Francesco Nava and Michele Piccione

Efficiency in repeated games with local interaction and uncertain local monitoring

Working paper

Efficiency in Repeated Games with Local Interaction and Uncertain Local Monitoring

Francesco Nava and Michele Piccione*

Third Version: November 2012

Abstract

The paper discusses community enforcement in infinitely repeated, two-action games with local interaction and uncertain monitoring. Each player interacts with and observes only a fixed set of opponents, of whom he is privately informed. The main result shows that when beliefs about the monitoring structure have full support, efficiency can be sustained with sequential equilibria that are independent of the players' beliefs. Stronger results are obtained when only acyclic monitoring structures are allowed or players have unit discount rates. These equilibria satisfy numerous robustness properties.

*London School of Economics.

1 Introduction

In many strategic environments, interaction is local and segmented. Competing neighborhood stores by and large serve different yet overlapping sets of customers, the behavior of the residents of an apartment block affects their contiguous neighbors to a larger extent than neighbors in a different block, a nation's foreign or domestic policy typically generates larger externalities for neighboring nations than for remote ones. One classic case is the private provision of local public goods in which the strategic interaction is modelled using either a prisoner's dilemma or by a hawk-dove game. For example, many forms of anti-social behavior are generally captured by the former whereas investments in common security, infrastructure or maintenance that yield benefits only when a fixed cut-off level is reached by the latter. In addition to local interaction, one notable feature of these environments is uncertain monitoring: whereas participants are aware of their own neighbors' identities and actions, they are not necessarily aware of the identity and actions of their neighbors' neighbors.

Within these strategic environments, it is of particular interest to study long run interaction, when incentives can only be provided locally in a decentralized manner. Our objective is to analyze such interaction within a repeated game framework that differs from the standard one by allowing actions to be observed only locally. Such framework, despite its plainness and its potential applications, has not yet produced significant results in the literature. A natural question that we will address is whether local community enforcement suffices to generate efficient behavior. The main obstacle to sustaining cooperation is that information about individuals' past behavior in a relationship is local: it is common knowledge within the relationship, but is not necessarily available to outsiders. The absence of publicly observable histories implies that punishment are no longer based on "simultaneous" coordination: by punishing a neighbor's deviation, a player can trigger subsequent punishments from different neighbors, who were not related to the original defector and were thus unable to observe the initial deviation. Thus, if a shop ceases to collude in order to punish defections by a neighboring competitor, it will affect the behavior of other neighboring shops that were not affected by the first defection. Moreover, as such defections spread through neighborhoods, they might return to one of the players who was either a source of such defection or had retaliated to it, and enter cycles. Naturally, in these circumstances the construction of equilibrium incentives for cooperative behavior and the derivation of equilibrium beliefs is a challenging task.

1.1 Summary

We study infinitely repeated two-action games. The setup consists of a finite number of players who choose in every period whether to cooperate or defect. A graph that represents the monitoring structure, the information network, is realized at the beginning of the game. Each player is privately informed of his neighborhood, namely the subset of players with whom he will interact in bilateral relationships for an infinite number of periods, but receives no information as to other players' neighborhoods. A player observes only the actions played by his neighbors and, crucially, cannot discriminate among them by choosing different actions. That is, in every period a player chooses one action that applies to all bilateral relationships in his neighborhood. All the players play the same game in all neighborhoods.

We show that, for sufficiently high discount rates and any beliefs with full support about the monitoring structure, sequential equilibria exist in which the efficient stage-game outcome is played in every period. It should be noted that standard results do not apply because bilateral enforcement may not be incentive compatible when punishments in one relationship affect outcomes in all the others. For instance, punishing a neighbor indefinitely with a grim trigger strategy is not viable if cooperation in other relationships is disrupted, and modifications as in Ellison (1994) work only for particular specifications of payoffs. Indeed, equilibrium strategies will be such that, after any history, players' believe that cooperation will eventually resume.

Our proofs are constructive, and exploit simple bounded-punishment strategies which are robust with respect to the players' priors about the monitoring structure. In particular, in the equilibria characterized only local information matters to determine players' behavior. Efficiency is supported by strategies that respond to defections with further defections. When the players' discount rate is smaller than one, the main difficulty in the construction of sequentially rational strategies that support efficiency is the preservation of short-run incentive compatibility after some particular histories of play. When defections spread through a network, two complications arise. The first occurs when a player expects future defection coming from a particular direction. Suppose that somewhere in a cycle, for example, a defection has occurred and reaches a player from one direction. If this player does not respond, he may expect future defections from the opposite direction caused by players who are themselves responding to the original defection. This player's short term incentives then depend on the timing and on the number of future defections that he expects. In such cases, the verification of sequential rationality and the calculation of consistent beliefs can be extremely demanding. We will circumvent this difficulty via the construction of consistent beliefs such that a player never expects future defections to reach him. Such beliefs are generated trivially when priors assign positive probability only to acyclic monitoring structures. More importantly, as we shall see, such beliefs can also be generated when priors have full support. The second complication arises when a player has failed to respond to a large number of defections. On the one hand, matching the number of defections of the opponent in the future may not be incentive compatible, say when this player is currently achieving efficient payoffs with a large number of different neighbors. The restriction that a player's action is common to all neighbors is of course the main source of complications here. On the other hand, not matching them may give rise to the circumstances outlined in the first type of complications, that is, this player may then expect future defections from a different direction. The former hurdle will be circumvented by bounding the length of punishments and the latter, as before, by constructing appropriate consistent beliefs.

The above difficulties do not arise when players are patient as short-term incentives are irrelevant and punishments need not be bounded. Indeed, stronger results are obtained for the case of limit discounting in which payoffs are evaluated according to Banach-Mazur limits. We will show that efficiency is resilient to histories of defections. In particular, there exists a sequential equilibrium such that, after any finite sequence of defections, paths eventually converge to the constant play of efficient actions in all neighborhoods in every future period. An essential part of the construction is that in any relationship in which defections have occurred the number of periods in which the inefficient actions are played is "balanced": as the game unfolds from any history, both players will have played the inefficient action an equal number of times before resuming the efficient play. Remarkably, such balanced retaliations eventually extinguish themselves and always allow the resumption of cooperation throughout the network.

Although our formal analysis will be restricted to uniform discount rates and symmetric stage games with deterministic payoffs, the equilibria characterized are robust with respect to heterogeneity in payoffs and discount rates, and with respect to uncertainty in payoffs and population size, as long as the ordinal properties of the stage games are maintained across the players. The above equilibria will obviously persist as babbling equilibria in setups with communication. In addition, these equilibria can be easily modified to accommodate monitoring structures in which players interact with fewer players than they observe.

Section 2 presents the setup and defines the relevant equilibrium properties. Section 3 considers games in which players are arbitrarily patient and proves the existence of cooperative equilibria. Such equilibria are shown to be independent of the players' beliefs on the monitoring structure, and to satisfy a desirable notion of stability and several other robustness properties. Section 4 considers games with impatient players and shows how cooperation can be achieved when prior beliefs have full support. The first part of the appendix shows that results trivially extend to games in which only acyclic monitoring structures are possible. All the proofs omitted from the main text appear in the second part of the appendix.

1.2 Related Literature

This paper fits within the literature on community enforcement in repeated games. A major strand pioneered by Kandori (1992) and Ellison (1994) has focussed on environments with random matching of players and shown that efficient allocations can be sustained as equilibria when players become arbitrarily patient. Subsequent contributions include Takahashi (2008) and Deb (2011). In our model, matching is not random but determined at the beginning of the game and fixed throughout the play.

A large, growing literature investigates community enforcement in environments in which players interact with and monitor different subsets of other players under a variety of different modelling assumptions. The advantage of our framework is that it does not rely on neighbor-specific punishments, communication, or knowledge of the global monitoring structure. Some notable studies allow players to choose neighbor specific actions, such as Ali and Miller (2008), Lippert and Spagnolo (2008), Mihm, Toth and Lang (2009), Fainmesser (2010), Jackson et al (2010), Fainmesser and Goldberg (2011), while others restrict attention to environments in which the monitoring structure is common knowledge and communication is possible, such as Ahn (1997), Vega-Redondo (2006), and Kinateder (2008). The vast majority of these studies focuses on prisoner's dilemma type interactions.

