GOVERNANCE AND THE EFFICIENCY
OF ECONOMIC SYSTEMS
GESY

Discussion Paper No. 28

Networks of Relations

Steffen Lippert*
Giancarlo Spagnolo**

November 2004

*Steffen Lippert, Universität Mannheim and Université Toulouse 1, steffen@econ.uni-mannheim.de
**Giancarlo Spagnolo, Department of Economics, Stockholm School of Economics, C.E.P.R. and Consip Spa., giancarlo.spagnolo@tesoro.it

Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

Sonderforschungsbereich/Transregio 15 · www.gesy.uni-mannheim.de
Universität Mannheim · Freie Universität Berlin · Humboldt-Universität zu Berlin · Ludwig-Maximilians-Universität München
Rheinische Friedrich-Wilhelms-Universität Bonn · Zentrum für Europäische Wirtschaftsforschung Mannheim

Speaker: Prof. Konrad Stahl, Ph.D. · Department of Economics · University of Mannheim · D-68131 Mannheim,
Phone: +49(0621)1812786 · Fax: +49(0621)1812785

Networks of Relations*

Steffen Lippert†
Université Toulouse 1
and Universität Mannheim

Giancarlo Spagnolo‡
Stockholm School of Economics,
C.E.P.R. and Consip

This version: November 03, 2004
First draft: December 2001

Abstract

We model networks of relational (or implicit) contracts, exploring how sanctioning power and equilibrium conditions change under different network configurations and information transmission technologies. In our model, relations are the links, and the value of the network lies in its ability to enforce cooperative agreements that could not be sustained if agents had no access to other network members' sanctioning power and information. We identify conditions for network stability and in-network information transmission as well as conditions under which stable subnetworks inhibit more valuable larger networks.
JEL Codes: L13, L29, D23, D43, O17
Keywords: Networks, Relational Contracts, Indirect Multimarket Contact, Social Capital.

*We have benefited enormously from discussions with and comments from Francis Bloch, Yann Bramoullé, María Paz Espinosa, Sanjeev Goyal, Matthew Jackson, Rachel Kranton, Benny Moldovanu, Patrick Rey, Pekka Sääskilahti, Sudipta Sarangi, Konrad Stahl, from participants of the ESEM 2002 in Venice, the EARIE Annual Meeting 2002 in Madrid, the 2002 Meeting of the German Economic Association in Innsbruck, the 2003 Spring Meeting of Young Economists in Leuven, the Econometric Society's NAWM 2004 in San Diego, the Conference on Networks: Theory and Application in Vaxholm 2004, the 2nd World Congress of the Game Theory Society 2004 in Marseille, the 13th WZB Conference on Markets and Political Economy and 2nd Workshop of the CEPR Research Training Network *Competition Policy in International Markets* on "Collusion and Cartels", as well as from seminar participants at the Universities of Bilbao, Mannheim, and Toulouse. The usual caveat applies. Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged. Spagnolo gratefully acknowledges finance from the European Commission (RTN/CPIM and COCOMBINE). Lippert gratefully acknowledges finance from the European Commission (MCTS and RTN/CPIM).

†Corresponding author. Université Toulouse 1, IDEI, Manufacture des Tabacs, MF 304, 21 Allée de Brienne, 31000 Toulouse, France. Email: steffen.lippert@univ-tlse1.fr.

‡Department of Economics, Stockholm School of Economics, C.E.P.R. and Consip Spa. Email: giancarlo.spagnolo@tesoro.it.

1 Introduction

Relational (or implicit) contracts, long-term informal cooperative arrangements sustained by repeated interaction are an important governance mechanism for most forms of economic and social exchange. When several long-term cooperative relationships link different agents in a group, these agents and their relationships form a *network of relations.* This paper is an attempt to characterize some of their features.

Sociologists have forcefully argued that, by ignoring the networks of social relationships in which economic transactions are "embedded", economists fail to understand important features of the economic process.[1] Like social relations, economic transactions themselves are seldom isolated exchanges. Most often, they are episodes of a history of exchanges of various type, embedded in a network of other economic and social relationships.[2] This is obviously the case for transactions within organizations – from employment to interactions between units and employees – but also for many of those between organizations, in particular supply relations, including financial ones.[3]

Networks of relational arrangements are not only crucial in developing economies, where explicit contracting is hard: in advanced economic environments, and most prominently in the fast changing one of high-tech industries firms often cooperate to share the high risk and return from their activities. In these industries, formal arrangements merely represent the tip of the iceberg "beneath which lies a sea of informal relations" (Powell et al. 1996). On the one hand, lacking contractibility over the main ingredients – investments into human capital and knowledge transfers – explicit contracts can only play a limited role.[4] On the other hand, the need for flexibility linked to the fast changing and highly unpredictable environment make rigid explicit contracts dangerous and vertical integration unattractive. High tech firms therefore often establish informal cooperative agreements with several other firms, and these arrangements link them in a common network of relations.[5]

[1] The work of Coleman (1988, 1990) and Granovetter (1985) is particularly relevant.For example, the latter writes *"The embeddedness argument stresses instead the role of concrete personal relations and structures (or "networks") of such relations in generating trust and discouraging malfeasance"* (1985, p. 490).

[2] Greif (1993) and Casella and Rauch (2002) discusses the importance of ethnic ties for trade in environments where other enforcement mechanisms are ineffective.

[3] Macaulay (1963) first drew attention on the crucial role plaid by relationships in the economic process; Klein and Leffler (1981) have stressed the importance of long term firm-customer relationships; cornerstones of the formal theory of implicit contracts are Bull (1987) and MacLeod and Malcomson (1989); Becker et al (2001), Levin (2003) and Rajo (2003) constitute important recent developments.

[4] The experimental work of Fehr et al. (2004) nicely documents the overwhelming importance of long term relationships for specific economic transactions.

[5] Saxenian (1994) reports a highly specialized network-like organization within the computer-industry in Silicon Valley. She argues that networks of informal cooperative information-sharing relations play a crucial role for the success of the district in comparison with Route 128, a competing district close to Boston. In her words, "While they competed fiercely, Silicon Valley's producers were embedded in, and inseparable from, these social and technical networks." It is noteworthy that the informal relations reported by Saxenian are not only of value on their own, they are of special value due to their being part of a network of such relations between engineers. Examining the

The interbank market can also be seen as a network of long term relationships, where the links that spread contagion among interdependent financial institutions also induce liquid banks to cooperate and privately bail out illiquid ones (see Leitner, 2004). And social networks have been recently show to have a pervasive - and often negative influence on corporate governance practices (e.g. Kramarz and Thesmar, 2004).

In fact, cooperation is often not for the good: corruption, illegal trade (in drugs, arms and people) and organized crime in general can only rely on relational contracts for the governance of illegal transactions, which therefore typically take place within networks of tight relations. Similarly, collusive agreements to increase prices or restrain output are a form of illegal (and common) relational contracts. Multiproduct firms at different levels of the production chain, meeting and cooperating/colluding in different input, geographical and/or product markets where they have different costs or capacities, form a network of relations that may link many apparently distant and unrelated firms, creating pro-collusive *indirect multimarket* contact where no multimarket contact is present.

In this paper we describe equilibrium conditions for different architectures of networks of relations under different informational regimes, paying special attention to differences between circular and non-circular architectures. Most of the dilemmas mentioned earlier, from hold-up situations in specific (legal or illegal) exchanges to cheating in cartel agreements or on public good contributions have the strategic features of a Prisoner's Dilemma game, so our basic model is a repeated game in which each agent interact in generic, asymmetric strategic situations with the structure of repeated Prisoner's Dilemmas and can form links – cooperative relationships – with a small subset of the other agents. In our model, the links are the relationships, the network is directed and the links' orientation captures the presence of net gains from cooperation (slack of enforcing power in the bilateral relation). We consider three informational assumptions: the benchmark case of complete information, where each agent observes the histories of play of all agents; the opposite case where no information can be transmitted from an agent to the others on their observed history of play; and the case where, while agents meet to transact, they can choose to exchange and pass on received information on the respective histories of play. In this last case we assume that time is required for information to travel from one agent to the other, and allow for different speeds of information circulation within the network. We begin by characterizing sustainable networks where agents can only have relations with two neighbors. We show that when relations are asymmetric, since an agent would only cooperate if she receives some incoming arrows, there is a kind of an "end-network effect" (resembling the "end-game effects" of finitely repeated games), and network structures such as trees are not sustainable. Circular networks overcome this problem, ensuring that all agents' defections would be met with punishment, which provide a clear and intuitive explanation to the

biotechnology industry, Powell et al. (1996) point out that the "development of cooperative routines goes beyond simply learning how to maintain a large number of ties. Firms must learn how to transfer knowledge across alliances and locate themselves in those network positions that enable them to keep pace with the most promising scientific or technological developments."

importance attributed by sociologists like Coleman (1988, 1990) to the "closure" of social networks.

We then show that the possibility of transmitting information about defections to other agents in the network is never used in equilibrium if enforcement relies on unrelenting "grim trigger" punishment strategies: when this is the case, once an agent deviates, a contagious process eliminates all prospect of future cooperation in the network, which removes all incentives to transmit information. With "forgiving" punishment strategies agents may instead choose to transmit information to keep on cooperating in the rest of the network while punishing multilaterally one deviator. We also find that under imperfect information and unrelenting punishment strategies, bilaterally enforceable relations between some agents may hinder the stability of larger networks containing these agents because these may not be willing to sacrifice their relation to perform their part in the punishment phase that could sustain the larger network. This problem, though, can also be overcome with the use of relenting punishments. In contrast to results in other literatures (e.g. Kranton, 1996; Spagnolo, 2002), in our model improved outside options, like a more efficient spot market, may under certain conditions foster cooperation by making the breakup of a relation in the case of a deviation a credible threat. Generalizing these results to more complex network architectures where agents may have more than two partners/neighbors, we provide a definition of individual and communities' "social capital". Doing so, we generalize the definition introduced in Spagnolo (1999a, 2000), which are based on Bernheim and Whinston's (1990) multimarket contact paper.

Related Literature. To study networks of relations we borrow from several as yet unrelated literatures, creating a link between them. Besides being related to the mentioned relational contracts literature, this paper contributes to the literature on the emergence and stability of networks. Prominent contributions to this literature - elegantly surveyed in Jackson (2003) - include, among others, Jackson and Wolinsky (1996), who model the emergence and stability of a social networks when agents choose to set up and maintain or destroy costly links using the notion of pairwise stability; Bala and Goyal (2000a) who consider the setup of directed and non-directed links by one agent only; Johnson and Gilles (2000), who introduce a spatial cost structure leading to equilibria of locally complete networks; Bala and Goyal (2000b), who explore the role of communication reliability in networks; and Kranton and Minehart (2000, 2001) who introduce investment and competition after in a buyer-seller network where buyers choose links in a the first stage. Belleflamme and Bloch (2003) model the formation of networks of market-sharing collusive relations between firms. These models focus on agents' decision whether to build and maintain a link or not. The common central question is: Given a value of a network, a sharing rule and the cost of maintaining a link, which networks will emerge in equilibrium, and are they efficient? The underlying game and enforceability problems are left out of consideration.[6] Our approach is complementary. We depart from this literature by explicitly modelling the underlying game, which allows us to study

[6] In a footnote of their introduction, Belleflamme and Bloch write: "In this paper, our focus is on the stability of market sharing agreements, and we assume that these agreements are enforceable. The issue of enforceability of market sharing agreements is an important one, which cannot be answered in traditional models of repeated oligopoly interaction. We leave it for further study." Our work can be seen as a first part of this further study.

the consequences of its features for the stability of network structures; by focussing on the equilibrium sustainability of network structures rather than on the process of network formation; and by showing that the condition for sustainability of each relation of which a network is composed is generally not independent of the network's architecture.

Related work that explicitly models enforcement problems in communities has mostly focused on random matching games. Kandori (1992), Ellison (1994) and others consider repeated random matching prisoner's dilemma games, showing how much cooperation can be sustained under no information transmission between agents. More recently, a similar framework is used by Dixit (2003a) to study the effects of different types of third-party enforcement, and in Dixit (2003b) to analyze the efficiency of relational vs. explicit governance systems when distance among agents differ, inviting in his conclusion to endogenize information transmission. Groh (2002) extends this approach by including an endogenous decision to pass on information to other agents, hence he is closest to our framework. In contrast to this literature, we consider situations where agents with potentially changing opponents establish long-term relationship with fixed partners (e.g. neighbors). This introduces an important forward induction element into strategic behavior when defecting. We keep Groh's endogenous choice whether to pass on information on past actions and introduce the further possibility to pass on informations received by partners in the underlying game.

Our work is probably closest to the simultaneous and independent work of Haag and Lagunoff (2002) and Vega Redondo (2003).[7] Haag and Lagunoff examine a planner's optimal choice of social linkages - or "neighborhood structure" - when each agent plays symmetric repeated prisoner's dilemma games with those other agents selected to be her neighbors, the agents' discount factors differ and are stochastically determined after the planners' choice, information is assumed tom flaw along the links, and agents sustain cooperation by a kind of stationary grim trigger strategies. Among other things, they find a trade off between suboptimal equilibrium punishment (due to imperfect monitoring) and excessive social conflict (linked to heterogeneous discount factors). Our approach is similar in so far that we also look at the effect of different network structures on the maximum level of cooperation sustainable. However, our approaches are very different in most other respects. In their model, as in Kranton and Minehart (2000, 2001), the presence of a link is a pre-condition for interaction hence for a cooperative relation. In our model, instead, the link is the relation and there is no link without cooperation. Moreover, we allow for general asymmetries in payoffs, so that the same agent can be very interested in cooperating with one agent but ready to cheat with the other, and consider in detail the effect of different strategies besides grim trigger. Finally, we endogenize information transmission and characterize the relation with different punishment strategies.

Vega Redondo models the evolution of a social network where social relations are idiosyncratic bilateral repeated prisoners' dilemmas with symmetric payoffs, subject to random shocks.[8] In

[7] We are grateful to Sanjeev Goyal who let us know about these interesting, complementary papers.

