a sequence of vectors $(\alpha^n)_{n \geq 1} \subseteq \mathbb{R}^N$ with $(\mathcal{T}^G)^n(\alpha^*)(i^1, \ldots, i^T) = \alpha^n_{i^T}$. Denote $\bar{a}^G = \lim_{n \to \infty} \alpha^n$. By construction, the sequence (α^n) satisfies $\alpha^n = \Phi(\alpha^{n-1}, G)$ with α^1 an upper bound for the set of fixed points of $\Phi(\cdot, G)$. As a result, $\bar{a}^G$ is the largest fixed point of the increasing function $\Phi(\cdot, G)$. Since $\bar{a}^G$ actually represents the largest fixed point of $\mathcal{T}^G$, it readily follows that for any $\alpha \leq \mathcal{T}^G(\alpha)$, $\alpha(i^1, \ldots, i^T) \leq \bar{a}^G_{i^T}$ for all $(i^1, \ldots, i^T)$.

To see the last part of the lemma, just note that $\bar{a}^G$ and $\bar{a}^{G^*}$ can be computed by iterative applications of $\Phi(\cdot, G)$ and $\Phi(\cdot, G^*)$. Starting both iterative procedures from a common upper bound $\bar{A}$, denote by $\bar{a}^n$ and $\bar{a^*}^n$ the corresponding sequences. It follows that in each round of iteration, $\bar{a}^n \leq \bar{a^*}^n$. The result follows by passing to the limit.

Having established some properties of the system of incentive constraints, we now prove Lemma 1, restated below.

Lemma 1 If G has some incomplete component, then for any $\sigma \in \Sigma(G)$,

$$\mathbb{P}[\alpha_\sigma(i^1, \ldots, i^t) < \bar{a}^G_{i^t} \text{ for some } t \geq 1] = 1.$$

To prove Lemma 1, let $\sigma \in \Sigma(G)$. Observe that if $\bar{a}^G_i$ is played on-path by player i, for all i, then a defection against some i must entail a loss in continuation values equal to

$$V(C) - V(D) = \frac{\delta}{1-\delta} \frac{1}{N} \sum_{j \in \bar{N}(i,G)} f(\bar{a}^G_j).$$

Since $V(C) = \frac{\delta}{1-\delta} \frac{1}{N} \sum_{j=1}^N f(\bar{a}^G_j)$, we deduce that the continuation value after a defection against i equals

$$V(D) = \frac{\delta}{1-\delta} \frac{1}{N} \sum_{j \notin \bar{N}(i,G)} f(\bar{a}^G_j).$$

Now, decompose the continuation value $V(D)$ following the defection against i as payoffs accruing from encounters with investors $j \in \bar{N}(i,G)$, $\bar{v}$, and payoffs accruing from encounters with investors $j \notin \bar{N}(i,G)$, $\hat{v}$. By definition $V(D) = \bar{v} + \hat{v}$. Since agents' payoffs are nonngative, $\bar{v} \geq 0$. Moreover, following a defection against i, it is feasible for the agent to cooperate in all remaining encounters against $j \notin \bar{N}(i,G)$ and therefore $\hat{v} \geq \frac{\delta}{1-\delta}\frac{1}{N}\sum_{j\notin \bar{N}(i,G)} f(\bar{a}_j^G)$. It follows that $\bar{v} = 0$ and $\hat{v} = \frac{\delta}{1-\delta}\frac{1}{N}\sum_{j\notin \bar{N}(i,G)} f(\bar{a}_j^G)$. Moreover, following a defection against i, continuation play must be such that (i) all investors $j \in \bar{N}(i,G)$ do not participate, and (ii) there is no contagion: investors $j \notin \bar{N}(i,G)$ keep investing according to $\bar{a}_j^G$ and the agent cooperates in all those encounters.

