

Repeated Games Played in a Network

Markus Kinateder*

Universitat Autònoma de Barcelona[†]

23 October 2006

Abstract

Delayed perfect monitoring in an infinitely repeated discounted game is modelled by letting the players form a connected and undirected network. Players observe their immediate neighbors' behavior only, but communicate over time the repeated game's history truthfully throughout the network. The Folk Theorem extends to this setup, although for a range of discount factors strictly below 1, the set of sequential equilibria and the corresponding payoff set may be reduced. A general class of games is analyzed without imposing restrictions on the dimensionality of the payoff space. This and the bilateral communication structure allow for limited results under strategic communication only. As a by-product this model produces a network result; namely, the level of cooperation in this setup depends on the network's diameter, and not on its clustering coefficient as in other models.

JEL classification numbers: C72, C73, L14

Keywords: Repeated Game, Delayed Perfect Monitoring, Network, Communication

1 Introduction

Repeated games are frequently used to model repeated strategic interaction between impatient economic agents. Usually, it is possible to sustain equilibria that do not arise in

*I am very grateful for the support received from my supervisor Jordi Massó. Additionally, I thank Toni Calvó-Armengol, Julio González-Díaz and Penélope Hernández for their advice and time. I benefited hugely from conversations with Elchanan Ben-Porath, Drew Fudenberg, Olivier Gossner, David Levine, Jérôme Renault, Tristan Tomala and Fernando Vega-Redondo, and from comments made by participants of workshops in Valencia, Girona and at UAB. Financial support from the Spanish Ministry of Education and Science in form of a scholarship (FPU) and through grant SEJ2005-01481 is acknowledged.

[†]Departamento de Economía y de Historia Económica, Universitat Autònoma de Barcelona, Edifici B, 08193 Bellaterra, Spain; email: mkinateder@idea.uab.es

a one-shot game by repeating it. The associated payoff vectors, moreover, can be Pareto superior to the ones achieved in all stage game equilibria. The well-known Folk Theorem states this result. This paper focuses on infinitely repeated discounted games for which Fudenberg, Levine and Takahashi (2006), thereafter FLT, obtain the subgame-perfect Folk Theorem. They dispose of any dimensionality condition previously imposed by Fudenberg and Maskin (1986) and by Abreu, Dutta and Smith (1994), thereafter ADS, and moreover, extend the result of Wen (1994) to (unobservable) mixed actions.

For simplicity, other strong assumptions are normally imposed, such as perfect monitoring. A player observes his opponents' behavior immediately and perfectly. To relax this assumption is the aim of the imperfect monitoring literature, in which each player receives an imperfect private or public signal of the action profile played. Under certain conditions, the set of sequential equilibria, or of other equilibrium concepts that extend subgame-perfectness to repeated games of imperfect information, is usually non-empty. In some cases even the Folk Theorem obtains. The interested reader is referred to a private monitoring survey by Kandori (2002) and Mailath and Samuelson's (2006) textbook.

The aim of this paper is to model delayed perfect monitoring by letting the players, that play an infinitely repeated discounted game, form a connected and undirected network. In each period, a player observes his neighbors' action choices and communicates non-strategically, that is truthfully, these observations and other information he has received before to all neighbors. In general, the players take decisions under imperfect information in any but the first period. Since therefore the only proper subgame of the repeated game is the entire game, the concept of sequential equilibrium is used. Nevertheless, the entire history of the repeated game spreads gradually throughout the network over time. The network gives a structure to this heterogeneous flow of information. It is also possible, however, to interpret the delay in information transmission as being due to the time it takes a player to process information or to react to new information.¹

In reality, impatient economic agents frequently form a network due to which the information flow is delayed. In many industries, such as the car industry, big producers are at the center of a large network of suppliers, which may be linked among themselves. Links are enforced by long-term contracts or relationships and high fines are levied on firms that break such a contract. (The fines must be credibly enforceable which motivates the use of sequential equilibrium.) The network is usually organized along the value chain

¹Hence, the network may be substituted by a matrix (of dimension "number of players" time "number of player"), in which each entry specifies the delay with which the corresponding two players obtain information about each other. All results are even valid when this matrix is non-symmetric, or equivalently, for directed networks.

and information about a firm's non-compliance with certain quality or service standards spreads only slowly throughout the network until it reaches the center. In turn, the big firm at the center of the network might communicate changes in quality requirements or product specifications to its suppliers. Sometimes it also imposes price reductions on their products. The suppliers decide whether to accept the proposed changes and if or how to enforce them on their suppliers, respectively. Information may thus flow back again to the center, for example, when a small firm in the network's periphery threatens to either accept a price reduction for its products and to go bankrupt thereafter or to continue as before. The theoretical model developed in this paper encompasses some of the key features just described, although it also abstracts from some of them. This model can be applied in several other contexts, some of which are mentioned in the conclusion.

Under the assumption of truth-telling, the Folk Theorem extends to the delayed perfect monitoring model, that is, any feasible and strictly individually rational payoff vector can be supported by a sequential equilibrium strategy profile when the players are sufficiently patient. Then, they do not mind to receive the repeated game's history of action profiles gradually over time. However, for a range of discount factors strictly below 1, the delay in information transmission caused by the network may trigger a player's deviation from some previously agreed sequence of play. The reduction in the set of sequential equilibria in comparison to the perfect monitoring case, which arises when players are impatient, seems to reflect many real situations well. Moreover, the concept of punishment reward is adapted to the network case and in order to analyze a general class of games, no restriction is imposed on the dimensionality of the payoff space.² As a consequence, the introduction of strategic communication becomes much more involved and the *effective minmax* concept has to be used. Finally, this model also contributes to the network literature in which the clustering coefficient, or similar measures of local connectedness, usually determine the level of cooperation sustainable in a network. In this model, to the contrary, the diameter of the network is decisive. The two measures are not related as is illustrated in an example.

The related literature can be roughly divided into three setups. In the first one, each pair of neighbors in a network plays a bilateral repeated game. A player's communication and observations are restricted to his neighborhood as well, that is, they are also bilateral. In the second group of models, all players play the same repeated game and a player observes an imperfect private or public signal of the action profile played, or a bilateral observation structure imposed by a network is assumed. Communication takes the form

²Whereas in a perfect monitoring model this assumption is purely technical and may be disposed of, in this setup it makes a huge difference due to the heterogeneity of the information flow.

of public announcements of past signal realizations or of own behavior in the past. Hence, all players are informed about the repeated game's history at the same time. All models in this group, additionally, assume a full-dimensional payoff space and allow for strategic communication. Finally, a setup as in this paper is characterized by a bilateral communication and observation network in which all players play the same repeated game, but never have to report their own action choices. Only Renault and Tomala (1998) also derive a model with these characteristics.

Nevertheless, two papers from the second group are also important since they assume a bilateral observation structure. Ben-Porath and Kahneman (1996) study sequential equilibria of infinitely repeated discounted games in which the players form a (not necessarily connected) network. They assume that players publicly announce their own action choices and observations made about their neighbors in a strategic way, that is, including lies. When each group contains strictly more than two players unilateral deviations are detectable, and hence, do not occur in equilibrium. In Ben-Porath and Kahneman (2003) this idea is extended. Since monitoring is costly, only one monitor is assigned to every player. After an incompatible announcement, which in equilibrium does not occur, both players are punished and the monitor is substituted. Renault and Tomala (1998), in turn, show how to sustain uniform Nash Equilibria—which is a weaker concept than sequential equilibrium—in finitely and infinitely repeated undiscounted games when the players form a 2-connected and directed graph. Since this implies that there are two distinct paths between any pair of players, lies are prevented in equilibrium. In their model, however, the payoff accumulation stops during a communication phase and, as in Ben-Porath and Kahneman (2003), the players, in general, do not receive the repeated game's history.

The next section introduces notation and definitions. The model without loss of generality is presented in pure actions. In section 3, the features of the model are demonstrated in an example. Section 4 is dedicated to derive two concepts, the information sharing process and the punishment reward phase. Both are prerequisites for the Folk Theorem, which is stated in section 5, along with conditions under which impatient players deviate from a given sequence of action profiles. In the same section, moreover, the model's extension to strategic communication is discussed and how it relates to the imperfect monitoring and the network literature, respectively. Before concluding, remarks about mixed actions follow.

2 Preliminaries

2.1 Stage Game

Each player i in the finite set of players $I = \{1, \dots, n\}$ has a finite and non-empty set of pure actions A_i ; a pure action a_i is an element of this set. The pure action space of the stage game is $A = \times_{i \in I} A_i$, with generic element a , called action profile. To emphasize the role of player i , pure action profile a is written as (a_i, a_{-i}) . For any non-empty set of players $S \subset I$, let $A_S = \times_{i \in S} A_i$, and denote by a_S an element of this set. Player i 's payoff function is a mapping $h_i : A \rightarrow \mathbb{R}$, and the payoff function $h : A \rightarrow \mathbb{R}^n$ assigns a payoff vector to each pure action profile. The stage game in normal form is then the tuple $G \equiv (I, (A_i)_{i \in I}, (h_i)_{i \in I})$. Finally, define the convex hull of the finite set of payoff vectors corresponding to pure action profiles in G as $co(G) = co\{x \in \mathbb{R}^n \mid \exists a \in A : h(a) = x\}$.

2.2 Network

The players in set I are the vertices of a network g , whose graph is defined as (I, E) , where $E \subseteq I \times I$ denotes the set of links or edges between them. A directed link from player i to player j is denoted by (i, j) . Graph (I, E) is undirected, that is, for all $i, j \in I$, (i, j) if, and only if, (j, i) . Being linked allows players to mutually observe each other's behavior and to communicate the actions chosen by other players. A specific description of this process requires afore the introduction of further notation.

Given network g , path s between a pair of distinct players i and j is denoted by p_{ij}^s and defined as a sequence of players i_1^s, \dots, i_r^s such that $i_1^s = i$, $i_r^s = j$, $i_m^s \neq i_l^s$ for all $1 \leq m < l \leq r$, and (i_{l-1}^s, i_l^s) , for all $1 < l \leq r$. The length of path p_{ij}^s is $r - 1$. It is denoted by l_{ij}^s . The network is assumed to be connected, or in other words, to consist of one component only; that is, each player is connected to at least one other player directly and to all remaining ones via paths of finite lengths. For any two distinct players i and j , there is a finite set $S_{ij} = \{p_{ij}^1, \dots, p_{ij}^{Q_{ij}}\}$ of different paths between them with lengths $l_{ij}^1, \dots, l_{ij}^{Q_{ij}}$, respectively. The number of links along the shortest among all these paths is called *distance* between players i and j . It is denoted by d_{ij} and defined as $d_{ij} = \min_{1 \leq q \leq Q_{ij}} l_{ij}^q$; obviously, $d_{ij} = d_{ji}$ and let $d_{ii} = 0$. Moreover, denote the *largest distance* between player i and any other player in the network by $d_i = \max_{j \in I} d_{ij}$, and define the *diameter* of network g as the maximal *largest distance* among all players, that is, $d = \max_{i \in I} d_i$. Finally, denote player i 's set of direct neighbors by $i(1) = \{j \in I \mid d_{ij} = 1\}$ and, in general, for any $1 \leq m \leq d_i$, define his set of m -neighbors as $i(m) = \{j \in I \mid d_{ij} = m\}$.