[8] See also Jackson and Watts (2002), who analize the process of network formation when agents interact in coordination games.

his model, links are created and destroyed by agents depending on the expected net gains from cooperation; information is assumed to flow across the network one link per period; and enforcement power is transmitted to non sustainable relations. As in Spagnolo (1999a, 2000), "social capital" is defined as the slack enforcement power from cooperative relations that can be used to enforce cooperation in other relations where bilateral cooperation is not sustainable. Vega Redondo is mainly interested in the formation and evolution of social networks. He assumes circulation of information in the network and focuses on symmetric situations and grim trigger strategies. In contrast, we do not deal with network formation and evolution but dig more in depth in terms of sustainability of given network structures, allowing for asymmetries, different punishment strategies and agents' choice of whether to pass or conceal information. Among other things, we show that a network of relations may sustain relations none of which is sustainable if agents rely only on bilateral punishment mechanisms; and that information transmission among agents is not consistent with the use of unforgiving strategies such as "grim trigger" or "Nash-reversion".

Finally, our work is also closely related to the theoretical literature on multimarket contact and collusive behavior sparkled by the seminal work of Bernheim and Whinston (1990). In their model, collusion between two firms is fostered by tying collusive behavior in one market to collusive behavior in the other thereby pooling asymmetries in incentive constraints in the two markets.[9] The closest paper within this strand of literature is probably Maggi (1999), who adapts and extends the multimarket contact framework modelling multilateral self-enforcing international trade agreements. We generalize and extend the work of Bernheim and Whinston by considering imperfect information and endogenous information transmission, and most importantly by showing that agents/firms can easily exploit *indirect multimarket contact* to sustain otherwise unfeasible cartels where absolutely no multimarket contact is present. We generalize and extend Maggi's work by considering generic strategic situations and generic number of agents and relations, and by characterizing the role of different information transmission mechanisms and punishment strategies on networks stability.

We proceed with defining of a network of relations in section 2. In section 3, we derive results for sustainable networks with the restriction of at most two neighbors. We extend these results to situations with more neighbors in section 4. Section 5 concludes.

2 The model

Interaction There is a set $N = \{1, ..., n\}$ of infinitely lived agents $i \in N$ able to interact in pairs according to a connection structure $\mathcal{C}$ of two element subsets of N, where $ij \in \mathcal{C}$, $i, j \in N$, if they are connected. Denote $\mathcal{C}_i$ the set of connections of agent i. In each period t, connected agents play

[9] Spagnolo (1999a) extends the setting to objective functions submodular in payoffs from different markets and shows that multimarket contact may facilitate collusion even in the absence of asymmetries. Matsushima (2001) introduces imperfect monitoring and shows that when firms meet in a sufficient number of markets efficient collusion can be sustained under almost the same conditions as with perfect monitoring.

according to a generic prisoners' dilemma with idiosyncratic payoffs given by the following matrix:

		agent j	
		C^{ji}	D^{ji}
agent i	C^{ij}	$c^{i,j}, c^{j,i}$	$l^{i,j}, w^{j,i}$
	D^{ij}	$w^{i,j}, l^{j,i}$	$d^{i,j}, d^{j,i}$

where $l^{i,j} < d^{i,j} < c^{i,j} < w^{i,j}$ and $l^{i,j} + w^{i,j} < 2c^{i,j}$, $\forall i, j \in N$, $i \neq j$. The stage game is assumed to be constant over time. Note that the assumptions on the payoffs imply the static Nash equilibrium characterized by $\left(D^{ij}, D^{ji}\right)$. One interpretation of agent i's actions C^{ij} and D^{ij} is that agent i is either taking a cooperative action C^{ij} with respect to j, or not taking it, i.e. taking no action at all, D^{ij}.

We can think at C^{ij} as "contributing" to any kind of local public good, "complying" with the terms of any relational agreement, or "colluding"; and to D^{ij} as "don't...". The asymmetric prisoner's dilemma structure captures the essential strategic features of most of the examples discussed in the introduction[10].

Agents are assumed to interact repeatedly. Time is discrete, and all agents are assumed to share a discount factor $\delta < 1$. For simplicity, we assume additive separability of agents' payoffs across interactions and across time[11]. Agents are assumed to choose actions which maximize their discounted utility.

Relations and relational networks In this subsection, we define what we mean by a relation and by a network of relations and give some definitions useful for analyzing these networks. We start by defining a relation:

Definition 1 (Relation) *Given a strategy profile, two agents i and j share a relation if they repeatedly play C^{ij}, C^{ji}.*

Let $R \subset \mathcal{C}$ denote the set of connections between agents who share a relation and $R_i = \{j | \, ij \in R\}$ the set of agents with whom i shares a relation.

For notational convenience, let g^{ij} denote player i's net expected discounted gains from the relation with player j, i.e. the difference between the discounted payoff from playing $\left(C^{ij}, C^{ji}\right)$ forever and defecting and playing the static Nash equilibrium $\left(D^{ij}, D^{ji}\right)$ forever after

$$g^{ij} \equiv c^{i,j} - \delta d^{i,j} - (1 - \delta)\, w^{i,j}.$$

In a standard bilateral repeated game setting both conditions, $g^{ij} \geq 0$ and $g^{ji} \geq 0$, are necessary for a cooperative relation to be sustainable in equilibrium, as the repeated prisoner's dilemma,

[10] Matsushima (2001) shows this in detail for *quantity setting oligopolies,* where firms simultaneously choose either a small amount of supply ("cooperation") or a large amount of supply ("defection").

[11] Removing this (standard) assumption, along the lines of Spagnolo (1999a, 1999b), would complicate the analysis but leave all qualitative results unaffected.

Friedman's (1971) grim trigger (or "unrelenting Nash reversion") strategies are optimal in the sense of Abreu (1988). Note also that if $g^{ij} > 0$ player i does not have an incentive to defect from a cooperative agreement in an infinitely repeated prisoners' dilemma where players use optimal punishment strategies; but $g^{ij} < 0$ does not mean that there is no gain for agent i from cooperation with agent j. It just means that agent i would like to deviate and bilateral cooperation is, therefore, not sustainable. We call a relation of player i with player j deficient for player i if $g^{ij} < 0$ and non-deficient for player i if $g^{ij} \geq 0$.

Definition 2 (mutual, unilateral, bilaterally deficient relation) *The relation ij is called mutual iff $g^{ij} \geq 0$ and $g^{ji} \geq 0$, it is called unilateral iff either $g^{ij} < 0$ and $g^{ji} \geq 0$ or $g^{ij} \geq 0$ and $g^{ji} < 0$, it is called bilaterally deficient iff $g^{ij} < 0$ and $g^{ji} < 0$.*

We are now in the position to define a network of relations.

Definition 3 (Relational network) *A relational network $\mathcal{N}^S = (N, R)$ is a graph consisting of the set of agents N and the set of relations R.*

Definition 4 (Sustainable relational network) *A relational network $\mathcal{N}^S = (N, R)$ is sustainable iff the strategy profile prescribing the relations in R is a sequential equilibrium.*

Definition 5 (Stable sustainable relational network) *A sustainable relational network $\mathcal{N}^S = (N, R)$ is strategically stable if it fulfills Kohlberg and Mertens' (1986) stability criteria.*

Graphical representation A simple way to represent relational networks is graphical, where a line or an arrow is drawn from agent j to agent i if $ij \in R$. This is standard in the literature. We would like to emphasize, however, that *our graphical representation of relational networks departs from the conventional graphical representation* in the networks formation literature. There, an arrow outgoing from a vertex i usually depicts a link sponsored or formed by vertex i. In our graphical representation, on the other hand, the presence of arrows conveys information on the sustainability of relations with optimal bilateral punishments, on each agent's net discounted gains from defecting from a bilateral relation: We depict a relation $ij \in R$ with $g^{ij} > 0$ by an *incoming arrow* to player i.

A unilateral relation, thus, is depicted by an arc originating from the agent for whom the relation in deficient. A mutual relation is depicted by an incoming arc to both players. A bilaterally deficient relation is just a line. If two agents i, j can take action w.r.t. each other, i.e. $ij \in \mathcal{C}$, but do not share a relation, i.e. $ij \notin R$, we depict this by a dotted line. Refer to figure 1: Agents 1 and 2 share a mutual relation, the relation between 2 and 3 is unilateral – it is deficient for player 2 and non-deficient for player 3 – and agents 1 and 3 share a bilaterally deficient relation. Finally, agents 4 and 1 are connected in the sense that $14 \in \mathcal{C}$, however $14 \notin R$, i.e. $4 \in \mathcal{C}_1$ but $4 \notin R_1$.

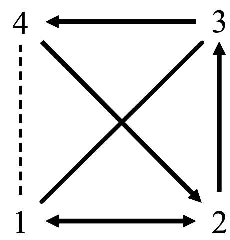

Figure 1: Graphical representation of a network of relations

Definition 6 (mutual, non-mutual, mixed relational network) *A relational network is mutual if it only consists of mutual relations; it is non-mutual if it does not contain mutual relations; and it is mixed if it consists of both, mutual and other relations.*

As we are going to use – to some (limited) extent – graph theoretical language, let us define the used concepts here. In the relational network, agents i and j are called adjacent from/to each other or *directly connected* if $ij \in R$. The set of agents with whom i shares relations are the *neighborhood* of i, denoted by R_i, and $j \in R_i \Leftrightarrow i \in R_j$. Given $\mathcal{N}^S = (N, R)$, the number of agents in N is called the *order* of $\mathcal{N}^S$ and the number of relations in R the *size* of $\mathcal{N}^S$. The number of arcs directed into agent i is called the *indegree* of agent i, denoted by $\operatorname{id} i$. The *degree* of vertex i is the number of edges of agent i, denoted $\deg i$. An agent of degree 1 is called end vertex. The network in figure 1 is of order 4 and size 5, there is no end vertex, and 2 is a vertex with $\deg 2 = 3$ and $\operatorname{id} 2 = 2$. A network is called an $i - j$ *path* if it consists of a finite alternating sequence of agents and links that begins with agent i and ends with agent j, in which each link in the sequence joins the agent that precedes it in the sequence to the agent that follows in the sequence, in which no agent is repeated. An $i - j$ path is called a *cycle* if $i = j$. A cycle of size c is called a $c-$*cycle.*

Information structures We will consider the following three informational assumptions. Let H^{ij} be the set of histories in the relation between agents i and j with $\left(a_t^{ij}, a_t^{ji}\right)_{t=1,\ldots,T} \in H^{ij}$.

(**I1**) *Complete Information:* At time τ, each agent $i \in \mathcal{N}^S$ observes

- $(a_t^{mn})_{t=1,\ldots,\tau} \in H^{mn}\ \forall m, n \in \mathcal{N}^S$.

Each agent observes the history of play of all other agents.

(**I2**) *No Information Transmission:* At time τ, each agent $i \in \mathcal{N}^S$ observes

- $\left(a_t^{ij}, a_t^{ji}\right)_{t=1,\ldots,\tau} \in H^{ij}\ \forall j \in R_i$.

Each agent only observes the history of (his own and) his direct opponents' play.

(**I3**) *Network Information Transmission:* At time τ, each agent $i \in \mathcal{N}^S$ observes

- $\left(a_t^{ij}, a_t^{ji}\right)_{t=1,\dots,\tau} \in H^{ij}\ \forall j \in R_i$ and
- $\left(a_t^{mn}, a_t^{nm}\right)_{t=1,\dots,int\left[\tau-\frac{l}{v}\right]} \in H^{mn},\ m \in R_n$, where $\min\left[l\left(i,m\right), l\left(i,n\right)\right] = l$ if there exists an $i-m$ path and if every agent on that path is *willing* to transmit information on their own history as well as messages received.

Under the *Network Information Transmission* mechanism, (**I3**), besides observing the history of his direct opponents' play, in each period each agent i can transmit and receive truthful messages - pieces of hard information - to/from each agent $j \in R_i$ about the histories of play and about messages they received. A message on past behavior can travel over v links per period. We assume that agents only meet when they cooperate, hence information can only be transmitted through existing cooperative relations/links.

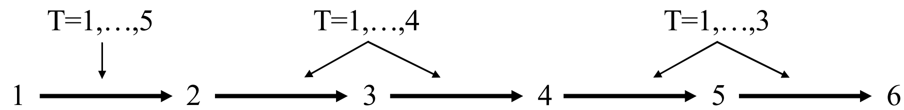

Figure 2: Agent 1's possible "observations"

For an illustration of the three informational assumptions, consider a non-circular network with 6 agents, call them agent 1 through 6, as in figure 2. Suppose first agents use the Network Information Transmission mechanism (**I3**), and let $v = 2$. Then in $t = 5$, agent 1 observes the full history of his own play starting at $t = 1$ through $t = 5$. Furthermore, he will receive messages from agent 2 about the play between 2 and 3 and thus "observe" actions $\left(a_t^{2,3}, a_t^{3,2}\right)_{t=1,\dots,4}$. The messages from 2 will also contain his received messages and thus agent 1 will "observe" actions $\left(a_t^{3,4}, a_t^{4,3}\right)_{t=1,\dots,3}$, $\left(a_t^{4,5}, a_t^{5,4}\right)_{t=1,\dots,3}$, and so on. Consider now the *Complete Information* case (**I1**). Each agent immediately knows everything that happened between every other two players, that is for example between agents 5 and 6 or between agents 2 and 3. This is of course also a degenerate form of Network Information Transmission mechanism (**I3**) where $\nu \to \infty$. With No Information Transmission (**I2**) each agent only knows the history of his own play, that is agent 1 only knows what happened between agents 1 and 2. This is also an extreme case of the network Information Transmission mechanism (**I3**) where $\nu = 0$.

Our information transmission mechanisms relate to the literature on perfect, public, and private monitoring in the following way. *Complete Information* (**I1**) implies perfect monitoring. *No Information Transmission* (**I2**) implies perfect monitoring for agents i and j on their bilateral history of play, but private monitoring for the same agents on the history of play of other agents and of their neighbors with other neighbors[12]. With the *Network Information Transmission* mechanism (**I3**), a

[12] See Mailath and Morris (1999) for an example of private monitoring where the private signal about the other players' actions is imperfect.

temporal modification of (**I2**) is assumed. Again, refer to figure 2 and let $v = 2$ and $t = 5$. There is perfect monitoring for all actions that happened more than 3 periods ago. Actions between agents 5 and 6 from period 4 are assumed to be private w.r.t. agent 1. They are perfectly monitored by agents 6, 5, 4, and 3. The network information transmission regime introduces therefore a space-time neighborhood structure into relational networks, in the sense that perfect monitoring may travel through the network with time. Note also that there is no public monitoring in any of our information structures[13].