Our framework is closely related to several works which, unlike our model, postulates no uncertainty about the monitoring structure. Ben Porath and Kahneman (1996) establish a sequentially rational Folk Theorem for general stage game payoffs when each player is observed by at least two other players, and when public communication and public randomization are allowed. Renault and Tomala (1998) establish a Nash Folk Theorem for special monitoring structures (in which the subgraphs obtained by suppressing any one player are still connected), general stage game payoffs, no discounting, and no explicit communication. Haag and Lagunoff (2006) consider games with prisoner's dilemma interactions and heterogeneous discount rates, and show for which monitoring structures cooperation can be sustained by local trigger strategies. Xue (2004) and Cho (2010 & 2011) also focus on the prisoner's dilemma. Cho (2010) considers acyclical networks and allows neighbors to communicate. Cho (2011) shows the existence of sequential equilibria in which players cooperate in every period and in which cooperation eventually resumes after deviations if public randomization is allowed. Xue (2004) restricts the analysis to linear networks.

Wolitzky (2012) investigates a setup similar to ours with uncertainty about the monitoring structure, and characterizes the maximal level of cooperation that can be enforced for fixed discount rates in a local public goods game with compact action sets. Unlike our model, the monitoring structure changes every period and is learned at the end of each period. This feature of the model plays an essential role in the equilibrium construction, and prevents any of his results to apply to our framework.

One significant point of departure of our paper from the above literature is the construction of equilibrium strategies. In particular, reciprocity will play a crucial role in the characterization of sequentially rational behavior. Our equilibria are somewhat evocative of the "trading favors" equilibria in Möbius (2001) and Hauser and Hopenhayn (2004), despite the frameworks bearing little resemblance. Notably, our players can be viewed "trading" punishment off the equilibrium path.

2 Setup And Equilibrium Properties

We first introduce the setup and the information structure. Then, we proceed to define the solution concept and equilibrium properties.

2.1 The Stage Game

Consider a game, the stage game, played by a set N of n players in which any player i interacts with a subset of players $N_i \subseteq N \backslash \{i\}$ of size n_i, which we call the *neighborhood* of player i. We assume that $j \in N_i$ if and only if $i \in N_j$. This structure of interaction defines an undirected graph (N, G) in which $ij \in G$ if and only if $j \in N_i$. We shall refer to G as the *information network*. Define a *path* to be an m tuple of players $(j_1, .., , j_m)$ such that $j_{k+1} \in N_{j_k}$, $k = 1, 2..., m-1$. If $j_m = j_1$, a path is a *cycle*. Given a neighborhood N_i for player i, let $\Gamma(N_i)$ be the set of information networks in which player i's neighborhood is N_i.

Players are privately informed about their neighborhood. The beliefs of player i regarding the information network, conditional upon observing his neighborhood, are derived from common prior beliefs f over the set of information networks.[1] We say that a prior f is *admissible* if, for any $i \in N$ and $M \subseteq N \backslash \{i\}$, $f(G) > 0$ for some G for which $N_i = M$. Admissibility ensures that posterior beliefs are well defined for any realization of the information network. We assume throughout the paper that priors are admissible. The set of admissible priors is denoted by Π^A.

The set of actions of player i is A_i and consists of only two actions labeled C and D. We will refer to action C as cooperation and to action D as defection. A player must

[1] The assumption that priors are common is inessential.

choose the same action for all his neighbors. That is, a player cannot discriminate across neighbors and his action must be played in his entire neighborhood. Given a subset M of players, let A_M denote $\times_{j \in M} A_j$ and a_M an element of A_M. We will often use $-i$ to denote $N \backslash \{i\}$. The payoff of any player is separable across relationships. Let η_{ij} define the *emphasis* of player i in the relationship with player j. The stage game payoff of player i is

$$v_i(a_i, a_{N_i}) = \sum_{j \in N_i} \eta_{ij} u_{ij}(a_i, a_j)$$

where $u_{ij}(a_i, a_j)$, the payoff of player i in the relationship $ij \in G$, is given by

$i \setminus j$	C	D
C	1	$-l$
D	$1+g$	0

For ease of notation, we assume that $\eta_{ij} > 0$ for any ij in G. Note that, if $\eta_{ij} = 0$ for $ij \in G$, player i observes the actions of player j but his payoff is not affected. All our results extend to the case in which some η_{ij}'s are equal to zero for $ij \in G$.

We adopt the convention that payoffs are equal to zero when N_i is empty. For simplicity, the above payoff matrix is common to all bilateral relationships. We will clarify along the analysis when this assumption can be dispensed with.

We restrict attention to stage games payoffs for which mutual cooperation is efficient. We will also assume that defection is a best response when the opponent cooperates to rule out the trivial case in which mutual cooperation is an equilibrium of the stage game. Such restrictions amount to the following assumption, which will be maintained throughout.

Assumption A1: $g - l < 1$, $g > 0$.

Payoffs are common knowledge. After the main results, we will discuss the extent to which this assumption is necessary. Naturally, if $l > 0$, the stage game has a unique Bayes Nash equilibrium in which all players play D. If instead $l < 0$, the stage game always possesses a mixed strategy Bayes Nash equilibrium.[2]

2.2 The Repetition

The players play the infinite repetition of the stage game. The information network is realized prior to the beginning of play and remains constant thereafter. In every period, a player observes only the past play of his neighbors. The set of possible histories for

[2]When $l < 0$, pure strategy equilibria also exist in some networks, as choosing actions different than their neighbors' can be a player's best reply. In particular, if beliefs are concentrated on networks with cycles of even length, pure equilibria exist, since players can successfully mis-coordinate actions with all their neighbors.

player $i \in N$ whose realized neighborhood is N_i is defined as

$$H_{i,N_i} = \{\emptyset\} \cup \{\cup_{t=1}^{\infty} \left[\times_{s=1}^{t} A_{N_i \cup \{i\}}\right]\}$$

where $\emptyset$ denotes the empty history. An interim strategy for player i with neighborhood N_i is a function σ_{i,N_i} that assigns to each history in H_{i,N_i} an action in $\{C, D\}$. The set of interim strategies of player i is Σ_{i,N_i}. A strategy σ_i of player i is a collection of interim strategies $\{\sigma_{i,M}\}_{M \subset N \setminus \{i\}}$.

Players discount the future with a common factor $\delta \leq 1$. To define the payoffs of the infinitely repeated game, fix a network G. Given a profile of strategies $\sigma_N = (\sigma_1, \sigma_2, .., \sigma_n)$, let $\{a_N^t\}_{t=0}^{\infty}$ be the sequence of stage-game actions generated by σ_N when the information network is G, and $\{v_i(a_i^t, a_{N_i}^t)\}_{t=1}^{\infty}$ be the sequence of stage game utilities of player i. Define

$$w_i^t(\sigma_N|G) = \sum_{s=1}^{t} \frac{v_i(a_i^s, a_{N_i}^s)}{t}$$

to be the average payoff up to period t and $w_i(\sigma_N|G) = \{w_i^t(\sigma_N|G)\}_{t=1}^{\infty}$ to be the sequence of average payoffs. Repeated game payoffs conditional on network G are defined as

$$\mathcal{V}_i(\sigma_N|G) = \begin{cases} (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} v_i(a_i^t, a_{N_i}^t) & \text{if} \quad \delta < 1 \\ \Lambda\left(w_i\left(\sigma_N|G\right)\right) & \text{if} \quad \delta = 1 \end{cases}$$

where $\Lambda(\cdot)$ denotes the Banach-Mazur limit of a sequence. If ℓ_∞ denotes the set of bounded sequences of real numbers, a Banach-Mazur limit is a linear functional $\Lambda : \ell_\infty \to \mathbb{R}$ such that: (i) $\Lambda(e) = 1$ if $e = \{1, 1, ...\}$; (ii) $\Lambda(x^1, x^2, ...) = \Lambda(x^2, x^3, ...)$ for any sequence $\{x^t\}_{t=0}^{\infty} \in \ell_\infty$ (see [4]). It can be shown that, for any sequence $\{x^t\}_{t=0}^{\infty} \in \ell_\infty$,

$$\liminf_{t\to\infty} x^t \leq \Lambda\left(\{x^t\}_{t=1}^{\infty}\right) \leq \limsup_{t\to\infty} x^t$$

