

Networks of Relations*

Steffen Lippert[†]

University of Mannheim

Giancarlo Spagnolo[‡]

University of Mannheim and Stockholm School of Economics

First Version February 08, 2002

Incomplete and preliminary, comments welcome

Abstract

We explore sanctioning power within a network of relational contracts under different information technologies depending on the shape of the network. The value of the relational network lies in the enforcement of cooperative agreements which would not be enforceable for the agents without access to the punishment power of other network members. We identify conditions for stability of such networks, conditions for transmission of information about past actions, and conditions under which self-sustainable subnetworks may actually inhibit a stable network.

JEL Codes: D23, D43, L13, L29, O17

Keywords: Networks, Relational Contracts, Collusion, Social Capital.

1 Introduction

In recent economic research, both, the emergence and stability of networks and relational governance mechanisms, have aroused the interest of many theoretical as well as experimental scholars. Being well informed, at the best from themselves well informed partners, is valuable. When agents set up costly links, thereby forming a network, a conflict between efficient and stable networks may arise. This line of research has been started by a seminal

*We have benefited from discussions with Konrad Stahl and seminar participants at the University of Mannheim.

[†]Department of Economics, Applied Microeconomics, University of Mannheim, L 7, 3-5, D - 68131 Mannheim, Germany. email: steffen@econ.uni-mannheim.de.

[‡]Department of Economics, Applied Microeconomics, University of Mannheim, L 7, 3-5, D - 68131 Mannheim, Germany. email: gianca@uni-mannheim.de.

paper by Aumann and Myerson (1988), who show that, even when a network forms sequentially, efficiency and stability are not compatible for their fixed allocation rule. Jackson and Wolinsky (1996), take up that strand of literature in their highly acclaimed paper, modeling the emergence and stability of a social information and communication network when agents choose to set up and maintain or destroy costly links, using the notion of pairwise stability. Considering a value function depending on the communication structure, they show that efficiency and stability are not compatible under mild assumptions, anonymity and component balancedness, for fixed allocation rules. Using this paper as a corner stone, models with architecture-dependent variations in the cost structure and the value function have been built. Bala and Goyal (2000a) consider the setup of a link by one agent only. A spatial cost structure leading to equilibria of locally complete networks has been introduced by Johnson and Gilles (2000). Communication reliability has been made an issue in Bala and Goyal (2000b), who show that given imperfect information transmission both, efficient and equilibrium networks exclusively feature redundant links. The strategic aspect in these models lies in the question of whether to maintain a link or not. A common feature of these models is the assumption of a value function, an allocation rule, and link establishment costs, and thereby the leaving out of the underlying game. We explicitly model an underlying game, which allows us to study consequences of its features for the stability of the network.

To our best knowledge, work that explicitly models the underlying game in the past has largely focused on random matching games. Kandori (1992) and Groh (2002) consider such a random matching repeated prisoners' dilemma situation. Both show how much cooperation is possible, Kandori without information processing and Groh with reliable and unreliable communication in a network. Groh introduces the endogenous decision of players to pass on information. In contrast to Kandori and Groh, we do not consider games between changing partners, but fixed neighbors. This introduces a forward induction element into strategic behavior when defecting. We keep Groh's endogenous decision of players to pass on information on past games' actions. We introduce the possibility to pass on informations received by partners in the underlying game.

Increasing evidence shows that relational arrangements are an important governance mechanism over interactions of economic agents. This is not only the case in developing economies but also in well developed economic frameworks, most prominently in the fast changing environment of high-tech industries. Especially in R&D-intensive industries, many firms enter collaborations in order to trade-off risk and return from these risky activities. But formal arrangements often 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 – excludes

market relations, the need for flexibility on the other hand excludes vertical integration. Annalee Saxenian (1994) reports a highly specialized, network-like vertical organization within the computer-industry in Silicon Valley within which informal relations play a crucial role for the success of the district in comparison with Route 128, a competing district close to Boston: "While they competed fiercely, Silicon Valley's producers were embedded in, and inseparable from, these social and technical networks." Examining 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." The networks themselves form when individuals establish relations. Using their position within the network, and therefore using the network itself for their interests, thus becomes a central issue for those firms.

There is a large literature modeling long-term cooperative relations as an alternative governance mechanism of vertical relations besides market structure and vertical integration. In a series of papers, Kranton and Minehart (2000, 2001) for example investigate the role of networks in vertical industrial relations. Buyers invest in links to sellers of an intermediate input and are thereby able to buy cheaper due to risk pooling on the sellers' aggregate demand or to improve their bargaining position. In most of the studies, though, (see for example North 1990, Williamson 1996, Li 1999, or Baker, Gibbons and Murphy 2001a,b) relational governance has been seen as an inferior mechanism which might be used when market and contractual governance are infeasible. North (1990), Li (1999) Kranton (1996), and Spagnolo (1999a) all predict that relational governance arrangements will break down once economic development is achieved and financial liberalization takes place. This seems to contrast the empirical evidence shown by Saxenian or Powell et al.

We contribute to the literature by combining these two strands of literature – the literature on stability of networks and the one on relational contracts – creating a model of a network of relations. Our work relates to the literature of multimarket contact à la Bernheim and Whinston (1990) and Spagnolo (1999b). In their papers, collusion between two agents is fostered by tying the actions from one relation to the ones in the other relation. Asymmetries of payoffs (Bernheim and Whinston) and payoff interdependencies due to concave payoff functions (Spagnolo 1999b) drive the result. In our model, agents exploit indirect multimarket contact. They maintain relations by using the network that not only consists of their own but also of other agents' relations and thereby pool asymmetries in payoffs.

In this paper, 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 á la Maggi (1999). We consider three information structures. We study first complete information, that is each agent observes the histories of the games for all agents. We consider secondly the case where no information can be transmitted at all. Here each player only observes the histories of his own games with his direct partners. And thirdly we consider two cases where information is being transmitted through the network of relations, that is in addition to observing his own history, in each period each agent transmits or receives a verified message to/from each of his partners about the histories of their games and about messages they received. We assume that it takes (a) one period or (b) a smaller number of periods for such an information to travel from one agent to the other, therefore with a delay, an agent may be informed about all other players' actions to whom he is connected in the network. However, we always require agents to be willing to pass on information, that is shouting – informing one's neighbor's neighbors is not allowed for.

We begin with sustainable network where agents can only have relations with two neighbors. We show that if agents cannot discipline themselves within a certain relation, "redundant" links, that is circularity, can sustain the relation. In contrast to Groh, the possibility to transmit information about the cheating of someone through the links in the network will not be an equilibrium action if enforcement relies on optimal punishment. Once an agent deviates, a contagious process eliminates cooperation in the network. With more complex punishment strategies, agents may use information transmission, thereby keeping on cooperating in the rest of the network while punishing the deviator. We show that, under the complete information assumption, bilaterally unsustainable relations in a network without "redundant" links, may can be sustained by having self-sustaining relations at the ends of the network while this does not work for the other informational assumptions. Thirdly we show that having self-sustaining relations in the network may actually hurt cooperation in the case without full information because agents might not be willing to perform the punishment. In this case a network may be sustainable if we use less severe punishments than grim trigger or by rewarding the punisher. As opposed to standard results in the literature, in our model, improved outside options, possibly by more efficient spot markets, for one player may under certain conditions actually foster cooperation by making the breakup of a relation in the case of a deviation a credible threat. The results are finally generalized to more complex network architectures where players may have more than two neighbors.

The paper starts with the definition of a network of relations in section 2. We then derive results for sustainable networks with the restriction of at most two neighbors in section 3, extend these results to situations with more neighbors in section 4. Section 5 concludes.

2 The Model

2.1 Interaction

There is a set of infinitely lived agents $\mathcal{N} = \{1, \dots, n\}$, with $i \in \mathcal{N}$. Each agent i chooses from the set of actions A^i he can take at time t

$$A^i = \times_{j \in \mathcal{N} \setminus i} A^{ij} \quad \text{with} \quad a^{ij} \in A^{ij},$$

where A^{ij} is the set of actions that i can take with respect to agent j . An action profile is then $A \equiv \times_{i \in \mathcal{N}} A^i$. The actions taken by each individual may, but do not have to be verifiable. An action profile at time t is denoted by A_t . Per period payoffs are a real valued function $U^i : A \mapsto \mathfrak{R}$, $i \in \mathcal{N}$

$$U^i(A), \quad i \in \mathcal{N}.$$

For simplicity, we assume the following restriction on the set of actions and the payoffs resulting from different action profiles.