There are many situations in which there does not exist an institution that gathers and disseminates immediately information on the behavior of network members, as assumed implicitly in the complete information case (**I1**). In the network information transmission regime (**I3**) we thus suppose that information can only be transmitted through personal contact of members of the network, and that transmission takes time. We assume that in networks of relations communicating besides interacting is not costly. This, we think, is a reasonable assumption since we have in mind chatting while carrying out one's daily business. We will see that an essential feature of this information structure is that, even though information transmission is costless in itself, agents must be given incentives to actually transmit information. Even with high speeds of information transmission, agents may prefer not to transmit information but rather deviate from their relations to reap short run deviation profits, in which case the potential higher speed of information transmission does not realize nor does it affect the sustainability of the network.

Specificity We assume fully specific relations, i.e. such that if a relation between two agents breaks down, these agents cannot substitute it with relations with other agents (i.e., it is not possible for an agent to substitute a partner with another one)[14]. Little changes (apart from notation) if agents are assumed imperfect substitutes, in the sense that a relation with an agent can be replaced at a finite but high cost with a relation with another agent.

Assumption 1 *We restrict our attention to relational networks (equilibria)* $\mathcal{N}^S = (N, R)$*, with* $R = \mathcal{C}$*.*

We allow for costless substitution in a related paper (Lippert and Spagnolo 2004), so that punishment through exclusion/replacement becomes an option, and find that the results of the present paper continue to apply: relational networks where defecting agents are excluded and the relations shared with them replaced by relations between the defecting agent's former neighbors are either not sustainable, or not strategically stable in the sense of Kohlberg and Mertens (1996).

[13] For an example of public monitoring, see Green and Porter (1984). They assume that players observe their own actions, but only an imperfect public signal about the actions of the other players.

[14] The most obvious examples of such situations are networks where a geography limits the set of potential partners of each agent, or where agents perform different functions (e.g. they supply different goods-services).

3 Analysis

Most insights can be gained by examining networks with a restricted number of neighbors. For the time being, therefore, we simplify the analysis by focussing on networks with nodes of a maximal *degree of two*, i.e. where each agent can have at most two neighbors[15].

Assumption 2 $\deg i \leq 2$.

In section 4, we will discuss how the results generalize to more complex networks.

3.1 Non-mutual networks

Mutual relations can be sustained by direct bilateral punishments, so if all relations are mutual, a network cannot improve on what agents can sustain bilaterally. A relational network plays a role when it allows to sustain unilateral or bilaterally deficient relations, i.e. relations that would not be sustainable in the absence of a network. In this section, we explore how relational networks can be sustainable even if they do not contain any relation sustainable in the absence of such a network (assumption 3). We will show how the network's ability to pool payoff asymmetry and redistribute sanctioning power and information improves on what agents could achieve through bilateral interaction.

Assumption 3 *Relational networks do not contain mutual relations.*

Let us start with a necessary condition for multilateral punishment mechanisms in a relational network:

Lemma 1 $\mathrm{id}\, i \geq 1\ \forall i \in R$ *is a necessary condition for a relational network to be sustainable.*

Proof. Suppose $i \in R$ and $\mathrm{id}\, i = 0$. Then $g^{ij} \leq 0\ \forall j \in R_i$ and i had an incentive to deviate from all her relations. ■

This is a straightforward generalization of the sustainability condition for a bilateral relational contract: For each contracting party, the net gain from cooperating has to be non-negative. The following proposition follows immediately:

Proposition 1 *End-network effect: The only sustainable non-mutual non-circular relational network is the empty one (independent of the discount factor and the information structure).*

As long as relations are not mutual, they are not sustainable by a multilateral mechanism within a non-circular network. Figure 3 illustrates this: Part (a) shows a network that is not sustainable. In that situation, agent 1 always has an incentive to deviate and the only sustainable network is empty, as shown in (b).

[15] This assumption may represent a time constraint: It is always possible not to take an action w.r.t. someone you are connected to, however, it takes time to indeed take a *cooperative* action.

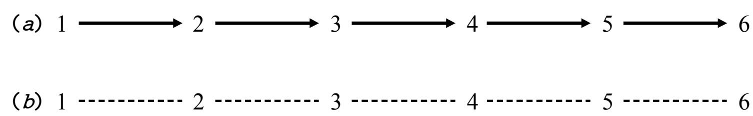

Figure 3: Only the empty network (b) is sustainable

Proposition 1 highlights an end-network effect much similar to the end-game effect of standard finite games and rather general. Relaxing assumption 2 but keeping assumption 1, it is straightforward to see that this effect generalizes to *trees* (see figure 4 for an intuition), *stars* and any other network forms where there are "end vertices" that have only outgoing arrows.

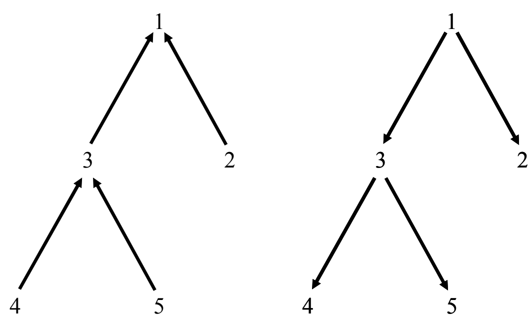

Figure 4: Trees

One way to ensure that the necessary condition from lemma 1 is satisfied in a non-mutual network is to "close" the network. If agents 1 and 6 from figure 3 shared a unilateral relation that is non-deficient for 1, as in Figure 5, then each agent in the network would have an incoming and an outgoing arrow, so that a multilateral punishment mechanism may exploit payoff asymmetries.

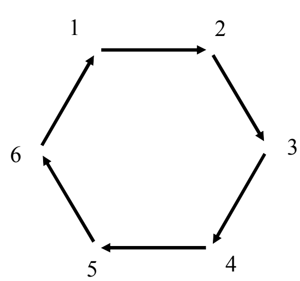

Figure 5: Circular unilateral network

To capture this effect, we define below the unrelenting strategy profiles (**S1**) for the *complete information* case (**I1**), and (**S2**) for both, the *no information transmission,* (**I2**) and the *network information transmission* case (**I3**). These strategy profiles can be thought of as a network versions of Friedman's (1971) "grim trigger" strategies.

Strategy profile (S1)

1. Each agent $i \in \mathcal{N}^S$ starts playing the agreed upon action vector C^{ij} $\forall i \in \mathcal{N}^{\mathcal{S}}, \forall j \in R_i$.

2. Each player i goes on playing C^{ij} $\forall j \in R_i$ as long as no deviation by any player in the network is observed.

3. Every agent i reverts to D^{ij} $\forall j \in R_i$ for ever otherwise.

Strategy and belief profile (**S2**) is a straightforward adaptation of the grim trigger like strategies (**S1**) to an environment without full information.

Strategy and belief profile (S2)

1. Each agent $i \in \mathcal{N}^S$ starts playing the agreed upon action vector C^{ij} $\forall i \in \mathcal{N}^{\mathcal{S}}, \forall j \in R_i$.

2. As long as player i observes every neighbor $j \in R_i$ play C^{ji} she goes on playing C^{ij} $\forall j \in R_j$.

3. If player i observes a neighbor j play D^{ji} in $t = \tau$ she reverts to D^{ij} $\forall j \in R_i$ $\forall t \geq \tau + 1$.

The beliefs players have after observing their neighbors – which we define formally in the appendix – are such that (*i*) and (*iv*) they believe that everybody in the network cooperated if they observe cooperation on both sides, (*ii*) and (*v*) they believe "anything" consistent if they observe cheating from a neighbor whose net gain from cooperating with them is positive, and (*iii*) and (*vi*) they assign an equal probability to the event that any of the other players was the first to deviate in case they observe Cooperate from their neighbor with a positive net gain from cooperation with them and Defect from the neighbor with a negative net gain from cooperating with them. As for parts (*iii*) and (*vi*) of the belief structure, a priori a player does not know anything else about any other player than that they are all symmetric w.r.t. their incentives to deviate in their respective bilateral relations. Then the observation that only one neighbor deviated does not provide any further knowledge. Following Bernoulli's "Principle of Insufficient Reason", we, therefore, assume that he assigns an equal probability of any of the other players to have been the first to deviate from ($S2$) point 1. This assumption in part (*iii*) of the belief structure is innocent as this observation is part of a dominated deviation.

We can then state the following.

Proposition 2 *If the relational network is a c-cycle and agents use unforgiving strategies, then:*

1. *under complete information* (**I1**)*, a non-mutual relational network is sustainable if and only if* $\forall i \in \mathcal{N}^S$ $g^{i,i-1} + g^{i,i+1} > 0$*;*

2. *under no information transmission* (**I2**)*, a non-mutual relational network is sustainable if and only if* $\forall i \in \mathcal{N}^S$ $\delta^{c-2} g^{i,i-1} + g^{i,i+1} > 0$*; and*

3. *under the network information transmission regime* (**I3**)*, a non-mutual relational network is sustainable if* $\forall i \in \mathcal{N}^S \;\; \delta^{c-2} g^{i,i-1} + g^{i,i+1} > 0$*, regardless of the speed of information transmission.*

For the proof of proposition 2, refer to figure 6. Also note that in a non-mutual network sustained by the above strategies and beliefs, unless there is perfect information the agents' optimal deviation is defecting immediately from deficient relations; and it is to postpone defections from non-deficient relations to the period before the punishment from that neighbor is expected to start.

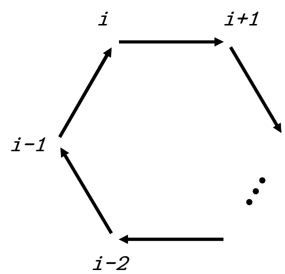

Figure 6: Circular unilateral network

Proof. *Part 1* of proposition 2: *Sufficiency*: Consider (**S1**). Since a deviator faces immediate Nash-reversion from both his neighbors, it is optimal to deviate on both neighbors, and the circular network is a Nash-Equilibrium if $\forall i \; g^{i,i-1} + g^{i,i+1} > 0$. In the punishment phase, the stage Nash equilibrium is played and therefore a best response. *Necessity*: Since during the punishment phase the agents play their minimax strategy, the punishment phase is infinitely long, and it starts immediately, this is the strongest punishment available to the agents. If cooperation is not possible with these strategies, it will not be possible with other ones.

Part 2 of proposition 2: *Sufficiency*: Consider (**S2**). The optimal deviation for an agent i is now first deviating on the deficient relation, that is from his relation with $i+1$, and as late as possible – since deviating from a bilaterally non-deficient relation is a cost – from his other relation. The second deviation should take place after $c-2$ periods. Therefore deviation will not be profitable if

$$\delta^{c-2} g^{i,i-1} + g^{i,i+1} \geq 0 \qquad \forall i \in \mathcal{N}^{\mathcal{S}} \text{ and } \{i-1, i+1\} = R_i.$$

Since every agent i in the network would want to deviate bilaterally from his relation with $i+1$, was it not for the threat of the loss of cooperation in her other relation, after losing this other relation for ever, "infecting" is optimal. This is true for *any* belief about the history of the game. *Necessity*: Since during the punishment phase the agents play their minimax strategy and the punishment phase is infinitely long, this is the strongest punishment available to the agents. As there is no possibility to transmit information on past behavior, it is also not possible to enter a punishment phase on both sides with a faster speed than one agent per period. If cooperation is not possible with these strategies, it will not be possible with other – less strong – punishments.

Part 3 of proposition 2: Assume the network information transmission regime (**I3**) and unforgiving strategies. Suppose agent i observes a deviation of his neighbor $i-1$ in his ($i-1$'s) deficient relation. Then, since, due to the unforgiving strategies, there will never be a return to cooperation with $i-1$, the best response of i in his (i's) remaining deficient relation would be to deviate from that relation. Therefore agent i will not make use of her ability to transmit information, leaving only room for the same strategies as under (**I2**).*Q.E.D.* ■

As we see from part 3 of proposition 2, an important feature of our model is that the design of the punishment paths interacts with agents' incentives to transmit information. One implication of this is that even though grim trigger strategies are optimal punishment strategies in all the bilateral relations (i.e. if they rely on bilateral punishment mechanisms), for non-mutual relational networks, the grim trigger-like strategies (**S1**) and (**S2**) are only *optimal* punishment strategies for the *complete information* (**I1**) and the *no information transmission* case (**I2**), respectively. They are optimal because punishment is as strong as possible on both sides, once it arrives there, *and* it arrives on both sides with the smallest possible delay. Under the *network information transmission* regime (**I3**) with high speeds of information transmission instead, i.e. in a world where information can be transmitted via links and this information travels more than one link per period, strategies (**S2**) are *not optimal* anymore. The potentially high speed of information transmission is – individually optimally – not being used, and therefore, punishment "on the other side" arrives later than necessary, reducing the enforcement power of the network. In section 3.2.3, we will introduce a forgiving punishment mechanism that uses information transmission and that we will show to be optimal.

A short comment on the circular form of the network is due. Even though proposition 2 is a statement on a particular network architecture, a c-cycle, of course this circular network could be embedded into bigger networks. The strategy profiles (**S1**) or (**S2**) we studied would not conflict with that. Our implicit assumption by concentrating on a c-cycle – if it is embedded into a bigger network – is that the multilateral punishment mechanism (**S1**) or (**S2**) is taken for that particular subnetwork only.

To give an example for circular networks (or subnetworks), one could think of firms located on a (Salop) circle, with different capacities in the left and right market, cooperating/colluding with their neighbors. Coleman (1990) insists on the importance of the "closure" (circularity) of social networks. Giving a graphical representation as in figure 7[16], he suggests that if parents (A and B), whose children (a and b) are friends, share a relation, too, as in figure 7 part (a), they have more "power" over their children – thanks to what Coleman calls "intergenerational closure" – than if they do not, as in figure 7 part (b). Lack of relations among parents makes it more difficult for them to successfully impose/enforce norms on/upon their children. He does not provide a game theoretical foundation for his claim, but our model fits precisely his story.