Remark 1 *For simplicity, we will restrict players to use pure strategies. Since player i's beliefs assign positive probability to a finite number of paths for any history in* H_{i,N_i}*, linearity ensures that the expectation of the Banach-Mazur limit is the same as the Banach-Mazur limit of the expectation. Our analysis can be extended to mixed strategies with infinite supports by using special Banach-Mazur limits, called medial limits, which can be shown to exists under the continuum hypothesis (see [1]).*

Define the set of histories for the entire game to be

$$H = \{\emptyset\} \cup \{\cup_{t=1}^{\infty} \left[\times_{s=1}^{t} A_N\right]\}$$

Given a history $h \in H$, the realization of an information network G, and a profile of strategies $\sigma_N = (\sigma_1, \sigma_2, .., \sigma_n)$, define the profile $\sigma^h_{N,G} = (\sigma^h_{1,N_1}, \sigma^h_{2,N_2}, .., \sigma^h_{n,N_n})$ induced by the history h and information network G in the standard way. A pair (G, h) will be referred to as a *node* of the dynamic game.[3] A pair (N_i, h_i) of a neighborhood and an observed history (or simply an observed history h_i as its components identify the neighbors of player i) is associated uniquely with *information set* $\mathcal{I}(h_i)$ and viceversa.[4] With some abuse of notation, we will sometimes use h_i to denote $\mathcal{I}(h_i)$.

A *system of beliefs* β defines at each information set $\mathcal{I}(h_i)$ of player i the conditional probability $\beta(G, h|h_i)$ of each node $(G, h) \in \mathcal{I}(h_i)$. The marginal belief of a network G is denoted by $\beta(G|h_i)$ and of a history h by $\beta(h|h_i)$.

2.3 Equilibrium Properties

In this section, we define three properties of strategies. The first requires a strategy profile to be a sequential equilibrium that is invariant with respect to any prior beliefs in a subset of admissible beliefs.

Definition (Π Invariant Equilibrium – Π-IE): A strategy profile is a Π-*invariant equilibrium,* $\Pi \subseteq \Pi^A$, if it is a sequential equilibrium for any prior beliefs in Π.

As strategies depend on the observed neighborhood, Π-invariance requires that the players' behavior is not affected by conditional beliefs about remote parts of the network derived from priors in Π. Naturally, the scope of this requirement depends on the choice of possible beliefs. Within the confines of such choice, invariance implies that local responsiveness suffices for sequential rationality and equilibrium behavior. Relatedly, Π-invariance also implies that prior beliefs need not be common, in so far as they belong to the set Π. All the equilibrium constructions presented in the paper will satisfy some form of invariance. We highlight this property in our analysis as it establishes that efficient behavior need not be fine-tuned to the exact beliefs about the global monitoring structure: the network structure itself is immaterial in that only local information matters for the determination of a player's incentives.

The second property is straightforward and selects strategies in which every player cooperates for any information network.

Definition (Collusive – C): A strategy profile is *collusive* if the sequence of stage-game actions generated for any information network is such that the players play C in every period.

[3] Throughout, the term *vertex* is used to refer to the nodes of the information network, whereas the term *node* is used to refer to the nodes of the extensive form game.

[4] Formally define $\mathcal{I}(\bar{h}_i) = \mathcal{I}(\bar{N}_i, \bar{h}_i) = \{(G, h) \,|N_i = \bar{N}_i \text{ and } h_i = \bar{h}_i\}$.

The final property characterizes the robustness of an equilibrium to occasional defections by players. This definition is similar to, yet marginally stronger than, the notion of *global stability* defined in Kandori (1992).

Definition (Π Stability – Π-S): A strategy profile satisfies Π-*stability*, $\Pi \subseteq \Pi^A$, if for any information network G such that $f(G) > 0$ for some $f \in \Pi$ and any history $h \in H$, there exists a period T_G^h such that all the players play C in all periods greater than T_G^h.

We deem equilibria satisfying Π-stability of interest as cooperation will always resume after any number of mistakes.

The main results of this paper establish the existence of collusive strategy profiles that are Π-invariant equilibria for various choices of Π, with Π-stability sometimes playing a role in the equilibrium construction. Several additional robustness properties will be discussed after each result. Obviously, the main hurdles are brought about by the restriction that a player's action applies to indiscriminately to his entire neighborhood. If players could choose a different action for each relationship, standard results would yield a Folk Theorem.

3 Patient Players

In this section, we show that when short-term incentives are inessential, as the players' payoffs equal the long-term average, cooperation can be achieved via a simple strategy profile that satisfies Π^A-invariance and Π^A-stability. In this profile, cooperation is "balanced": as the game unfolds from any history, in each relationship a player will have defected for the same number of periods as his opponent, before reverting to permanent cooperation.

This case is obviously of interest in and of itself when long-run payoffs are the sole players' motive in the strategic interaction. More importantly, it brings into focus two considerations. First, retaliatory punishments that are balanced, although propagating through the information network, always extinguish themselves in aggregate either by reaching a player with only one neighbor or by neutralizing themselves when reaching a player simultaneously from different directions. Second, such retaliatory behavior can be made consistent with sequential rationality because of the irrelevance of short-term incentives. If in each relationship a player will have ultimately defected for the same number of periods as his opponent, there does not exist a finite bound that applies to all histories on the number of the defections that a player expects from his opponent. Thus, there may not be a discount rate sufficiently large to neutralize short term incentives after any history. As we shall see in the next section, when the discount factor is less than unity, we induce short-term incentive compatibility by abandoning balanced retaliations and bounding punishments at the expense of Π^A-stability.

To formulate the equilibrium strategies, first define a pair of state variables $(d_{ij}, d_{ji}) \in \mathbb{N}_+^2$ for each relationship $ij \in G$. Both state variables depend only on the history of past play within the relationship and are therefore common knowledge for players i and j. The number d_{ij} represents the number of periods in which player i will have to play D as a consequence of the past play in relationship ij. The state variables' transitions are constructed so that (i) unilateral deviations to D are punished with an additional D by the opponent; (ii) unilateral deviations to C are punished with an additional D both by the player and by his opponent; (iii) joint deviations to the same action are not punished whereas joint deviations to different actions are punished as unilateral deviations. Thus, the transition rule for (d_{ij}, d_{ji}) is defined as follows. In the first period, $d_{ij} = 0$ for any $ij \in G$. Thereafter, for any history $h \in H$ leading to state (d_{ij}, d_{ji}) in the relationship ij, if actions (a_i, a_j) are chosen by players i and j, the states evolve according to the following table, where Δd_{ij} denotes the change in the variable d_{ij} and the $+$ sign a strictly positive value:

d_{ij}	0	0	0	0	0	0	0	0	$+$	$+$	$+$	$+$
d_{ji}	0	0	0	0	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
a_i	D	D	C	C	D	D	C	C	D	D	C	C
a_j	D	C	D	C	D	C	D	C	D	C	D	C
Δd_{ij}	0	0	1	0	0	1	0	1	-1	0	1	0
Δd_{ji}	0	1	0	0	0	2	-1	1	-1	1	0	0

(1)

Let $d_{ij}(h_i)$ denote the value of d_{ij} following a history $h_i \in H_{i,N_i}$. We will often abuse notation and define $d_{ij}(h)$ for a history $h \in H$, where the terms not in h_i enter vacuously. Define the interim strategy $\zeta_{i,N_i} : H_{i,N_i} \to \{C, D\}$ as

$$\zeta_{i,N_i}(h_i) = \begin{cases} C & \text{if} \quad \max_{j \in N_i} d_{ij}(h_i) = 0 \\ D & \text{if} \quad \max_{j \in N_i} d_{ij}(h_i) > 0 \end{cases}$$

This interim strategy instructs each player i to defect if and only if at least one of his "required" number of defections d_{ij} is positive. The strategy ζ_i of player i is the collection interim strategies $\{\zeta_{i,M}\}_{M \subset N \setminus \{i\}}$. A profile of such strategies will be denoted by ζ_N.