Now, assume that G is not of complete components and let i and k be within distance 2. Consider a defection against i in $t = 1$ and consider the incentives the agent faces if he meets k in $t = 2$. From the previous paragraph, the path of play in encounters $j \notin \bar{N}(i)$ should involve cooperation. Following an argument similar to that in the proof of Lemma 2, we derive the incentive constraint for cooperation when facing k after a defection against i

$$g(\bar{a}_k^G) \leq \frac{\delta}{1-\delta}\frac{1}{N}\Big(\sum_{j\notin \bar{N}(i)} f(\bar{a}_j^G) - \sum_{j\notin \bar{N}(i)\cup\bar{N}(k)} f(\bar{a}_j^G)\Big)$$

Since i and k are within distance 2, it follows that $\sum_{j\notin \bar{N}(i)} f(\bar{a}_j^G) - \sum_{j\notin \bar{N}(i)\cup\bar{N}(k)} f(\bar{a}_j^G) < \sum_{j\in \bar{N}(k)} f(\bar{a}_j^G)$ and therefore $\bar{a}_k^G < \Phi_k(\bar{a}^G)$, yielding a contradiction and concluding the proof.

To finally prove Theorem 1, let $\sigma \in \Sigma(G)$, construct $\alpha_\sigma(i^1,\ldots,i^T)$ as the action implemented by σ on the path of play given the sequence $(i^1,\ldots,i^T)$. Note that $\alpha_\sigma(i^1,\ldots,i^T) \leq \bar{a}_{i^T}^G$ from Lemma 2, where the inequality is strict for some history due to Lemma 1. Noting that $\bar{a}^{G^*}$ can actually be implemented using trigger strategies in G^* and $\bar{a}^G \leq \bar{a}^{G^*}$, the result follows since v is strictly increasing and f is nondecreasing. ■

Proof of Corollary 2. To prove this result, we build upon the arguments presented in the proof of Theorem 1. We first show that under Assumption 1, the upper bound for equilibrium paths assumes a simple form.

Lemma 4 *Under Assumption 1, for all $\delta > \bar{\delta}$, and all G*

$$\bar{a}_i^G = g^{-1}\Big(\frac{\delta}{1-\delta}\frac{|\bar{N}(i,G)|}{N}\bar{f}\Big).$$

where $\bar{f} = \max\{f(a) \mid a \geq 0\}$.

To see why this lemma holds, define $\tilde{a}$ as $\tilde{a}_i = g^{-1}\big(\frac{\delta}{1-\delta}\frac{|\bar{N}(i,G)|}{N}\bar{f}\big)$ and note that $\bar{a}_i^G \leq \tilde{a}_i$ for all i. Since $\tilde{a}_i \geq \min(\arg\max f)$ for all i, $\Phi_i(\tilde{a}) = g^{-1}(\frac{\delta}{1-\delta}\frac{|\bar{N}(i,G)|}{N}\bar{f}) = \tilde{a}_i$. Therefore, for all $\delta > \bar{\delta}$, $\tilde{a}$ is a fixed point of Φ_i. Since $\bar{a}^G$ is the largest fixed point, $\tilde{a}_i \leq \bar{a}_i^G$ for all i and we conclude that $\tilde{a} = \bar{a}^G$. This completes the proof of the Lemma.

Let $\rho = \sum_{i=1}^N |\bar{N}(i, G^*)|$. From Lemma 1, for any $\sigma \in \Sigma(G)$

$$\alpha_\sigma(i^1, \ldots, i^t) \leq g^{-1}\big(\frac{\delta}{1-\delta}\frac{\bar{N}(i^t, G)}{N}\bar{f}\big)$$

with at least some inequality strict. Taking expectations and summing

$$\sum_{i=1}^N u_i(\sigma, G) < \sum_{i=1}^N v \circ g^{-1}\big(\frac{\delta}{1-\delta}\frac{\bar{N}(i^t, G)}{N}\bar{f}\big).$$

Further, since G^* has complete components, we can implement $\bar{a}^{G^*}$ on the path of play. As a result, there exists $\sigma^* \in \Sigma(G^*)$ such that

$$\sum_{i=1}^N u_i(\sigma^*, G^*) = \sum_{i=1}^N v \circ g^{-1}\big(\frac{\delta}{1-\delta}\frac{\bar{N}(i, G^*)}{N}\bar{f}\big).$$