As already mentioned, links between players serve as means of communication when the stage game is played repeatedly. In each period, a player first chooses an action, in a way specified below, and then makes observations and communicates with his neighbors. He observes the actions chosen by his immediate neighbors, before receiving the information they received one period earlier from their neighbors. Similarly, he reveals to any direct neighbor the action he plays, before communicating him the information he received one period ago. Hence, information flows one link per period and with a $d_i - 1$ period lag player i gets to know the repeated game's entire history.³ It is assumed that communication is non-strategic, or in other words, that players always truthfully reveal what their neighbors did and told them. How to relax this assumption is discussed later.

Additionally, a player has perfect recall. Hence, for any player $i \in I$ at any $t \geq 1$, there is a *set of observations* denoted by Ob_i^t , that includes all histories of observations that player i may have made at the end of period t . It is defined recursively as

$$\begin{aligned} Ob_i^1 &= A_i \times A_{i(1)}, \\ Ob_i^2 &= A_i^2 \times A_{i(1)}^2 \times A_{i(2)}, \\ &\vdots \\ Ob_i^t &= A_i^t \times A_{i(1)}^t \times A_{i(2)}^{t-1} \times \dots \times A_{i(d_i)}^{t-d_i+1} \end{aligned}$$

for all $t \geq d_i$, where for any $1 \leq m \leq d_i$ and any $t \geq 1$, $A_{i(m)}^t = (\times_{j \in i(m)} A_j)^t$. An observation made by player i at time t is denoted by $ob_i^t \in Ob_i^t$. Given G and g , a sequence of action profiles $\{a^t\}_{t=1}^\infty$, where $a^t \in A$ for all $t \geq 1$, generates a sequence of observations for player i ,⁴

$$\begin{aligned} ob_i^1 &= (a_i^1, a_{i(1)}^1), \\ ob_i^2 &= (a_i^1, a_{i(1)}^1, a_{i(2)}^1, a_i^2, a_{i(1)}^2), \\ &\vdots \\ ob_i^t &= (\{a_i^s\}_{s=1}^t, \{a_{i(1)}^s\}_{s=1}^t, \{a_{i(2)}^s\}_{s=1}^{t-1}, \dots, \{a_{i(d_i)}^s\}_{s=1}^{t-d_i+1}) \end{aligned}$$

for all $t \geq d_i$. At any $1 \leq t < d_i$, player i is not yet informed about the behavior of at least one other player. At $t = d_i$, $ob_i^{d_i}$ contains the actions chosen by all players in period one. Abusing notation, this is referred to as $a^1 \in ob_i^{d_i}$ (since a^1 belongs to A). At any $t > d_i$, the action profiles a^1, \dots, a^{t-d_i+1} are identified by player i , and hence, in an abuse of

³At the end of any $t \geq d_i$, for example, player i knows the actions played in period t by himself and all players in $i(1)$, the actions played by himself and all players in $i(1)$ and $i(2)$ at $t - 1, \dots$, and finally the actions played by all players at $t - d_i + 1$ and at any point in time before.

⁴Equivalently, this setup can be interpreted as follows. Each action profile a^t generates a public signal with a delay of $d - 1$ periods and certain private signals in all periods s , where $t \leq s < t + d - 1$.

terminology, said to be elements of ob_i^t . Thus, at any $t \geq 1$, the sequence of action profiles generates an observation profile $ob^t \in Ob^t$, where $Ob^t = \times_{i \in I} Ob_i^t$. The players organized in this way play an infinitely repeated discounted game.

2.3 Repeated Game with Delayed Perfect Monitoring

In the infinitely repeated discounted game played on the fixed network g , thereafter called repeated network game, at each point in discrete time, $t = 1, 2, \dots$, the stage game G is played. Set I is assumed to contain at least three players since otherwise the analysis of the network case is trivial.

Let player i 's set of strategies be $F_i = \{\{f_i^t\}_{t=1}^\infty \mid f_i^1 \in A_i, \text{ and for all } t > 1, f_i^t : Ob_i^{t-1} \rightarrow A_i\}$. At any $t \geq 1$, player i 's strategy $f_i = \{f_i^t\}_{t=1}^\infty$ prescribes him to choose some action. For $t > 1$, this prescription is a mapping from his *set of observations* to his action set. The cartesian product of all players' strategy sets $F = \times_{i \in I} F_i$, constitutes the strategy space of the repeated network game. A strategy profile $f = (f_1, \dots, f_n)$ is an element of F . To emphasize player i 's role, it is written as (f_i, f_{-i}) . At any $t \geq 1$, each $f \in F$ recursively generates a pure action profile $a^t(f) = (a_1^t(f), \dots, a_n^t(f))$ and a corresponding observation profile $ob^t(f) = (ob_1^t(f), \dots, ob_n^t(f))$: for any player i , let $a_i^1(f) = f_i^1$ and $ob_i^1(f) = (a_i^1(f), a_{i(1)}^1(f))$, and for $t > 1$ given $ob_i^{t-1}(f) \in Ob_i^{t-1}$, $a_i^t(f) = f_i^t(ob_i^{t-1}(f))$ and $ob_i^t(f)$ is defined accordingly. Each $f \in F$ thus generates a sequence of action profiles $\{a^t(f)\}_{t=1}^\infty$, which in turn generates a sequence of observation profiles $\{ob^t(f)\}_{t=1}^\infty$.

Given a common discount factor $\lambda \in (0, 1)$, the function $H^\lambda : F \rightarrow \mathbb{R}^n$ assigns a payoff vector to each strategy profile of the repeated network game. Given $f \in F$, player i 's payoff, $H_i^\lambda(f) = (1 - \lambda) \sum_{t=1}^\infty \lambda^{t-1} h_i(a^t(f))$, is the $(1 - \lambda)$ -normalized discounted sum of stage game payoffs. The repeated network game associated with stage game G , discount factor λ and network g is then defined as the normal form game $G^{g,\lambda} \equiv (I, (F_i)_{i \in I}, (H_i^\lambda)_{i \in I})$.

When g is complete, $i(1) = I \setminus \{i\}$ for all $i \in I$ and $G^{g,\lambda}$ is identical to the infinitely repeated discounted game, referred to as G^λ . In this case, f_i simplifies: for any $t > 1$ it is now a mapping from $A^{t-1} = (\times_{i \in I} A_i)^{t-1}$ to A_i , that is, each player conditions his action choice on the history of action profiles chosen by the players between periods 1 and $t - 1$.

Moreover, each player has complete information about the game to be played, the form of the network and the strategy choices available to all players. Finally and importantly, each player i observes his payoff with a delay of $d_i - 1$ periods in order to prevent him from deducing other players' behavior by observing his payoff. At any $t \geq d_i$, however, player i knows the action profiles played between periods 1 and $t - d_i + 1$, and hence, he can calculate or equivalently observe his payoff for all these periods.

2.4 Payoff Vectors Generated by Sequential Equilibria

2.4.1 Individual Rationality without Full-Dimensional Payoff Space

A player's individually rational payoff is the lowest payoff to which he can be forced in a stage game. It obtains when a player maximizes his payoff while all others minimize it, and hence, is called *minmax* payoff. For any player $i \in I$, the *minmax* payoff in pure actions is defined as

$$\bar{v}_i \equiv \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} h_i(a_i, a_{-i}). \quad (1)$$

ADS use the *minmax* payoff to define a player's individually rational payoff in any repeated (network) game,⁵ in which the dimension of the payoff space is equal to the number of players, or at most of one dimension less. They show that this dimensionality condition holds whenever no two players have equivalent payoff functions in the corresponding stage game. Such games fulfill the NEU-condition (of non-equivalent utilities).⁶

Two distinct players i and j have equivalent utilities (EU), when one player's payoff function is the positive affine transformation of the other's, that is, there are $\alpha > 0$ and $\beta \in \mathbb{R}$ such that for all $a \in A$,

$$h_i(a) = \alpha h_j(a) + \beta. \quad (2)$$

This relation between EU-players i and j is denoted by $i \sim j$. When (2) is violated for two distinct players i and j , they have non-equivalent utilities, denoted by $i \not\sim j$. The EU-players are partitioned into U sets, S_1, \dots, S_U , such that $i \sim j$ holds for all $i, j \in S_u$, $1 \leq u \leq U$. Let $S \equiv \cup_{u \in U} S_u$, then $i \not\sim j$ holds for all $i \in S_u, j \in S_{u'}$ such that $u \neq u'$, and for all $i \notin S, j \in I \setminus \{i\}$. Finally, assume that no player is universally indifferent among all action profiles, that is, for all $i \in I$, there are $a, a' \in A$ such that $h_i(a) \neq h_i(a')$.

When a stage game does not fulfill the NEU-condition, that is, $S \neq \emptyset$, a player's *effective minmax* payoff is his individually rational payoff in the corresponding repeated (network) game. Following Wen (1994), the *effective minmax* payoff in pure actions of

⁵All results mentioned in this section extend without loss of generality to the repeated network game.

⁶Intuitively, two players have non-equivalent utilities when the projection of the payoff space on the corresponding two player plane yields an ellipse or a line with negative slope. Conversely, a positively-sloped line arises when one player's payoff increases monotonically in the other's. A stage game fulfills the NEU-condition if there is no pair of players whose payoff space is a positively-sloped line and at most one for which it is a negatively-sloped line.

any player $i \in S_u$ is defined as⁷

$$\nu_i \equiv \min_{a \in A} \max\{h_i(a_j, a_{-j}) \mid j \in S_u, a_j \in A_j\}. \quad (3)$$

In each EU-group, a reference player is selected whose maximization yields any member of the group who is minimized the largest possible payoff. The *effective minmax* payoff of an EU-player, therefore, is larger than or equal to his *minmax* payoff, while for all other players $i \notin S$ the two payoffs are identical.

Denote the vector of *effective minmax* payoffs in pure actions by ν , and the pure action profile forcing player i to his *effective minmax* payoff by \bar{a}^i . It is one solution to the optimization problem on the right-hand-side of (3), on which the players agreed. Without loss of generality the *effective minmax* payoff of all players is normalized to 0, that is, for all $i \in I$, $h_i(\bar{a}^i) \equiv 0$. All players with equivalent utility to i 's obtain a payoff of 0 as well when he is forced to his *effective minmax* payoff.

In a perfect monitoring model, the decisions of all players in an EU-group are identical since they are based on the same information—the commonly observed history of the repeated game. Hence, one player could represent the entire group. Conversely, in the repeated network game, each member of an EU-group chooses an action based on the observations he made thus far, and usually, these do not coincide.

2.4.2 Set of Feasible and Strictly Individually Rational Payoff Vectors

The *set of feasible payoff vectors* of the repeated (network) game is defined as

$$\mathcal{F} = \{x \in \mathbb{R}^n \mid \exists \{a^t\}_{t=1}^{\infty} : \forall t \geq 1, a^t \in A, \text{ and } \forall i \in I, x_i = (1 - \lambda) \sum_{t=1}^{\infty} \lambda^{t-1} h_i(a^t)\}.$$

Following Sorin (1986) and Fudenberg, Levine and Maskin (1994), any payoff vector in $co(G)$ is feasible for $\lambda \in (1 - \frac{1}{z}, 1)$, where z is the number of vertices of $co(G)$. For any discount factor in this range, the sets \mathcal{F} and $co(G)$ coincide. Moreover, any feasible payoff vector is achievable by a sequence of pure action profiles in the repeated (network) game.