Note that, if $d_{ij} > d_{ji}$, the states return to $(0, 0)$ after d_{ji} periods of (D, D) and $d_{ij} - d_{ji}$ periods of (D, C). Hence, d_{ij} may be interpreted as the number of defections that players i and j require from player i in the future to return to the initial state. The next theorem shows that such a strategy profile satisfies the three properties of Section 2.3.

Theorem 1 *If $\delta = 1$, the strategy profile ζ_N satisfies C, Π^A-IE, and Π^A-S.*

The proof of Theorem 1 exploits two crucial attributes of the above strategies. First, the strategy profile ζ_N satisfies Π^A-stability. For a crude intuition, consider Figures 1 and 2. The number next to each vertex inside the graph denotes a player, the outside letter the actions, and the outside numbers on each edge the pair (d_{ij}, d_{ji}). Consider the pentagon in Figure 1. A deviation of player 1 spreads along the cycle and is stopped by the simultaneous play of D by players 3 and 4. Consider now the hexagon. Defections stop spreading because they reach player 4 simultaneously. Note how the play of D which originates from player 1, moves away from player 1 in both directions. That is, player 1 is a "source" of D's. In the pentagon, after players 2 and 5 play D, the play of D moves way from these players as well, that is, players 2 and 5 become sources. Our proof strategy generalizes this observation: there always exists a source player and the set of source players expands. Figure 2 provides additional intuition about the "annihilation" of D's that occurs when players conform to the profile ζ_N. Note that the graph has two cycles. Consider a history of length 10 in which player 1 deviates in the first period only, player 2 does not respond and does play C for the first 10 periods, and all other players always conform to the profile ζ_N. The first plot of Figure 2, depicts the state of play at the beginning of period 10 when player 2 plays his final deviation to C. By period 15, $d_{21} = d_{23}$ and no player except player 2 plays D. Thus, defections will die out in 5 periods. Notice one additional feature of ζ_N: when the play reverts to cooperation in all relationships, all connected players will have played the same number of D's.

Second, the retaliatory nature of the profile ζ_N is such that, in any relationship, a play of (D, C) is always matched by a later play of (C, D). Hence, a payoff of $1 + g$ is followed by a payoff of $-l$. As we shall see, this is the reason why A1 and Π^A-stability guarantee that, after any history, conforming to the profile ζ_N yields an *average* payoff at least as large as the average payoff from any deviation.

We first establish that the strategy profile ζ_N satisfies Π^A-stability. For any history $h \in H$, define the "excess defection" in a relationship to be $e_{ij}(h) = d_{ij}(h) - d_{ji}(h)$. Fix an information network G and, for any history $h \in H$ and any path $\pi = (j_1, .., , j_m)$, define

$$E_\pi(h) = \sum_{k=1}^{m-1} e_{j_k j_{k+1}}(h)$$

to be the sum of the excess defections along the path. Let P_{if} be the set of paths with initial vertex i and terminal vertex f and P_{ii} the set of cycles with initial vertex i. Finally, let $S(h)$ denote the set of players such that the aggregate excess defection on any path

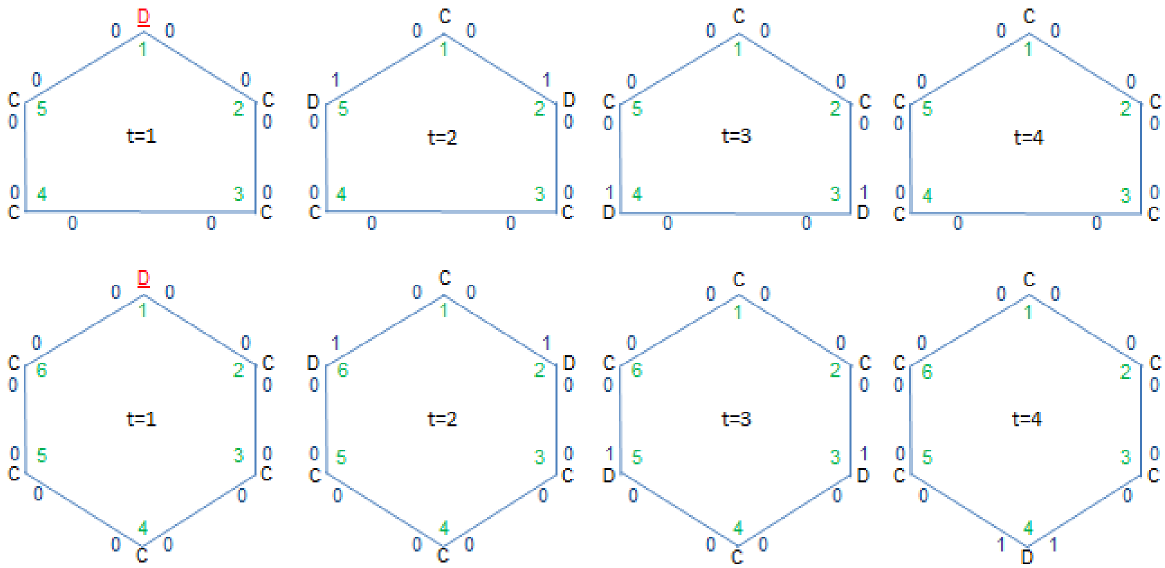

Figure 1: The time period is denoted by t. The number next to a vertex inside the graph denotes the player, the letter next to a vertex outside the graph denotes the action chosen in period t (the letter is underlined if the player is deviating), and the outside numbers on an edge denote the pair (d_{ij}, d_{ji}) at the beginning of the period.

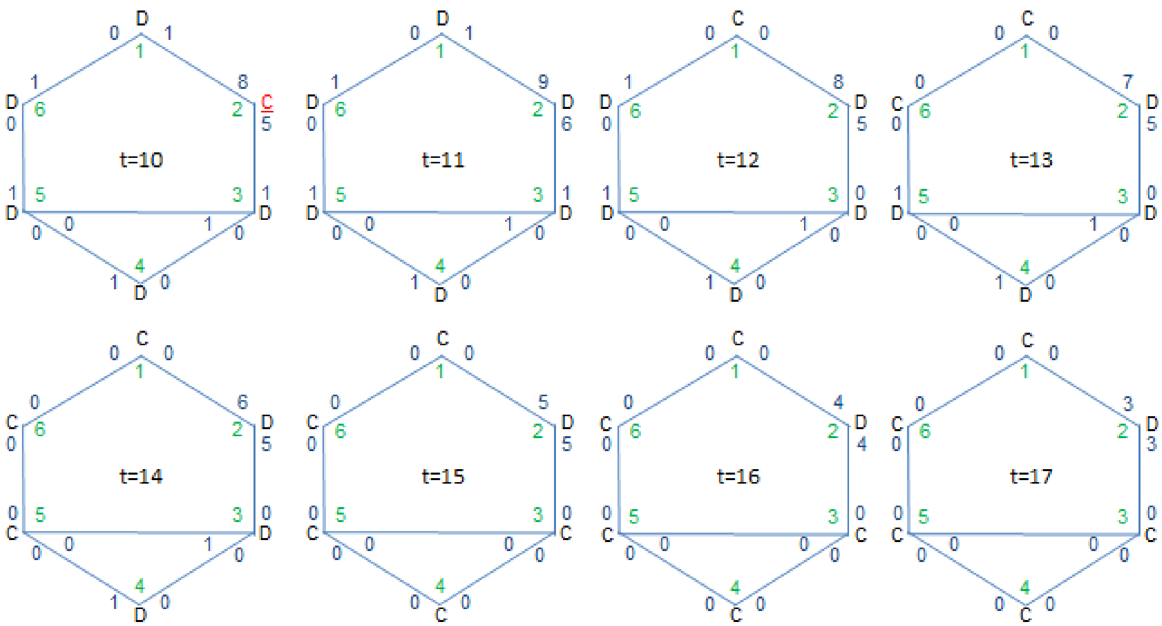

Figure 2: The time period is denoted by t. The number next to a vertex inside the graph denotes the player, the letter next to a vertex outside the graph denotes the action chosen in period t (the letter is underlined if the player is deviating), and the outside numbers on an edge denote the pair (d_{ij}, d_{ji}) at the beginning of the period.

departing from them is non-positive, that is,

$$S(h) = \{i \in N : E_\pi(h) \leq 0 \text{ for any } \pi \in P_{if}, \text{ for any } f \in N\}$$

Such players can be interpreted as the sources of D's in the network in that defections travel away from players in $S(h)$. The next lemma shows that aggregate excess defections along paths depend only on the initial and terminal vertices and that $S(h)$ is non-empty for any history h. Let the function $\mathbb{I}(\cdot)$ denote the indicator function.