[16] Note that his representation differs from ours by using two arrows to describe one relation.

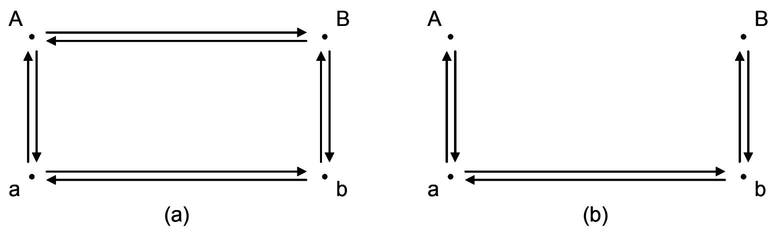

Figure 7: Representation of two communities: (a) with and (b) without intergenerational closure (from Coleman, 1990).

3.2 Mixed networks

In this section, we will relax assumption 3 that excluded mutual relation from the relational networks under consideration. We will explicitly allow for them (assumption 4), and study their impact on the sustainability and of the various types of relational networks. We proceed examining the consequences of an increase of the stage game payoff $c^{i\cdot}$ of an agent i such that one of his relations becomes a mutual one. Increasing the cooperation payoff $c^{i\cdot}$ of an agent i increases both, the *profitability* of cooperating for this agent as well as the *sustainability* of the relative relation with a bilateral punishment mechanism.

After demonstrating a cooperation-enhancing effect for non-circular networks under information structure (**I1**), we will show that a circular network's ability to pool payoff asymmetry and redistribute sanctioning power under information structures (**I2**) and (**I3**) decreases if the unforgiving punishments from section 3.1 are used. When the increase in $c^{i\cdot}$ transforms a non-mutual relation into a mutual one, agent i may lose the incentive to exercise the multilateral punishment strategy, which sustained the network and thus the other bilaterally non-sustainable relations in the network. Subsequently, we will show that forgiving strategies overcome the problem for information structure (**I3**).

Assumption 4 *Relational networks contain both, mutual and other relations.*

3.2.1 Non-circular networks with unforgiving punishments

Proposition 1 states that there does not exist a non-circular non-mutual network other than the empty one. This is true because there would be an agent having only deficient relations and, thus, just an incentive to deviate. If one increases the cooperation payoff $c^{\cdot\cdot}$ of that agent, so that his relation becomes mutual, this incentive to deviate of vanishes. Under *full information* a multilateral punishment like (**S1**) can then sustain such a network. Part 1 of proposition 3 states that. Part 2 shows that the negative result of Proposition 1 remains for the other information transmission mechanisms. And Part 3 shows that the equilibrium in Part 1 does not satisfy reasonable stability criteria put forward by Kohlberg and Mertens (1986). In particular, the equilibrium (**S1**) does

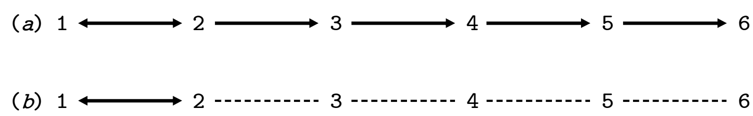

Figure 8: Sustainable networks under (a) info structure (**I1**), (b) info structure (**I2**) and (**I3**)

not satisfy their *Iterated Dominance* and *Admissibility* criteria and gives thus rise to a forward induction problem.

Proposition 3 *Suppose* $\deg i \leq 2$*. Then*

1. *under information structure* (**I1**)*, a non-circular relational network* $\mathcal{N}^{\mathcal{S}}$ *is sustainable if*

   (a) $\mathrm{id}\, i\,|_{\deg i=1} = 1$ *and*

   (b) for all other agents in the relational network $g^{i,i-1} + g^{i,i+1} > 0$*, and*

2. *under information structures* (**I2**) *and* (**I3**)*, there exists no sustainable non-circular mixed relational network.*

3. *If the relational network under* (**I1**) *relies on unforgiving punishments, it is not strategically stable.*

Proof. Parts 1 and 2 of proposition 3 are straightforward. Part 3 of proposition 3: *Unforgiving* punishment in our framework means to play according to (**S1**), i.e. to play D on both sides *forever* if a deviation occurred in the network. Ruling out the play of strictly dominated strategies gives rise to a profitable deviation for each agent i of the mutual subnetwork who is also part of a non-mutual subnetwork. Let agent 2 in figure 8 (a) play $D^{2,3}$ and $C^{2,1}$ in a period t. Then reverting to $D^{2,3}$ and $D^{2,1}$ for ever in $t+1$ is part of a strictly dominated strategy for 2. It is strictly dominated by $D^{2,3}$ and $D^{2,1}$ in a period t and reverting to $D^{2,3}$ and $D^{2,1}$ for ever in $t+1$. Thus, if agent 1 observes $D^{2,3}$ and $C^{2,1}$ in t, he can conclude that a rational agent 2 does not want to stick to the multilateral punishment mechanism. Given that 2 played $C^{2,1}$, there exists a focal equilibrium. This focal equilibrium is to switch to a bilateral punishment mechanism, the normal grim trigger strategy. The resulting – stable – equilibrium is the same as the one under ($I2$) and ($I3$), sketched in figure 8 (b). This gives rise to a profitable deviation for agent 2. *Q.E.D.* ■

Figure 8 illustrates proposition 3. Under the *full information* assumption (**I1**), every agent knows the history of every other player and can, thus, enter into a punishment phase. Given this, figure 8 (a) is an equilibrium. Under the other information transmission mechanisms, this is not the case, figure 8 (a) is not an equilibrium network, while figure 8 (b) is.

The sustainability of 1's relation in the absence of a network enables cooperation in the network. However, according proposition 3 the resulting network under (**I1**) is not strategically stable. The

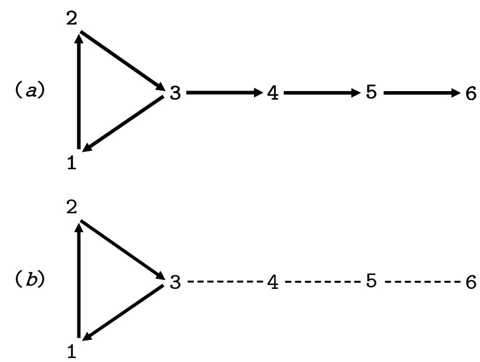

Figure 9: Non-circular network with a (possibly) sustainable subnetwork at one end.

mutual interest in cooperation, which made cooperation of all agents in the non-circular network an equilibrium, puts it on weak feet as it makes it unlikely to be selected as the equilibrium played.

Relaxing assumption 2 ($\deg i \leq 2$), it is straightforward to see that Proposition 3 generalizes to lines that are adjacent to subnetworks which are sustainable in autarky. Assume in figure 9 that the subnetwork $(\{1,2,3\},\{12,23,31\})$ is sustainable in autarky, i.e. without making use of possible relations $34, 45, 56$. Then, under (**I1**), strategies (**S1**) make (a) a sustainable network if, in addition, $g^{31}+g^{32}+g^{34}\geq 0$, $g^{43}+g^{45}\geq 0$, and $g^{54}+g^{56}\geq 0$, whereas network (b) is the only sustainable one under (**I2**) and (**I3**), irrespective of the payoffs in the relations 34, 45, and 56.

Remark. *All statements made on mutual relations also apply to subnetworks that are sustainable in the absence of the rest of the network.*

3.2.2 Mixed circular networks with unforgiving punishments

We now turn to circular networks. We will proceed in the same way we did in subsection 3.2.1: Again, we will increase the cooperation payoff $c^{i,\cdot}$ of an agent i's deficient unilateral relation such that it becomes mutual. As in subsection 3.2.1, we will discuss the impact of this change on the sustainability of a network.

Under *full information*, (**I1**), we will retain the results found so far. The equilibrium given by strategies (**S2**) however, relied on each agent cheated upon by a neighbor having an incentive to carry out the punishment on the deficient side. If we introduce a mutual subnetwork, there exist agents *who do not have* a deficient relation. Contrary to the full information environment (**I1**), and given that with (**S2**) it is not optimal for agents to transmit information, under the other two information regimes it is not possible to identify the initial deviator. Agents, who are part of a mutual subnetwork, may therefore be reluctant to enter into an punishment phase immediately if they observe a deviation on only one side: They only expect their neighbor to enter the punishment phase with a certain probability. This leads to proposition 4.

Proposition 4 *In a non-mutual circular relational network of size c with* $g^{i,i+1} \leq 0$ *and* $g^{i,i-1} \geq 0$ $\forall i \in \mathcal{N}^S$, *let* $\underline{\delta} \equiv \left\{ \delta | \, g^{i,i+1} + \delta^{c-2} g^{i,i-1} = 0 \right\}$. *For agent k increase* $c^{k,k+1}$ *such that* $g^{k,k+1} > 0$, *so that the network becomes mixed.*

1. *Then, under information structure* (**I1**),

   (a) *the resulting relational network is still sustainable*

   (b) *but not strategically stable.*

2. *Denote with* $\underline{\underline{\delta}}$ *the minimum discount factor necessary to sustain the resulting network under* (**I2**) *and* (**I3**) *with strategy and belief profiles* (**S2**). *Then*

   (a) *for sufficiently low* $l^{i,i+1}$ ***or*** *sufficiently high* $w^{i,i+1}$, $\underline{\underline{\delta}} = \underline{\delta}$.

   (b) *for too high* $l^{i,i+1}$ *and too low* $w^{i,i+1}$, (**S2**) *does not result in a sustainable network.*

   (c) *a too low* $w^{i,i+1}$ *results in strategic instability of the network.*

Proof. Part 1 (a): The optimality of the actions during a punishment phase proposed in part 1 of the proof of proposition 2 only depend on the strategies played by the deviator and his neighbors being a stage-game Nash equilibrium for the bilateral interaction. Since we have full information, everybody knows everybody else's history and expecting the other to stick to the prescribed strategy (**S1**), would lead to playing D^{ij} whenever a deviation is observed.

Part 1 (b): The proof parallels the one for proposition 3 part 3.

Part 2 (a) through (c) we relegate to the appendix. *Q.E.D.* ■

The intuition for parts 2 (a) and (b) is the following (refer to figure 10): With the beliefs specified in (**S2**), if agent i in figure 10 observes $D^{i-1,i}$ and $C^{i+1,i}$ in $t = \tau$, he assigns probability $\frac{1}{c-1}$ to the event that any of the other agents in the network deviated first. Then, the bigger the network becomes, the more likely it is a priory that the agent that started the contagious process is an agent other than $i+1$ and $i+2$. Since in this case, $i+1$ will not play $D^{i+1,i}$ until $t = \tau + 2$, and since the net gain from cooperating with $i+1$ is positive for i, for a big size of the network, it is not a best response to play $D^{i,i+1}$ in $t = \tau + 1$. However, for agent i, with probability $\frac{1}{c-1}$ agent $i+1$ started. Because of that, if the loss from playing $C^{i,i+1}$ if $i+1$ plays $D^{i,i+1}$, $l^{i,i+1}$ is high enough, the expected payoff from carrying out the punishment may be higher than the one from going on cooperating for one more period. Furthermore, for agent i, with probability $\frac{1}{c-1}$ agent $i+1$ started. In that case, agent i expects $D^{i+1,i}$ from $t = \tau + 2$ on. Then, if the payoff from playing $D^{i,i+1}$ in $t = \tau + 1$, i.e. $w^{i,i+1}$, is very high in comparison to the payoff from playing $C^{i,i+1}$, agent i might also prefer to punish immediately.

The intuition for part 2 (c) is the following: Strategic stability rules out the belief that agent $i+1$ started and then sticks to the multilateral punishment since this is strictly dominated by having played $D^{i+1,i}$ in $t = \tau$. This only leaves a high $w^{i,i+1}$ as a reason to carry out punishments immediately.

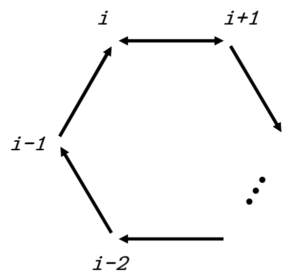

Figure 10: Circular network with a mutual relation

Proposition 4 shows a trade-off between profitability and sustainability of cooperation in networks: An agent, who benefits (too) strongly from relations with everybody he is connected to, may hurt cooperation between other agents because he may be unwilling to punish deviants.

3.2.3 Mixed circular networks with forgiving punishments

In this subsection, we will show that harsh, but forgiving punishments lead agents to use the so far unused possibility to transmit information through links (under *network information transmission*, (**I3**)). For high speed of information transmission, these strategies will give rise to equilibrium networks not sustainable with the unrelenting grim trigger-type strategies studied so far. We find that these forgiving punishments are *optimal strategies* under (**I3**) , while grim trigger-type strategies are not even though they are in the repeated prisoner's dilemma.

Remember that under strategy profile (**S2**) agents do not exploit the possibility to transmit information offered by (**I3**), independent of the speed ν. Because of this, the results under (**I2**) and (**I3**) do not differ. Transmitting information cannot be an equilibrium choice with (**S2**) because the punishment phase lasts forever. A defection leads then to a complete breakdown of the relational network during the punishment phase[17], and agents prefer to "grab what they can" before the collapse of the network by defecting/infecting rather than maintaining the relation and transmitting information. The potential of high speed information transmission is therefore left unused.

Rewarding agents for transmitting information instead of infecting her neighbor, it becomes possible to avoid the breakdown of cooperation and to make use of high speeds of information transmission, thereby, relaxing the agents' incentive constraint and allowing a sustainable network for a lower δ than (**S2**). Proposition 5 shows this.

For that end, let us define the following strategy profile:

Strategy profile S3

1. Agents start by playing C^{ij} $\forall i \in \mathcal{N}^{\mathcal{S}}, \forall j \in R_i$.

[17] That holds also if one considers a change in (**S2**) such that the reversion to the stage Nash equilibrium does not last forever but only for T periods.