The *set of feasible and strictly individually rational payoff vectors* is denoted by \mathcal{F}^* . It contains all feasible payoff vectors that are larger than $\nu = (0, \dots, 0)$ and is defined as

$$\mathcal{F}^* = \mathcal{F} \cap \{x \in \mathbb{R}^n \mid x > \nu\}.$$

Any payoff vector in this set is a candidate to be supported by a sequential equilibrium.

⁷Wen (1994) defines the *effective minmax* concept in mixed actions, assuming however, that a player's deviation within the support of his mixed action is observable to the other players. For the general case, which includes unobservable deviations from mixed actions, this concept is defined by FLT.

2.4.3 Sequential Equilibrium

Although Kreps and Wilson's (1982) definition of sequential equilibrium is for finite extensive form games of imperfect information, it extends to such games when they are infinite. A strategy profile and a system of beliefs are a sequential equilibrium if they are sequentially rational and consistent, respectively. In the repeated network game, for any equilibrium strategy profile at least one consistent system of beliefs exists, in which the limit of the players' beliefs coincides with the (already known or still unknown) history of the repeated network game. Hence, beliefs are not modelled explicitly and a sequential equilibrium is said to exist when the condition of sequential rationality is fulfilled.

Definition 1. A strategy profile $\dot{f} \in F$ is a sequential equilibrium of $G^{g,\lambda}$, if for all $t \geq 1$ and given any $ob^t \in Ob^t$, $\{f^\tau(ob^{\tau-1})\}_{\tau=t+1}^\infty$ is such that for all $i \in I$ and all $f_i \in F_i$,

$$(1 - \lambda) \sum_{s=t+1}^{\infty} \lambda^{s-1} h_i(a^s(\dot{f})) \geq (1 - \lambda) \sum_{s=t+1}^{\infty} \lambda^{s-1} h_i(a^s(f_i, \dot{f}_{-i})).$$

When g is complete this definition includes G^λ and the concepts of sequential and subgame-perfect equilibrium coincide. For simplicity, equilibria of $G^{g,\lambda}$ and G^λ are called sequential when Definition 1 is satisfied, and the corresponding sets of sequential equilibrium strategy profiles are denoted by $SE(G^{g,\lambda})$ and $SE(G^\lambda)$, respectively. A strategy profile is a sequential equilibrium if, and only if, any player's finite unilateral deviation at any point in time is not profitable.⁸

3 The network makes a difference

The following example illustrates how imposing a network on a set of players may affect the set of sequential equilibria of a repeated game. Let $\hat{G} = (I, A, h)$ be a generalized Prisoner's Dilemma game, where $n > 2$. At each point in time, a player chooses between two pure actions: C which stands for *cooperate* and D which stands for *defect*. The payoff function of any player $i \in I$ is defined as follows: for each $a \in A$,

$$h_i(a) = \begin{cases} 3 & \text{if } a_j = C, \forall j \in I \\ 0 & \text{if } a_i = C \text{ and } \exists j \in I \setminus \{i\} \text{ s.t. } a_j = D \\ 4 & \text{if } a_i = D \text{ and } a_j = C, \forall j \in I \setminus \{i\} \\ 2 & \text{if } a_i = D, \exists j \in I \setminus \{i\} \text{ s.t. } a_j = D \text{ and } \exists l \in I \setminus \{i, j\} \text{ s.t. } a_l = C \\ 1 & \text{if } a_j = D, \forall j \in I. \end{cases}$$

⁸Since $\lambda < 1$, a player's gain from a deviation of infinite length can be approximated by that of a finite deviation. Therefore, unilateral deviations of finite length from a strategy profile are not profitable if, and only if, it is a sequential equilibrium of the repeated network game (Mailath and Samuelson (2006)).

The unique Nash Equilibrium of stage game \hat{G} is the action profile in which all players choose D , since it is a strictly dominant action. In the repeated Prisoner's Dilemma, however, it is possible to sustain strategy profiles that yield a higher payoff to all players and are sequential equilibria under certain conditions, such as the trigger strategy profile. It prescribes each player to cooperate as long as all other players cooperate and to defect forever if any other player defected. Given any network g , the trigger strategy of player i , denoted by $\hat{f}_i \in F_i$, is defined as follows: $\hat{f}_i^1 = C$, and for $t \geq 1$, given $ob_i^t \in Ob_i^t$,

$$\hat{f}_i^{t+1}(ob_i^t) = \begin{cases} D & \text{if } \exists 1 \leq \tau \leq t \text{ such that for } a^\tau \in ob_i^t, a_j^\tau = D, \text{ while } a_{-j}^\tau = C \\ C & \text{otherwise.} \end{cases}$$

Given $\hat{f} \in F$, observe that for all $i \in I$ and all $t \geq 1$, first $a_i^t(\hat{f}) = C$, and second, $ob_i^t(\hat{f})$ is such that for all $a_j^\tau \in ob_i^t(\hat{f})$, $a_j^\tau = C$ as well for all $1 \leq \tau \leq t$ and all $j \in I$. Hence, for all $i \in I$, $H_i^\lambda(\hat{f}) = (1 - \lambda) \sum_{t=1}^{\infty} \lambda^{t-1} h_i(a^t(\hat{f})) = (1 - \lambda) \sum_{t=1}^{\infty} \lambda^{t-1} 3 = 3$.

3.1 The Players form a Star

Consider a *star* with $n = 3$, where the graph of g is $E = ((1, 2), (2, 1), (2, 3), (3, 2))$, represented in Figure 1. (Figure 2 represents \hat{G} for $n = 3$, where player 1 chooses rows, player 2 columns and player 3 matrices.) The trigger strategy profile is a sequential equilibrium of $\hat{G}^{g,\lambda}$ if, and only if, all players are patient enough, that is, λ is higher than some threshold value. Then, none of them ever deviates. Corresponding conditions on λ must be found for the truncation of the repeated network Prisoner's Dilemma after any point in time, and therefore, given any observation profile. However, to keep this example simple, only unilateral deviations are considered, that is, simultaneous deviations of two or more players, by convention, do not occur.⁹ Then, three classes of unilateral deviations can be identified. Any deviation that may arise in the course of play can be uniquely allocated to one class. The three classes are

- 1) initial unilateral deviations,
- 2) subsequent unilateral deviations (before the initial is known by all players), and
- 3) unilateral deviations when the punishment takes place.

Obviously, unilateral deviations during the punishment are not profitable since all players play D . The resulting action profile is the stage game Nash Equilibrium in strictly

⁹For example, player 1 in Figure 1 cannot distinguish between a unilateral deviation by 2 and a multilateral (simultaneous) one by 2 and 3 until he knows the action profile of the period in which 2 deviated. Multilateral deviations are abstracted from here, but taken into account from section 4 on.

1—2—3

Figure 1: Three players form a Star

3

	C		D
1-2	C	D	
C	3, 3, 3	0, 4, 0	
D	4, 0, 0	2, 2, 0	

	C		D
1-2	C	D	
C	0, 0, 4	0, 2, 2	
D	2, 0, 2	1, 1, 1	

Figure 2: Prisoner's Dilemma for three players

dominant actions. Hence, every player plays his best-reply independently of g and of λ . For the same reason, no player can deviate profitably from the trigger strategy profile in part 2. After a player's initial deviation, he and any player who knows about it are best-off to play D forever (rather than to deviate and to choose C at any point in time).

It remains to show that no player has a profitable unilateral deviation from the trigger strategy profile when all players play C . Given λ , player 2 (who is directly observed by 1 and 3) does not deviate in any period τ if, and only if,

$$(1 - \lambda) \sum_{t=1}^{\infty} 3\lambda^{t-1} \geq (1 - \lambda) \sum_{t=1}^{\tau-1} 3\lambda^{t-1} + 4(1 - \lambda)\lambda^{\tau-1} + (1 - \lambda) \sum_{t=\tau+1}^{\infty} 1\lambda^{t-1},$$

$$(1 - \lambda) \sum_{t=\tau+1}^{\infty} 2\lambda^{t-1} \geq (1 - \lambda)\lambda^{\tau-1},$$

$$2\lambda^{\tau+1} \geq (1 - \lambda)\lambda^{\tau},$$

$$\lambda \geq \frac{1}{3}.$$

The value of $\frac{1}{3}$ is not only the threshold value for player 2 in this example but also the one for all players in a complete network. The network affects, however, the threshold value of the remaining two players in this example. Given λ , player 1 (and similarly 3) does not deviate from the trigger strategy profile in any period τ if, and only if,

$$(1 - \lambda) \sum_{t=1}^{\infty} 3\lambda^{t-1} \geq (1 - \lambda) \sum_{t=1}^{\tau-1} 3\lambda^{t-1} + 4(1 - \lambda)\lambda^{\tau-1} + 2(1 - \lambda)\lambda^{\tau} + (1 - \lambda) \sum_{t=\tau+2}^{\infty} 1\lambda^{t-1},$$

$$(1 - \lambda)\lambda^\tau + (1 - \lambda) \sum_{t=\tau+2}^{\infty} 2\lambda^{t-1} \geq (1 - \lambda)\lambda^{\tau-1},$$

which can be simplified to $2\lambda + \lambda^2 - 1 \geq 0$. The only positive solution for λ in this quadratic equation is approximately 0.414. Hence, in part 1 of the sequential equilibrium conditions the requirement on λ , or the patience of the players, is higher in the *star* with three players considered here than in a complete network. This is due to the one period delay with which players 1 and 3 obtain information about each other's action choice.

This example extends to the case where $n > 3$ and the players form a *star*. The player at the center of the *star* has the same role as player 2 in this example, and for all other players the same conditions apply as for players 1 and 3 in this example.

3.2 The Repeated Prisoner's Dilemma Played in any Network

A similar result can be derived for any network. Suppose that $n > 3$ and that all players in a network follow the trigger strategy profile. Then, an analogous calculation to the one for players 1 and 3 in the above example yields a condition such that no player $i \in I$ deviates. The corresponding expression is $2\lambda + \lambda^{d_i} - 1 \geq 0$. Although it depends on d_i , even in very large networks the threshold value for λ is bounded above by $\frac{1}{2}$. Hence, for "moderately patient" players, the trigger strategy profile is a sequential equilibrium in any repeated network Prisoner's Dilemma when there are no multilateral deviations.

Another general result for the Prisoner's Dilemma as defined before can be obtained—still abstracting from multilateral deviations. Given g and λ , it is possible to determine for any sequence of action profiles, and not only the one generated by the trigger strategy profile, whether it can be supported by a sequential equilibrium strategy profile. The key step is to calculate each player's *worst payoff* which he can ensure himself by playing D forever from any point in time on. A player's *worst payoff* is determined by the *largest distance* between him and any other player in the network. This is the time it takes until all players punish him, thereby best-replying to his deviation. It also depends on the sequence of action profiles played by the other players until they are informed about his deviation. A given sequence of action profiles can be generated by a sequential equilibrium strategy profile, if it yields each player at any point in time a continuation payoff that is larger than the player's corresponding *worst payoff* at that point in time.