Lemma 2 *Consider an information network G. For any history $(h, a) \in H$ in which a history $h \in H$ is followed by stage-game action profile $a \in A_N$:*
(1) If $\pi \in P_{if}$

$$E_\pi(h, a) = E_\pi(h) + \mathbb{I}(a_i \neq a_f)[\mathbb{I}(a_i = C) - \mathbb{I}(a_i = D)]$$

(2) If $\varkappa \in P_{ii}$

$$E_\varkappa(h) = 0$$

(3) If $\pi, \pi' \in P_{if}$

$$E_\pi(h) = E_{\pi'}(h)$$

(4) $S(h)$ is non-empty.

The next result uses Lemma 2 to establish that the strategy profile ζ_N satisfies Π^A-stability. The main idea of the proof is that the set $S(h)$ expands when players play according to the strategy profile ζ_N. The intuition follows by observing that first, when deviations "travel away" from a player $i \in S(h)$, (d_{ij}, d_{ji}), $j \in N_i$, declines, and second, if a player i is in $S(h)$ and has a neighbor j such that $(d_{ij}(h), d_{ji}(h)) = (0, 0)$, then player j is also in $S(h)$.

Lemma 3 *The strategy profile ζ_N satisfies Π^A-S.*

We will use Lemmas 2 and 3 to prove Theorem 1. The intuition for the final leg of this result follows from the profile ζ_N being such that, in any relationship, the outcome (D, C) is always matched by the outcome (C, D). The difficulty consists in evaluating the payoff of sequences for which no limit exists and in which deviations occur an infinite number of times, as the one shot deviation principle is inapplicable. Too see how these complications are resolved consider any history. The strategy ζ_N specifies a future play for the remainder of the game that leads to cooperation within finite time. Moreover, within any finite horizon, the number of periods in which a player can gain g in any relationship by deviating from ζ_N can be larger than the number of period in which he will incurs $-l$ by at most one. This follows as any deviation to defection is always met by an immediate defection and as cooperation is restored only after the deviating player has incurred $-l$. Then, as a direct consequence of A1, a player cannot strictly gain from deviating as the time horizon grows large. Indeed, an infinite number of deviations brings the payoff strictly below the cooperative payoff.

Proof of Theorem 1. The profile ζ_N trivially satisfies C. We will now show that, for any history $h \in H$,

$$\mathcal{V}_i(\zeta^h_{N,G}|G) \geq \mathcal{V}_i(\theta_i, \zeta^h_{-i,G}|G)$$

for any interim strategy $\theta_i \in \Sigma_{i,N_i}$, any $G \in \Gamma(N_i)$, and any $i \in N$. One can easily verify that Π-IE then follows.

Consider any history $h \in H$ of length $z-1$. Notice that by Π^A-S, (ii) in the definition of Banach-Mazur limits, and linearity

$$\mathcal{V}_i(\zeta^h_{N,G}|G) = \textstyle\sum_{j\in N_i} \eta_{ij}$$

Hence, ζ_N is Π^A-IE if and only if for any player $i \in N$ and for any interim strategy $\theta_i \in \Sigma_{i,N_i}$

$$\textstyle\sum_{j\in N_i} \eta_{ij} \geq \mathcal{V}_i(\theta_i, \zeta^h_{-i,G}|G) \text{ for any } G \in \Gamma(N_i)\,.$$

Let $\{\overline{a}^t_N\}^{\infty}_{t=z}$ be the sequence of stage-game actions generated by $(\theta_i, \zeta^h_{-i,G})$ after history h when the information network is G. Define $\overline{h}^t$, $t \geq z-1$, to be the history of length t generated by the strategy profile $(\theta_i, \zeta^h_{-i,G})$ after history h, that is, $\overline{h}^{z-1} = h$ and, for any $t \geq z$, $\overline{h}^{t+1} = (\overline{h}^t, \overline{a}^{t+1}_N)$. Consider any relationship $ij \in G$. Omitting some dependent variables for notational convenience, define a variable which counts how many times an action profile (a_i, a_j) has been played by the pair ij between periods s and $s+T$ in history $\overline{h}^{s+T}$, $s \geq z$,

$$n^s_{ij}(a_i, a_j|T) = \textstyle\sum^{s+T}_{t=s} \mathbb{I}\left(\overline{a}^t_i = a_i\right) \mathbb{I}\left(\overline{a}^t_j = a_j\right).$$

Then, from Table (1) and the definition of $e_{ij}(\cdot)$, for any $s \geq z$,

$$n^s_{ij}(D, C|0) - n^s_{ij}(C, D|0) = e_{ij}(\overline{h}^{s-1}) - e_{ij}\left(\overline{h}^s\right)$$

which trivially implies that

$$\begin{aligned} n^z_{ij}(D, C|T) - n^z_{ij}(C, D|T) &= \textstyle\sum^{T+z}_{t=z} \left(n^t_{ij}(D, C|0) - n^t_{ij}(C, D|0)\right) = \\ &= e_{ij}(\overline{h}^{z-1}) - e_{ij}\left(\overline{h}^{T+z}\right) \equiv \Delta^z(T) \end{aligned}$$

Notice that $e_{ij}\left(\overline{h}^{t}\right)<0$ implies that $d_{ji}\left(\overline{h}^{t}\right)>0$, which implies that $\overline{a}_j^{t+1}=D$, which finally implies that $e_{ij}\left(\overline{h}^{t+1}\right)\geq e_{ij}\left(\overline{h}^{t}\right)$. Thus, when player j plays according to ζ_j after history h, it must be the case that, for any T, $e_{ij}\left(\overline{h}^{T+z}\right)\geq -1$, if $e_{ij}(\overline{h}^{z-1})>0$; and $e_{ij}\left(\overline{h}^{T+z}\right)\geq e_{ij}(\overline{h}^{z-1})$, if $e_{ij}(\overline{h}^{z-1})<0$. Hence, for some $M^z>0$, $\Delta^z(T)\leq M^z$ for every T. It follows that the payoff of player i in relationship ij must satisfy

$$\begin{aligned}\textstyle\sum_{t=z}^{T+z} u_{ij}(\overline{a}_i^t,\overline{a}_j^t) &= n_{ij}^z(C,C|T)+(1+g)n_{ij}^z(D,C|T)-ln_{ij}^z(C,D|T)= \\ &= n_{ij}^z(C,C|T)+\frac{1+g-l}{2}2n_{ij}^z(C,D|T)+(1+g)\Delta^z(T)\end{aligned}$$

Note that

$$n_{ij}^z(C,C|T)+2n_{ij}^z(C,D|T)+n_{ij}^z(D,D|T)+\Delta^z(T)=T+1$$

and that, by A1, $1+g-l<2$. Then, since $\Delta^z(T)\leq M^Z$ for every T,

$$\limsup_{T\to\infty}\frac{\sum_{t=z}^{T+z} u_{ij}(\overline{a}_i^t,\overline{a}_j^t)}{T+1}\leq 1$$

Therefore, the Banach-Mazur limit satisfies

$$\Lambda\left(\left\{\frac{\sum_{t=z}^{T+z} u_{ij}(\overline{a}_i^t,\overline{a}_j^t)}{T+1}\right\}_{T=0}^{\infty}\right)\leq 1$$

The claim follows as Banach Mazur limits are linear. ■

Comments

Theorem 1 applies to several extensions of the baseline model. First, it is trivially robust to uncertainty on the number of players. Second, payoffs can be heterogeneous and allowed to depend on each relationship as long as A1 holds in all relationships. Indeed, Theorem 1 works even if payoffs are private information as long as they satisfy A1 in all possible realizations. Second, nowhere in the proof of Theorem 1 was it assumed that $\eta_{ij}>0$ for any $ij\in G$. Indeed, the arguments hold when $\eta_{ij}=0$ for some $ij\in G$. Thus, this result extend to the case in which the set of players observed by another player is larger than the set of players that affect this player's payoff.