2. As long as any agent i observes C^{ji} $\forall j \in R_i$, and as long as no message containing D^{mn} for any $m \in \mathcal{N}^S$, agent i goes on playing C^{ij} $\forall j \in R_i$.

3. If agent i observes D^{ji} for any $j \in R_i$ and she received no message about an earlier defection of j, agent i then sends a message about the deviation to her other neighbor and plays D^{ij} until j and i played D^{ij}, C^{ji} for T_j periods. After that i sends her other neighbor a message about the end of the punishment phase for player j and they go back to 2. thereafter. Each agent truthfully passes on the messages.

4. If a neighbor k of j receives a message about j's initial deviation, she plays D^{kj} until both, she receives the message that D^{ij}, C^{ji} has been played for T_j periods and D^{kj}, C^{jk} has been played for T_j periods. She returns to 2. thereafter.

5. If agent j played D^{ji}, she plays C^{ji} for the next T_j periods, D^{jk} in the period when k receives the information on her initial deviation and C^{jk} for the next T_j periods. She returns to 2. thereafter.

6. If some agent deviates from the actions in 3. – 6., the punishment starts against this agent.

Proposition 5 *In a non-mutual circular network of size c with $g^{i,i+1} \leq 0$ and $g^{i,i-1} \geq 0$ $\forall i \in \mathcal{N}^S$, let $\underline{\delta} \equiv \left\{ \delta | \, g^{i,i+1} + \delta^{c-2} g^{i,i-1} = 0 \right\}$. Let $\widetilde{\Delta}$ be the set of δ for which – together with an appropriate T_j, $\forall j \in \mathcal{N}^S$ – (**S3**) constitutes a sustainable non-mutual network with $g^{i,i+1} \leq 0$ and $g^{i,i-1} \geq 0$ under (**I3**) and $\widetilde{\delta} = \min \left\{ \widetilde{\Delta} \right\}$. Then*

(i) $\widetilde{\delta} \leq \underline{\delta}$ with a strict inequality for high speeds of information transmission (for $v > 1$).

(ii) if one substitutes non-mutual subnetworks with mutual ones the network is still sustainable and strategically stable $\forall \delta \in \widetilde{\Delta}$ for any l.

For the proof, which we relegate to appendix , there are four incentive constraints to consider:

1. Every agent has to have an incentive to stick to C^{ij} $\forall j \in R_i$ as long as neither he observes D^{ji} for a $j \in R_i$ nor he receives a message containing D^{mn} for an $m \in \mathcal{N}^S$. $\left(IC^{CI}\right)$

2. Given one neighbor m of i played $D^{m,i}$, each agent j has to have an incentive to send a message containing $D^{m,i}$ her other neighbor n and stick to $C^{i,n}$. $\left(IC^{CII}\right)$

3. Every neighbor of an original cheater has to have an incentive to carry out the punishment. $\left(IC^{P}\right)$

4. Every original cheater has to agree to be punished. $\left(IC^{LP}\right)$

We first show that $\left(IC^{CII}\right)$ and $\left(IC^{P}\right)$ are never binding. Using $\left(IC^{LP}\right)$ and $\left(IC^{CI}\right)$, we then show that, for a speed of $v = 1$, it is possible to choose a length T_j, $\forall j \in \mathcal{N}^S$, of the punishment period for each agent such that the punishment payoff for her is equivalent to minimaxing her on both sides forever[18], i.e. the strength of the punishment is equivalent to the one for (**S2**). Increasing the speed of information transmission reduces the delay of the punishment and, thus, relaxes $\left(IC^{LP}\right)$ which in turn gives room to make it more severe. This establishes *(i)*. Since agents are being rewarded for punishing their neighbor, they always have an incentive to do so during a punishment phase even if they want to cooperate bilaterally, which establishes *(ii)*.

Corollary 1 *Under* network information transmission (**I3**) *and assumptions 1, 2, and 4, for high enough ν it is possible to find a T_j $\forall j \in \mathcal{N}^S$ such that* (**S3**) *is an optimal punishment mechanism whereas* (**S2**) *is not.*

Proof. Two elements determine the strength of the multilateral punishment mechanism in the network: the payoff after punishment starts on each side, and the promptness with which this punishment starts on each side after a deviation. It is always possible to adjust the length of the punishment phase T_j for each player j such that he receives an punishment payoff equivalent to minimax forever. Furthermore, according to assumption 1 $R = \mathcal{C}$, the other neighbor of an agent that first defects can "get to know" about the defection and start the punishment phase at the earliest with the information that travelled through the network. This means that (**S3**) is an optimal punishment mechanism. As for high ν, information transmission is faster than contagion, (**S2**) is not an optimal punishment mechanism for high ν. ■

Punishment with (**S3**) is as strong as possible and as fast as possible, therefore these are the optimal (punishment) strategies in our network. Proposition 5 also shows that it is not necessary to have a complete breakdown of cooperation in the network in case of a deviation if information about past actions can be transmitted. The equilibrium is, thus, more robust (against e.g. mistakes) and increases welfare during punishment phases.

Since under perfect information (**I1**) the agent that defects first is known, the complete breakdown of the network in a punishment phase can be avoided through punishments as in (**S3**). These strategies[19] result in a critical discount factor as for (**S1**), as punishment was immediate on both sides already with (**S1**).

While strategy profile (**S3**) avoids the breakdown of the network due to mutual subnetworks for (**I3**), it can not be used under (**I2**) since it makes use of information transmission. Without

[18] To avoid divisibility problems, one can always assume a public randomization device giving the end of the punishment period for each agent such that in expectation the punishment payoff of the initial deviator is equivalent to minimaxing him forever.

[19] All neighbors $j \in N_i$ of an initial cheater i start playing $D^{j,i}$ until i has played $C^{i,j} \forall j \in N_i$ for T periods and then they go back to plaing $C^{i,j}, C^{j,i}$. In all other games in the network, the players go on playing the cooperative action during the punishment phase for player i. As the initial cheater can always get his minimax payoff forever, which is the payoff from the punishment in (**S1**), the biggest T, for which this strategy profile is an equilibrium, gives him exactly this payoff.

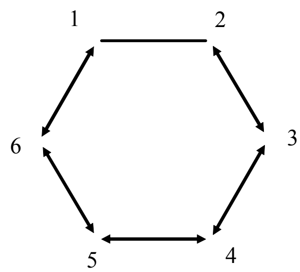

Figure 11: Mixed relational network containing only mutual relations except one bilaterally deficient one

information transmission it is impossible to know who deviated first from the equilibrium path and a targeted punishment of only the agent that defects first becomes unattainable.

Up to now, we have not explicitly considered bilaterally deficient relations. It should however be clear at this point that a mixed circular relational network containing bilaterally deficient relations – as for example the network in figure 11 – is sustainable with the same strategies discussed above under the same conditions given.

4 Higher degree networks and social capital

In this section we show that there are generalizations of the results we obtained for the simple relational networks above allowing for more than two neighbors[20]. For this end, we will use a c-cycle as a basic structure and add a link. We will show how networks of relations that generate "slack enforcement power" for some agents may enable these to sustain cooperation on additional deficient relations and even in *one shot* prisoner's dilemma interactions. We then offer an interpretation of this use of networks of relations as cooperation-enforcement/governance devices for new social dilemmas in terms of the highly debated but somewhat vague concept of "social capital".

In our model, establishing a link always increases the discounted payoff of the agents creating it, as it is always profitable to cooperate. However, regarding the sustainability of the network, though, adding a non-mutual relation has two effects: On the one hand, adding any relation that is not sustainable in autarky uses scarce enforcement power. Thus, there is a limit to adding them. On the other hand, if information travels with delay along the links of the network, or where information cannot "travel" and strategies rely on contagion, new links shorten paths making multilateral punishments faster.

In the remainder of the section, we consider for each of the three informational regimes, (**I1**) – (**I3**), the effects of adding to a non-mutual circular network a bilaterally deficient, a unilateral, and a mutual relation, one at a time.

[20] We have done so already in the sections before when we looked at trees, stars, or non-circular networks, one end node of which was an autarkically sustainable subnetwork.

Full information (**I1**) It is straightforward to generalize proposition 2 part 1 and we state without proof:

Proposition 6 *Assume* (**I1**) *and the strategy profile* (**S1**)*. Then a network is sustainable iff*

$$\sum_{j \in R_i} g^{ij} > 0 \quad \forall i \in \mathcal{N}^{\mathcal{S}}. \tag{1}$$

As long as (1) is satisfied, also bilaterally deficient relations can be sustained in equilibrium. Consider for example figure 12 (c). Agents i's and k's being part of the network helps them sustain a bilaterally deficient relation if the sum of the net gains from cooperating for i and k are big enough.

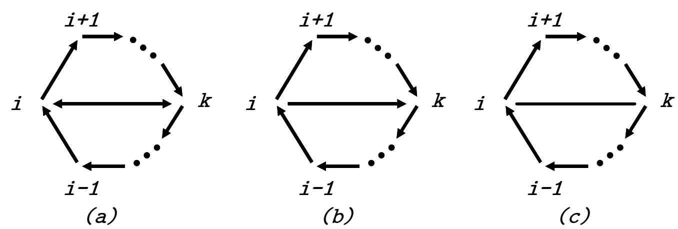

Figure 12: Adding a relation to a circular, non-mutual network

While "grim trigger" strategies (**S1**) are an equilibrium, the forward induction argument of Proposition 3 part 3 also applies here as long as (a) there are subnetworks that are sustainable without the rest of the network, and (b) there is a "rest" of the network that is not, i.e. as long as the relation ik that is added to $\mathcal{N}^{\mathcal{S}} \setminus ik$ is not sustainable outside $\mathcal{N}^{\mathcal{S}}$.

To see this, consider first figure 12, networks (b) or (c). Since ik is a deficient relation for i, $\mathcal{N}^{\mathcal{S}}$ is only sustainable with (**S1**) if $\mathcal{N}^{\mathcal{S}} \setminus ik$ is sustainable in autarky. If this is the case, then the same stability argument made for mutual subnetworks apply. If e.g., agent i deviates only from her relation with agent k, but not from his other two relations, induces speculation on future play as under the current strategy profile the deviation is strictly dominated by a simultaneous deviation on all relations. Furthermore there is an equilibrium – $\mathcal{N}^{\mathcal{S}} \setminus ik$ – which (i) Pareto-dominates the continuation equilibrium in the punishment phase of (**S1**) and which is (ii) a focal point after this deviation. This is a profitable deviation, given the agents coordinate on $\mathcal{N}^{\mathcal{S}} \setminus ik$, since $g^{ik} < 0$.

Consider now network (a) with strategy profile (**S1**). If we add a mutual relation ik to a circular network $\mathcal{N}^{\mathcal{S}} \setminus ik$ that is *not sustainable* because $g^{i,i-1} + g^{i,i+1} < 0$ and/or $g^{k,k-1} + g^{k,k+1} < 0$, and if g^{ik} and g^{ki} are big enough st. $\mathcal{N}^{\mathcal{S}}$ is sustainable with (**S1**), the stability argument from proposition 3 part 3 applies: agents i and k had a "profitable deviation" from $\mathcal{N}^{\mathcal{S}}$ leaving them with ik (refer to figure (a) for this).

If instead we add the mutual relation ik to a *sustainable* network $\mathcal{N}^{\mathcal{S}} \setminus ik$, both subnetworks are sustainable in autarky and there is no need to combine them into one multilateral punishment mechanism. Furthermore, under (**I1**), every member of $\mathcal{N}^{\mathcal{S}} \setminus ik$ immediately observes the play of every other player so that there is no delay in punishment that can be reduced by shortening paths through the new relation ik. However, even if players agreed on (**S1**) including ik, the sustainability of both subnetworks rules out the stability argument from Proposition 3 part 3.

As in previous sections, with more sophisticated forgiving punishment strategies (**S3**), this forward induction argument vanishes since punishments phases are followed by a return to cooperation that, together with rewards for the punishers provide incentives to pass on information and punish.

No information transmission (**I2**) Under the no information transmission assumption (**I2**) we now study sustainable networks when agents use the contagion strategies (**S2**).

Refer to figure 12, first considering network (a). Obviously, if both subnetworks ik and $\mathcal{N}^{\mathcal{S}} \setminus ik$ were sustainable in autarky, treating the subnetworks separately and adding ik to $\mathcal{N}^{\mathcal{S}} \setminus ik$ results in a sustainable network.

If, on the other hand, $\mathcal{N}^{\mathcal{S}} \setminus ik$ is not sustainable on its own, adding ik might help sustain the network for two reasons. First, if $\mathcal{N}^{\mathcal{S}} \setminus ik$ is not sustainable because $g^{i,i+1} + \delta^{c-2} g^{i,.i-1} < 0$ and if $g^{i,i+1} + \delta^{m-2} g^{i,k} + \delta^{c-2} g^{i,.i-1} > 0$, where m is the size of the subnetwork $\{i, i+1, ..., k\}$, adding ik will result in a sustainable network if both, i and k have, given their beliefs, an incentive to contribute to a multilateral punishment using their mutual relation. Second, if $\mathcal{N}^{\mathcal{S}} \setminus ik$ is not sustainable because $g^{j,j+1} + \delta^{c-2} g^{j,j-1} < 0$, adding ik may result in a sustainable network under the same condition because the delay with which the punishment reaches j is shorter.