It is possible to calculate an upper and a lower bound to a player's *worst payoff*. For any $f \in SE(\hat{G}^{g,\lambda})$, the *worst payoff* of any player i in the repeated network Prisoner's Dilemma lies between the two identified bounds. The lower bound is identical to player i 's

(*effective*) *minmax* payoff $\bar{v}_i = 1$. (The two concepts coincide in the Prisoner's Dilemma since it fulfills the NEU-condition.) It is obtained when all players play D forever after his deviation (and it is independent of the network and the discount factor). The upper bound depends on a player's position in the network and on the discount factor. It is achieved, for example, for the trigger strategy profile. In this case, a deviator can gain most since all players play C until they become aware of his deviation.

The trigger strategy profile $\hat{f} \in F$ generates a sequence of action profiles such that for all $i \in I$ and any $t \geq 1$, $a_i^t(\hat{f}) = C$. After deviating unilaterally player i receives

$$(1 - \lambda)[2 + 2\lambda + \dots + 2\lambda^{d_i-2} + 1\lambda^{d_i-1} + \dots] =$$

$$(1 - \lambda)\left[\sum_{t=1}^{d_i-1} 2\lambda^{t-1} + \sum_{t=d_i}^{\infty} 1\lambda^{t-1}\right] =$$

$$2 - \lambda^{d_i-1}.$$

This upper bound of a player's *worst payoff* is strictly larger than 1, unless the network is complete, that is, $d_i = 1$ for all players. For different values of λ and depending on a player's position in the network it lies between 1 and 2, as depicted in Figure 3. For small values of λ , it is close to 2 even when player i 's *largest distance* is small. Conversely, for λ close to 1, the upper bound of a player's *worst payoff* is close to 1 even in large networks.

Hence, the network may reduce the set of discount factors for which a strategy profile is a sequential equilibrium, and moreover, for a given discount factor the set of sequential equilibrium strategy profiles and the corresponding set of payoff vectors may be strictly smaller in the repeated network game than in the version with complete network. The next step is to extend this result to repeated network games based on any stage game.

4 Information Sharing and Punishment Reward

In general, the conditions for sequential equilibria are not as simple as in the repeated network Prisoner's Dilemma. First, the action profile forcing a player to his *effective minmax* payoff does not coincide with a stage game Nash Equilibrium in strictly dominant actions, and hence, punishment is non-symmetric and may be costly for some players. Second, the players' behavior after multilateral deviations has to be determined. The approach taken here, is to let the players wait until everyone knows whether a deviation was uni- or multilateral. This allows the players, moreover, to coordinate their punishment. In this section both issues are dealt with starting with the second one.

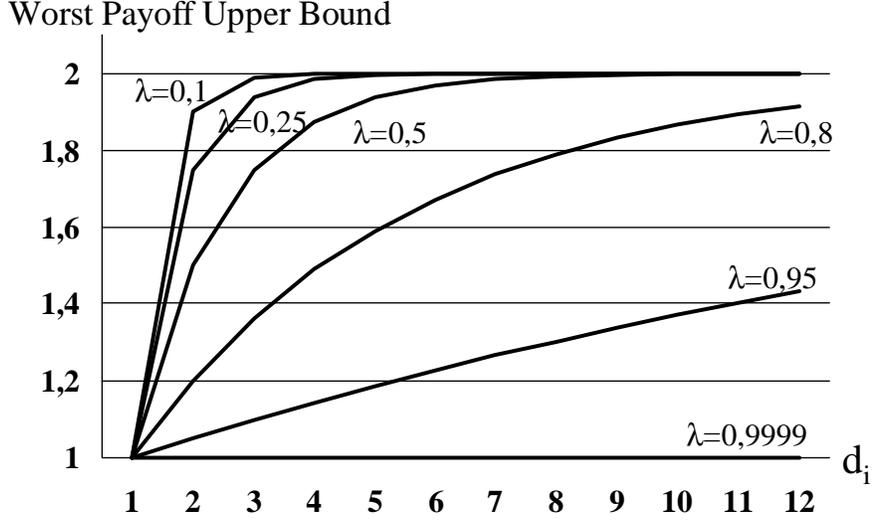


Figure 3: Upper bound of a player's worst payoff in the Prisoner's Dilemma

Until all players in the network know about an initial deviation, they are required to follow the sequence of action profiles, although the deviator may continue to deviate or subsequent deviations by other players may occur. Once all players have identified the initial deviator, they start to punish him. In case the initial deviation was multilateral, however, the players ignore it.¹⁰ This phase of information transmission is called Information Sharing Process (*ISP*). Note, that the *ISP*-payoff is not normalized by $(1 - \lambda)$.

Definition 2. Given $f \in F$, the Information Sharing Process payoff of player i following an initial deviation in period t' only is defined as

$$ISP_i^{t'} = h_i(a^{t'+1}(f)) + \dots + \lambda^{d-2} h_i(a^{t'+d-1}(f)).$$

Note, that an action profile is known by all players after $d - 1$ periods. The *ISP* can be extended easily to cover a deviation of finite length by any player. Any subsequent unilateral deviator with non-equivalent utility to the initial one starts a new *ISP* which, however, may overlap with the ongoing one. Once every player has identified the last deviator, he is forced to his *effective minmax* payoff at least until his entire gain from deviating is taken away or until another subsequent deviator is punished. All players that contribute to the punishment may incur a loss in their own payoff as long as it

¹⁰In this case, the limit of the players' beliefs is as follows. A player believes that all others follow the strategy profile, unless he observes an initial deviation. Then, he either believes that it was unilateral or multilateral or any average of both until knowing the truth. In equilibrium, any such belief is consistent.

lasts. Hence, punishment should be restricted to a minimal amount of time, and the punishers should be rewarded thereafter. Obviously, the reward must not be beneficial for the deviator—otherwise the punishment would be reversed again.

Assume without loss of generality that all NEU-players occupy positions $1, \dots, \hat{i}$ in I , and that thereafter all players in the distinct EU-groups S_1 to S_U follow. In analogy to ADS, given any feasible and strictly individually rational target payoff vector $x \in \mathcal{F}^*$, for all NEU-players there are player-specific punishment reward payoff vectors denoted by $\omega^1, \dots, \omega^{\hat{i}}$. They can be achieved by sequences of pure action profiles and have the following properties. For any player $i \notin S$, $x_i > \omega_i^i > 0$, and for two distinct players $i \approx j$, $\omega_i^i < \omega_j^j$, that is, the i -th component of vector i is strictly smaller than that of any other one.

For EU-players the punishment reward phase is simpler. Some time after player $i \in S_u$ deviated, all members of S_u are subjected to the same punishment reward since their payoffs are equivalent to i 's. Hence, it is enough to define one punishment reward payoff vector for each EU-group. A cascade of deviations by players in S_u is prevented by taking away the gain each of the deviators in this EU-group obtained, that is, by forcing the players in S_u to their *effective minmax* payoff for a long enough amount of time. Thereafter, the group's punishment reward phase is played. Hence, for each group S_u , $1 \leq u \leq U$, there is one punishment reward payoff vector ω^{S_u} .

Given any target payoff vector $x \in \mathcal{F}^*$, the punishment reward payoff vectors $\omega^1, \dots, \omega^{\hat{i}}, \omega^{S_1}, \dots, \omega^{S_U}$, have the following properties:

- i) for all $i \notin S$, $x_i > \omega_i^i > 0$,
and for any $1 \leq u \leq U$ and all $i \in S_u$, $x_i > \omega_i^{S_u} > 0$.
- ii) a) For all $i \neq j$, $i, j \notin S$, $\omega_i^i < \omega_j^j$,
b) for any $1 \leq u \leq U$, all $i \in S_u$ and all $j \notin S$, $\omega_i^{S_u} < \omega_j^j$ and $\omega_j^j < \omega_j^{S_u}$,
c) for all $i \in S_u$, $j \in S_{u'}$ such that $u \neq u'$, $\omega_i^{S_u} < \omega_i^{S_{u'}}$ and $\omega_j^{S_{u'}} < \omega_j^{S_u}$,
d) and for any $1 \leq u \leq U$, and all $i, j \in S_u$, there are $\alpha > 0$ and $\beta \in \mathbb{R}$
such that $\omega_i^{S_u} = \alpha \omega_j^{S_u} + \beta$.

The conditions in part i) are target payoff vector domination and individual rationality. The ones in part ii) ensure that a player is worst off during his or his EU-group's punishment reward phase, but that he can be rewarded otherwise.

The existence of the punishment reward payoff vectors for any $x \in \mathcal{F}^*$ follows from ADS, who construct them explicitly and give the following geometric interpretation, graphically illustrated in Figure 4. For two distinct players $i \approx j$, the projection of

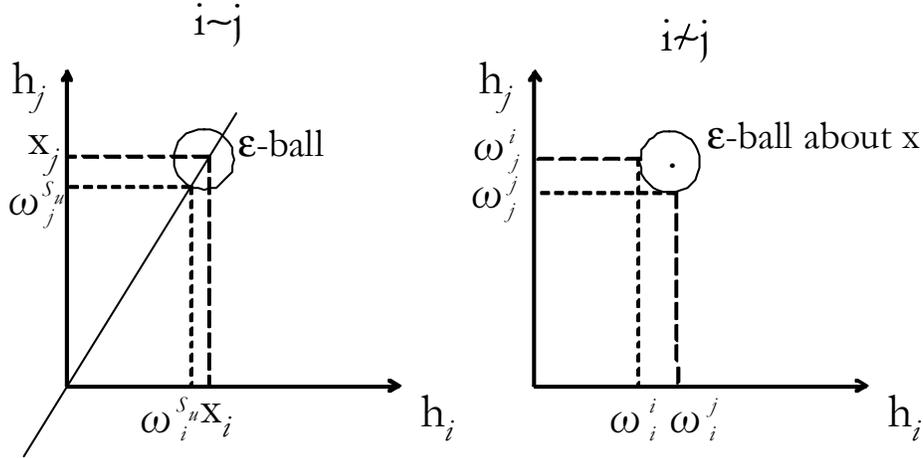


Figure 4: Construction of Punishment Reward Payoff Vectors for EU- and NEU-Players

the payoff space on the corresponding two player plane yields an ellipse or a line (with negative slope), whereas for all others it is a line (with positive slope). In the first case, the smallest i - and j -coordinates on a ball with arbitrarily small radius $\varepsilon > 0$ about the target payoff vector gives the payoff that the corresponding player receives in his punishment reward phase. In any other case, the EU-group's punishment reward payoff vector is the lowest point, in which the line and the ε -ball about the target payoff vector intersect. This intersection determines the punishment reward payoff of each player in the group.

5 The Results

5.1 Folk Theorem

As explained above, a strategy profile in the repeated network game is a sequential equilibrium if, and only if, given any observation profile, no player's unilateral deviation from the continuation strategy profile is profitable. Since each observation profile that may arise can be uniquely allocated to one of a small number of classes of observation profiles, it is necessary and sufficient to show for each class that any player's finite unilateral deviation is not profitable. The outline and proof of the Folk Theorem (which can be found in appendix A) adapt some arguments of ADS and Wen to the network case.