Assumption 1 *For each pair of sets of actions $(A^{ij}, A^{ji} | \bullet)$ we assume the agents to play a prisoners' dilemma that is $A^{ij} \equiv \{C^{ij}, D^{ij}\}$ with D^{ij} being the dominant action and $(D^{ij}, D^{ji} | \bullet)$ the static Nash equilibrium and*

$$U^i(C^{ij}, D^{ji} | \bullet) < U^i(D^{ij}, D^{ji} | \bullet) < U^i(C^{ij}, C^{ji} | \bullet) < U^i(D^{ij}, C^{ji} | \bullet), \quad \forall i, j \in \mathcal{N}.$$

Agents are assumed to interact repeatedly with a positive probability of continuation. Time is discrete, and all agents are assumed to share a discount factor δ , meant to capture both, time preferences and a potential breakdown of a relation.

Assumption 2 *We assume time separability of agents' payoffs, that is, in the repeated game, agents choose action which maximize their average discounted payoff (utility)*

$$V^i = \sum_{t=0}^{\infty} \delta^t U_t^i(A_t).$$

In addition to additive separability of payoffs across time, let us suppose additive separability of payoffs across "relations", that is

$$U^i(A) = U^{i1}(a^{i1}, a^{1i}) + \dots + U^{in}(a^{in}, a^{ni}).$$

Notation 1 For notational convenience let us define

$$g^{ij} \equiv U^i (C^{ij}, C^{ji} | \bullet) - \delta U^i (D^{ij}, D^{ji} | \bullet) - (1 - \delta) U^i (D^{ij}, C^{ji} | \bullet)$$

and

$$g_T^{ij} \equiv g^{ij} - \delta^{T+1} [U^i (D^{ij}, D^{ji} | \bullet) - U^i (C^{ij}, C^{ji} | \bullet)].$$

A standard interpretation of g^{ij} is the net gains for i from cooperating with j considering grim trigger strategies according to Friedman (1971). Therefore, if $g^{ij} > 0$, i does not have an incentive to deviate in an infinitely repeated prisoners' dilemma with trigger strategies. However, a $g^{ij} < 0$ 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, unfeasible. The term g_T^{ij} expresses the equivalent with a T -period reversion to the static Nash equilibrium.

2.2 Relations and Networks

Definition 1 (*relation*) Agents i and j are connected by a relation if and only if they repeatedly cooperate in the stage game.

Cooperation is defined as the coordination on actions superior to the static Nash equilibrium. We call a relation deficient if it is not sustainable on a bilateral basis.

Definition 2 (*network*) A network $\mathcal{N}^S = (\overline{\mathcal{N}}, R)$ is a directed graph¹ consisting of the finite nonempty subset $\overline{\mathcal{N}}$ of agents in \mathcal{N} together with a set R of two element ordered subsets of $\overline{\mathcal{N}}$, where $(i, j) \in R$ iff i and j are connected by a relation and $g^{ij} > 0$.

A convenient form to represent such a network \mathcal{N}^S is an $n \times n - (0, 1)$ -adjacency matrix. Each column and each row of this matrix corresponds to a distinct agent in \mathcal{N}^S . The (i, j) -th element of the adjacency matrix notation of \mathcal{N}^S is denoted by

$$n^{ij} = \begin{cases} 1 & \text{if } (i, j) \in R \\ 0 & \text{otherwise} \end{cases}.$$

We will comply with the convention that agents cannot have a relation with themselves, that is $(i, i) \notin R$ or $n^{ii} = 0 \forall i = 1, 2, \dots, n$. A second way to represent such a network is of course graphical, where an arc is drawn from agent j to agent i if $(i, j) \in R$.

¹A directed **graph** $G = (V, E)$ is a finite nonempty set V of elements called **vertices**, together with a set E of two element ordered subsets of V called **edges or arcs**.

Figure 1 shows such a network of relations where $g^{12}, g^{21}, g^{24}, g^{32}$ and $g^{43} > 0$ and the corresponding adjacency matrix

$$\mathcal{N}^S = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

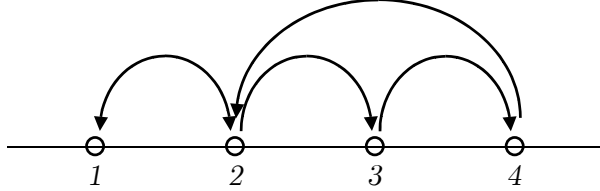


Figure 1: Relational network with four agents

Agent i is called adjacent from agent j and agent j adjacent to agent i if $(i, j) \in R$. Two agents are called directly connected in the social network (or adjacent) if either $(i, j) \in R$ or $(j, i) \in R$ or both. The set of agents adjacent to or from i are the neighborhood of i , denoted by N_i , and $j \in N_i \Leftrightarrow i \in N_j$.

Given $\mathcal{N}^S = (\overline{\mathcal{N}}, R)$, the number of agents in $\overline{\mathcal{N}}$ 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 away from agent i is called the outdegree of agent i and denoted by $\text{od } i$, and the number of arcs directed into agent i is called the indegree of agent i , denoted by $\text{id } i$. The degree of vertex i is the number of arcs directed away or into agent i , denoted $\text{deg } i = \text{od } i + \text{id } i$. An agent of degree 1 is called end vertex. In figure 1, 1 is an end vertex, $\text{deg } 2 = 3$, $\text{id } 2 = 2$, $\text{od } 2 = 1$.

Let i and j be two agents of \mathcal{N}^S . A $i - j$ (*directed*) walk in \mathcal{N}^S is a finite alternating sequence of agents and links that begins with agent i and ends with agent j and in which each link (*arc*) in the sequence joins the agent that precedes it in the sequence to the agent that follows in the sequence. The number of links in an $i - j$ walk is length $l(i, j)$ of the $i - j$ walk. A path connecting i_1 and i_k is an $i_1 - i_k$ walk in which no agent is repeated. An $i - j$ walk is closed if $i = j$ and open otherwise. A closed path is a cycle. A network of order c that consists only of a cycle is called the c -cycle. If a network contains no cycles, it is called acyclic. The network \mathcal{N}^S is called circular if there exists a path $\{i_1, i_2, \dots, i_k\}$ with $k = 1$.

A network is strongly connected if for any agent in \mathcal{N}^S there exists a directed path from that agent to any other agent in \mathcal{N}^S . It is weakly connected if when we remove the orientation from the arcs in \mathcal{N}^S , there exists a path between every pair of agents in the underlying undirected network. \mathcal{N}^S is called complete on \mathcal{N} if there exists a direct connection between each member of \mathcal{N} .

Definition 3 (*sustainable*) A relational network $\mathcal{N}^S = (\overline{\mathcal{N}}, R)$ is sustainable iff all relations are simultaneously supportable in Subgame Perfect Nash Equilibrium between the agents in $\overline{\mathcal{N}}$.

2.3 Information structures

We will consider the following mechanisms for the transmission of information between agents. Let H^{ij} be the set of histories in the relation between agents i and j with $(a_t^{ij}, a_t^{ji})_{t=1, \dots, T} \in H^{ij}$.

(I1) Complete Information: At time τ , each agent $i \in \mathcal{N}^S$ observes $(a_t^{mn})_{t=1, \dots, \tau} \in H^{mn}$
 $\forall m, n \in \mathcal{N}^S$.

Each agent observes the histories of the games for all agents.

(I2) No Information Transmission: At time τ , each agent $i \in \mathcal{N}^S$ observes $(a_t^{ij}, a_t^{ji})_{t=1, \dots, \tau} \in H^{ij} \forall j \in N_i$.

Each agent only observes the histories of his own games with his direct opponents.

Kandori (1992) showed for a random matching game where agents play a prisoners' dilemma, that there exists a contagious equilibrium given the players are sufficiently patient and the loss from meeting a cheating player when playing "Cooperate" is large enough. In this equilibrium, the information that a player deviated is "transmitted" to future matches by defecting on them. Deviation is deterred by the loss of cooperation gains with future matches due to the contagious process.

Our connected network differs from Kandori's setup in assigning probabilities of being matched to a certain partner of either one or zero: Either partners are connected, which means they will interact again with a positive probability, or they are not, which means future interaction with probability zero. When we are dealing with a circular network or with circular subnetworks, there may be deterrence of deviation by the fact that non-cooperation to one neighbor may translate into loss of future cooperation gains with the other due to a contagious process.