Proposition 7 *Let a network $\mathcal{N}^{\mathcal{S}}$ consist of a non-mutual circular network of size c, $\mathcal{N}^{\mathcal{S}} \setminus ik$, with $g^{i,i+1} \leq 0$ and $g^{i,i-1} \geq 0$ $\forall i \in \mathcal{N}^{\mathcal{S}} \setminus ik$ and a mutual relation ik between two non-adjacent agents. Let $\underline{\delta} \equiv \left\{ \delta | \, g^{i,i+1} + \delta^{c-2} g^{i,i-1} = 0 \right\}$ $\forall i \in \mathcal{N}^{\mathcal{S}} \setminus ik$. Let $\widehat{\Delta}$ be the set of δ for which $\mathcal{N}^{\mathcal{S}}$ is sustainable with* (**S2**) *and beliefs specified in appendix E and let $\widehat{\delta} = \min \left\{ \widehat{\Delta} \right\}$. Then* for $l^{i,k}$ *and* $l^{k,i}$ small enough *or* $w^{i,k}$ *and* $w^{k,i}$ big enough*, $\widehat{\delta} < \underline{\delta}$.*

Proof. Assume (**S2**) and the beliefs specified in appendix E. As in the proof of proposition 4, by assuming $l^{i,k}$ and $l^{k,i}$ low enough or $w^{i,k}$ and $w^{k,i}$ big enough, i's (k's) expected profit from playing C^{ik} (C^{ki}) after having observed agent $i-1$ ($k-1$) deviate is smaller than if they not only play $D^{i,i+1}$ ($D^{k,k+1}$), i.e. infect agent $i+1$ (agent $k+1$), but also $D^{i,k}$ ($D^{k,i}$), i.e. infect also agent k (agent i). Therefore punishment sets in earlier and a lower discount factor is needed to sustain $\mathcal{N}^{\mathcal{S}}$. *Q.E.D.* ■

Again, if i's (k's) loss from playing C^{ik} (C^{ki}) if k (i) plays D^{ki} (D^{ik}) or the gain from playing D^{ik} (D^{ki}) if k (i) plays C^{ki} (C^{ik}) is big, the expected payoff from not punishing is relatively low and the agents sharing the mutual relation are willing to contribute to a collective punishment mechanism.

Consider now networks (b) and (c). Here, adding the relation ik, which is unilateral (bilaterally deficient), involves a trade-off. On the one hand, punishment will be faster, which relaxes the incentive constraint for each agent in $\mathcal{N}^{\mathcal{S}} \setminus ik$ and makes the network sustainable for lower discount factors. On the other hand, one agent (two agents) will have to sustain one deficient relation more, which tightens the incentive constraint for this agent (these agents). The set of discount factors for which the network is sustainable may therefore expand or shrinks with the addition the new relation, depending on parameter values.

The conditions for sustainability of the network, which we give together with the belief structure in appendix E, are a straightforward generalization of the conditions we had for the simple network with $\deg(i) \leq 2$.

Network information transmission (I3) Consider first network (a) from figure 12. Given the feasibility of information transmission, consider strategies (**S3**) which make use of it. For network (a) to be sustainable, the incentive constraints for agents other than i and k, are equivalent to the ones given in appendix D with one change: Since the ways are shorter, the delay with which punishment sets in is shorter as well, making it easier to sustain the network. As an example for the incentive constraints for agents i and k, we give the ones for i in appendix F. Again, the sustainability conditions from appendix D generalize.

Consider networks (b) and (c). Again, adding the relation ik, which is unilateral (bilaterally deficient), involves a trade-off. On the one hand, punishment will be faster, which relaxes the incentive constraint for each agent in the network and makes the network sustainable for lower discount factors. On the other hand, one agent (two agents) will have to sustain one deficient relation more, which tightens the incentive constraint for this agent (these agents). It is, thus, not clear whether the set of discount factors for which the network is sustainable increases or shrinks with adding the additional relation.

Social Capital Consider again figure 12 (c). We stated above that agents i's and k's being part of the network may help them sustain a bilaterally deficient relation between them. This is the case if the sum of the net gains from cooperating for i and k from their other relations are large enough, i.e. if they dispose of sufficient *slack enforcement power* to enforce the additional relation.

Suppose the circular network $\{i, i+1, ..., k-1, k, k+1, ..., i-1, i\}$ is a social network, i.e. the relations in it are *social relations*, and suppose the bilaterally deficient relation between i and k is a *one-shot prisoner's dilemma,* say an occasional business transaction where each agent can "hold up" the other. Then the slack enforcement power from our social network, used to govern a one-shot business interaction, is much like what Coleman (1990) defines *social capital*:

> Social capital is defined by its function. It is not a single entity, but a variety of different entities having two characteristics in common: They all consist of some aspect of social structures, and they facilitate certain actions of individuals who are within that structure. Like other forms of capital, social capital is productive, making possible the achievement of certain ends that would not be attainable in its absence. Like physical

capital and human capital, social capital is not completely fungible, but is fungible with respect to certain activities. A given form of social capital that is valuable in facilitating certain actions may be useless or even harmful for others. Unlike other forms of capital, social capital inheres the structure of relations between persons and among persons. It is lodged neither in individuals nor in physical implements of production.

"*...social capital* inheres the *structure of relations* between persons and among persons" and it makes "possible the achievement of certain ends that would not be attainable in its absence." This is a micro-perspective on social capital. Our model allows for a formal definition for social capital à la Coleman:

Definition 7 *(Social capital à la Coleman): Take a sustainable social network $\mathcal{N}^S$ with $i, k \in \mathcal{N}^S$. Then we define the* ***individual social capital*** *i and k can draw upon for a one-shot business interaction ik as*

$$sc_{ik} = \left(\max\left\{w^{ik} - c^{ik}, w^{ki} - c^{ki}\right\}\middle| C^{ik}, C^{ki} \textit{ is equilibrium in a MPM containing } \mathcal{N}^S \textit{ and } ik\right).$$

The social capital agent i can draw on from being part of a social network is defined as the slack enforcement power usable to enforce cooperation-compliance in other interactions in need of governance through an *MPM* (multilateral punishment mechanism). With complete information (**I1**), this is only a player specific definition as it is equivalent to the sum of his net gains from cooperation in all his social relations $sc_{ik} = \min\left\{\sum_{j\in R_i} g^{ij}, \sum_{j\in R_i} g^{kj}\right\}$. For the other information regimes, the extent to which existing relations in a social network can facilitate "the achievement of certain ends" for an agent depends not only on his net gains from cooperation, i.e. how much he has to loose in his social relations. Since the delay with which an eventual punishment sets in matters, it also depends on *partners' locations* in the network.

Robert Putnam (1995) takes another perspective on social capital. For him, the concept "refers to the collective value of all 'social networks' and the inclinations that arise from these networks to do things for each other." This is a macro-perspective on social capital, which, translated into our model, lead to the following formal definition:

Definition 8 *(Social capital à la Putnam): Take a sustainable social network $\mathcal{N}^S$ with $i, k \in \mathcal{N}^S$. Then we define the* ***social capital of a society*** *as the average individual social capital in that society*

$$\frac{1}{n\,\mathrm{card}\,(R_i)} \sum_{i\in\mathcal{N}^S} \sum_{k\in R_i} sc_{ik}.$$

The conclusion to be drawn from our model for the construction of aggregate measures of social capital is: If there is full information about the actions of economic agents, it suffices to have a measure of the average sum of the net gains from cooperation per person from social relations in the economy. However, if this is not the case, as in most real world situations, in addition, a measure of the density of the network should be used.

Information transmission as social capital We would like to emphasize that the value of the social network may also rest in the enforcement of the transmission of information on the history of interactions with outsiders. If the outsiders interact repeatedly with changing members of the network, transmission of information on the history of the play in these interactions through the network may help facilitate cooperation in them. In that sense, our model is a microfoundation of Kandori's (1992) *attaching a label to a cheater* by the members of the social network. The fact that such a transmission of information in a society is of economic value has been shown in variuos studies, among others in Acemoglu and Zilibotti (1999).

5 Conclusion

Each of us is involved in a network of long term relationships of different kinds and with different parties. Networks of social and economic relations include colluding firms, industrial districts, interbank markets as well as criminal/terrorist organizations. In this paper we have tried to clarify how the structure of such networks of relations affects the feasible equilibrium pattern of interaction.

In our model, agents maintain long term self-enforcing relations thanks to the information circulation and the enforcement/sanctioning power ensured by a network of such relations. We identify equilibrium conditions for different architectures of such networks, paying special attention to differences in these conditions for circular and non-circular architectures. The basic framework is that of repeated games between fixed partners with three basic information structures: complete information, no information, and information transmission through the network's links.

We show that if agents cannot discipline themselves within a certain relation, the pooling of asymmetries in payoffs across the network may allow to sustain the relation under all three informational assumptions. We find an end-network effect, i.e. that a non-circular network or subnetwork is not sustainable. We find that the possibility to transmit information about a defection through the links in the network is not exploited in equilibrium if enforcement relies on unforgiving punishment phases. More complex punishment strategies induce agents to use information transmission, and to keep on cooperating in the rest of the network while punishing a defection (which increases efficiency and decreases the discount factor necessary to sustain the network). If information can be transmitted via the network, grim trigger strategies, therefore, cease to be optimal punishments as they do not use the possibility to transmit information to punish cheaters faster. Having self-sustaining relations in the network turns out to hurt cooperation with imperfect information, because agents may then not be willing to perform the prescribed punishment after a defection. When information can be transmitted, the network may be sustained using strategies that reward the punisher and encourage information transmission.

We model relations as cooperative agreements in generic infinitely repeated prisoners' dilemmas forming the links of the network of relations. The model is general enough to capture numerous economic and social situations. We provide a microfoundation to Granovetter's (1985) idea of "embeddedness" according to which, by ignoring the social background in which economic transactions are embedded, economists fail to understand important features of the economic process. Our end-network effect, i.e. the finding that a non-circular network or subnetwork is not sustainable, provides a clear explanation of why ”closure“ of social networks is so important for social capital, as argued by Coleman (1988) and (1990). Finally, we drew some conclusions about sensible measures of social capital in a network of relations, both on an individual and an aggregate level.

Immediate applications of our model include the organization of inter-firm relations in industrial districts, the enforcement of collusive behavior in business networks, interbank relations and the effects of ”social capital” on the governance of economic and social interactions (as discussed by Coleman (1988, 1990), Putnam (1993) and Greif (1993) and formalized by Spagnolo (1999b)). In her much acclaimed book, Saxenian (1994) attributes a large part of Silicon Valley's success to a special culture of cooperation in that industrial district, which stems from a common background of the early workforce in that area. We believe our model offered a complementary explanation how a cooperative social networks may help enforce information exchanges and circulation in a community.

References

[1] Abreu, D., (1988): ”On the Theory of Infinitely Repeated Games with Discounting”, *Econometrica*, **56**, 383-396.

[2] Acemoglu, D. and Zilibotti, F., (1999): ”Information Accumulation in Development”, *Journal of Economic Growth*, **4**, 5-38.

[3] Baker, G., Gibbons, R. and Murphy, K. J. (2001): ”Relational Contracts and the Theory of the Firm,” *Quarterly Journal of Economics*, **117**, 39-84.

[4] Bala, V. and Goyal, S. (2000a): ”A Noncooperative Model of Network Formation,” *Econometrica*, **68**, 1181-1229.

[5] Bala, V. and Goyal, S. (2000b): ”A Strategic Analysis of Network Reliability,” *Review of Economic Design*, **5**, 205-228.

[6] Belleflamme, P. and Bloch, F. (2004): "Market Sharing Agreements and Collusive Networks," *International Economic Review*, **45**, 387-411.

[7] Bernheim, B. D. and Whinston, M. D. (1990): ”Multimarket contact and collusive behavior,” *RAND Journal of Economics*, **21**, 1-26.

[8] Brown, M., Falk, A. and Fehr, E. (2004): "Relational Contracts and the Nature of Market Interaction," *Econometrica*, **72**, 747-781.

[9] Bull, C. (1987): ”The Existence of Self-Enforcing Implicit Contracts,” *Quarterly Journal of Economics*, **102**, 147-59.

[10] Casella, A. and Rauch, J. (1988): "Anonymous Market and Group Ties in International Trade," *Journal of International Economics*, **58**, 19-47.

[11] Coleman, J. (1988): "Social Capital in the Creation of Human Capital," *The American Journal of Sociology*, **94**, S95-S120.

[12] Coleman, J. (1990): *Foundations of Social Theory*, Cambridge MA: Harvard University Press.

[13] Dixit, A. (2003a): "On Modes of Economic Governance," *Econometrica*, **71**, 449-481.

[14] Dixit, A. (2003b): "Trade Expansion and Contract Enforcement," *Journal of Political Economy,* **111**, 1293-1317.

[15] Ellison, G. (1994): "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching," *Review of Economic Studies*, **61**, 567-588.

[16] Friedman, James W. (1971): "A Non-Cooperative Equilibrium for Supergames," *Review of Economic Studies*, **38**, 1-12.

[17] Granovetter, M., (1985): "Economic action and social structure: the problem of embeddedness." *American Journal of Sociology*, **23**, 481-510.

[18] Greif, A., (2003): "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders Coalition." *American Economic Review*, **83**, 523-548.

[19] Groh, C. (2002): "Communication Networks and Cooperation," University of Bonn, mimeo.

[20] Haag, M. and Lagunoff, R. (2002): "Social Norms, Local Interaction and Neighborhood Planning", Warwick and Georgetown Universities, mimeo.

[21] Jackson, M. O. (2003): "A Survey of Models of Network Formation: Stability and Efficiency," California Institute of Technology, mimeo.

[22] Jackson, M. O. and Wolinsky, A. (1996): "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, **71**, 44-74.

[23] Jackson, M. O. and Watts, A. (2002): "On the Formation of Interaction Networks in Social Coordination Games," *Games and Economic Behavior*, **41**, 265-291.

[24] Johnson, C. and Gilles, R. P. (2000): "Spatial Social Networks," *Review of Economic Design*, **5**, 273-299.

[25] Kandori, M. (1992): "Social Norms and Community Enforcement," *Review of Economic Studies*, **59**, 63-80.

[26] Klein, B. and Leffler, K. (1981): "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, **89**, 615-641.

[27] Kohlberg, E. and Mertens, J.-F. (1986): "On the Strategic Stability of Equilibria," *Econometrica*, **54**, 1003-1038.

[28] Kramarz, F., and Thesmar, D. (2004): "Social Networks in the Boardroom," manuscript, CREST-INSEE, Paris. and IZA, Bonn.

[29] Kranton, R. E. (1996): "Reciprocal Exchange: a Self-Sustaining System," *American Economic Review*, **86**, 63-80.

[30] Kranton, R. E. and Minehart, D. F. (2001): "A Theory of Buyer-Seller Networks," *American Economic Review*, **91**, 485-508.

[31] Kranton, R. E. and Minehart, D. F. (2000): "Networks versus Vertical Integration," *RAND Journal of Economics*, **31**, 301-331.