Theorem 1. *Let G and g be given. Then, for all $x \in \mathcal{F}^*$, there is $\dot{\lambda} < 1$ such that for each $\lambda \in (\dot{\lambda}, 1)$, there is a corresponding $\tilde{f} \in F$ such that $\tilde{f} \in SE(G^{g,\lambda})$ and $H^\lambda(\tilde{f}) = x$.*

Intuitively, strategy profile \tilde{f} prescribes the players to punish a unilateral deviator once all of them know that he deviated at least until his entire gain is taken away. Thereafter, his or his EU-group's punishment reward phase is played. Given any observation profile, unilateral deviations from \tilde{f} are not profitable when λ is close enough to 1. The Folk Theorem can be proved as well, possibly even for lower discount factors than $\hat{\lambda}$, using other strategy profiles, as discussed in the next subsection. However, most of them are technically and intuitively more involved than \tilde{f} .

Patient players in the network do not mind to obtain the history of the repeated game gradually over time. Immediate punishment or punishment that sets in after a finite delay are equivalently strong threats for the players in this case. In the limit, the network specific effects disappear and the same set of payoff vectors can be generated by sequential equilibria in the repeated game and in its network version.

Corollary 1. *Let G and g be given. Then, there is $\bar{\lambda} < 1$ such that for all $\lambda \in (\bar{\lambda}, 1)$ and all $x \in \mathcal{F}^*$, there are $\tilde{f} \in SE(G^{g,\lambda})$ and $\bar{f} \in SE(G^\lambda)$ such that $\{a^t(\tilde{f})\}_{t=1}^\infty \equiv \{a^t(\bar{f})\}_{t=1}^\infty$, and $H^\lambda(\tilde{f}) = H^\lambda(\bar{f}) = x$.*

There is also a lower bound of the discount factor $\underline{\lambda}$, and the corollary holds as well for all $\lambda \in [0, \underline{\lambda}]$. For this range of discount factors, only sequences of action profiles that prescribe the infinite repetition of stage game Nash Equilibria can be supported by sequential equilibria in both games. Another Folk Theorem follows from Theorem 1.

Corollary 2. *Let G , g and $f \in F$ be given and assume there is $\hat{\lambda} < 1$ such that $f \in SE(G^{g,\hat{\lambda}})$. Then, for all $\lambda \in [\hat{\lambda}, 1)$, $f \in SE(G^{g,\lambda})$ and $H^\lambda(f) > 0$.*

Since network g is assumed to be undirected, a simple structure on the information transmission obtains. As already hinted in the introduction, however, the players may not be able to obtain information about each other simultaneously. The Folk Theorem extends to repeated games played on directed networks that are connected since each player still gets to know the repeated game's history with a finite delay. Apart from this, the observation and the communication network may not coincide. A player may observe a neighbor, but not be able to communicate with him. Denote by (I, E^{Ob}) and (I, E^{Com}) the *observation* and the *communication* graph of the observation network g^{Ob} and the communication network g^{Com} , respectively. The two graphs are defined as (I, E) . However, both may be directed and fulfill the following connectedness property. Each player is observed by at least one other player. The players communicate their observations via a directed network g^{Com} such that all of them obtain the repeated network game's

history after a finite delay.¹¹ For any network g^{OC} , consisting of an observation network g^{Ob} and a communication network g^{Com} , the Folk Theorem holds.

Corollary 3. *Let G and g^{OC} be given. Then, for all $x \in \mathcal{F}^*$, there is $\ddot{\lambda} < 1$ such that for each $\lambda \in (\ddot{\lambda}, 1)$, there is a corresponding $\ddot{f} \in F$ such that $\ddot{f} \in SE(G^{g^{OC}, \lambda})$ and $H^\lambda(\ddot{f}) = x$.*

Finally, note that for a given set of players the network in which the delay after which punishment starts is largest in a *tree*, that is, a line of length $n - 1$. In this case, the *diameter* among all networks that can be formed from the set of players is maximal. Given G , let \hat{g} be an arbitrary *tree* network formed by the players in set I . Then, the following corollary follows from Theorem 1.

Corollary 4. *Let G , \hat{g} and $f \in F$ be given. Assume that $f \in SE(G^{\hat{g}, \lambda})$ for all $\lambda \in (\hat{\lambda}, 1)$. Then, for any g formed by set I and all $\lambda \in (\hat{\lambda}, 1)$, $f \in SE(G^{g, \lambda})$ and $H^\lambda(f) > 0$.*

In other networks than *trees* the *diameter* is lower, and hence also the requirement on the players' level of patience. In general, f may be a sequential equilibrium even for lower discount factors when the players form any other network than a *tree*.

5.2 Impatient Players

For impatient players, or in other words, for a range of discount factors strictly below 1, the network may make a difference. For the Prisoner's Dilemma this was shown in section 3 abstracting, however, from multilateral deviations. The aim in this section is to derive a similar result for any stage game, any network and including multilateral deviations. Ideally, it should state that for all discount factors larger than $\underline{\lambda}$, identified after Corollary 1, and smaller than or equal to $\dot{\lambda}$, identified in the Folk Theorem, the set of payoff vectors generated by sequential equilibria in the repeated game is a strict superset to the corresponding payoff set in its network version. However, as already mentioned, the Folk Theorem may hold for lower discount factors than $\dot{\lambda}$ when other strategy profiles than \tilde{f} are used. To identify them allows to reduce the effects due to the imposition of the network on a repeated game. Two profiles which achieve this are described. Under both, the players use the information they receive earlier than under \tilde{f} .

Given any network, a player can start to punish a deviator, for example, when he knows the action profile of the period, in which the deviation occurred. Until then, he cannot rule out that the deviation was multilateral. Hence, with respect to any player i , the time delay, with which the players can identify player i 's unilateral deviation, induces

¹¹I am very grateful to Elchanan Ben-Porath who suggested the idea of two separate networks.

a partition of the set of players such that all players in a group observe the action profile played in the period of i 's deviation with the same delay. After some initial delay, during which i 's deviation is unpunished, the players which first identified i 's deviation start to punish him. Since the network is connected the group of punishers, thereafter, grows strictly in each period until it comprises all players, d periods after i 's deviation. At the end of the previous period, all players know that the deviation occurred.

Alternatively, all players may immediately punish any deviating neighbor.¹² In subsequent periods the group of punishers grows strictly until it comprises all players. The delay until this is the case is determined by the deviator's *largest distance*. For at least one pair of players, this coincides with the network's *diameter* (which is the maximal *largest distance* between any pair of players). If a player becomes aware that the initial deviation was multilateral, he resumes playing the sequence of action profiles.

In both cases, a consistent system of beliefs exists, for example, as described in footnote 10. To illustrate both ideas, consider a network whose graph is as depicted in Figure 1 and any stage game G . A unilateral deviation by player 1 (and similarly by player 3) is immediately identified by player 2 since he also observes player 3's action choice in the period of player 1's deviation. Hence, from the subsequent period on, player 2 punishes player 1. Player 3 contributes to the punishment only from one period afterwards on.

In the repeated network game, the *diameter* of the network thus determines when the group of punishers comprises all players. Only then, punishment can be as effective as in a complete network already one period after the deviation. Hence, the threat of punishment in any network is always equally or less strong than in a complete one.

It remains to determine a form of punishment which eliminates the deviator's gain entirely but at the same time minimizes the loss the punishers may incur. In general, this is impossible without specifying the stage game, the network, the discount factor and the sequence of action profiles, since it is not obvious, if it is better to "minmax" a deviator, to start a punishment reward phase or some other sequence of action profiles. However, in a sequential equilibrium strategy profile, unilateral deviations cannot be ignored, and hence, the time delay caused by the imposition of a network on a repeated game may reduce the set of sequential equilibria. This is expressed formally in Corollary 5, which is complementary to Corollary 1 and the condition stated thereafter.

Corollary 5. *Let G and g be given. Then, there are $0 < \underline{\lambda} \leq \bar{\lambda} < 1$ such that for all $\lambda \in (\underline{\lambda}, \bar{\lambda}]$, $\{\{a^t(f)\}_{t=1}^\infty \mid f \in SE(G^{g,\lambda})\} \subset \{\{a^t(\bar{f})\}_{t=1}^\infty \mid \bar{f} \in SE(G^\lambda)\}$.*

¹²Suppose $S = \emptyset$, that i deviates at t' from $\{a^t\}_{t=1}^\infty$ and that any deviator is "minmaxed" at once. Then, one period after, i 's payoff is $\min_{a_{i(1)} \in \times_{j \in i(1)} A_j} \max_{a_i \in A_i} h_i(a_i, a_{i(1)}, a_{-(i \cup i(1))}^{t'+1}) \leq \max_{a_i \in A_i} h_i(a_i, a_{-i}^{t'+1})$. In case $S \neq \emptyset$, an analogous condition can be found.

When the network is complete, the lower and upper bound of λ coincide and the corollary is trivially true. Otherwise, it is easily proved by induction with the arguments given above. The interplay between the delay in information transmission caused by the network and the patience of a player may reduce the set of sequential equilibria and the corresponding payoff set, although for large enough discount factors the Folk Theorem holds. Comparative statics on g have similar effects on $SE(G^{g,\lambda})$, when due to the removal or addition of a link the network's *diameter* or some player's *largest distance* changes.

Finally, formal conditions are identified under which the network reduces the set of sequential equilibria for impatient players. Given G , λ and g , assume that $\bar{f} \in SE(G^\lambda)$ and let $\{\hat{a}^t\}_{t=1}^\infty \equiv \{a^t(\bar{f})\}_{t=1}^\infty$. Say that the *network has an impact with respect to \bar{f}* , as defined in Theorem 1, if \bar{f} does not support $\{\hat{a}^t\}_{t=1}^\infty$ as a sequential equilibrium of $G^{g,\lambda}$. (Note, however, that this does not rule out that there is some other strategy profile $f \neq \bar{f}$ such that $f \in SE(G^{g,\lambda})$ and $\{a^t(f)\}_{t=1}^\infty = \{\hat{a}^t\}_{t=1}^\infty$). Suppose that player i can gain

$$\beta_i^\tau \equiv \sum_{t=\tau}^{\tau+d-1} \lambda^{t-\tau} [\max_{a_i \in A_i} h_i(a_i, \hat{a}_{-i}^t) - h_i(\hat{a}^t)]$$

by a deviation (of length $d-1$) from $\{\hat{a}^t\}_{t=1}^\infty$ that starts at τ . If $i \notin S$, let

$$\delta_i^\tau(T) \equiv \sum_{t=\tau+d}^\infty \lambda^{t-\tau-1} h_i(\hat{a}^t) - (1-\lambda)^{-1} \lambda^T \omega_i^i$$

for $T \geq 2d-2$. (It takes $d-1$ periods until all players know about i 's deviation, and $2d-2$ periods after it, all of them know if i deviated again one period before his punishment started.) An analogous expression can be obtained when $i \in S$. Then, Proposition 1 identifies conditions under which the *network has an impact with respect to \bar{f}* .

Proposition 1. *Let G , $\lambda < 1$ and g be given. Suppose there is $\bar{f} \in SE(G^\lambda)$, $i \in I$ and $\tau \geq 1$, such that for all positive integers $T \geq 2d-2$, $\beta_i^\tau > \delta_i^\tau(T)$. Then, the network has an impact with respect to \bar{f} .*

Appendix B contains the proof of Proposition 1. Intuitively, player i deviates from $\{\hat{a}^t\}_{t=1}^\infty$, if the punishment threat prescribed by strategy profile \bar{f} is discounted by too much, and hence, not strong enough to prevent i 's deviation. Therefore, the strategy profile defined in Theorem 1 does not support the sequence of action profiles $\{\hat{a}^t\}_{t=1}^\infty$ as a sequential equilibrium of $G^{g,\lambda}$, and the *network has an impact with respect to \bar{f}* . Similar conditions can be identified for any other strategy profile than \bar{f} .