(I3) Network Information Transmission: At time τ , each agent $i \in \mathcal{N}^S$ observes $(a_t^{ij}, a_t^{ji})_{t=1, \dots, \tau} \in H^{ij} \forall j \in N_i$ and $(a_t^{mn}, a_t^{nm})_{t=1, \dots, \tau-l} \in H^{mn}$, $m \in N_n$, where $\min[l(i, m), l(i, n)] = l$ if there exists an $i - m$ path.

In information structure (I3), in addition to observing his own history, in each period each agent i transmits or receives verifiable messages to/from each agent $j \in N_i$ about the

histories of their relations or about messages they received. It takes 1 period for such an information to travel from one agent to the other.

In a relational context, it seems natural to assume that there does not exist an institution that gathers and disseminates truthfully any information concerning the behavior of network members. Suppose instead, that information can only be transmitted through personal contact of members of the network and that each transmission takes time, e.g. one period. Information transmission is being delayed and therefore punishment will set in at a later point in time. Therefore, a higher discount factor δ , that is more patience of agents, will be necessary to sustain the network. In a sense one could think of this as the cost of using the network for information transmission, the cost being that using a contact to transmit information consumes time.

It might sound more realistic that agents can only meet one neighbor at a time, that is per period, and exchange information only with one neighbor while meeting. However, the main essence coming from that – being that agent who transmit information have to have an incentive to do so and that there is a delay – is captured by this more simple information transmission structure as well, which is why we kept it.

(I4) Fast Network Information Transmission: At time τ , each agent $i \in \mathcal{N}^S$ observes $(a_t^{ij}, a_t^{ji})_{t=1, \dots, \tau} \in H^{ij} \forall j \in N_i$ and $(a_t^{mn}, a_t^{nm})_{t=1, \dots, \tau - \text{int}[\frac{1}{v}]} \in H^{mn}$, $m \in N_n$, where $\min[l(i, m), l(i, n)] = l$ if there exists an $i - m$ path.

As opposed to **(I3)**, under information structure **(I4)** it takes only $\text{int}[\frac{1}{v}]$ periods for such an information to travel from one agent to the other or – put differently – a message can travel over v links per period.

To assume that in networks of relations communicating besides interacting is not costly is reasonable since very often chatting next to everyday business – if anything – gives pleasure to agents.

3 Sustainable networks of degree 2

In order to gain a first insight let us restrict the number of neighbors a player may have and the payoff space to be identical across individuals.

Assumption 3 Assume $\deg i \leq 2 \forall i \in \mathcal{N}^S$, that is each agent can have at most 2 neighbors.

Definition 4 (mutual, non-mutual, mixed) A relational network of degree less or equal than 2, is (T-) non-mutual iff $g_{(T)}^{ij} > 0 \Leftrightarrow g_{(T)}^{ji} \leq 0$. It is (T-) mutual iff $g_{(T)}^{ij} > 0 \forall i \in \mathcal{N}^S \forall j \in N_i$. It is (T-) mixed if it consists of both, (T-) mutual and (T-) non-mutual subnetworks.

Definition 5 (*directed*) A relational network of degree less or equal than 2, is (T-) directed if $g_{(T)}^{i,i-1} \leq 0 \Rightarrow g_{(T)}^{i,i+1} > 0$.

3.1 Non-mutual networks

In the theory of repeated games it is stated that in two-player repeated prisoners' dilemmas, in order to sustain a cooperative outcome as a SPNE, it is necessary that the gain from deviating net of the loss from punishment must be outweighed by the gain agents incur from cooperating for ever. Now, the condition of non-mutuality rules out such cooperation in bilateral games. Agents would not cooperate with each other since the most severe bilateral punishment available is not strong enough and there are no other agents to discipline them.

3.1.1 Benchmark results – non-circular directed networks

Lemma 1 Under assumptions 1 – 3, no non-mutual non-circular network is supportable in SPNE, independent of the discount factor and the information structure.

Proof. A network has been defined non-circular if no agent $i_1 \in \mathcal{N}^S$ there exists a path $\{i_1, i_2, \dots, i_k\}$ 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 $\text{od } e = 1$ or an agent m in the middle with $\text{od } m = 2$. Since we assumed $\text{deg } i \leq 2$, 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. ■

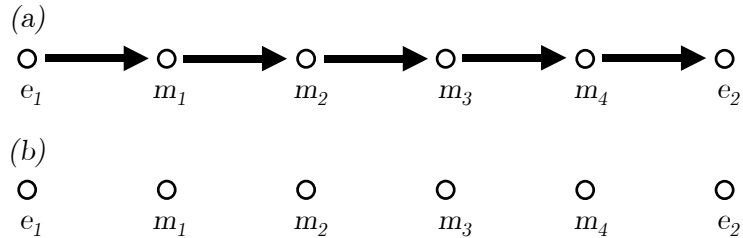


Figure 2: Only the empty network (b) is sustainable

Figure 2 illustrates Lemma 1. Part (a) shows a network out of equilibrium. In that situation, agent e_1 always has an incentive to deviate and the only sustainable network is empty, as shown in (b).

3.1.2 Circular directed networks

We will first be considering circular networks in which there is a one and only one cycle. We define the following strategies:

(S1) Each agent $i \in \mathcal{N}^S$ starts by playing the agreed upon action vector $C^{ij} \forall i \in \mathcal{N}^S, \forall j \in N_i$, and goes on playing that action towards every $\forall j \in N_i$ as long as no deviation is observed.

Every agent i reverts to the stage Nash equilibrium in his interactions for ever if a deviation in the network occurred.

(S2) Each agent $i \in \mathcal{N}^S$ starts by playing the agreed upon action vector $C^{ij} \forall i \in \mathcal{N}^S, \forall j \in N_i$, and goes on playing that action towards every $\forall j \in N_i$ as long as no one in his neighborhood deviated towards him.

Every agent i reverts to the static Nash equilibrium in all his interactions if a deviation is experienced.

Proposition 1 *Suppose the network is a c-cycle and assumptions 1 - 3 apply. Then*

1. *under information structure (I1), a directed non-mutual relational network is supportable if and only if $\forall i \in \mathcal{N}^S g^{i,i-1} + g^{i,i+1} > 0$,*
2. *under information structures (I2), a directed non-mutual relational network is supportable if and only if $\forall i \in \mathcal{N}^S \delta^{c-2} g^{i,i-1} + g^{i,i+1} > 0$, where $g^{i,i+1} < 0$.*

Figure 3 visualizes a directed non-mutual network which is a cycle. Every agent has

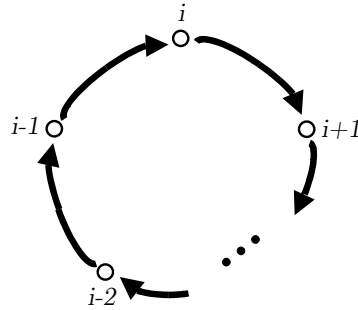


Figure 3: Circular directed non-mutual network

two neighbors, $\deg j = 2, \text{od } j = 1, \text{id } j = 1 \forall j \in \mathcal{N}^S$, which implies both, directedness and non-mutuality.

Proof. Part 1: Sufficiency: Consider strategies (S1) Since a deviator faces immediate Nash-reversion from both his neighbors, no matter whether she deviates towards one partner or both, she can just as well deviate from both her relations. Therefore, the network is a Nash-Equilibrium in a circular network if $\forall i g^{i,i-1} + g^{i,i+1} > 0$. It is subgame perfect since in the punishment phase, the stage Nash equilibrium is played.

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. If cooperation is not possible with these strategies, it will not be possible with other – less strong – punishments. Therefore $g^{i,i-1} + g^{i,i+1} > 0 \forall i \in \mathcal{N}^S$ is also necessary for the relational network to be supportable.

Part 2: Sufficiency: Consider strategies **(S2)**. An agent might want to deviate only towards one neighbor in the first period and continue cooperating with the other neighbor until the period in which this other neighbor is being communicated the deviation of i in his interaction with the first neighbor. If at all, the agent would sensibly first deviate from his 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. This would be after $c - 2$ periods. Therefore deviation will not be profitable if

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

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 rational and the equilibrium is subgame perfect.