[32] Leitner, Y. (2004): "Financial Networks: Contagion, Commitment, and Private-Sector Bailouts" manuscript, Federal Reserve Bank of Philadelphia.

[33] Levin, J. (2003):,"Relational Incentive Contracts," *American Economic Review*, **93**, 835-857.

[34] Lippert, S. and Spagnolo, G. (2004): "Exclusion in Networks of Relations," mimeo.

[35] Macaulay, S. (1963): "Non Contractual Relations in Business: A Preliminary Study," *American Sociological Review*, **28**, 55-67.

[36] MacLeod, B. and Malcolmson, J. (1989): "Implicit contracts, Incentive Compatibility and Involuntary Unemployment," *Econometrica,* **57**, 447-480.

[37] Maggi, G. (1999): "The Role of Multilateral Institutions in International Trade Cooperation," *American Economic Review*, **89**, 190-214.

[38] Matsushima, H. (2001): "Multimarket Contact, Imperfect Monitoring, and Implicit Collusion," *Journal of Economic Theory*, **98**, 158-178.

[39] North, D. C. (1990): *Institutions, Institutional Change and Economic Performance*, Cambridge UK: Cambridge University Press.

[40] Okuno-Fujiwara, M. and Postlewaite, A. (1995) "Social Norms in Random Matching Games", *Games and Economic Behavior,* **9**, 79-100.

[41] Powell, W. W., Koput, K. W., and Smith-Doerr, L. (1996): "Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology," *Administrative Science Quarterly*, **41**, 116-145.

[42] PUTNAM, R., (2002): *Making Democracy Work: Civic Traditions in Modern Italy,* Princeton: Princeton University Press.

[43] PUTNAM, R., (1995): "Bowling Alone: Americas Declining Social Capital" *Journal of Democracy*, **6**, 65-78.

[44] RAJO, L. (2003): "Relational Team Incentives and Ownership", manuscript, University of Chicago.

[45] SAXENIAN, A. (1994): "*Regional advantage: Culture and competition in Silicon Valley and Route 128"*, Harvard University Press, Cambridge, MA.

[46] SPAGNOLO, G. (1999a): "On Interdependent Supergames: Multimarket Contact, Concavity, and Collusion," *Journal of Economic Theory*, **89**, 127-139.

[47] SPAGNOLO, G. (1999b): "Social Relations and Cooperation in Organizations," *Journal of Economic Behavior and Organization*, **38**, 1-25.

[48] SPAGNOLO, G. (2000): *Norme Sociali, Incentivi ed Efficienza Organizzativa: una Analisi del Profit-Sharing*, Bologna: CLUEB.

[49] SPAGNOLO, G. (2002): "Globalization and Cooperative Relations," *C.E.P.R. Discussion Paper No. 3522*

[50] TELSER, L., (1981): "A Theory of Self-Enforcing Agreements," *Journal of Business,* **LIII**, 27-44.

[51] VEGA-REDONDO, F. (2002): "Building Up Social Capital in a Changing World," mimeo, University of Alicante.

[52] WILLIAMSON, O. E. (1996): *The Mechanisms of Governance*, Oxford University Press, New York, NY.

A Strategy and belief profile (S2)

1. Each agent $i \in \mathcal{N}^S$ starts playing the agreed upon action vector C^{ij} $\forall i \in \mathcal{N}^{\mathcal{S}}, \forall j \in R_i$.
2. As long as player i observes every neighbor $j \in R_i$ play C^{ji} she goes on playing C^{ij} $\forall j \in R_j$.
3. If player i observes a neighbor j play D^{ji} in $t = \tau$ she reverts to D^{ij} $\forall j \in R_i$ $\forall t \geq \tau + 1$.

(*i*) $a_t^{j,j+1}, a_t^{j+1,j} = C^{j,j+1}, C^{j+1,j}$ and $a_t^{j,j-1}, a_t^{j-1,j} = C^{j,j-1}, C^{j-1,j}$ $\forall t = 1, ..., \tau$, they believe $a_t^{k,l}, a_t^{l,k} = C^{k,l}, C^{l,k}$, $\forall kl \in R$, $\forall t = 1, ..., \tau$,

(*ii*) $a_\tau^{j+1,j} = D^{j+1,j}$ and $a_\tau^{j-1,j} = D^{j-1,j}$ or $a_\tau^{j+1,j} = C^{j+1,j}$ and $a_\tau^{j-1,j} = D^{j-1,j}$ with $g^{j,j+1} < 0$ they can have any belief consistent with this observation,

(iii) $a_\tau^{j+1,j} = D^{j+1,j}$ and $a_\tau^{j-1,j} = C^{j-1,j}$ with $g^{j,j+1} < 0$, they assign an equal probability $\Pr\left(a_t^{k,l} = D^{k,l} \wedge a_t^{k,l} = C^{m,n} \forall m \neq k\right), t \leq \tau, \forall k \neq j$.

For agents j with $\mathrm{id}(j) = 2$, beliefs are such that if they observe[21]

(iv) $a_t^{j,j+1}, a_t^{j+1,j} = C^{j,j+1}, C^{j+1,j}$ and $a_t^{j,j-1}, a_t^{j-1,j} = C^{j,j-1}, C^{j-1,j}$ $\forall t = 1, ..., \tau$, they believe $a_t^{k,l}, a_t^{l,k} = C^{k,l}, C^{l,k}$, $\forall kl \in R$, $\forall t = 1, ..., \tau$,

(v) $a_\tau^{j+1,j} = D^{j+1,j}$ and $a_\tau^{j-1,j} = D^{j-1,j}$ they can have any belief consistent with this observation,

(vi) $a_\tau^{j+1,j} = D^{j+1,j}$ and $a_\tau^{j-1,j} = C^{j-1,j}$ or $a_\tau^{j+1,j} = C^{j+1,j}$ and $a_\tau^{j-1,j} = D^{j-1,j}$, they assign an equal probability $\Pr\left(a_t^{k,l} = D^{k,l} \wedge a_t^{k,l} = C^{m,n} \forall m \neq k\right), t \leq \tau, \forall k \neq j$.

B Proposition 1

Proof. A network has been defined non-circular if for no agent $i_1 \in \mathcal{N}^S$ there exists a path $\{i_1, i_2, ..., i_k\}$ with $i_1 = i_k$. It has been defined non-mutual if $g^{ij} > 0 \Leftrightarrow g^{ji} \leq 0$. In such a network, there would have to be either an agent e at the end vertex with $\mathrm{od}\, e = 1$ or an agent m in the middle with $\mathrm{od}\, m = 2$. Since we assumed $\deg i \leq 2$, there will not be any punishment from *other* neighbors and agent $e's$ or agent $m's$ dominant strategy is to defect from the relation. *Q.E.D.* ■

C Proposition 4

First we proof that with an unforgiving punishment, cooperation may break down if we replace a unilateral relation with a mutual one. We then show that for $U^i\left(C^{ij}, D^{ji}\right)$ in the mutual relation small enough, the set of equilibria will not shrink.

Proof. Part 2 (a) and (b). Consider strategies (**S2**) and beliefs as outlined above. Suppose, we are in the situation of figure 10 with agents i and $i+1$ forming a mutual subnetwork. Consider the following defection: Agent $i+1$ plays $D^{i+1,i+2}$ and after $c-2$ periods goes on playing $C^{i+1,i}$. After $c-2$ periods, say in period $t = \tau$, agent i observes $D^{i-1,i}$ and $C^{i+1,i}$. Playing $D^{i,i+1}$ in $t = \tau + 1$ is rational for agent i only if she expects $i+1$ to play $D^{i+1,i}$ in $t = \tau + 1$. Whether she expects this to happen, depends on her beliefs on who started the deviation. Agent i may have three possible beliefs about who defected initially.

(a) Agent $i+1$ started and deviated only from his relation with $i+2$. If agent $i+1$ after his initial deviation sticks to the strategies prescribed, he will play $D^{i+1,i}$ in $t = \tau + 1$. Then it is in i's best interest to play $D^{i,i+1}$ as well. In the expected discounted payoff, this receives a bigger weight, the lower $l^{i,i+1}$.

[21] We will need this part of the belief structure only when we consider mixed networks. In unilateral networks, by definition there are no agents with an indegree of two.

(b) Agent $i+2$ started: Then $i+2$ would infect $i+1$ in $t=\tau+1$, thus, no matter what agent i plays in $t=\tau+1$, agent $i+1$ will play $D^{i+1,i}$ in $t=\tau+2$. Therefore it is better to have a deviation profit in $t=\tau+1$ and play $D^{i,i+1}$. In the expected discounted payoff, this receives a bigger weight, the higher $w^{i,i+1}$.

(c) An agent $m \in \mathcal{N}^S \setminus \{i, i+1, i+2\}$ started: The earliest period when $i+1$ would be infected by $i+2$ would be $\tau+2$. Thus i will expect $i+1$ to play $C^{i+1,i}$ at least until $t=\tau+2$. Since we assumed $g^{i,i+1}>0$, for this belief it is *not* a best response to play $D^{i,i+1}$ in $t=\tau+1$.

Since agent i does not have any information, a consistent belief is that cases (a) and (b) have occurred with probability $\frac{1}{c-1}$ and case (c) with probability $\frac{c-3}{c-1}$. If c gets large, therefore, the expected payoff for agent i from deferring the punishment phase by one period may become positive.

This in turn delays the expected punishment date of an initial deviator, which leads to a breakdown of the network if $l^{i,i+1}$ is not small and $w^{i,i+1}$ is not big.

Part 2 (c). The proof parallels the one for proposition 3 part 3. *Q.E.D.* ■

D Proposition 5

For notational convenience the following definition will be useful.

Definition 9 *We define a function*

$$\theta(c,v) \equiv \begin{cases} \max\left\{\frac{c-2}{v}, 1\right\} & \text{if } int\left(\frac{c-2}{v}\right) = \frac{c-2}{v} \\ \max\left\{int\left(\frac{c-2}{v}+1\right), 1\right\} & \text{if } int\left(\frac{c-2}{v}\right) \neq \frac{c-2}{v} \end{cases}.$$

This function maps the order of the cycle c and the speed of information transmission v into the strictly positive natural numbers and indicates the period in which an information about play between agents i and $i+1$ in period 0 reaches agent $i-1$.

In the proof we first consider the incentive constraints for agents in the network not to deviate from cooperation in phase I (IC^{CI}), from cooperation with their other neighbor in phase II that is if one neighbor cheated (IC^{CII}), from punishing the original cheater in phase II (IC^{P}), and from letting the others punish when she deviated in the first place (IC^{LP}). In a second step we show that $\widetilde{\delta} \leq \underline{\delta}$. It is shown that IC^{CII} and IC^{P} are never binding, so we can concentrate on IC^{CI} and IC^{LP}. For a speed of $v=1$, by an appropriate choice of the length of the punishment, the conditions for cooperation can be made equivalent to the ones for (**S2**). Increasing the speed then relaxes IC^{LP} which gives room to make punishment more severe, which establishes (i): $\widetilde{\delta} \leq \underline{\delta}$. Since agents are being rewarded for punishing their neighbor, they always have an incentive to do so during a punishment phase even if they want to cooperate bilaterally, which establishes (ii). If T is chosen such that punishment is as hard as playing minimax strategies with both neighbors forever, this is the hardest punishment possible. Since here information transmission is used, every mean to decrease the delay before punishment on both sides sets in is used. This establishes the corollary.

Proof. The following incentive constraints are to be satisfied:

1. (IC^{CI}) For each agent i, playing $D^{i,i+1}$ in $t=0$ and $D^{i,i-1}$ in $t=\theta(c,v)$, which is her best deviation, yields $w^{i,i+1}$ in $t=0$, $l^{i,i+1}$ for the following T_i periods and $c^{i,i+1}$ thereafter, as well as $c^{i,i-1}$ until $t=\theta(c,v)-1$, $w^{i,i-1}$ in $t=\theta(c,v)$, $l^{i,i-1}$ for the following T_i periods and $c^{i,i-1}$ thereafter. Playing $C^{i,i+1}$ and $C^{i,i-1}$ forever yields $\frac{1}{1-\delta}\left(c^{i,i+1}+c^{i,i-1}\right)$. Summing up leads to $\left(IC^{CI}\right)$, which is the condition for (**S3**) to be a Nash equilibrium.

$$\begin{aligned} IC^{CI} \equiv &\left(c^{i,i+1}-w^{i,i+1}\right)+\sum_{t=1}^{T_i}\delta^t\left(c^{i,i+1}-l^{i,i+1}\right) \\ &+\delta^{\theta(c,\nu)}\left(c^{i,i-1}-w^{i,i-1}\right)+\sum_{t=\theta(c,\nu)+1}^{\theta(c,\nu)+T_i}\delta^t\left(c^{i,i-1}-l^{i,i-1}\right)\geq 0 \\ &\forall i\in N^{\mathcal{S}},\ i+1,i-1\in R_i. \end{aligned}$$

2. (IC^{CII}) Suppose that in period $t=0$, agent $i-1$ played $D^{i-1,i}$.