We allow a pair (d_{ij},d_{ji}) to grow unbounded to prevent D's from cycling around the graph. Intuitively, suppose that ij is a relationship on a cycle. If player i fails to respond once to a play of (C,D) in relationship ij, D propagates only in one direction and enter a cycle. To "extinguish" this D, player i must play D so that D travels in the opposite direction as well. Although the network is finite, local information prevents the players from finding the smallest number of "counterbalancing" D's that prevent periodicity of punishments. As strategies only rely on local information, all D's propagating in one direction must be offset by the same number of D's in the opposite direction.

4 Impatient Players

This section studies games with players having discount factors below one. The first subsection introduces strategies and proves some preliminary results. The strategies constructed here are variants of the strategy discussed in Section 3. Punishments remain contagious and spread through the information network, but the maximal number of defections expected by any neighbor is bounded. Thus, retaliations are no longer balanced in the sense discussed in the previous section. To see why the profile ζ_N needs to be modified when the discount factors are below one, suppose that the information network is a large star network. Take a history of length T in which one peripheral player has always played D and the remaining players always C. It straightforward to check that, the longer T, the larger δ must be for the central player to comply with ζ_N and that no lower bound smaller than one exists for such δ.

Since retaliations are not balanced, inducing incentive compatibility runs into the problem that defections can cycle. In particular, players may expect defections to reach them in the future even when cooperation has resumed in each of their relationships. Checking sequential rationality in such cases is extremely demanding. It is possible to circumvent this difficulty with a rather direct approach that restricts the set of information networks. This section shows how to extend such an approach to our general framework. In appendix 5.1, we prove that, if priors assign positive probability only to acyclic information networks, a simple Π-invariant equilibrium exists that satisfies C and Π-stability. This result is a stepping stone for the main theorem presented here, which establishes that, if prior beliefs have full support, the very same strategy profile satisfies sequential rationality for an appropriate selection of a consistent system of beliefs. Numerous robustness properties of these bounded-punishment strategies are discussed after the main result.

4.1 Strategies and Preliminary Results

This subsection introduces the strategy profile ξ_N that differs from the one in Section 3 in that the maximal number of defections expected from any player is bounded by 2. As before, two state variables (d_{ij}, d_{ji}) characterize the state of each relationship $ij \in G$ and require each player i to defect if and only if at least one of his "required" number of defections d_{ij} is positive. Thus, for $h_i \in H_{i,N_i}$,

$$\xi_{i,N_i}(h_i) = \begin{cases} C & \text{if } \max_{j\in N_i} d_{ij}(h_i) = 0 \\ D & \text{if } \max_{j\in N_i} d_{ij}(h_i) > 0 \end{cases}$$

where $d_{ij}(h_i)$ is the value of d_{ij} after history h_i.

The transitions for the state variables (d_{ij}, d_{ji}) differ from Section 3 and depend on the sign of the payoff parameter l.

Case $l > 0:$ In the first period, $d_{ij} = 0$ for any $ij \in G$. Given a state (d_{ij}, d_{ji}) and actions (a_i, a_j) for the relationship ij, the state in the next period is determined by the following transition rule

d_{ij}	0	0	0	0	0	0	0	0	+	+	+	+
d_{ji}	0	0	0	0	+	+	+	+	+	+	+	+
a_i	D	D	C	C	D	D	C	C	D	D	C	C
a_j	D	C	D	C	D	C	D	C	D	C	D	C
Δd_{ij}	0	0	2	0	0	d_{ji}	0	d_{ji}	-1	0	0	0
Δd_{ji}	0	2	0	0	0	0	-1	0	-1	0	0	0

where Δd_{ij}, as before, denotes the change in variable d_{ij} and the + sign a strictly positive value.

Case $l < 0:$ In the first period, $d_{ij} = 0$ for any $ij \in G$. Given a state (d_{ij}, d_{ji}) and actions (a_i, a_j) for the relationship ij, the state in the next period is determined by the transition rule

d_{ij}	0	0	0	0	0	0	0	0	+	+	+	+
d_{ji}	0	0	0	0	+	+	+	+	+	+	+	+
a_i	D	D	C	C	D	D	C	C	D	D	C	C
a_j	D	C	D	C	D	C	D	C	D	C	D	C
Δd_{ij}	0	0	1	0	0	0	0	2	-1	$2-d_{ij}$	$2-d_{ij}$	$2-d_{ij}$
Δd_{ji}	0	1	0	0	-1	-1	-1	$2-d_{ji}$	-1	$2-d_{ji}$	$2-d_{ji}$	$2-d_{ji}$

where Δd_{ij}, again, denotes the change in variable d_{ij} and the + sign a strictly positive value.

Case $l = 0:$ Choose either transition rule.

We denote a profile of such strategies by ξ_N.[5] To achieve incentive compatibility

[5] We omit the dependence on parameter l for simplicity.

at every information set, (d_{ij}, d_{ji}) is bounded by $(2,2)$ in all cases. Note that, when the stage game is the prisoner's dilemma, equilibrium punishments following a deviation from the efficient play last for two periods. To see why, consider a player who needs to punish the opponent in one relationship but to cooperate in a second relationship in which his opponent's is expected to play D. If this player delays the punishment in the first relationship by one period, and thus temporarily restores cooperation in the second, he will have to defect in next period to restore cooperation in the first. Such action will then be a new deviation in the second relationship and thus trigger a two-period punishment. One can easily see that if a one-period punishment was instead triggered, delaying the punishment by one period in the first relationship can yield a higher payoff in the second when $1+g-l>0$.

The following result is instrumental to the proof of the main theorems of this section. It provides sufficient conditions for player i never to expect his neighbors to play D because of the past play in relationships to which player i does not belong. These conditions are: (i) all deviations have occurred in player i's neighborhood; (ii) no two neighbors of player i are connected by a path.

Given a history $h \in H$ of length T and a network G, let $\mathcal{D}(G,h,t)$ denote the set of players who deviate from the strategy profile ξ_N in period $t \leq T$. Further, define

$$\mathcal{D}(G,h) = \bigcup_{t=1}^{T} \mathcal{D}(G,h,t).$$

Again, let $d_{ij}(h)$ be the value of d_{ij} following history h. A *component* of an undirected graph is a maximal subgraph in which any two vertices are connected to each other by a path. A relationship $ij \in G$ is a *bridge* in G if its deletion from G increases the number of components.

Lemma 4 *Consider a network G, a player $i \in N$, and a history $h \in H$ such that:*

(i) $\mathcal{D}(G,h) \subseteq N_i \cup \{i\}$;

(ii) If $j \in \mathcal{D}(G,h) \backslash \{i\}$, the relationship ij is a bridge in G.

Then, $d_{jk}(h) = 0$ for any $j \in N_i$ and $k \in N_j \backslash \{i\}$.

The proof proceeds by induction. It shows that if all deviations have occurred in player i's neighborhood, and if there is no cycle that includes player i and his deviating neighbors, then player i never expects anyone of his neighbors to defect in response to behavior outside their relationship, regardless of his actions. Intuitively, since defections spread outwards in the information network, they can only return to player i if there is a cycle connecting i to a deviating player.

4.2 Full Support

This section establishes that the strategy profile ξ_N is a Π-invariant equilibrium satisfying C whenever prior beliefs have full support. Some of arguments developed here rely on the analysis of acyclic networks which appears in appendix 4.1. Let Π^{FS} be the set of prior beliefs having full support, that is, if $f \in \Pi^{FS}$ then $f(G) > 0$ for any G. The main idea of the proof consists in constructing a consistent system of beliefs such that all deviations are "local" and do not spread. That is, beliefs will be such that, following a deviation by a neighbor, a player believes that this neighbor is isolated. Naturally, the assumption of full support is crucial for this task. The perturbations of the equilibrium strategies needed in the construction of our consistent system of beliefs are chosen to converge pointwise to the equilibrium strategy.