Now, when $v \circ g^{-1}$ is linear, the result is obvious. When $v \circ g^{-1}$ is strictly concave, consider the relaxed problem

$$\max_{(n_1,\ldots,n_N)} \sum_{i=1}^N v \circ g^{-1}\big(\frac{\delta}{1-\delta}\frac{n_i + 1}{N}\bar{f}\big).$$

subject to

$$n_i \in \{1, 2, \ldots, N\}, \; \sum_{i=1}^N n_i \leq 2(\rho - N).$$

This problem can be thought of as the problem of assigning a total of $2(\rho - N)$ pennies between N persons with the purpose of maximizing average utility. Since $v \circ g^{-1}$ is strictly concave and all components of G^* are of the same size $\gamma \geq 2$, this problem has a single solution $n_i^* = \gamma - 1$ for all i. Noting that $(n_i)_{i=1}^N$ and $(n_i^*)_{i=1}^N$ defined as $n_i = |N(i, G)|$ and $n_i^* = |N(i, G^*)| = 2(\rho - N)/N$

are, respectively, feasible and optimal for the maximization above, it follows that

$$\sum_{i=1}^{N} v \circ g^{-1}\big(\frac{\delta}{1-\delta}\frac{\bar{N}(i,G)}{N}\bar{f}\big) \leq \sum_{i=1}^{N} v \circ g^{-1}\big(\frac{\delta}{1-\delta}\frac{\bar{N}(i,G^*)}{N}\bar{f}\big).$$

■

Proof of Corollary 3. Given G^*, construct the following strategy σ_i^* for each investor i. Player $i \in V$ invests $\bar{a}_i^{G^*}$ on the path of play, any defection against some $j \in V$ triggers no further participation; after successive defections against a subset of investors $S \subseteq N(i,G^*) \setminus V$ player i keeps participating and investing an amount $\bar{a}_i^{G^* \setminus S}$.[19] Player $i \in \cup_{j \in V} N(j,G^*) \setminus V$ participates and invests an amount $\eta > 0$ so that $g(\eta) \leq \frac{\delta}{1-\delta}\frac{1}{N}f(\eta)$; any defection observed by i triggers no further participation. Finally, a player $i \notin \cup_{j \in V} N(j,G^*) \setminus V$ participates and chooses action $\tilde{a}$ such that $g(\tilde{a}) \leq \frac{\delta}{1-\delta}\frac{\max f}{N}$, ignores defections against $j \neq i$, and no further participates following a defection against i.

We will show that there is a sequential best reply for the agent, σ_0^*, such that $\sigma^* \in \Sigma(G^*)$. First, note that following a defection against $i \in V$, all players in $\bar{N}(i,G^*)$ do not participate in all subsequent rounds, while players in $\cup_{j \neq i, j \in V} N(j,G^*) \setminus \bar{N}(i,G^*)$ keep choosing η and it is in the interest of the agent to cooperate in all those encounters. A defection against $i \in V$ is deterred, therefore, provided

$$g(\bar{a}_i^{G^*}) \leq \frac{\delta}{1-\delta}\frac{|\bar{N}(i,G^*)|}{N}\bar{f}.$$

which holds with equality. Now, consider a defection against $i \in N(j,G^*) \setminus V$ with $j \in V$. Following a defection against i, i does not participate in all subsequent rounds, players $k \in N(i,G^*) \cap V$ participate and invest $\bar{a}_k^{G^* \setminus \{i\}}$, while players in $\cup_{j \in V} \bar{N}(j,G^*) \setminus \bar{N}(i,G^*)$ keep participating and investing according to $\bar{a}^{G^*}$. A defection against i does not trigger further defections and, since $g(\eta) < \frac{\delta}{1-\delta}\frac{1}{N}\bar{f}$, it is in the interest of the agent to cooperate when facing i. To characterize σ_0^* after further deviations, observe that the construction of strategies $(\sigma_i^*)_{i=1}^N$ ensures that after any number of defections, it is in the interest of the agent to cooperate in remaining encounters. We conclude that $\sigma^* \in \Sigma(G^*)$ implements $\bar{a}_i^{G^*}$ for all $i \in V$ on the path of play.

Finally, if G is such that $G\mid_V$ has some incomplete component, then the argument in Lemma 1 shows that no equilibrium can implement $\bar{a}_i^G$ for all $i \in V$ in network G. The result follows noting that $\bar{a}_i^{G^*} \geq \bar{a}_i^G$ can be implemented for all $i \in V$ in G^*. ■

[19] $G^* \setminus S$ denotes the network in which all links to players in S are deleted.