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. If cooperation is not possible with these strategies, it will not be possible with other – less strong – punishments. Therefore $\delta^{c-2}g^{i,i-1} + g^{i,i+1} > 0 \forall i \in \mathcal{N}^S$ is also necessary for the relational network to be supportable. ■

If one does not assume non-mutuality but a mixed network instead, part 2 of proposition 1 does not hold anymore. We will explore this in the next paragraph on mutual and mixed networks.

Corollary 1 *Under (I2) and (I3), transmitting information over the network is not an equilibrium action if we restrict to grim trigger punishment.*

Proof. Information transmission from i would start in the event of a deviation of agent $i - 1$. Since there will never be a return to cooperation with $i - 1$, the best response of i in his remaining deficient relation would be to deviate from that relation. ■

Let us define the following strategy:

(S3) Agents start by playing $C^{ij} \forall i \in \mathcal{N}^S, \forall j \in N_i$, and go on playing $C^{ij} \forall j \in N_i$ as long as no $j \in N_i$ played D^j .

If agent $i - 1 \in N_i$, $g^{i,i-1} < 0$, plays $D^{i-1,i}$, i tells $i + 1$ about it and plays $D^{i,i-1}$ until $i - 1$ and i played $(D^{i,i-1}, C^{i-1,i})$ for T periods. After that i tells $i + 1$ about it and they go back to playing $C^{i,i-1}, C^{i-1,i}$.

If agent receives an information about past actions of other players, he passes this on to his other neighbor.

If an agent $i - 2$ receives the information that $i - 1 \in N_{i-2}$ deviated from his other relation, she plays $D^{i-2,i-1}$ until she receives the information that $(D^{i,i-1}, C^{i-1,i})$ has been played for T periods. She returns to playing $C^{i-2,i-1}$ afterwards.

If agent $i - 1 \in N_{i-2}$, $g^{i-2,i-1} > 0$ plays $D^{i-1,i-2}$, and agent $i - 2$ did not receive a message about $i - 1$'s deviation from his relation with i , agent $i - 2$ assumes a mistake and plays D for T periods.

Proposition 2 *Strategy (S3) constitutes a sustainable network if*

$$g_T^{i,i+1} + (\delta - \delta^{T+2}) [U^i (D^{i,i+1}, D^{i+1,i}) - U^i (C^{i,i+1}, D^{i+1,i})] + \delta^{c-2} g_T^{i,i-1} > 0 \quad \forall i \in \mathcal{N}^S, i+1, i-1 \in N_i$$

and

$$g_T^{i,i+1} + (\delta - \delta^{T+2}) [U^i (D^{i,i+1}, D^{i+1,i}) - U^i (C^{i,i+1}, D^{i+1,i})] + \delta^{c-2} (1 - \delta^{T+1}) [U^i (D^{i,i-1}, C^{i-1,i}) - U^i (D^{i,i-1}, D^{i-1,i})] + \delta^{c-2} \delta^{T+1} (1 - \delta) [U^i (C^{i,i-1}, C^{i-1,i}) - U^i (D^{i,i-1}, D^{i-1,i})] > 0. \quad \forall i \in \mathcal{N}^S, i+1, i-1 \in N_i$$

Proof. For each agent i , playing $D^{i,i+1}$ in $t = 0$ and $D^{i,i-1}$ in $t = c - 2$, which is her best deviation, yields $U^i (D^{i,i+1}, C^{i+1,i})$ in $t = 0$, $U^i (C^{i,i+1}, D^{i+1,i})$ for the following T periods and $U^i (C^{i,i+1}, C^{i+1,i})$ thereafter, as well as $U^i (C^{i,i-1}, C^{i-1,i})$ until $t = c - 3$, $U^i (D^{i,i-1}, C^{i-1,i})$ in $t = c - 2$, $U^i (D^{i,i-1}, D^{i-1,i})$ for the following T periods and $U^i (C^{i,i-1}, C^{i-1,i})$ thereafter. Playing $C^{i,i+1}$ and $C^{i,i-1}$ forever yields $\frac{1}{1-\delta} [U^i (C^{i,i+1}, C^{i+1,i}) + U^i (C^{i,i-1}, C^{i-1,i})]$. Summing up leads to the first inequality in Proposition 2, which is the condition for (S3) to be a Nash equilibrium.

After observing $D^{i-1,i}$ in $t = 0$, a deviation, that is playing $D^{i,i+1}$, yields the same payoffs from the interactions with $i + 1$ as in the non-punishment subgame. Thus the first part of the second inequality in proposition 2. The consequence of this deviation for the interactions with $i - 1$ are the loss of the deviation profits for $T - 1$ periods and the loss of cooperation profits for one period. Thus the second part of the second inequality in proposition 2. Transmission of information received does not alter the payoffs of an agent, that is not doing

it is not profitable. Therefore under the condition stated, strategy **(S3)** is subgame perfect.

■

Proposition 2 shows that it is not necessary to have a complete breakdown of cooperation in the network if information about past actions can be transmitted. However, since under **(I3)** information travels with the same speed as contagion, the possibility to transmit it does not increase the set of sustainable equilibria as opposed to **(S2)**. Assuming **(I4)**, one should easily be able to come up with a modified strategy profile **(S3')** that does support more cooperative equilibria than **(S2)**.

3.2 Mutual and mixed networks

We state without proof:

Lemma 2 *Any mutual network is sustainable.*

Starting from lemma 2, we will now check what happens to the negative results in lemma 1 and proposition 1 if we allow for subnetworks to be mutual. As we will see, there will be both, cooperation-enhancing and cooperation-repressing effects coming from mutual subnetworks.

3.2.1 Mixed non-circular directed networks

Corollary 2 *Suppose the conditions stated in proposition 1 hold and there exists a mutual subnetwork in \mathcal{N}^S . Then under information structure **(I1)**, a directed relational network is supportable as a SPNE under the condition given in proposition 1.*

Proof. The optimality of the actions during a punishment phase proposed in part 1 of the proof of proposition 1 only depended on the fact that the strategies played by the deviator and his neighbors were in fact a stage Nash equilibrium. Thus optimality of the actions during a punishment phase did not depend on intertemporal incentives of any agent and so mutuality cannot destroy subgame perfection. ■

Let us first come to situations where mutual subnetworks enable cooperation that would otherwise not be possible.

Proposition 3 *Suppose assumptions 1 and 3 hold. Then*

1. *under information structure **(I1)**, a directed non-circular network \mathcal{N}^S is supportable in SPNE if $\text{id } i |_{\text{deg } i=1} = 1$ and for all other agents in the network $g^{i,i-1} + g^{i,i+1} > 0$ holds, and*

2. under information structures **(I2)** and **(I3)**, no non-circular network containing non-mutual subnetworks is supportable.

Proof. Part 1: Consider again **(S1)**. The property of directedness together with assumption 3 rules out the possibility that an agent has $\text{od } i > 1$. Therefore, all agents with $\text{deg } i = 2$ face immediate punishment after deviating from both sides and have no incentive to deviate if $g^{i,i-1} + g^{i,i+1} > 0$. The only agents that might have an incentive to deviate then are the ones with $\text{deg } i = 1$, the end vertices of the network, which have no incentive to deviate if their indegree is 1.

Part 2: Under **(I2)** or **(I3)**, enforcement relies on contagion or transmission of information about past actions through the agents. In a non-mutual subnetwork of a non-circular network, no agent i would get punished by another agent j than the one from whose relation she is deviating. Agent j will not be infected or be informed about the deviation by anyone, respectively. This is because i is the only one who could infect or inform j , respectively. Therefore it is a dominant strategy of any agent $i \in \mathcal{N}^S$ to defect to any neighbor $k \in N_i$ if $g^{ik} \leq 0$. ■

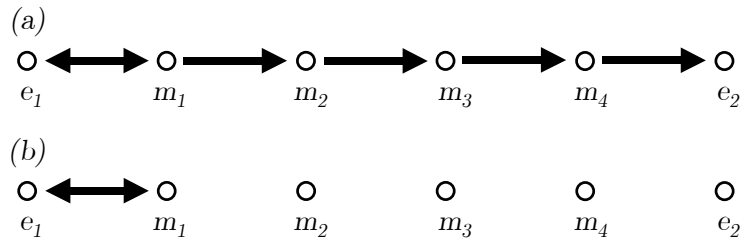


Figure 4: Sustainable networks under (a) info structure **(I1)**, (b) info structure **(I2)** and **(I3)**

Proposition 3 is illustrated by figure 4. If agent e_2 has the possibility to tell m_3 about m_4 having deviated and deviating to both e_2 and m_3 is not profitable for m_4 , this network is supportable. This is the case under **(I1)**, thus part 1 of proposition 3 says given **(I1)**, figure 4 (a) is an equilibrium. It is not the case under **(I2)** and **(I3)**, thus part 2 of proposition 3 says given **(I2)** or **(I3)**, figure 4 (a) is not an equilibrium. The equilibrium network in that case would be figure 4 (b).