5.3 Strategic Communication and Related Literature

Although in certain cases strong social or legal norms may impose truth-telling on impatient economic agents, this assumption should be relaxed. The players could either decide whether to transmit information or not, or even lie. A player can be easily prevented from stopping the information transmission by the threat to punish him as if he had deviated. The second type of deviation, therefore, is more interesting but also more involved. In a sequential equilibrium initial as well as subsequent deviations have to be prevented and a sequence of different liars and deviators may be impossible to disentangle for a player.

To assume that players may lie is standard in the literature. Compte (1998) and Kandori and Matsushima (1998), for example, model imperfect private monitoring in repeated games (without network) as follows. Each player receives a distinct distorted private signal of the period's action profile. By publicly announcing these private observations every $K > 0$ periods, the players restore a public history on which they condition their action choices and a Folk Theorem obtains. Without communication the players' beliefs about where in the game tree they are might diverge and some player's profitable deviation may be undetectable. Kandori (2003) uses a similar idea in the case of imperfect public monitoring in which all players observe the same imperfect signal of the period's action profile and publicly announce their own action choices. Under strategic communication a Folk Theorem obtains under weaker conditions than without communication.¹³ In all cases, a payoff transfer mechanism provides incentives for the players to make truthful announcements. A player's payoff increases or decreases depending on his announcement. A similar mechanism induces the players in Ben-Porath and Kahneman (1996 and 2003) to truthfully announce their own and any neighbor's action choice publicly. These constructions require a full-dimensional payoff space which is even stronger than the NEU-condition.

In the repeated network game, the payoff of an EU-group may have to be increased and decreased at the same time under such a payoff transfer mechanism. This, however, is impossible. The presence of EU-players may also create the following problem. Suppose that two or three players that monitor each other belong to the same EU-group. Then, cooperation immediately breaks down because all other players anticipate a sequence of deviations by the EU-players which these will, obviously, not reveal when communicating their mutual observations of each other. Since in this setup a player does not communicate his own action choice, but only those of his neighbor(s), the problem can be solved by isolating the EU-players. They would, for example, occupy the places of players 1 and 3

¹³The Folk Theorem under imperfect public monitoring without communication in Fudenberg, Levine and Maskin (1994) holds if the public signal allows the players to statistically detect unilateral deviations.

in the graph depicted in Figure 1, while a NEU-player would take player 2's.¹⁴

Nevertheless, due to the bilateral communication structure, strategic communication is difficult to introduce to the repeated network game. Each player receives different information gradually over time, and hence, all players can never simultaneously condition their action choices on the same (communicated) information. To prevent lies, therefore, requires, apart from isolating the EU-players in a network, to adapt the punishment reward phase. In order not to tell the lie that another player deviated, for each player $i \in I$, let $x_i > \omega_i^j$ for all $j \notin S$, and $x_i > \omega_i^{S_u}$ for all S_u . By lying a player makes himself worse off under this condition. To induce a player to truthfully reveal any neighbor's deviation, additionally, for all $i \in I$, let $\omega_i^k > \max_{a_k \in A_k} h_i(a_k, a_{-k}^t)$ for all $k \in i(1)$ with $k \notin S$ and any $t \geq 1$. The analogous condition must hold when $k \in S$. Then, by tolerating k 's deviation i is worse off than by reporting it. The information that k deviated flows throughout the network only, if this incentive constraint holds sequentially for all players in $k(1)$, $k(2)$, and so on. Then, lies are unprofitable for λ close enough to 1 and restricting the players to initial deviations only. After a history that includes lies and deviations, the construction of the punishment reward payoff vectors has to be revised in order to maintain the incentives for truthtelling and complying with the strategy profile. The punishment reward phase thus becomes history dependent and to show that it can be adapted adequately is very complicated, if not impossible without making further assumptions.

However, there is a possibility to prevent lying in this setup. Since each player $i \in I$ observes his payoff with a delay of $d_i - 1$ periods, the players can revert to a stage game Nash Equilibrium forever, if this observation is not compatible with the payoff calculated using the prescribed sequence of action profiles. When each player's deviation changes the payoff of all players, all payoff vectors which dominate that of a stage game Nash Equilibrium can be supported by sequential equilibria in the repeated network game with strategic communication. If players can, moreover, (statistically) detect which player deviated, Theorem 1 obtains and truthtelling is achieved endogenously.

To allow for strategic communication is appealing for two reasons. First, imperfect private monitoring, which so far is imposed exogenously in many models, could be made endogenous in a general class of games. Instead of letting each player obtain a probabilistically determined amount of information, more realistically, this amount should depend on strategic decisions of other players. Second, information asymmetries that arise in repeated strategic interaction, such as hidden actions or hidden knowledge, could be modelled in this way. Therefore, this seems a promising direction for further research.

¹⁴Cooperation can be sustained in any *star* as well when the player at the center has a constant payoff.



Figure 5: a) two triads

b) no triad

5.4 Network Analysis

The result that in a repeated game played on a (fixed) network its *diameter* determines whether cooperation is sustainable, for a given discount factor, is new to the network literature. Conversely, various results in this literature emphasize that cooperation depends on the clustering coefficient, which gives the ratio of triads or circles of three players in a network relative to all possible combinations of three players in I . Whereas the *diameter* is a global measure, the clustering coefficient measures local connectedness. Its importance in the network literature is due to two sociology papers. Granovetter (1973) defines the concept of *strong links* which exist, for example, between three friends when they form a triad (or a circle). This facilitates cooperation since the three friends mutually observe each other's behavior. Coleman (1988), in turn, develops the concept of *closures*, which are circles of connected people as well but not necessarily of size three.

To see that a lower *diameter* in a network need not imply a higher clustering coefficient, consider the two networks depicted in Figure 5; for both $n = 6$. The network in part a) has 7 links and two triads, whereas the *wheel* in part b) has 9 links and no triad. The *diameter* of network a) is 3 and the one in part b) is 2. The clustering coefficient of the *wheel* is zero, whereas the other network's one is positive. Hence, the relationship between the clustering coefficient and the *diameter* in a network need not be monotonic. (Obviously, one could construct other examples in which the monotonic relation holds.)

Cooperation in the setup of this paper can be sustained more easily in network b). Hence, given any stage game, for a certain range of discount factors, there are sequential equilibria in network b) which generate sequences of action profiles that do not arise from sequential equilibria in network a). (However, as the Folk Theorem implies, for patient players this difference disappears.) Conversely, in the network literature, network a) would fare much better in terms of sustaining cooperation than the one in part b) of Figure 5.

The importance of the clustering coefficient is emphasized, for example, in Lippert and Spagnolo (2005), who model relational contracts by letting each linked pair of players play a bilateral repeated discounted Prisoner's Dilemma until one player deviates, which

severs the link. They analyze different informational setups, including a case in which players can choose not to transmit information, and conclude that closures are crucial to sustain cooperation, modelled in form of sequential equilibria. Another example is Vega-Redondo, Marsili, and Slanina (2005), who let each linked pair of players play a bilateral Prisoner's Dilemma, in which the payoffs are stochastically decaying over time. A player severs a link once his payoff falls below some threshold, although as a consequence he is punished by all mutual neighbors the two players have. However, a player can create new links in each period. This yields a dynamic process whose parameter choice influences the form of the network in the long-run. Both papers are examples of setups, in which the repeated game played as well as the communication and observation process are bilateral.

Since usually a player's payoff depends not only on his and his neighbor's decisions but also on decisions of other players in the network, even if they are "far away", it seems realistic to consider repeated games played on a fixed network.¹⁵ The three papers most closely related to this, however, also obtain that closures are decisive to sustain equilibria. A crucial condition for Ben-Porath and Kahneman (1996) to sustain a sequential equilibrium in their repeated game with public announcements is that there are at least three players in each group. Then, any liar can be detected in equilibrium. This is exactly identical to strong links. In their paper with costly monitoring, Ben-Porath and Kahneman (2003) require a similar condition to hold. In Renault and Tomala (1998), in which strategic communication includes lying as well, cooperation can be sustained only in networks that are 2-connected. Intuitively, this requires two distinct paths to exist between any pair of players and is just the formal description of a closure.

The delayed perfect monitoring model yields a different result since it assumes bilateral communication. Players become informed about the repeated network game's history gradually over time. To the contrary, after a public announcement in Ben-Porath and Kahneman (1996 and 2003), all players can immediately punish any deviator. In models, in which communication, observations and the repeated games played are bilateral, punishment is also immediate. In Renault and Tomala (1998), play is interrupted until all players know who has cheated. Simultaneously, the payoff accumulation stops, which is unimportant since the repeated game is undiscounted. In the repeated network game, the impatient players—except of the deviator—may suffer from the delay, during which punishment is less effective than in a complete network. The *diameter* of the network captures this delay and determines together with the discount factor whether a strategy

¹⁵However, a network formation game might precede the repeated network game, and as mentioned in subsection 5.2, comparative statics on the topology of the network are straightforward. Hence, this setup can be extended to explicitly take into account incentives to form or maintain links.

profile is a sequential equilibrium. This result extends to cases where the players can lie but truth-telling prevails in equilibrium. Other assumptions may also be responsible for the different outcomes. In particular, a deeper analysis of the matrices that contain the distinct networks might yield interesting results.¹⁶

6 Final Remarks

6.1 Mixed Actions

The extension of the Folk Theorem to mixed actions is straightforward in the complete network. Additionally, a player's deviation within the support of his mixed action, which is not observed by the other players, must be prevented. FLT achieve this in the complete network setting by making future play dependent on the realized action profile today. By letting a high payoff today follow a low one tomorrow, and vice versa, FLT can make each player exactly indifferent among all pure actions in the support of a mixed one.¹⁷

The Folk Theorem for the repeated network game can be extended to mixed actions using FLT's idea. A player who is punished would be forced to his *effective minmax* payoff in mixed actions. After the number of periods equivalent to the *diameter* of the network has passed, every player knows the pure action profile generated by the mixed action in the first punishment period. Punishment continues, anyway, at least until this period, and then, FLT's strategy can be used to compensate the players for their choices in the first punishment period. Thereafter, the second punishment period is compensated, and so on. This process stops in finite time. The main advantage of this extension is that a larger set of payoff vectors can be sustained by sequential equilibria. However, patient players can achieve first best outcomes already with pure actions.

6.2 Conclusion

In this paper, delayed perfect monitoring in an infinitely repeated discounted game is modelled by letting the players form a connected (and undirected) network. The Folk Theorem obtains since patient players do not mind to receive the repeated game's history

¹⁶Network g can be expressed in a symmetric matrix which is of dimension $n \times n$. Each entry in the off-diagonal represents the distance between the corresponding two players and all entries in the diagonal, by convention, are set equal to 1. Similar matrices can be generated for the other models.