Proof of Proposition 1. Let a strategy profile σ sustains cooperation in network G. Following a defection, a feasible (but not necessarily optimal) strategy for the agent is to cooperate in remaining encounters and therefore the incentive condition for cooperation on the path of play implies $g \leq \frac{\delta}{1-\delta}\frac{\bar{N}(i,G)}{N}$ for all i. Since network G^* is denser than network G, $g \leq \frac{\delta}{1-\delta}\frac{\bar{N}(i,G^*)}{N}$ for all i. This condition is necessary and sufficient for trigger strategies to sustain cooperation in network G^*. ■

Proof of Proposition 2. We proceed as in the proof of Theorem 1. Denote $P^t(i,j;G) = \mathbb{P}_G[i^t = j \mid i^0 = i]$ and derive a set of necessary conditions for a stationary profile $a \in \mathbb{R}^N$ to be implementable on the path of play

$$g(a_i) \leq \sum_{t=1}^{\infty} \sum_{j \in \bar{N}(i,G)} \delta^t P^t(i,j;G) f(a_j)$$

This system of inequalities has a largest solution $\bar{a}^G$. We claim that $\bar{a}^G \leq \bar{a}^{G^*}$. To derive this result, it will be enough to establish the following lemma.

Lemma 5 *For all $t \geq 1$, all i and all j,*

$$\max_G \{P^t(i,j;G) \mid i \in \bar{N}(j,G),\ G \text{ a } \kappa\text{-regular network } \}$$

is attained at a network G^* such that $\bar{N}(i,G^*) = \bar{N}(j,G^*)$.

We will prove this lemma by induction. The result is obvious if $t = 1$. Assume the result holds for $t \geq 1$, and let us prove it for $t+1$. Take $j \in \bar{N}(i,G)$ and write

$$P^{t+1}(i,j;G) = p(\kappa+1) \sum_{k \in \bar{N}(j,G)} P^t(i,k;G) + q(\kappa+1) \sum_{k \notin \bar{N}(j,G)} P^t(i,k;G).$$

Using the induction hypothesis, we see that to maximize the first sum it is enough to form a network G^* in which $\bar{N}(j,G^*) = \bar{N}(i,G^*)$. This proves the lemma.

Lemma 5 implies that $\bar{a}^G \leq \bar{a}^{G^*}$. Lemma 1 concludes the result. ■

B A Proof for Section 4

Proof of Theorem 2. The proof of this result is similar to that of Theorem 1. We first derive an upper bound for on-path actions and then prove that the upper bound can be attained only in networks of complete components.

Lemma 6 *Let $\sigma \in \Sigma^{cf}(G)$, $(i^1, j^1 \dots, i^T, j^T) \in \{1, \dots, N\}^{2T}$ be any selection of randomly chosen investors, and $a^T_{i^T,j^T}$ be the action chosen by i^T in round T on the path of play. Then,*

$$g(a^T_{i^T,j^T}) \leq \sum_{t=T+1}^{\infty} \delta^{t-T} \sum_{k \in \bar{N}(j^T),\ k \neq i^T} \Big\{ \frac{1}{N(N-1)} \mathbb{E}_\sigma \big[v(a^t_{i^t,j^t}) \mid i^1, j^1 \dots, i^T, j^T, i^t = i^T, j^t = k \big] + \frac{1}{N(N-1)} \mathbb{E}_\sigma \big[v(a^t_{i^t,j^t}) \mid i^1, j^1 \dots, i^T, j^T, i^t = k, j^t = i^T \big] \Big\}$$

and

$$g(a^T_{i^T,j^T}) \leq \sum_{t=T+1}^{\infty} \delta^{t-T} \sum_{k \in \bar{N}(i^T),\ k \neq j^T} \Big\{ \frac{1}{N(N-1)} \mathbb{E}_\sigma \big[v(a^t_{i^t,j^t}) \mid i^1, j^1 \dots, i^T, j^T, i^t = j^T, j^t = k \big] + \frac{1}{N(N-1)} \mathbb{E}_\sigma \big[v(a^t_{i^t,j^t}) \mid i^1, j^1 \dots, i^T, j^T, i^t = k, j^t = j^T \big] \Big\}.$$

To see why this claim holds, consider the incentives player i^T has to cooperate right after action $a^T_{i^T,j^T}$ has been made. The payoff to player i^T from interactions with players outside $\bar{N}(j^T)$ does not depend on whether he cooperates or defects. The only loss player i^T may suffer in continuation values after a defection is the loss in payoffs from players who are neighbors of j^T, and such a loss is bounded above by the right hand side of the first inequality. The second inequality is the incentive constraint for player j^T to cooperate.