Generally, agents who are not connected by a path additional to their direct link cannot utilize punishment power of other agents in the network unless there exists some central information processing agency and thus any cooperative action vector between these two agents has to be bilaterally self sustaining.

3.2.2 Mixed circular directed networks

We have just seen that more mutual relations might help. We have also seen that in circular directed non-mutual networks, it is possible to sustain cooperation when there is no information processing available to agents, that is under **(I2)** and **(I3)**. However, under these informational assumption it turns out to be necessary, that there are no mutual subnetworks in the cycle, whereas under **(I1)** this is not the case. The intuition is the following: If an agent benefits from both his neighbors, she might have an incentive to unilaterally deviate from infecting his neighbor when she should so.

Notation 2 Let \mathcal{N}_m^S be a connected mutual subnetwork of \mathcal{N}^S .

We define the following strategy:

(S4) Each agent $i \in \mathcal{N}^S$ starts by playing the agreed upon action vector $C^{ij} \forall i \in \mathcal{N}^S, \forall j \in N_i$, and goes on playing that action towards every $\forall j \in N_i$ as long as no one in his neighborhood deviated towards him.

If a deviation is experienced, every agent i plays D^{ij} in his interaction with the deviator for T periods and with his other neighbor for $T + 1$ periods.

Afterwards the agents return to $C^{ij} \forall i \in \mathcal{N}^S, \forall j \in N_i$.

Proposition 4 Suppose a c -cycle directed network and assumptions 1-3 to hold. Then under **(I2)** and **(I3)**,

1. grim trigger strategies **(S2)** cannot support cooperation if there exists a mutual subnetwork \mathcal{N}_m^S .
2. strategies **(S4)** constitute a sustainable network if there exists a T such that the network becomes T -nonmutual and $\forall i \in \mathcal{N}^S \delta^{c-2} g_T^{i,i-1} + g_T^{i,i+1} > 0$, where $g_T^{i,i+1} < 0$.

Proof. Part 1: Consider strategies **(S2)**. Suppose, we are in the situation of figure 5 with agents i and $i + 1$ forming a mutual subnetwork. Suppose agent i observes a deviation from his left neighbor $i - 1$ in period $t = \tau$ and no deviation from $i + 1$.

Agent i may have two possible beliefs about who defected initially. (a) agent $i + 1$ started: Since $i + 1$ did not deviate from his relation with i in period $t = \tau$, (which would be his best deviation iff he ever intended to deviate from his relation with $i + 1$), agent i should assume, by a simple forward induction argument, that $i + 1$ did not cheat because he would want to stick to cooperation in his relation with i . Since i will play the static Nash equilibrium with $i - 1$ for ever – no matter what he plays with $i + 1$ – not defecting from his relation with

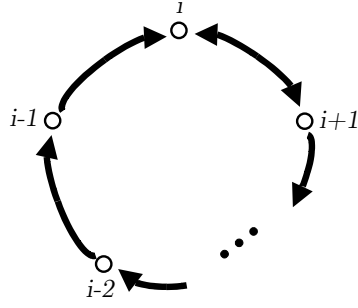


Figure 5: Subgame perfection under (I2) and (I3)

$i + 1$ is a dominant action. (b) another agent than $i + 1$ started: Since $g^{i,i+1} > 0$, for agent i it represents a cost to deviate from his relation with $i + 1$. This cost should be incurred as late as possible. Since $i + 1$ will be cheated on by $i + 2$ in $t = \tau + 2$ the earliest, infecting $i + 1$ in $t = \tau + 1$ is not profitable for i .

Part 2: The proof is analogous to the proof of proposition 1 part 2. ■

Due to his mutual relation with $i + 1$, agent i does not have an incentive to infect $i + 1$ since this means the loss of cooperation with him forever. Under (I2) and (I3), for given payoffs, a low δ is necessary to actually infect in $t = \tau + 1$ or, to put the result bluntly, when there does not exist an information processing agency, an agent, who benefits from everybody, hurts cooperation because he might be unwilling to punish, given there is no full information.

The main idea in part 2 is, by changing the punishment strategies, to transform the network into a non-mutual one and thereby giving each member an incentive to infect their neighbor. The conditions for the network to be sustainable resemble the ones for grim trigger strategies.

Let us now remove the assumption of a simple penal code. Consider the situation in which there does not exist a T such that the network becomes a sustainable T -nonmutual directed network. By rewarding the punisher, it may be possible to transform the network into a directed non-mutual one, given that strategy. Let us define the following strategies:

(S5) Each agent $i \in \mathcal{N}^S$ starts by playing the agreed upon action vector $C^{ij} \forall i \in \mathcal{N}^S, \forall j \in N_i$, and goes on playing that action towards every $\forall j \in N_i$ as long as no one in his neighborhood deviated towards him.

Agents $j \notin \mathcal{N}_m^S$ play strategies (S2).

If a player $i \in \mathcal{N}_m^S$ observes a player $i - 1 \notin \mathcal{N}_m^S, i - 1 \in N_i$ play $D^{i-1,i}$, she reverts to $D^{i,i-1}$ forever and to $D^{i,i+1}$ towards the other neighbor $i + 1 \in \mathcal{N}_m^S$ until they played

$(D^{i,i+1}, C^{i+1,i})$ for T^m consecutive periods.

If a player $i + 1 \in \mathcal{N}_m^S$ observes a player $i \in \mathcal{N}_m^S$ play $D^{i,i+1}$, she plays $C^{i+1,i}$ for T^m periods and $D^{i,i+1}$ forever if $i + 1 \notin \mathcal{N}_m^S$ or until they played $(D^{i,i+1}, C^{i+1,i})$ for T^m consecutive periods if $i + 1 \in \mathcal{N}_m^S$.

If a player $i \in \mathcal{N}_m^S$ observes both, $D^{i-1,i}$ with $i - 1 \notin \mathcal{N}_m^S$ and $D^{i+1,i}$ with $i + 1 \in \mathcal{N}_m^S$, she plays $D^{i,i-1}$ forever and $D^{i,i+1}$ for T^m periods.

After these punishment periods, the members of the mutual subnetwork return to cooperative behavior.

(S6) Each agent $i \in \mathcal{N}^S$ starts by playing the agreed upon action vector $C^{ij} \forall i \in \mathcal{N}^S, \forall j \in N_i$, and goes on playing that action towards every $\forall j \in N_i$ as long as no one in his neighborhood deviated towards him.

Agents $j \notin \mathcal{N}_m^S$ play strategies **(S2)**.

If a player $i \in \mathcal{N}_m^S$ observes a player $i - 1 \notin \mathcal{N}_m^S$ play $D^{i-1,i}$ and $g^{i-1,i} < 0$ and $C^{i+1,i}$ with $i + 1 \in \mathcal{N}_m^S$ she reverts to $D^{i,i-1}$ forever and to $D^{i,i+1}$ for T periods. After T periods she returns to $C^{i,i+1}$.

If any other agent $i + 1 \in \mathcal{N}_m^S$ observes an agent $i \in \mathcal{N}_m^S$ play $D^{i,i+1}$ he plays $D^{i+1,i}$ for the next $T - 1$ periods and then returns to $C^{i+1,i}$.

If a player $i \in \mathcal{N}_m^S$ observes both $D^{i-1,i}$ and $D^{i+1,i}$ with $i - 1 \notin \mathcal{N}_m^S$ and $i + 1 \in \mathcal{N}_m^S$ she reverts to $D^{i,i-1}$ and to $D^{i,i+1}$ forever.