¹⁷The main difficulty arises for players with positively and negatively related payoff functions since their payoff space is a line. Since it is only possible to move "up" and "down" the line in the future, it is non-trivial to define a continuation strategy which makes each player indifferent among the support of the mixed action. However, FLT construct a strategy profile which achieves this task in finite time.

gradually over time. To achieve truth-telling endogenously is complicated, however, due to the bilateral communication structure and the less than full-dimensional payoff space. For impatient players the network may make a difference which need not be big, as shown for the Prisoner's Dilemma. The interplay between the *diameter* of the network and the patience of the players leads to the reduction in the set of sequential equilibria. This paper also contributes to the network literature, which so far emphasized the importance of the clustering coefficient for the level of cooperation sustainable in a network.

As already mentioned in the introduction, this setup can be applied to various specific contexts. Not only companies, but also impatient people form networks and interact strategically over time, such as within a company, in any other organization, or in society at large. As long as all of them are on the same hierarchical level, this model applies. Also macroeconomic applications can be thought of. The players in a network thus might be all the companies in an economy and a deviation could be interpreted as one of them going bankrupt.¹⁸ The network effects in repeated strategic interaction can also be observed in financial markets. For example, innovative financial strategies, such as those used by hedge funds, spread throughout a network over time. Whereas at the beginning only few players use a certain strategy, over time everyone adopts a successful one.

Finally, extending the idea of *worst payoff* introduced in section 3, an upper and a lower bound of a player's payoff after deviating can be calculated and a recursive structure in this setup may be obtained. By applying dynamic programming arguments, it may then be possible to obtain a superset and a subset of the set of equilibrium payoffs. Also, noise can be introduced into this setup by letting the players only observe distorted signals of each other's behavior. Moreover, an imperfect signal's informativeness might be decaying over time, that is, along the shortest path between two players the quality of the signal about each other's action choice deteriorates.

References

- Abreu, D., Dutta, P. K. and Smith, L. (1994), "The Folk Theorem for Repeated Games: A New Condition," *Econometrica* 62, 939-948.
- Ben-Porath, E. and Kahneman, M. (1996), "Communication in Repeated Games with Private Monitoring," *Journal of Economic Theory* 70, 281-297.

¹⁸I am very grateful to Albert Marcet who suggested this example.

- Ben-Porath, E. and Kahneman, M. (2003), "Communication in Repeated Games with Costly Monitoring," *Games and Economic Behavior* 44, 227-250.
- Coleman, J. S. (1988), "Social Capital in the Creation of Human Capital," *American Journal of Sociology* 94, S95-S120.
- Compte, O. (1998), "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica* 66, 597-626.
- Fudenberg, D., Levine, D. K. and Maskin, E. (1994), "The Folk Theorem with Imperfect Public Information," *Econometrica* 62, 997-1039.
- Fudenberg, D., Levine, D. K. and Takahashi, S. (2006), "Perfect Public Equilibria when Players are Patient," *mimeo*.
- Fudenberg, D. and Maskin, E. (1986), "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica* 54, 533-554.
- Granovetter, M. S. (1973), "The Strength of Weak Ties," *American Journal of Sociology* 78, 1360-1380.
- Kandori, M. (2002), "Introduction to Repeated Games with Private Monitoring," *Journal of Economic Theory* 102, 1-15.
- Kandori, M. (2003), "Randomization, Communication, and Efficiency in Repeated Games with Imperfect Public Monitoring," *Econometrica* 71, 345-353.
- Kandori, M. and Matsushima, H. (1998), "Private Observation, Communication and Collusion," *Econometrica* 66, 627-652.
- Kreps, D. M. and Wilson, R. (1982), "Sequential Equilibria," *Econometrica* 50, 863-894.
- Lippert, S. and Spagnolo, G. (2005), "Networks of Relations and Social Capital," *mimeo*.
- Mailath, G. J. and Samuelson, L. (2006), *Long-Run Relationships*, Oxford University Press.
- Renault, J. and Tomala, T. (1998), "Repeated Proximity Games," *International Journal of Game Theory* 27, 539-559.
- Sorin, S. (1986), "On Repeated Games with Complete Information," *Mathematics of Operations Research* 11, 147-160.

Vega-Redondo, F., Marsili, M. and Slanina, R. (2005), “Clustering, Cooperation, and Search in Social Networks,” *Journal of the European Economic Association* 3, 628-638.

Wen, Q. (1994), “The ‘Folk Theorem’ for Repeated Games with Complete Information,” *Econometrica* 62, 949-954.

Appendix A Proof of Theorem 1

Given G and g , fix $x \in \mathcal{F}^*$ and note that x , as well as any other payoff vector in $co(G)$, is feasible when $\lambda \in (1 - \frac{1}{z}, 1)$, where z is the number of vertices of $co(G)$ —see subsection 2.4.2. Hence, let $\dot{\lambda} = \max\{\tilde{\lambda}, 1 - \frac{1}{z}\}$, where $\tilde{\lambda} < 1$ is determined below. Then, for each $\lambda \in (\dot{\lambda}, 1)$, there is a corresponding sequence of pure action profiles $\{a^s\}_{s=1}^{\infty}$ which yields x . When λ changes, the sequence of action profiles that generates x may differ. Hence, strategy profile $\tilde{f} \in F$, which thereafter is defined and shown to be a sequential equilibrium of $G^{g,\lambda}$ for any $\lambda \in (\dot{\lambda}, 1)$, may prescribe a different sequence of action profiles for each λ , although its structure is unchanged. For any $j \in I$, define $\tilde{f}_j \in F_j$ as follows:

$\tilde{f}_j^1 = a_j^1$, and for $t > 1$, given $ob_j^{t-1} \in Ob_j^{t-1}$, in a slight abuse of notation, let $\tilde{f}_j^t(ob_j^{t-1}) =$

- 1) a_j^t , unless there is $1 \leq t' < t$ such that for $\hat{a}^{t'} \in ob_j^{t'-1}$, $\hat{a}_i^{t'} \neq a_i^{t'}$, while $\hat{a}_{-i}^{t'} = a_{-i}^{t'}$. In this case, switch to phase 2 at $t' + d_j$ and let $\tilde{a}_j^s = a_j^s$, for all $s \geq 1$.
- 2) \tilde{a}_j^t , if $t' + d_j \leq t < t' + d$, unless player l , where $l \neq i$ and $l \notin S_u$ if $i \in S_u$, deviates at any t'' , where $t' + d > t'' > t'$. Then, restart phase 2, set $t' = t''$ and choose \tilde{a}_j^s accordingly. Otherwise, switch to phase 3 at $t' + d$.
- 3) \bar{a}_j^t , if $t' + d \leq t \leq t' + T$, where T is determined below. If any player l deviates at any \bar{t} , where $t' + T \geq \bar{t} \geq t' + d$, restart phase 2, set $t' = \bar{t}$ and choose \tilde{a}_j^s accordingly. Otherwise, switch to phase 4 at $t' + T + 1$.
- 4) c_j^s , if $t \geq t' + T + s$, where $\{c^s\}_{s=1}^{\infty}$ is the sequence of action profiles that yields either ω^i if $i \notin S$, or ω^{S_u} if $i \in S_u$. If any player l deviates at any $\tau > t' + T$, restart phase 2, set $t' = \tau$ and choose \tilde{a}_j^s accordingly. If $l = i$ or $i, l \in S_u$, restart $\{c^s\}_{s=1}^{\infty}$ where it was truncated by l 's deviation, once phase 4 is reached again.

Phase 2 corresponds to the *ISP*, phase 3 to the *effective minmax* punishment of the last deviator, and phase 4 to the punishment reward phase.

After any subsequent unilateral deviation, the phase in which the game is at the time

of the deviation prescribes the play of the following $d - 1$ periods—in general, phase 2 is restarted. Then, the new deviator is punished. In case, the same player deviates again in phase 2 (and no other one does), however, this phase is not restarted, but his punishment begins d periods after his first deviation. His entire gain is eliminated, that is, he is forced to his *effective minmax* payoff for at least $d - 1$ periods, or longer, if necessary. After $d - 1$ periods of punishment, all players know if he deviated again in the period before his punishment started, and hence, for how long it has to continue in order to eliminate his entire gain from deviating. A similar argument applies for several unilateral deviations by distinct players of an EU-group during phase 2. After punishing the initial deviator (for at least $d - 1$ periods), the gain of the subsequent one(s) is eliminated.

By construction, the players can ignore multilateral deviations from \tilde{f} , and it remains to show that no player's unilateral deviation from \tilde{f} is ever profitable for large enough λ . The Folk Theorem holds trivially when a^t is a stage game Nash Equilibrium for all t , and hereafter, only strategy profiles that do not generate such sequences of action profiles are considered. Finally, a consistent system of beliefs, given \tilde{f} , is specified in footnote 10.

The proof is organized as follows. The result for phase 2 is shown first since it introduces arguments used thereafter to prove the results of phases 4, 1 and 3. Note, that the following 6 combinations of players' deviations have to be shown to be unprofitable; for the first four $i \approx j$ holds, whereas for the remaining two $i \sim j$ holds: $i \neq j$ and either $i, j \notin S$; or $i \in S$, but $j \notin S$; or $j \in S$, but $i \notin S$; or $i \in S_u$, $j \in S_{u'}$ such that $u \neq u'$; and finally, $i, j \in S_u$, or $i = j$. For each phase, the proof proceeds in this order.

PHASE 2

Figure 6 illustrates the order of time periods in phase 2. Suppose player $i \notin S$ deviated at t' . During the *ISP* player $j \neq i$, $j \notin S$, receives $ISP_j^{t'}$. By deviating at t'' , where $t' < t'' < t' + d$, he can maximally gain $b_j = \max_{a \in A} [\max_{\bar{a}_j \in A_j} h_j(\bar{a}_j, a_{-j}) - h_j(a)]$, since his remaining *ISP*-payoff is unchanged. However, from period $t'' + d$ on, he is forced to his *effective minmax* payoff of 0, and then, his punishment reward phase is played. Player j 's deviation at t'' is not profitable when for some positive integer \hat{T}_2 , where $t'' + d \leq t' + \hat{T}_2$,

$$(1 - \lambda)b_j + \lambda^{\hat{T}_2} \omega_j^j - (1 - \lambda) \sum_{t=t''+d}^{t'+\hat{T}_2} \lambda^{t-t''-1} h_j(\bar{a}^i) - \lambda^{t'+\hat{T}_2-t''} \omega_j^i < 0,$$

$$(1 - \lambda)b_j - (1 - \lambda) \sum_{t=t''+d}^{t'+\hat{T}_2} \lambda^{t-t''-1} h_j(\bar{a}^i) < \lambda^{t'+\hat{T}_2-t''} \omega_j^i - \lambda^{\hat{T}_2} \omega_j^j. \quad (4)$$

Substituting $\lambda^{t'+\hat{T}_2-t''}$ with $\lambda^{\hat{T}_2}$ makes the right-hand-side of (4) smaller. Since $t'' > t'$,