Using arguments similar to those in the proof of Theorem 1, we deduce that for any network G, there exists a largest solution $\bar{a}^G \in \mathbb{R}^{N \times N}_+$, of the system of equations

$$a_{ij} = \Phi_{ij}(a), \ \forall i \neq j$$
$$a_{ii} = 0, \forall i$$

with

$$\Phi_{ij}(a) = g^{-1}\Big(\frac{\delta}{1-\delta}\frac{1}{N(N-1)}\min\Big\{\sum_{k\in\bar{N}(j),\ k\neq i}\big(v(a_{ik})+v(a_{ki})\big),\sum_{k\in\bar{N}(i),\ k\neq j}\big(v(a_{jk})+v(a_{kj})\big)\Big\}\Big),\quad i\neq j$$

such that for any $\sigma \in \Sigma^{cf}(G)$ and at any on-path history

$$a^t_{i^t,j^t} \leq \bar{a}^G_{i^t j^t}.$$

Moreover, $\bar{a}^G_{ij} = \bar{a}^G_{ji}$ for all $i \neq j$ and, for G and G^* as in the statement of Theorem 2, $\bar{a}^G \leq \bar{a}^{G^*}$. It is also not hard to see that if there exists i such that $N(i,G) < \kappa$, then $\bar{a}^G < \bar{a}^{G^*}$,[20] and Theorem 2 follows. It is therefore enough to assume that G is κ-regular, in which case $\bar{a}^G = \bar{a}^{G^*}$. Denote by $\alpha_\sigma(i^1, j^1, \ldots, i^t, j^t)$ the action chosen by i^t on the path of play given selected investors $(i^1, j^1, \ldots, i^t, j^t)$.

Lemma 7 *Let G be any κ-regular network having some incomplete component, then for any $\sigma \in \Sigma(G)$,*

$$\mathbb{P}[\alpha_\sigma(i^1, j^1 \ldots, i^t, j^t) < \bar{a}^G_{i^t,j^t} \textit{ for some } t \geq 1] = 1.$$

To prove this result, suppose there exists $\sigma \in \Sigma^{cf}(G)$ implementing $\bar{a}^G$ on the path of play, but G is a κ-regular network that has some component which is not complete. Let i, k be such that $\text{dist}(i, k; G) = 2$ and, since $|N(i)| \geq 2$, we can find a player $j \in N(i)$ such that $N(i) \cap N(k) \setminus \{j\} \neq \emptyset$. Since $\sigma \in \Sigma^{cf}(G)$ implements $\bar{a}^G$ on the play-path, after a defection by j against i in round $t = 1$, all players $l \in \bar{N}(i, G) \setminus \{j\}$ do not participate in all subsequent rounds when facing j. Define $h = (i, j, \bar{a}^G_{ij}, C, C)$, $h' = (i, j, \bar{a}^G_{ij}, C, C, k, j, \bar{a}^G_{kj}, C, C)$, $\bar{h} = (i, j, \bar{a}^G_{ij}, C, D)$, and $\bar{h}' = (i, j, \bar{a}^G_{ij}, C, D, k, j, \bar{a}^G_{kj}, C, D)$. Note that $h \in \tilde{H}_\sigma$ and $h' \succeq_\sigma h$. Since $\bar{h}$ and $\bar{h}'$ coincidentally select investors but they do not have equivalent outcomes at (k, j), $\bar{h}' \not\succeq_\sigma \bar{h}$. We will therefore derive a contradiction by showing that $\bar{h}' \succeq_\sigma \bar{h}$.