Proposition 5 *If the network is a directed c-cycle, then under assumptions 1-3, **(I2)** and **(I3)**,*

1. *strategies **(S5)** constitute a sustainable network if for all agents $i + 1, i \in \mathcal{N}_m^S$ and $i - 1, i - 2$, and $i + 2 \notin \mathcal{N}_m^S$*

$$(1 - \delta^{T^m+1}) U^{i+1} (C^{i+1,i}, D^{i,i+1}) + \delta^{T^m+1} U^{i+1} (C^{i+1,i}, C^{i,i+1}) - U^{i+1} (D^{i+1,i}, D^{i,i+1}) > 0, \quad (2)$$

$$(1 - \delta^{T^m+1}) [U^i (C^{i,i+1}, C^{i+1,i}) - U^i (D^{i,i+1}, C^{i+1,i})] + \delta^{c-2} g^{i,i-1} > 0, \quad (3)$$

$$g^{i+1,i+2} + \delta^{c-2} g_{T^m}^{i+1,i} > 0, \quad \text{and} \quad (4)$$

$$g^{i-1,i} + \delta^{c-2} g^{i-1,i-2} > 0, \quad (5)$$

where $g^{i-1,i} < 0$ and there exists a path $\{i - 2, i - 1, i, i + 1, i + 2\}$.

2. strategies **(S6)** constitute a sustainable network if for agents $i + 1, i \in \mathcal{N}_m^S$ and $i - 1, i - 2$, and $i + 2 \notin \mathcal{N}_m^S$

$$g^{i-1,i} + \delta^{c-2} g^{i-1,i-2} > 0, \quad (6)$$

$$g_T^{i,i+1} + \delta^{c-2} g^{i,i-1} > 0, \quad \text{and} \quad (7)$$

$$g^{i+1,i+2} + \delta^{c-2} [g_T^{i+1,i} - (1 - \delta) [U^{i+1}(D^{i+1,i} C^{i,i+1}) - U^{i+1}(C^{i+1,i}, C^{i,i+1})]] > 0, \quad (8)$$

and inequality (9) from Appendix A.1 hold, where $g^{i-1,i} < 0$ and there exists a path $\{i - 2, i - 1, i, i + 1, i + 2\}$.

Proof. Part 1: Consider figure 5 again. Assume strategies **(S5)**.

Playing $D^{i,i+1}$ in $t = 0$ and $D^{i,i-1}$ in $t = c - 2$, which is agent i 's best defection, yields $U^i(D^{i,i+1}, C^{i+1,i})$ for T^m periods, $U^i(D^{i,i-1}, C^{i-1,i})$ in $t = c - 2$ and $U^i(D^{i,i-1}, D^{i-1,i})$ forever starting in $t = c - 1$. Playing $C^{i,i+1}$ and $C^{i,i-1}$ forever yields $U^i(C^{i,i+1}, C^{i+1,i})$ and $U^i(C^{i,i-1}, C^{i-1,i})$ forever. Therefore, the incentive constraint for agent i to cooperate with $i + 1$ and not deviate is given by (3).

Playing $D^{i+1,i+2}$ in $t = 0$ and $D^{i+1,i}$ in $t = c - 2$, which is agent $i + 1$'s best defection, yields $U^{i+1}(D^{i+1,i+2}, C^{i+2,i+1})$ in $t = 0$, $U^{i+1}(D^{i+1,i+2}, D^{i+2,i+1})$ from $t = 1$ on, $U^{i+1}(D^{i+1,i}, C^{i,i+1})$ in $t = c - 2$, $U^{i+1}(D^{i+1,i}, D^{i,i+1})$ for T^m periods starting in $t = c - 1$, and $U^{i+1}(C^{i+1,i}, C^{i,i+1})$ thereafter forever. Thus $i + 1$'s incentive constraint for cooperation with $i + 2$ is given by (4).

Agent $i - 1 \notin \mathcal{N}_m^S$ plays grim trigger strategies with his neighbors, therefore his incentive constraint is given by (5).

In the punishment subgame, it is always profitable for i to play $D^{i,i+1}$ since he gains from $(D^{i,i+1}, C^{i+1,i})$ and $D^{i,i-1}$ since $(D^{i,i+1}, D^{i+1,i})$ is a stage Nash equilibrium. Furthermore, in the punishment subgame, it is profitable for $i - 1 \notin \mathcal{N}_m^S$ to play $D^{i-1,i}$ and $D^{i-1,i-2}$ since these are the stage Nash equilibrium strategies.

Suppose agent $i + 1$ observes $D^{i,i+1}$ and he sticks to **(S5)**. This yields $U^{i+1}(C^{i+1,i}, D^{i,i+1})$ for T^m periods starting in $t = 0$ and $U^{i+1}(C^{i+1,i}, C^{i,i+1})$ thereafter forever. A one period deviation instead yields $U^{i+1}(D^{i+1,i}, D^{i,i+1})$ in $t = 0$, $U^{i+1}(C^{i+1,i}, D^{i,i+1})$ for T^m periods starting in $t = 1$ and $U^{i+1}(C^{i+1,i}, C^{i,i+1})$ thereafter forever. This is equilibrium behavior if (2) holds.

Part 2: Consider figure 5 again. Assume strategies **(S6)**.

Agent $i - 1$ plays grim trigger with his neighbors. Therefore his incentive constraint for cooperation is given by (6).

Agent i plays grim trigger with $i - 1$ and T -period Nash reversion with $i + 1$. Thus his incentive constraint for cooperation is (7).

Agent $i + 1$ will always prefer not to anticipate the punishment, that is he will play $C^{i+1,i}$ in $t = c - 2$ if he defected from his relation with agent $i + 2$ in $t = 0$, since

$$g^{i+1,i} > g_T^{i+1,i} - (1 - \delta) [U^{i+1}(D^{i+1,i}C^{i,i+1}) - U^{i+1}(C^{i+1,i}, C^{i,i+1})].$$

Therefore his incentive constraint for cooperation is given by (8). He will, however, start the punishment phase in period $c - 1$ since he now expects i to play $D^{i,i+1}$ given this is optimal for i .

It will be optimal for i to play $D^{i,i+1}$ in $t = c - 1$, if (9) from appendix A.1 holds, which is shown in appendix A.1. ■

Part 1: members of the mutual subnetwork gain enough from cooperating that they would be willing to incur a strong punishment or

Part 2: members of the mutual subnetwork sufficiently fear the loss from their neighbor in the mutual subnetwork playing D . his expected loss from playing $C^{i,i+1}$ where $i + 1$ plays $D^{i+1,i}$ is high enough which is the case either if the probability that $i + 1$ started the contagious process is high or if $U^i(C^{i,i+1}, D^{i+1,i})$ is very low.

4 Sustainable networks of higher degree

[work in progress]

The underlying structure of the stage game is a prisoners' dilemma and maintaining a relation as such is not costly. This means that the utility agents receive from having a relation, as compared to not having it, is always bigger. If agents could only choose between the stage Nash equilibrium and the outcome where both agents play cooperatively, they'd choose the latter. Thus, adding a relation to the network benefits the agents who add this relation. Furthermore, the higher the discount factor of agents in a network, the more difficult is it to sustain a network of relations where information travels with delay. In these networks it disciplines to have cycles of smaller order and thereby shorten ways. Therefore, if we allow for a higher degree of each agent, agents will have an incentive to add relations, including even *bilaterally deficient* ones, as long as this results in a sustainable network, given a basis structure.

4.1 Full information (I1):

The full information case is the easiest. A simple generalization of proposition 1 part 1 is the following corollary stated without proof.

Corollary 3 *Let us assume (I1) and the strategy profile (S1). Then the network is sustainable iff*

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

It always pays to deviate from non-deficient relations as late as possible. This includes never if possible. We have shown equilibrium conditions for a circular network of maximal degree two under (I2) and (I3). In order to generalize that result for networks with nodes of degree higher than two, let us start from exactly that situation and add one link according to figures 6-8.

4.2 No information transmission (I2):

A slight modification of strategy (S2) leads us to a workable definition of a strategy under (I2) in a setting with a network of higher degree:

(S7) Each agent $i \in \mathcal{N}^S$ starts by playing the agreed upon action vector $C^{ij} \forall i \in \mathcal{N}^S, \forall j \in N_i$, and goes on playing that action towards every $\forall j \in N_i$ as long as no one in his neighborhood deviated towards him.

Every agent i immediately plays D^{ij} forever with a deviator j .

Every agent i immediately deviates forever from his deficient relations to the static Nash equilibrium if a deviation is experienced.

Every agent i deviates forever from his non-deficient relations to the static Nash equilibrium in the period the neighbor in question gets infected by another of his neighbors if a deviation is experienced.