   (a) Suppose $\theta(c,v)\geq T_{i-1}-1$. Then nothing changes in the trade-off in his interactions with $i+1$ from IC^{CI}. In his interactions with $i-1$, i will already have returned to the cooperative phase, which means he will give up $c^{i,i-1}$ for T_i periods by infecting $i+1$. Thus, i is in the same situation as if he never had been cheated on by $i-1$, which means $IC^{CII}=IC^{CI}$.

   $$IC^{CII}=IC^{CI} \qquad if\ \theta(c,v)\geq T_{i-1}-1,$$

   (b) Suppose now $\theta(c,v)<T_{i-1}-1$. Again nothing changes in the trade-off in his interactions with $i+1$ from IC^{CI}. Thus the first line of IC^{CII} coincides with the first line in IC^{CI}. If in $t=1$, agent i plays $D^{i,i+1}$ instead of sticking to cooperation and just sending a message, this results in agent $i+1$ sending a message that reaches agent $i-1$ in $t=\theta(c,v)+1$. This yields agent i a utility of $l^{i,i-1}$ until $t=\theta(c,v)+T_i+2$. By sticking to cooperation, she would have had a utility of $w^{i,i-1}$ from $t=\theta(c,v)+1$ until $t=T_{i-1}$ and of $c^{i,i-1}$ from $t=T_{i-1}+1$. This difference constitutes the second and third line of IC^{CII}.

$$\begin{aligned} IC^{CII} \equiv &\left(c^{i,i+1}-w^{i,i+1}\right)+\sum_{t=1}^{T_i}\delta^t\left(c^{i,i+1}-l^{i,i+1}\right) \\ &+\sum_{t=\theta(c,\nu)+1}^{T_{i-1}-1}\delta^t\left(w^{i,i-1}-l^{i,i-1}\right)+\sum_{t=T_{i-1}}^{\theta(c,\nu)+T_i}\delta^t\left(c^{i,i-1}-l^{i,i-1}\right)\geq 0 \\ &\forall i\in N^{\mathcal{S}},\ i+1,i-1\in R_i \qquad if\ \theta(c,v)<T_{i-1}-1, \end{aligned}$$

Since

$$IC^{CI} - IC^{CII} = \begin{cases} \sum_{t=\theta(c,\nu)}^{T_{i-1}-1} \delta^t \left(c^{i,i-1} - w^{i,i-1}\right) < 0 & \forall \theta(c,v) < T_{i-1} - 1 \\ 0 & \forall \theta(c,v) \geq T_{i-1} - 1 \end{cases},$$

whenever IC^{CI} holds, IC^{CII} is satisfied.

3. (IC^P) Suppose agent i receives the message that agent $i+1$ deviated in their relation with one of their other neighbors. Then agent i has to have an incentive to punish him. Since $w^{i,j} > c^{i,j}$ together with $\left(IC^{CI}\right)$, this is always the case.

4. (IC^{LP}) Suppose in period $t = 0$, agent i played $D^{i,i+1}$. Then he has to agree to playing $\left(C^{i,i+1}, D^{i+1,i}\right)$ for T_i periods instead of his minimax strategy forever. After having played $D^{i,i+1}$ in $t = 0$, for agent i sticking to punishment strategies means incurring $l^{i,i+1}$ for T_i periods and $c^{i,i+1}$ thereafter. It furthermore means $w^{i,i-1}$ in $t = \theta(c,v)$, $l^{i,i-1}$ for the following T_i periods and $c^{i,i-1}$ thereafter. Deviating from punishment strategies yields $d^{i,i+1}$ forever, $w^{i,i-1}$ in $t = \theta(c,v)$ and $d^{i,i-1}$ forever thereafter. The difference between these utilities is represented by $\left(IC^{LP}\right)$.

$$\begin{aligned} IC^{LP} \equiv & \sum_{t=0}^{T_i-1} \delta^t \left(l^{i,i+1} - d^{i,i+1}\right) + \sum_{t=T_i}^{\infty} \delta^t \left(c^{i,i+1} - d^{i,i+1}\right) \\ & + \sum_{t=\theta(c,\nu)}^{\theta(c,\nu)+T_i} \delta^t \left(l^{i,i-1} - d^{i,i-1}\right) + \sum_{t=\theta(c,\nu)+T_i+1}^{\infty} \delta^t \left(c^{i,i-1} - d^{i,i-1}\right) \geq 0 \\ & \forall i \in \mathcal{N}^{\mathcal{S}},\ i+1, i-1 \in R_i. \end{aligned}$$

Constraint $\left(IC^{CI}\right)$ consists of addends that are either strictly increasing in δ or strictly positive. Constraint $\left(IC^{LP}\right)$ is strictly increasing in δ for $\delta \in (0,1)$. Both conditions do not hold for a δ close to 0. They do hold strictly for a δ close enough to 1, thus there exists a $\widetilde{\delta}$ for which both constraints hold. Therefore under the conditions stated, strategy (**S3**) is subgame perfect for $\delta > \widetilde{\delta}$.

Since $l^{i,j} < d^{i,j}$, it is possible to fix a T_i $\forall i$ such that $IC^{LP} = 0$[22]. Given that T_i, assume $v = 1$, such that $\theta(c,v) = c - 2$. For this, IC^{CI} is satisfied for all δ that satisfy $\delta^{c-2} g^{i,i-1} + g^{i,i+1} \geq 0$. Now consider $v > 1$. Again, it is possible to fix a T_i $\forall i$ such that $IC^{LP} = 0$. That ensures the same strength of the punishment. But now the punishment in the non-deficient relation sets in earlier which reduces the value of the deviation and therefore for $v > 1$, $\widetilde{\delta} < \underline{\delta}$.

Since agents are being rewarded for punishing their neighbor, they always have an incentive to do so during a punishment phase even if they want to cooperate bilaterally, which establishes (ii).

If T_i is chosen for each agent i such that punishment is as hard as playing minimax strategies with both neighbors forever, this is the hardest punishment possible. Since here information transmission

[22] That means that the punishment is as strong as if the deviator was punished with infinite reversion to the static Nash equilibium.

is used, every mean to decrease the delay before punishment on both sides sets in is used. This establishes the corollary. *Q.E.D.* ■

E Belief structure and sustainability conditions for section 4, information regime (I2)

For networks (a), (b), and (c) from figure 12, we assume the following beliefs:

For agents $j \notin \{i, k\}$, beliefs are such that

(i) if they observe cooperation on both sides, they believe that all agents in the network cooperated so far,

(ii) if they observe a deviation on both sides, they believe that the neighbor with whom they share their deficient relation was the first to deviate, and

(iii) if they observe a deviation only from the agent with whom they share their non-deficient relation, they give an equal probability to the event that any of the other players was the first to deviate.

For agents i and k, beliefs are such that

(iv) if they observe cooperation from all neighbors, they believe that all agents in the network cooperated so far,

(v) if they observe a deviation by all neighbors, they believe that everybody in the network deviated,

(vi) if i (if k) observes agent $i-1$ (agent $k-1$) deviate, but the other neighbors cooperate, agent i (agent k) gives an equal probability to the event that any agent $j \in \{k, k+1, ..., i-1\}$ (any agent $j \in \{i, i+1, ..., k-1\}$) was the first to deviate,

(vii) if i (if k) observes agents $i-1$ and k (agents $k-1$ and i) deviate, but the other neighbor cooperate, he believes that agent k (agent i) was the first to deviate,

$(viii)$ if i (if k) observes agent k, agent $i+1$, or both, agents k and $i+1$, (agent i, agent $k+1$, or both, agents i and $k+1$) deviate, but the other neighbors cooperate, agent i (agent k) gives an equal probability to the event that any agent $j \in \{i+1, i+2, ..., k\}$ (any agent $j \in \{k+1, k+2, ..., i\}$) was the first to deviate, and

(ix) if i (if k) observes agents $i-1$ and $i+1$ (agents $k-1$ and $k+1$) deviate, but the other neighbor cooperate, agent i (agent k) gives an equal probability to the event that any agent $j \in \mathcal{N}^{\mathcal{S}} \setminus i$ (any agent $j \in \mathcal{N}^{\mathcal{S}} \setminus k$) was the first to deviate.

Let $\mathcal{N}^{\mathcal{S}} \setminus ik$ be of size c and the subnetwork $\{i, i+1, ..., k-1, k, i\}$ be of size m. Then for the beliefs given, information structure (**I2**), and $l^{i,k}$ and $l^{k,i}$ low $\mathcal{N}^{\mathcal{S}}$ is sustainable iff

$$g^{i,i+1} + \delta^{m-2}\left(g^{i,k} + \delta^{c-m} g^{i,i-1}\right) \geq 0$$

$$g^{k,k+1} + \delta^{c-m}\left(g^{k,i} + \delta^{m-2} g^{k,k-1}\right) \geq 0$$

$$g^{j,j+1} + \delta^{m-2} g^{j,j-1} \geq 0 \quad \forall j \in \{i+1, ..., k-1\}$$

$$g^{j,j+1} + \delta^{c-m} g^{j,j-1} \geq 0 \quad \forall j \in \{k+1, ..., i-1\}$$

F Sustainability conditions for agent i in section 4, information regime (I3)

Refer to the figure in appendix E. We give the conditions exemplary for agent i.

1. $\left(IC_i^{CI}\right)$ During a cooperation phase, it must be profitable for i to play $C^{i,i+1}, C^{i,k}, C^{i,i-1}$ at any time, which yields $c^{i,i+1}$, $c^{i,k}$, and $c^{i,i-1}$ in each period, instead of choosing his best deviation ("static" best reply), which would be to play $D^{i,i+1}$ in $t=0$, $D^{i,k}$ in $t=\theta(m,\nu)$, and $D^{i,i-1}$ in $t=\theta(c,\nu)$ and then to face a T_i- period punishment during which he has to endure payoffs of only $l^{i,i+1}$, $l^{i,k}$, and $l^{i,i-1}$. Such a deviation is not profitable iff

$$\begin{aligned} IC_i^{CI} \equiv \left(c^{i,i+1} - w^{i,i+1}\right) &+ \sum_{t=1}^{T_i} \delta^t \left(c^{i,i+1} - l^{i,i+1}\right) \\ &+ \delta^{\theta(m,\nu)}\left(c^{i,k} - w^{i,k}\right) + \sum_{t=\theta(m,\nu)+1}^{\theta(m,\nu)+T_i} \delta^t \left(c^{i,k} - l^{i,k}\right) \\ &+ \delta^{\theta(c,\nu)}\left(c^{i,i-1} - w^{i,i-1}\right) + \sum_{t=\theta(c,\nu)+1}^{\theta(c,\nu)+T_i} \delta^t \left(c^{i,i-1} - l^{i,i-1}\right) \geq 0. \end{aligned}$$

2. $\left(IC_i^{CII}\right)$ Suppose that agent $i-1$ deviated in $t=-1$. Agent i has to have an incentive to pass on this information in $t=0$ to both his neighbors, $i+1$ and k, instead of infecting his neighbors $i+1$ in $t=0$ and k in $t=\theta(m,\nu)$ and then facing the punishment prescribed against himself. Again, we have to distinguish two cases depending on the speed of information transmission.

   (a) If $T_{i-1} - 1 < \theta(c,\nu)$, then the information that i did not pass on the info, but cheated instead against $i+1$, reaches $i-1$ *after* i and $i-1$ have gone back to cooperation. Therefore,
   $$IC^{CII} = IC^{CI} \qquad \forall \theta(c,v) \geq T_{i-1} - 1.$$

   (b) If $T_{i-1} - 1 \geq \theta(c,v)$, then the information that i did not pass on the info, but cheated instead against $i+1$, reaches $i-1$ *after* i and $i-1$ have gone back to cooperation.

That means that i looses punishment profits $w^{i,i-1}$ for a number of periods equal to the difference between $T-1$ and $\theta(c,\nu)$. Therefore,

$$
\begin{aligned}
IC_i^{CII} \equiv \left(c^{i,i+1} - w^{i,i+1}\right) + \sum_{t=1}^{T_i} \delta^t \left(c^{i,i+1} - l^{i,i+1}\right) \\
+ \delta^{\theta(m,\nu)} \left(c^{i,k} - w^{i,k}\right) + \sum_{t=\theta(m,\nu)+1}^{\theta(m,\nu)+T_i} \delta^t \left(c^{i,k} - l^{i,k}\right) \\
+ \sum_{t=\theta(c,\nu)+1}^{T_{i-1}-1} \delta^t \left(w^{i,i-1} - l^{i,i-1}\right) + \sum_{t=T_{i-1}}^{\theta(c,\nu)+T_i} \delta^t \left(c^{i,i-1} - l^{i,i-1}\right) \geq 0 \\
\forall \theta(c,v) < T_{i-1} - 1.
\end{aligned}
$$

Again, we see that

$$
\left(IC^I - IC^{II}\right) = \begin{cases} \sum_{t=\theta(c,\nu)}^{T_{i-1}-1} \delta^t \left(c^{i,i-1} - w^{i,i-1}\right) < 0 & \forall \theta(c,v) \geq T_{i-1} - 1 \\ 0 & \forall \theta(c,v) < T_{i-1} - 1 \end{cases}.
$$

Thus, $\left(IC^I\right)$ holds implies that $\left(IC^{II}\right)$ holds. Agent i also always has an incentive to punish a deviator immediately, thus, the equivalent to $\left(IC^P\right)$ always holds. We have to verify that $\left(IC^{LP}\right)$ holds.

3. $\left(IC^P\right)$ Suppose agent i receives the message that agent $i+1$ (agent k) deviated in their relation with one of their other neighbors. Then agent i has to have an incentive to punish them. Since $w^{i,j} > c^{i,j}$ together with $\left(IC^{CI}\right)$, this is always the case.

4. $\left(IC^{LP}\right)$ Lastly, agent i has to have an incentive to let his neighbors carry out the punishment on him if he deviated. He can ensure himself a payoff of $d^{i,i+1}$, $d^{i,k}$, and $d^{i,i-1}$ forever by playing $D^{i,i+1}$, $D^{i,k}$, and $D^{i,i-1}$ forever. This limits the punishment available to the community.

$$
\begin{aligned}
IC_i^{LP} \equiv \sum_{t=0}^{T_i-1} \delta^t \left(l^{i,i+1} - d^{i,i+1}\right) + \sum_{t=T_i}^{\infty} \delta^t \left(c^{i,i+1} - d^{i,i+1}\right) \\
+ \sum_{t=\theta(m,\nu)+1}^{\theta(m,\nu)+T_i} \delta^t \left(l^{i,k} - d^{i,k}\right) + \sum_{t=\theta(m,\nu)+T_i+1}^{\infty} \delta^t \left(c^{i,k} - d^{i,k}\right) \\
+ \sum_{t=\theta(c,\nu)+1}^{\theta(c,\nu)+T_i} \delta^t \left(l^{i,i-1} - d^{i,i-1}\right) + \sum_{t=\theta(c,\nu)+T_i+1}^{\infty} \delta^t \left(c^{i,i-1} - d^{i,i-1}\right) \geq 0
\end{aligned}
$$

By choosing an appropriate T_i, the punishment can again be made as hard as in the contagious equilibrium (with strategies (**S2**) and the respective beliefs). With $\nu > 1$, due to a faster punishment, the discount factor necessary to sustain the network will again be lower than with (**S2**).