Consider agent j's incentives at history $(i, j, \bar{a}^G_{ij}, C, D, k, j, \bar{a}^G_{kj})$ and let V^C and V^D be agent j's total expected discounted sum of payoff after cooperation and defection, respectively. At the largest solution to the system of incentive constraints, all the on-path incentives constraints bind. Since σ is contagion-free, $V^C = v(\bar{a}^G_{jk}) + \frac{\delta}{1-\delta}\frac{2}{N(N-1)}\sum_{l\notin\bar{N}(i),l\neq j} v(\bar{a}^G_{jl})$. After playing D, agent

[20] To see this property, write the fixed point equation characterizing $\bar{a}^G$ as $a = \Phi(a, G)$. Since $N(i, G) < \kappa$ for some i, $\bar{a}^G = \Phi(\bar{a}^G, G) < \Phi(\bar{a}^G, G^*)$ and therefore $\bar{a}^{G^*} = \sup\{a \mid a \leq \Phi(a, G^*)\} > \bar{a}^G$.

j can secure his on-path payoff in all subsequent encounters against agents not in $\bar{N}(i) \cup \bar{N}(k)$. Such strategy is feasible as, after such second defection, all players $l \notin \bar{N}(i) \cup \bar{N}(k)$ will have observed no off path behavior. In particular,

$$V^D \geq v(\bar{a}^G_{jk}) + g(\bar{a}^G_{jk}) + \frac{\delta}{1-\delta}\frac{2}{N(N-1)} \sum_{l \notin \bar{N}(i)\cup\bar{N}(k), l\neq j} v(\bar{a}^G_{jl}).$$

As σ is contagion-free, $V^C \geq V^D$ and therefore

$$\frac{\delta}{1-\delta}\frac{2}{N(N-1)}\Big(\sum_{l \notin \bar{N}(i), l\neq j} v(\bar{a}^G_{jl}) - \sum_{l \notin \bar{N}(i)\cup\bar{N}(k), l\neq j} v(\bar{a}^G_{jl})\Big) \geq g(\bar{a}^G_{jk}).$$

Noting that $N(i) \cap N(k) \setminus \{j\} \neq \emptyset$, it follows that

$$g(\bar{a}^G_{jk}) < \frac{\delta}{1-\delta}\frac{2}{N(N-1)} \sum_{l \in \bar{N}(k),\ l\neq j} v(\bar{a}^G_{jl})$$

and therefore $\bar{a}^G$ is not the largest solution to the system of incentive constraints. This is a contradiction and completes the proof of the Lemma.

To conclude the proof of the Theorem, just note that in a network of complete components G^*, trigger strategies can implement the largest solution $\bar{a}^{G^*}$ and are contagion-free. ■

C Welfare Maximizing Networks

In Section 3 we showed conditions under which networks of complete components maximize total welfare. In this Appendix, we show that when $v \circ g^{-1}$ is sufficiently convex, a star maximizes total welfare. Intuitively, when $v \circ g^{-1}$ is sufficiently convex, the map from the number of links to the utility each player has when the upper bound is attained favors the formation of unequal neighborhoods. While in a star the upper bound cannot be attained, we show that the loss is small compared to the asymmetry gains.

Proposition 3 *Under Assumption 1, let $\delta > \bar{\delta}$, write $\Delta(x) = v \circ g^{-1}(\frac{\delta}{1-\delta}\frac{x+1}{N}\max f) - v \circ g^{-1}(\frac{\delta}{1-\delta}\frac{x}{N}\max f)$ and assume that*

a. $\Delta(x+1) + \Delta(1) \geq 2\Delta(x)$ for all $x \geq 1$; and *b.* $\Delta(x) \geq (x-2)\Delta(1) + \Delta(2) + v \circ g^{-1}(\frac{\delta}{1-\delta}\frac{3}{N}\max f)$ *for all* $x \geq 3$.

Let G^* *be a star formed by* E *links, with* $N-1 \geq E \geq 3$ *and* G *be any network formed by* E *links which is not a star. Then,* G^* *gives strictly higher total welfare than* G.