4.2.1 Adding a bilaterally non-deficient relation

Assume strategies (S7) and the network shown in figure 6. Then under (I2): Play C is Nash equilibrium if

$$\begin{aligned} g^{j,j+1} + \delta^{c-2} g^{j,j-1} &> 0 && \text{for } \forall j \in \{i+1, \dots, k-1\} \\ g^{k,k+1} + \delta^{d-2} g^{k,i} + \delta^{c-2} g^{k,k-1} &> 0 && \text{for } k \\ g^{j,j+1} + \delta^{d-2} g^{j,j-i} &> 0 && \text{for } \forall j \in \{k+1, \dots, i-1\} \\ g^{i,i+1} + \delta^{c-d} g^{i,k} + \delta^{c-2} g^{i,i-1} &> 0 && \text{for } i \end{aligned}$$

However, we have to make sure, that this is subgame perfect which it will *not* be for i and k . Modify (S7) such that for i and k no deviation will be necessary from their relation, then we are essentially in the situation of proposition 1 with the respective incentive constraints. We define the following strategy:

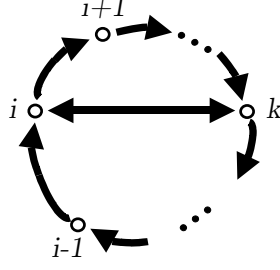


Figure 6: Bilaterally sustainable relation added

(S7') In addition to (S7), assume that agents i and k sharing the bilateral relation, which has been added to the directed non-mutual c -cycle, never play $(D^{i,k}, D^{k,i})$ except if an agent $j \in \{i, k\}$ deviated.

Assume strategies (S7') and the network shown in figure 6. Then under (I2), for an agent j to play C in all his interactions is a Nash equilibrium if

$$g^{j,j+1} + \delta^{c-2} g^{j,j-1} > 0 \quad \text{for } \forall j$$

This essentially means giving up the idea of transferring enforcement power from an autarkically sustainable relation onto other relations. However, with (S2''), transmitting info over the mutual relation becomes an option. We will explore this in section 4.3 below.

4.2.2 Adding a relation that is non-deficient only for one agent

Assume strategies (S7) and the network shown in figure 7. Then under (I2), playing C is Nash equilibrium if

$$\begin{aligned} g^{j,j+1} + \delta^{c-2} g^{j,j-1} &> 0 && \text{for } \forall j \in \{i+1, \dots, k-1\} \\ g^{k,k+1} + \delta^{d-2} g^{k,i} + \delta^{c-2} g^{k,k-1} &> 0 && \text{for } k \\ g^{j,j+1} + \delta^{d-2} g^{j,j-i} &> 0 && \text{for } \forall j \in \{k+1, \dots, i-1\} \\ g^{i,i+1} + \delta^{c-d} g^{i,k} + \delta^{c-2} g^{i,i-1} &> 0 && \text{for } i \end{aligned}$$

Subgame perfection in higher degree networks of relations The result from proposition 4 generalizes to

Corollary 4 *Suppose a non-mutual network and assumptions 1 and 2 to hold. Then under (I2) and (I3), simple penal codes and grim trigger strategies (S2') cannot support cooperation if there exists a subnetwork that is supportable in autarky.*

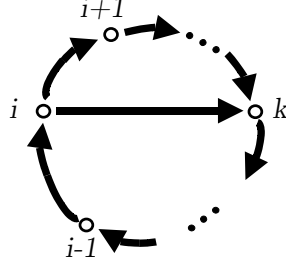


Figure 7: Adding a relation that is deficient for one and non-deficient for the other agent

Proof. The proof is analogous to that of proposition 4. ■

For an intuition, it is useful to have a look at figure 7. If $g^{k,k+1} + \delta^{d-2}g^{k,i} > 0$, then it represents a cost for agent k to deviate from the subnetwork $\{k, k + 1, \dots, i, k\}$. If agent $k - 1$ plays $D^{k-1,k}$ in $t = \tau$, and all other agents $j \in N_k$ play $C^{j,k}$, agent k does not have an incentive to incur that cost since the subnetwork is sustainable in autarky.

4.2.3 Adding a relation that is bilaterally deficient

Assume strategies (S7) and the network shown in figure 8. Then under (I2): to play C is a Nash equilibrium if

$$\begin{array}{ll}
 g^{j,j+1} + \delta^{c-d-2}g^{j,j-1} > 0 & \text{for } \forall j \in \{i + 1, \dots, k - 1\} \\
 g^{k,k+1} + \delta^{d-c}g^{k,i} + \delta^{c-2}g^{k,k-1} > 0 & k \\
 g^{j,j+1} + \delta^{d-2}g^{j,j-i} > 0 & \text{for } \forall j \in \{k + 1, \dots, i - 1\} \\
 g^{i,i+1} + \delta^{c-d}g^{i,k} + \delta^{c-2}g^{i,i-1} > 0 & i
 \end{array}$$

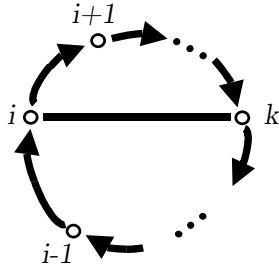


Figure 8: Adding a bilaterally deficient relation

Proposition 6 Under (I2), strategy (S7) does not constitute a sustainable network if subtracting a sustainable subnetwork leaves the remainder unsustainable.

Proof. Denote the sustainable subnetwork with \mathcal{N}_S^S and the remaining, unsustainable subnetwork with \mathcal{N}_N^S . If \mathcal{N}^S were to be sustainable, then no agent in \mathcal{N}_N^S would want to deviate. This means their implicit contracts must be enforced by agents from \mathcal{N}_S^S by a deviation from \mathcal{N}_S^S , which is not profitable. ■

4.3 Network information transmission (I3):

Conjecture 1 *It is possible to use a subnetwork, which is sustainable in autarky, for transmission of information.*

We define the following strategy:

(S8) In addition to (S7), assume that agents $j \in \{i, k, k+1, \dots, i-1, i\}$, which is a sustainable subnetwork formed by adding the link ik to the directed non-mutual c-cycle, never play $(D^{i,k}, D^{k,i})$ except if an agent $j \in \{i, k, k+1, \dots, i-1, i\}$ deviated.

Furthermore, assume that if an agent $h \notin \{i, k, k+1, \dots, i-1, i\}$ deviated from a relation with an agent $j \in \{i, k, k+1, \dots, i-1, i\}$, j tells all his neighbors who in the period following the information transmission play D forever in all their deficient relations with agents $h' \notin \{i, k, k+1, \dots, i-1, i\}$.

[to be extended]

4.3.1 Adding a bilaterally non-deficient relation

Let us have a look at figure 6. In this situation, the possibility to transmit the information that an agent has deviated and therefore a distinction between (I2) and (I3) becomes relevant. If $k-1$ deviated, k can without fearing the breakup of the relation with i tell her that $k-1$ deviated.

4.3.2 Adding a relation that is non-deficient only for one agent

[to be extended]

4.4 Combined star-wheel network

A generalization of the network from figure 6 is a combined wheel-star-network from figure 9. Here information flows with high speed through a central agent c . Thereby cooperation is being fostered to a high extent. Such a central agent would serve as a "central information processing agency" for the agents $i \in \mathcal{N}^S - c$.

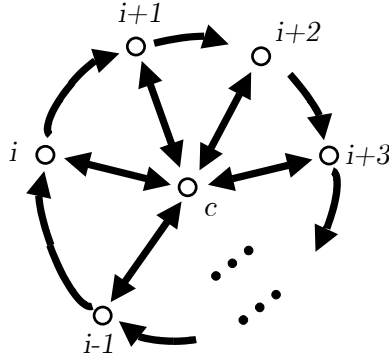


Figure 9: Combined star-wheel-network

5 Conclusion

In our model, agents maintain relations by using the network that, in addition to their own relations, consists of other agents' relations. We identified 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 always require agents to be willing to pass on information.

We showed that if agents cannot discipline themselves within a certain relation, "redundant" links, that is circularity, can sustain the relation. In contrast to previous literature, the possibility to transmit information about the cheating of someone through the links in the network has not been an equilibrium action if enforcement relied on optimal punishment. With optimal punishment, the deviation of an agent starts a contagious process that eliminates cooperation in the network. We showed that with more complex punishment strategies, agents may use information transmission, and thereby kept on cooperating in the rest of the network while punishing the deviator – which increased efficiency. We showed that, under the complete information assumption, bilaterally unsustainable relations in a network without "redundant" links, can be supported by having self-sustaining relations at the ends of the network while this does not work for the other informational assumptions. We also showed that having self-sustaining relations in the network may actually hurt cooperation in the case without full information because agents might not be willing to perform the punishment. In this case a network may be sustainable if agents use less severe punishments than grim trigger or by rewarding the punisher. These results were finally generalized to more complex network architectures.