Fix a player i with a neighborhood N_i. Let G_i^* denote the network in which $N_j = \{i\}$ for any player $j \in N_i$, and $N_j = N \backslash \{N_i \cup \{i, j\}\}$ for any $j \notin N_i \cup \{i\}$. That is, G_i^* consists of an incomplete star network, in which player i is the center and the players in N_i are the periphery, and a disjoint, totally connected component.[6] Consider the strategy ξ_N. Given a history h_i observed by player i when i's neighborhood is N_i, let $h^*(h_i)$ be the history such that $(G_i^*, h^*(h_i)) \in \mathcal{I}(h_i)$ and every player $j \notin N_i \cup \{i\}$ plays according to ξ_N (i.e. plays C) in every period. Hence, at the node $(G_i^*, h^*(h_i))$ all deviations are local in that they have occurred only in player i's relationships. We say that player $j \in N_i$ *i-deviates* from ξ_N at the observed history h_i if

$$j \in \mathcal{D}(G_i^*, h^*(h_i))$$

that is, if player j does not play according to ξ_N on the path to h_i when the network is G_i^*.

The next lemma shows that it is possible to construct a consistent belief system such that for any player i: (i) whenever a player j i-deviates, player i believes that player j's neighborhood contains only player i; (ii) player i believes that all deviations occur in his relationships. This is achieved by assuming that trembles are such that a deviation by a player with a singleton neighborhood is infinitely more likely than a deviation by a player with a larger neighborhood, and such that, as in the proof of Theorem 7, more recent deviations are infinitely more likely than less recent ones.

Lemma 5 *If priors beliefs are in Π^{FS}, there exists a system of beliefs β consistent with strategy profile ξ_N such that, for any player $i \in N$ and observed history h_i of length T,*

(a) *if player $j \in N_i$ i-deviates, then $\beta(G, h|h_i) = 0$ for any $(G, h) \in \mathcal{I}(h_i)$ for which G*

[6] The particular form of the latter component is inessential.

is such that $N_j \neq \{i\}$*;*

(b) *if* $(G,h) \in \mathcal{I}(h_i)$ *and for some* $t \leq T$*,*

$$\mathcal{D}(G,h,t) \neq \mathcal{D}(G_i^*, h^*(h_i), t),$$

then $\beta(G,h|h_i) = 0$*.*

The proof of the main result of this subsection follows from the preceding lemma and Lemma 4.

Theorem 6 *If* δ *is sufficiently close to one, the strategy profile* ξ_N *satisfies C and* Π^{FS}*-IE.*

Proof. The strategy profiles clearly satisfy C. We now establish Π^{FS}-IE. In particular it will be shown that given the system of beliefs β characterized in Lemma 5, it is sequentially rational to comply with the equilibrium strategy for any profile of prior beliefs satisfying A3. Fix: a player $i \in N$; a history h_i of length T observed by player i; and node (G,h) such that $\beta(G,h|h_i) > 0$. By Lemmas 4 and 5, for $j \in N_i$ and $k \in N_j \backslash \{i\}$, $d_{jk}(h') = 0$ for any history h' which has h as a subhistory and $\mathcal{D}(G,h') \backslash \mathcal{D}(G,h) \subseteq \{i\}$. Any player i believes that for any neighbor $j \in N_i$, $d_{jk}(h') = 0$ for any $k \in N_j \backslash \{i\}$. Consequently, player i believes that the action of a neighbor $j \in N_i$ at any history h' is solely determined $d_{ji}(h')$. Thus, the verification of sequential rationality is identical to the case in which networks are acyclic, and appears in Theorem 7 below. Property Π^{FS}-IE follows immediately as the strategies are independent of the prior beliefs. ■

Comments

The strategy profile of Theorem 6 is such that all players believe that defections spread away and never return, and that cooperation is restored permanently within two periods. This follows immediately from the above proof noting that no player expects defections to cycle and that the number of defections expected from a player in any of his relationships is bounded by two. Of course, such stability in "belief" may or may not be coexist with the actual systemic robustness of a permanent reversion to cooperation within finite time. Nevertheless, it does point out that it is possible to construct sequential equilibria in which incentives are always perceived as local. In such equilibria, defections are reactive and never anticipatory, that is, players do not defect in anticipation of forthcoming defections.

Several the robustness properties of the equilibrium strategy of Section 3 are satisfied by the equilibrium strategy of this section provided that the ordinal properties of the games are the same across all relationships. Uncertainty about the number of players, heterogeneity in payoffs, and uncertainty about payoffs consistent with A1 can be allowed for without compromising the results. The equilibrium in this section is also robust to heterogeneity in discount rates. The above theorem can also be extended to the case in which $\eta_{ij} = \eta_{ji} = 0$ for some $ij \in G$. This is again achieved by using the same system of beliefs as in Theorem 6 but modifying the strategies so that $d_{ij} = 0$ in any relationship ij for which $\eta_{ij} = 0$, that is, deviations in relationship ij are ignored. The intuition follows from such deviations being irrelevant for the immediate payoffs and not being expected to return via a different path.

The assumption of full support can be dispensed with when $l > 0$ by adapting an argument first used by Ellison (1994).[7] Note that a simple grim trigger strategy sustains cooperation for values of δ in some interval $\left(\underline{\delta}, \overline{\delta}\right)$. Then, cooperation can be extended to any $\delta \in \left(\underline{\delta}/\overline{\delta}, 1\right)$ by partitioning the game into $T-1$ independent games played every T periods and by playing according to grim trigger strategies in each of the independent games. The number T is chosen so that implied discount rate δ^T is in $\left(\underline{\delta}, \overline{\delta}\right)$. The equilibrium profile, however, is not robust to heterogeneous stage-game payoffs and, in particular, to heterogeneous discount rates since all players must partition the repeated game into independent games of identical length. Moreover, a player who defects in one of the $T-1$ games never returns to cooperation in that game. Play eventually settles on constant defection in the component in which this player resides. Thus, such equilibria never satisfy Π-stability.

The full support assumption is helpful in establishing theorem 6, as it allows sufficient flexibility in the determination of appropriate posterior beliefs. In particular, in the proof posterior beliefs are concentrated on networks that never lead to cycles of defections in histories in which deviations were observed. In a network environment, McBride (2006) exploits an analogous flexibility in posteriors by adopting the notion of conjectural equilibrium in Gilli (1992).

References

[1] Abdou J. and Mertens J., "Correlated Effectivity Functions", *Economic Letters* 30, 1989.

[2] Ahn I., "Three Essays on Repeated Games without Perfect Information", mimeo 1997.

[3] Ali N. and Miller D, "Enforcing Cooperation in Networked Societies", mimeo 2009.

[4] Aliprantis C. and Border K., "Infinite Dimensional Analysis", Springer, 2005.

[7] See Nava and Piccione (2011).

[5] Ben-Porath E. and Kahneman M., "Communication in Repeated Games with Private Monitoring", *Journal of Economic Theory* 70, 1996.

[6] Cho M., "Public Randomization in the Repeated Prisoner's Dilemma Game with Local Interaction" *Economic Letters*, forthcoming 2011.

[7] Cho M., "Cooperation in the Prisoner's Dilemma Game with Local Interaction and Local Communication, working paper, 2010.

[8] Deb J., "Cooperation and Community Responsibility: A Folk Theorem for Random Matching Games with Names", mimeo 2011.

[9] Ellison G., "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching", *Review of Economic Studies* 61, 1994.

[10] Fainmesser I., "Community Structure and Market Outcomes: A Repeated Games in Networks Approach", *American Economic Journal: Microeconomics*, 4, 2012.

[11] Fainmesser I. and Goldberg D., "Cooperation in Partly Observable Networked Markets", mimeo 2012.

[12] Gilli M., "On non-Nash Equilibria", *Games and Economic Behavior*, 27, 1999.

[13] Haag M. and Lagunoff R., "Social Norms Local Interaction, and Neighborhood Planning", *International Economic Review*, 47, 2006.

[14] Hopenhayn H. A. and Hauser C., "Trading Favors: Optimal Exchange and Forgiveness", Meeting Papers 125, Society for Economic Dynamics 2004.

[15] Jackson M., Rodriguez-Barraquer T. and Tan X., "Social Capital and Social Quilts: Network Patterns of Favor Exchange", mimeo, 2010.