Proof. Define $u(x) = v \circ g^{-1}(\frac{\delta}{1-\delta}\frac{x+1}{N}\max f)$. Observe first that in the star G^*, the strategy profile σ^* described in the proof of Corollary 3 yields total welfare which is bounded below by $u(E) + (N-1)u(0)$. Now, consider the maximization problem

$$\bar{W} = \max_G \sum_{i=1}^{N} u(|N(i,G)|) \tag{C.1}$$

subject to

$$\sum_{i=1}^{N} |N(i,G)| \leq 2E, \text{ and } |\{i \mid |N(i,G)| \geq 2\}| \geq 2.$$

Since G is a star iff $\{i \mid |N(i,G)| \geq 2\} = 1$, problem (C.1) yields an upper bound for total equilibrium welfare over all networks G different from a star and having less than E links. Let $\bar{G}$ be optimal for (C.1) and define $I = \{i \mid |N(i,\bar{G})| \geq 2\}$. Assume first that $|I| \geq 4$ and let $i^* = \arg\max_i |N(i,\bar{G})|$. Observe that if $ij \in \bar{G}$, then $ji^*, ii^* \in \bar{G}$ as otherwise it would be worthwhile to delete link ij and connect i^* to some l such that $N(l,\bar{G}) = \emptyset$. Indeed, such new network is feasible (as $|I| \geq 4$) and yields strictly higher welfare provided

$$\Big(u(|N(i^*,\bar{G})|+1) + u(1) + u(|N(i,\bar{G})|-1) + u(|N(j,\bar{G})|-1)\Big) \\ -\Big(u(|N(i^*,\bar{G})|) + u(0) + u(|N(i,\bar{G})|) + u(|N(j,\bar{G})|)\Big) \geq 0.$$

Since $\Delta(x) = u(x) - u(x-1)$, this inequality can be written

$$\Delta(|N(i^*,\bar{G})|+1) + \Delta(1) \geq \Delta(|N(i,\bar{G})|) + \Delta(|N(j,\bar{G})|).$$

As Δ is increasing, the inequality holds as $\Delta(x+1)+\Delta(1) \geq 2\Delta(x)$. It then follows that $|I| \leq 3$. Consider first the case $|I| = 3$ and let $I = \{i^*, j_1, j_2\}$. We claim that $N(j_n,\bar{G}) = \{i^*, j_{3-n}\}$ as otherwise, and following the construction above, deleting $j_n l$, with $l \notin I$, to form i^*m, with $N(m,\bar{G}) = \emptyset$, would be worthwhile. If $|I| = 3$, the value of (C.1) equals

$$\bar{W} = u(E-1) + (E-3)u(1) + 2u(2) + (N-E+1)u(0).$$

Now, assume $|I| = 2$ and let $I = \{i^*, j\}$. Observe that $|N(j, \bar{G})| = 2$ as otherwise we could remove one of the nodes connected to j and form a new link to increase the objective function. If $j \in N(i^*, \bar{G})$, then it is relatively easy to see that the objective function at $\bar{G}$ is less than or equal to $u(E-1) + (E-3)u(1) + 2u(2) + (N-E+1)u(0)$ and therefore $j \notin N(i^*, \bar{G})$. Take $l, m \in N(j, \bar{G})$ and note that deleting jm and connecting j to i^*, deleting lj and connecting l so that in the new network the distance between l and i^* equals 2, one can increase the objective. We deduce that $|I| = 3$ and no equilibrium in a network which is not a star can give total payoffs above

$$\bar{W} = u(E-1) + (E-3)u(1) + 2u(2) + (N-E+1)u(0).$$

By assumption $\Delta(E) \geq (E-2)(u(1) - u(0)) + \Delta(2) + u(2)$ and thus

$$u(E) + (N-1)u(0) \geq u(E-1) + (E-3)u(1) + 2u(2) + (N-E+1)u(0)$$

proving the result. ■

References

D. Abreu, D. Pearce, and E. Stacchetti. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 58(5):1041–1063, 1990.

I. Ahn and M. Suominen. Word-of-mouth communication and community enforcement. *International Economic Review*, 42:399–415, 2001.

S.N. Ali and D. Miller. Enforcing cooperation in networked societies. Technical report, UC San Diego, 2013.

F. Balmaceda and J.F. Escobar. Self governance in social networks of information transmission. Technical report, University of Chile, 2011.

J. Bendor and D. Mookherjee. Norms, third-party sanctions, and cooperation. *Journal of Law, Economics, & Organization*, 6(1):33–63, 1990.

L. Bernstein. Opting Out of the Legal System: Extralegal Contractual Relations in the Diamond Industry. *The Journal of Legal Studies*, 21(1):115–157, 1992.