Possible applications of our model or of modifications thereof, include the organization of inter-firm relations in industrial districts, social capital or collusive behavior that is enforced in networks of very different players. In her much acclaimed book cited above, 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. Our model may help explain what Saxenian calls a "culture of cooperation" with the means of economics and game theory as a network of long-term relations, each of which perhaps might not be sustainable on a bilateral basis.

References

- [1] AUMANN, ROBERT J. AND ROGER B. MYERSON (1988): "Endogenous Formation of Links between Players and Coalitions: An Application of the Shapley Value." In: Roth, A. (ed.): *The Shapley Value*. Cambridge University Press, Cambridge.
- [2] BAKER, GEORGE, ROBERT GIBBONS, AND KEVIN J. MURPHY (2001a), "Bringing the Market Inside the Firm?" *American Economic Review*, Vol. 91 (2), pp. 212-218.
- [3] BAKER, GEORGE, ROBERT GIBBONS, AND KEVIN J. MURPHY (2001b), "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics*, Vol 117 (1), pp. 39-84.
- [4] BALA, VENKATESH AND SANJEEV GOYAL (2000a): "A Noncooperative Model of Network Formation," *Econometrica*, Vol. 68 (5), pp. 1181-1229.
- [5] BALA, VENKATESH AND SANJEEV GOYAL (2000b): "A Strategic Analysis of Network Reliability," *Review of Economic Design*, Vol. 5, pp. 205-228.
- [6] BERNHEIM, B. DOUGLAS AND MICHAEL D. WHINSTON (1990): "Multimarket contact and collusive behavior," *RAND Journal of Economics*, Vol. 21 (1), pp. 1 - 26.
- [7] FRIEDMAN, JAMES W. (1971): "A Non-Cooperative Equilibrium for Supergames," *Review of Economic Studies*, Vol. 38, pp. 1-12.
- [8] GROH, CHRISTIAN (2002): "Communication Networks and Cooperation," University of Mannheim, mimeo.
- [9] JACKSON, MATTHEW O. AND ASHER WOLINSKY (1996): "A Strategic Model of Social and Economic Networks." *Journal of Economic Theory*, Vol. 71, pp. 44-74.

- [10] JOHNSON, CATHLEEN AND ROBERT P. GILLES (2000): "Spatial Social Networks," *Review of Economic Design*, Vol. 5, pp. 273 - 299.
- [11] KANDORI, MICHIHIRO (1992): "Social Norms and Community Enforcement," *Review of Economic Studies*, Vol. 59, pp. 63-80.
- [12] KRANTON, RACHEL E. (1996): "Reciprocal Exchange: A Self-Sustaining System." *American Economic Review*, Vol. 86 (4), pp. 830 - 851.
- [13] KRANTON, RACHEL E. AND DEBORAH F. MINEHART (2000): "Networks versus Vertical Integration," *RAND Journal of Economics*, Vol. 31 (3), pp. 301-331.
- [14] KRANTON, RACHEL E. AND DEBORAH F. MINEHART (2001): "A Theory of Buyer-Seller Networks," *American Economic Review*, Vol. 91 (3), pp. 485-508.
- [15] LI, SHUHE (1999): "The Benefits and Costs of Relation-based Governance: An Explanation of the East Asian Miracle and Crisis," Working Paper, Hong Kong University
- [16] MAGGI, GIOVANNI (1999): "The Role of Multilateral Institutions in International Trade Cooperation," *American Economic Review*, Vol. 89 (1), pp. 190-214.
- [17] NORTH, DOUGLAS C. (1990): "*Institutions, Institutional Change and Economic Performance*," Cambridge University Press, Cambridge UK.
- [18] POWELL, WALTER W., KENNETH W. KOPUT, AND LAUREL SMITH-DOERR (1996): "Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology," *Administrative Science Quarterly*, Vol 41, pp. 116-145.
- [19] SAXENIAN, ANNALEE (1994): "*Regional advantage: Culture and competition in Silicon Valley and Route 128*", Harvard University Press, Cambridge, MA.
- [20] SPAGNOLO, GIANCARLO (1999a): "Markets and Cooperation," Working Paper, available from <http://www.ssrn.com/>
- [21] SPAGNOLO, GIANCARLO (1999b): "On Interdependent Supergames: Multimarket Contact, Concavity, and Collusion," *Journal of Economic Theory*, Vol. 89 (1), pp. 127-39.
- [22] WILLIAMSON, OLIVER E. (1996): "*The Mechanisms of Governance*," Oxford University Press, New York, NY.

A Appendix

A.1 Condition for optimality of infecting under (S6)

Define $p_{(i+1)}$ to be i 's subjective probability of the event that $i + 1$ started the contagious process and $p_{(i+2|\neg i+1)}$ the conditional probability that $i + 2$ started the contagious process given that $i + 1$ did not start it. Suppose in $t = 0$, agent $i - 1$ plays $D^{i-1,i}$. Then sticking to $D^{i,i+1}$ in $t = 1$ yields actions

1	2	3	(...)	T	T+1	T+2	(...)	w/ prob.
(D, D)	(D, D)	(D, D)	(...)	(D, D)	(C, C)	(...)	(...)	$p_{(i+1)}$
(D, C)	(D, D)	(D, D)	(...)	(D, D)	(C, C)	(...)	(...)	$[1 - p_{(i+1)}]$

where all action vectors are $(a^{i,i+1}, a^{i+1,i})$. Not sticking to it yields actions

1	2	3	(...)	T	T+1	T+2	(...)	w/ prob.
(C, D)	(D, D)	(D, D)	(...)	(D, D)	(D, D)	(D, D)	(...)	$p_{(i+1)}$
(C, C)	(D, D)	(D, D)	(...)	(D, D)	(D, D)	(C, C)	(...)	$[1 - p_{(i+1)}] \cdot p_{(i+2 \neg i+1)}$
(C, C)	(D, C)	(D, D)	(...)	(D, D)	(D, D)	(C, C)	(...)	$[1 - p_{(i+1)}] \cdot [1 - p_{(i+2 \neg i+1)}]$

In addition, since not sticking to $D^{i,i+1}$ when $i + 1$ indeed started the contagious process results in playing the static Nash equilibrium between i and $i + 1$ forever, there will not be a return to $(C^{i,i-1}, C^{i-1,i})$ with probability $p_{(i+1)}$. Therefore the incentive constraint for i to infect $i + 1$ when being cheated by $i - 1$ only is

$$\begin{aligned}
& [U^i (D^{i,i+1}, D^{i+1,i}) + \delta U^i (D^{i,i+1}, D^{i+1,i})] \cdot p_{(i+1)} \\
& + [U^i (D^{i,i+1}, C^{i+1,i}) + \delta U^i (D^{i,i+1}, D^{i+1,i})] \cdot [1 - p_{(i+1)}] \\
& \quad + \frac{\delta^{T+1}}{1 - \delta} U^i (C^{i,i+1}, C^{i+1,i}) \\
& \quad - [U^i (C^{i,i+1}, D^{i+1,i}) + \delta U^i (D^{i,i+1}, D^{i+1,i})] \cdot p_{(i+1)} \\
& \quad - [U^i (C^{i,i+1}, C^{i+1,i}) + \delta U^i (D^{i,i+1}, D^{i+1,i})] \cdot [1 - p_{(i+1)}] p_{(i+2|\neg i+1)} \\
& - [U^i (C^{i,i+1}, C^{i+1,i}) + \delta U^i (D^{i,i+1}, C^{i+1,i})] \cdot [1 - p_{(i+1)}] \cdot [1 - p_{(i+2|\neg i+1)}] \\
& \quad - \frac{\delta^{T+1}}{1 - \delta} U^i (D^{i,i+1}, D^{i+1,i}) \cdot p_{(i+1)} \\
& - \left[\delta^{T+1} U (D^{i,i+1}, D^{i+1,i}) - \frac{\delta^{T+2}}{1 - \delta} U^i (C^{i,i+1}, C^{i+1,i}) \right] \cdot [1 - p_{(i+1)}] \\
& \quad + \frac{\delta^{c-1+T}}{1 - \delta} [U^i (C^{i,i-1}, C^{i-1,i}) - U^i (D^{i,i-1}, D^{i-1,i})] \cdot p_{(i+1)} > 0. \quad (9)
\end{aligned}$$