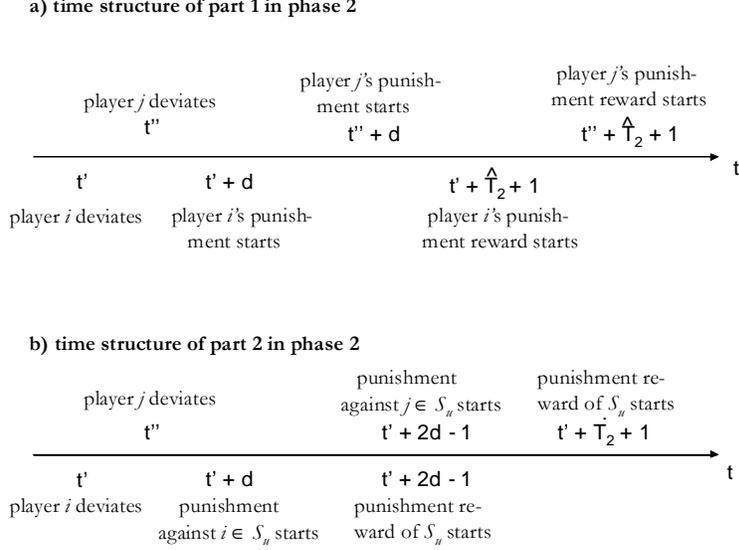


Figure 6: Order of time periods in phase 2

$\lambda^{t'+\hat{T}_2-t''} > \lambda^{\hat{T}_2}$ holds for all $\lambda < 1$. Hence, (5) implies (4) and it suffices to show (5).

$$(1 - \lambda)b_j - (1 - \lambda) \sum_{t=t''+d}^{t'+\hat{T}_2} \lambda^{t-t''-1} h_j(\bar{a}^i) < \lambda^{\hat{T}_2} [\omega_j^i - \omega_j^j] \quad (5)$$

As λ converges to 1, (5) is fulfilled: its left-hand-side converges to zero whereas its right-hand-side is strictly positive since $\omega_j^i > \omega_j^j$. This may hold for several distinct pairs of discount factor and strictly positive integer. (The last inequality is fulfilled trivially when player j 's gain from punishing player i is larger than b_j). An analogous argument holds, whenever $i \approx j$. The case $t'' + d > t' + \hat{T}_2$ is simpler since the sum on the left-hand-side of (5) drops out as well as j 's payoff in the first period(s) of i 's punishment reward phase, which for λ close to 1 is negligible.

For $i, j \in S_u$, after player j 's deviation at any t'' , where $t' < t'' < t' + d$, the *ISP* about i 's deviation continues. Once all players know about i 's deviation, \bar{a}^i is played for at least $d - 1$ periods, that is, at least until period $t' + 2d - 2$. Then, $\bar{a}^j \equiv \bar{a}^i$ is played until period $t' + \hat{T}_2$, to take away player j 's gain from deviating at t'' . Since j 's punishment lasts at least one period, $\hat{T}_2 > 2d - 2$. Thereafter, the EU-group's punishment reward phase is played. Player j 's deviation at t'' is not profitable, if for some positive integer $\hat{T}_2 > 2d - 2$,

$$(1 - \lambda)b_j + \lambda^{t'+\hat{T}_2-t''} \omega_j^{S_u} - \lambda^{t'+2d-2-t''} \omega_j^{S_u} < 0,$$

$$(1 - \lambda)b_j < (\lambda^{t'+2d-2-t''} - \lambda^{t'+\hat{T}_2-t''}) \omega_j^{S_u},$$

$$b_j < \lambda^{t'+2d-2-t''}(1-\lambda)^{-1}(1-\lambda^{\hat{T}_2-(2d-2)})\omega_j^{S_u}.$$

When λ converges to 1, the right-hand-side converges to $(\hat{T}_2 - 2d + 2)\omega_j^{S_u} > 0$, by l'Hospital. Since b_j is a fixed positive number, the inequality is fulfilled for a large enough \hat{T}_2 . A similar argument applies when several distinct players with equivalent utility to i 's deviate sequentially during the *ISP* about i 's deviation or when $i = j$, that is, one player deviates in several (subsequent) periods. Finally, select a large enough, strictly positive integer T_2 such that no player can deviate profitably in phase 2.

PHASE 4 and PHASE 1

The result for phase 4 is stated first since it implies the result for phase 1. Suppose that player $i \neq j$, that $i, j \notin S$, and that i is the last deviator. Player j does not deviate at τ , the first period of i 's punishment reward phase, if for some positive integer \hat{T}_4 ,

$$(1-\lambda)\max_{a_j \in A_j} h_j(a_j, c_{-j}^1) + \lambda(1-\lambda)ISP_j^\tau + \lambda^{\hat{T}_4}\omega_j^j - \omega_j^i < 0,$$

$$(1-\lambda)\max_{a_j \in A_j} h_j(a_j, c_{-j}^1) + \lambda(1-\lambda)ISP_j^\tau < \omega_j^i - \lambda^{\hat{T}_4}\omega_j^j.$$

When λ converges to 1, the left-hand-side of the last inequality converges to zero whereas the right-hand-side is strictly positive (since $\omega_j^i > \omega_j^j$, and for any $\lambda < 1$, $\lambda^{\hat{T}_4} < 1$). The same argument holds whenever $i \approx j$, and when player j deviates in any other than the first period of player i 's punishment reward phase since for λ close to 1, the payoff obtained at the beginning of any punishment reward phase is negligible.

If $i = j$, player i cannot deviate profitably in the $\hat{\tau}$ th period of his own punishment reward phase, if there is a positive integer \hat{T}_4 such that

$$(1-\lambda)b_i + \lambda(1-\lambda)ISP_i^\tau + \lambda^{\hat{T}_4}\omega_i^i|_{s=\hat{\tau}+1}^\infty - \omega_i^i|_{s=\hat{\tau}+1}^\infty < 0,$$

where $\tau \equiv t' + \hat{T}_4 + \hat{\tau}$ and $\omega_i^i|_{s=\hat{\tau}+1}^\infty \equiv (1-\lambda)\sum_{s=\hat{\tau}+1}^\infty \lambda^{s-1}h_i(c^s)$. This simplifies to

$$(1-\lambda)b_i + \lambda(1-\lambda)ISP_i^\tau < \omega_i^i|_{s=\hat{\tau}+1}^\infty - \lambda^{\hat{T}_4}\omega_i^i|_{s=\hat{\tau}+1}^\infty,$$

$$b_i + \lambda ISP_i^\tau < \frac{(1-\lambda^{\hat{T}_4})}{(1-\lambda)}\omega_i^i|_{s=\hat{\tau}+1}^\infty. \quad (6)$$

When λ converges to 1, the left-hand-side of (6) is bounded above by a positive number and the right-hand-side, by l'Hospital, converges to $\hat{T}_4\omega_i^i|_{s=\hat{\tau}+1}^\infty > 0$. (Although,

$\omega_i^i|_{s=\hat{\tau}+1}^\infty$ differs from ω_i^i , for λ close to 1, this difference is negligible and $\omega_i^i|_{s=\hat{\tau}+1}^\infty$ has the same properties as ω_i^i . For \hat{T}_4 large enough, (6) holds. A similar argument applies when $i, j \in S_u$, and j deviates in the punishment reward phase of his EU-group. This argument together with the one used in phase 2 above demonstrates that any player's unilateral deviation of finite length is neither profitable in phase 4. Finally, let T_4 be the smallest positive integer such that no player can deviate profitably in phase 4.

The result of phase 4 extends to phase 1 since by assumption any player's target payoff is strictly larger than his punishment reward payoff. Moreover, neither finite deviations by one player nor subsequent deviations by distinct players in an EU-group are profitable in phase 1. Hence, also for phase 1 there is a discount factor $\lambda < 1$ and a positive integers T_1 such that no player can deviate profitably from strategy profile \tilde{f} .

PHASE 3

Suppose player i is forced to his *effective minmax* payoff because he deviated at t' . By definition, neither player i nor any player $j \sim i$ can deviate profitably in this phase. Hence, suppose $i, j \notin S$. Player j does not deviate at any \bar{t} , where $t' + d \leq \bar{t} \leq t' + T_3$, if

$$(1 - \lambda)b_j + \lambda(1 - \lambda)ISP_j^{\bar{t}} + \lambda^{T_3}\omega_j^j - (1 - \lambda)\sum_{t=\bar{t}}^{T_3}\lambda^{t-\bar{t}}h_j(\bar{a}^i) - \lambda^{t'+T_3-\bar{t}}\omega_j^i < 0,$$

$$(1 - \lambda)b_j + \lambda(1 - \lambda)ISP_j^{\bar{t}} - (1 - \lambda)\sum_{t=\bar{t}}^{T_3}\lambda^{t-\bar{t}}h_j(\bar{a}^i) < \lambda^{t'+T_3-\bar{t}}\omega_j^i - \lambda^{T_3}\omega_j^j. \quad (7)$$

Proceeding as in phase 2, that is, substituting on (7)'s right-hand-side $\lambda^{t'+T_3-\bar{t}}$ with λ^{T_3} (for any $\lambda < 1$, $\lambda^{T_3-(\bar{t}-t')} > \lambda^{T_3}$ since $\bar{t} > t'$) and taking the limit of λ converging to 1, fulfills (7) for at least one pair of discount factor $\lambda < 1$ and strictly positive integer T_3 . An analogous argument holds for deviations, or a sequence of deviations, by EU- and NEU-players. Choose T_3 large enough to prevent any such deviation.

Let $T = \max\{T_1, T_2, T_3, T_4\}$, and let $\tilde{\lambda}$ be the lowest discount factor, for which, given T , no player can deviate profitably in any phase. (If there are several pairs of T and λ for which the proof holds, the pair with the lowest discount factor is selected.) Finally, let $\dot{\lambda} = \max\{\tilde{\lambda}, 1 - \frac{1}{z}\}$. Then, for any $\lambda \in (\dot{\lambda}, 1)$, \tilde{f} is a sequential equilibrium strategy profile of $G^{g,\lambda}$ and $H^\lambda(\tilde{f}) = x$.

Appendix B Proof of Proposition 1

Fix G , $\lambda < 1$ and g . Select $\bar{f} \in SE(G^\lambda)$ that generates the sequence of action profiles $\{a^t(\bar{f})\}_{t=1}^\infty \equiv \{\dot{a}^t\}_{t=1}^\infty$. Take a strategy profile with the same structure as \bar{f} , defined in Theorem 1, to support this sequence of action profiles as a sequential equilibrium of $G^{g,\lambda}$. Then, the *network has an impact with respect to \bar{f}* if some player can deviate profitably. Consider first, that for some $i \notin S$, some $\tau \geq 1$, and all positive integers $T \geq 2d - 2$,

$$(1 - \lambda) \sum_{t=\tau}^{\tau+d-1} \lambda^{t-\tau} \max_{a_i \in A_i} h_i(a_i, \dot{a}_{-i}^t) + \lambda^T \omega_i^i > (1 - \lambda) \sum_{t=\tau}^{\infty} \lambda^{t-\tau} h_i(\dot{a}^t),$$

$$\sum_{t=\tau}^{\tau+d-1} \lambda^{t-\tau} [\max_{a_i \in A_i} h_i(a_i, \dot{a}_{-i}^t) - h_i(\dot{a}^t)] + (1 - \lambda)^{-1} \lambda^T \omega_i^i > \sum_{t=\tau+d}^{\infty} \lambda^{t-\tau-1} h_i(\dot{a}^t).$$

Subtracting $(1 - \lambda)^{-1} \lambda^T \omega_i^i$ from both sides yields $\beta_i^\tau > \delta_i^\tau(T)$. The *network has an impact with respect to \bar{f}* if either the last inequality holds for some $i \notin S$ or an analogous condition for some $i \in S$. In the second case, ω_i^i is substituted with $\omega_i^{S_u}$.