F. Bloch, G. Genicot, and D. Ray. Informal Insurance in Social Networks. *Journal of Economic Theory*, 143(1):36–58, 2008.

M. S.-Y. Chwe. Structure and strategy in collective action 1. *American Journal of Sociology*, 105(1):128–156, 1999.

J.S. Coleman. *Foundations of Social Theory*. Belknap Press, 1990.

A. Dixit. *Lawlessness and Economics. Alternative Modes of Governance*. Oxford University Press, 2006.

G. Ellison. Cooperation in the Prisoner's Dilemma with Anonymous Random Matching. *The Review of Economic Studies*, 61(3):567–588, 1994.

J. Ely and J. Välimäki. A robust folk theorem for the prisoners dilemma. *Journal of Economic Theory*, 102(1):84–105, 2002.

J.C. Ely, J. Hörner, and W. Olszewski. Belief-Free Equilibria in Repeated Games. *Econometrica*, 73(2):377–415, 2005.

I.P. Fainmesser. Community structure and market outcomes: A repeated games-in-networks approach. *American Economic Journal: Microeconomics*, 4(1):32–69, 2012.

J.D. Fearon and D.D. Laitin. Explaining Interethnic Cooperation. *American Political Science Review*, pages 715–735, 1996.

A. Greif. Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition. *The American Economic Review*, 83(3):525–548, 1993.

A. Greif. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge University Press, 2006.

M. Haag and R. Lagunoff. Social norms, local interaction and neighborhood planning. *International Economic Review*, 47(1):265–296, 2006.

J.E. Harrington. Cooperation in a one shot prisoners' dilemma. *Games and Economic Behavior*, 8(2):364–377, 1995.

M.O. Jackson. *Social and Economic Networks*. Princeton University Press, 2008.

M.O. Jackson, T. Rodriguez-Barraquer, and X. Tan. Social capital and social quilts: Network patterns of favor exchange. *The American Economic Review*, 102(5):1857–1897, 2012.

M. Kandori. Social Norms and Community Enforcement. *Review of Economic Studies*, 59(1): 63–80, 1992.

M. Kandori. Introduction to repeated games with private monitoring. *Journal of Economic Theory*, 102(1):1–15, 2002.

D. Karlan, M. Mobius, T. Rosenblat, and A. Szeidl. Trust and Social Collateral. *Quarterly Journal of Economics*, 124(3):13071–361, 2009.

D.M. Kreps. *Corporate Culture and Economic Theory*. Oxford University Press, 1996.

S. Lippert and G. Spagnolo. Networks of relations and word-of-mouth communication. *Games and Economic Behavior*, 72(1):202–217, 2011.

G.J. Mailath and L. Samuelson. *Repeated Games and Reputations*. Oxford University Press, Oxford, 2006.

J. McMillan. Reorganizing vertical supply relationships. *Trends in Business Organization: Do Participation and Cooperation Increase Competitiveness*, pages 203–222, 1995.

M. Mihm, R. Toth, and C. Lang. What goes around comes around: A theory of indirect reciprocity in networks. Technical report, CAE Working Paper, 2009.

P.R. Milgrom, D.C. North, and B.R. Weingast. The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs. *Economics and Politics*, 2(1): 1–23, 1990.

S. Morris. Contagion. *Review of Economic Studies*, 67(1):57–78, 2000.

F. Nava and M. Piccione. Efficiency in repeated games with local interaction and uncertain local monitoring. *Theoretical Economics*, 2013.

M. Okuno-Fujiwara and A. Postlewaite. Social Norms and Random Matching Games. *Games and Economic Behavior*, 9(1):79–109, 1995.

E. Ostrom. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press, 1990.

W. Raub and J. Weesie. Reputation and Efficiency in Social Interactions: An Example of Network Effects. *American Journal of Sociology*, 96(3):626–654, 1990.

J.E. Rauch. Business and social networks in international trade. *Journal of Economic Literature*, pages 1177–1203, 2001.

B. Uzzi. The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect. *American Sociological Review*, 61(4):674–698, 1996.

O.E. Williamson. Transaction-cost economics: The governance of contractual relations. *Journal of Law and Economics*, 22(2):233–261, 1979.

A. Wolitzky. Cooperation with network monitoring. *The Review of Economic Studies*, 80(1): 395–427, 2013.