[16] Kandori M., "Social Norms and Community Enforcement", *Review of Economic Studies*, 59, 1992.

[17] Kinateder M., "Repeated Games Played on a Network", mimeo, 2008.

[18] Lippert S. and Spagnolo G., "Networks of Relations and Social Capital", *Games and Economic Behavior* 72, 2011.

[19] McBride M., "Imperfect Monitoring in Communication Networks", *Journal of Economic Theory*, 126, 2006.

[20] Mihm M., Toht R. and Lang C., "What Goes Around Comes Around: a Theory of Indirect Reciprocity in Networks", mimeo, 2009.

[21] Mobius M., “Trading Favors”, mimeo, 2001.

[22] Nava F. and Piccione M., “Efficiency in Repeated Two-Action Games with Uncertain Local Monitoring”, Sticerd working paper, 2011.

[23] Renault J. and Tomala T., “Repeated Proximity Games”, *International Journal of Game Theory* 27, 1998.

[24] Takahashi S., “Community Enforcement when Players Observe Past Partners' Play”, *Journal of Economic Theory* 145, 2010.

[25] Vega-Redondo F., “Building Social Capital in a Changing World”, *Journal of Economic Dynamics and Control* 30, 2006.

[26] Wolitzky A., “Cooperation with Network Monitoring”, mimeo 2012.

[27] Xue J. *Essays on Cooperation, Coordination, and Conformity*, PhD Thesis, Pennsylvania State University, 2004.

5 Appendix

5.1 Acyclic Networks

In this subsection, we circumvent the problem of cycling defections by restricting the class of information networks. In particular, we prove that, if priors assign positive probability only to acyclic information networks, the profile of strategies introduced in section 4.1 is a Π-invariant equilibrium satisfying C and Π-stability. That efficiency can be easily obtained with relatively simple strategies in any acyclic network is of interest in cases in which a planner chooses the information network as in Haag and Lagunoff (2006). Moreover, this result is a stepping stone for theorem 6 which establishes that, if prior beliefs have full support, the very same strategy profile satisfies sequential rationality for an appropriate selection of a consistent system of beliefs. Let Π^{NC} be the set of admissible beliefs such that if $f \in \Pi^{NC}$ and $f(G) > 0$, then G is acyclic.

Theorem 7 *If δ is sufficiently close to one, the strategy profile ξ_N satisfies C, Π^{NC}-IE, and Π^{NC}-S.*

We first establish that the equilibrium strategy satisfies Π^{NC}-stability and then we prove the general theorem.

Lemma 8 *The strategy profile ξ_N satisfies Π^{NC}-S.*

Proof. Suppose that G is a tree and consider any history. For notational simplicity, assume that G is connected. If the players play according to the profile ξ_N, the possible transitions are given by

if $l \geq 0$

d_{ij}	0	0	0	0	0	0	+
d_{ji}	0	0	0	0	+	+	+
a_i	D	D	C	C	D	C	D
a_j	D	C	D	C	D	D	D
Δd_{ij}	0	0	2	0	0	0	-1
Δd_{ji}	0	2	0	0	0	-1	-1

if $l \leq 0$

d_{ij}	0	0	0	0	0	0	+
d_{ji}	0	0	0	0	+	+	+
a_i	D	D	C	C	D	C	D
a_j	D	C	D	C	D	D	D
Δd_{ij}	0	0	1	0	0	0	-1
Δd_{ji}	0	1	0	0	-1	-1	-1

We will prove the claim by induction on the number of players. It is easily verified that Π^{NC}-stability holds for $n = 2$. Suppose that $n > 2$. Consider a relationship ij such that player i is the unique neighbor of player j (player j is a terminal vertex). First note that, if $d_{ij} = 0$, it will remains so for the remainder of the game. Consequently, if $d_{ij} = 0$, the relationship ij is superfluous for the play of player i as player i plays D if and only if $d_{ik} > 0$ for some neighbor $k \neq j$. Hence, by induction, there exists a period t such that the play of all the players in the network in which the relationship ij is removed is C in all periods greater than t. Obviously, the same will hold for player j for some period $t' \geq t$. Conversely, if $d_{ij} > 0$, since player j's only neighbor is player i, d_{ij} will become zero after a finite number of periods and the above argument applies again. ■

The proof of Theorem 7 exploits Π^{NC}-stability to establish that the strategy profile ξ_N is a Π^{NC}-invariant equilibrium. In the first part of the argument, we construct consistent beliefs such that players believe that deviations occur only in their neighborhood. This is achieved by defining trembles for which more recent deviations to D are infinitely more likely than less recent deviations. Such beliefs imply that any player i believes that the action of a neighbor $j \in N_i$ at any history h is determined exclusively by $d_{ji}(h)$. For example, consider the prisoner's dilemma and a linear information network with three players in which player 1 is connected to player 2 who is connected to player 3. If player 1, upon observing a defection believes that it originated with player 3 two period earlier, he expects player 2 to defect twice. If instead he believes that the defection originated with player 2, he expect no further defections. In our construction, consistent beliefs correspond to the latter case. The second part of the argument is a tedious step-by-step verification that sequential rationality holds given such a system of beliefs.

Comments

Acyclic graph allow us to bound punishments since deviations do not cycle even if retaliations are not balanced. Thus, we are able to obtain Π^{NC}-stability. Furthermore, at any history cooperation is restored after no more than $3n$ periods. All the robustness properties of the equilibrium strategy of Section 3 are satisfied by the equilibrium strategy of this section provided that the ordinal properties of the games are the same across all relationships. Uncertainty about the number of players, heterogeneity in payoffs, and uncertainty about payoffs consistent with A1 can be allowed for without compromising the results. The equilibrium in this section is also robust to heterogeneity in discount rates. The above theorem can be easily extended to the case in which $\eta_{ij} = \eta_{ji} = 0$ for some $ij \in G$. This is achieved by using the same beliefs as in Theorem 7, but modifying the strategies so that deviations in a relationship ij for which $\eta_{ij} = 0$ are not punished, that is, $d_{ij} = 0$. Such deviations are inconsequential for players i and j as they do not affect current payoffs and never return.

Proof of Theorem 7

We begin with a preliminary lemma.

Lemma 9 *If the prior beliefs are in Π^{NC}, there exists a system of beliefs β consistent with strategy profile ξ_N such that, for any history $h_i \in H_{i,N_i}$ observed a player $i \in N$, if $\beta(G,h|h_i) > 0$ for some $(G,h) \in \mathcal{I}(h_i)$, then $\mathcal{D}(G,h) \subseteq N_i \cup \{i\}$.*

Proof. Consider trembles such that (i) a deviation to D by player i in period t when $\max_j d_{ij} = 0$ occurs with probability ε^{α^t}, where $1 > n\frac{\alpha}{1-\alpha}$; (ii) a deviation to C by player i in period t when $\max_j d_{ij} > 0$ occurs with probability ε^2. As $\varepsilon \to 0$, any finite number of deviations to D is infinitely more likely than a single deviation to C and any finite number of recent deviations to D is infinitely more likely than one earlier deviation to D. Given the sequence of completely mixed behavior strategy profiles ξ_N^ε obtained by adding these trembles to the profile ξ_N, let $\theta^\varepsilon(G,h)$ be the probability of node (G,h). The strategy ξ_N^ε is such that, for every information set $\mathcal{I}(h_i)$ of player i, the conditional belief of node $(G,h) \in \mathcal{I}(h_i)$

$$\beta^\varepsilon (G,h|h_i) = \frac{\theta^\varepsilon(G,h)}{\sum_{(G',h') \in \mathcal{I}(h_i)} \theta^\varepsilon(G',h')}$$

converges as $\varepsilon \to 0$, since each $\theta^\varepsilon(G,h)$ is a polynomial.
Consider an acyclic network G for which $f(G) > 0$ and a player i and a neighbor $j \in N_i$. Consider any history $h_i \in H_{i.N_i}$ and let $h^+(h_i) \in H$ denote the unique history of play $(G,h^+(h_i)) \in \mathcal{I}(h_i)$ in which all players, but for players in $N_i \cup \{i\}$ comply with the equilibrium strategy, that is, all the deviations observed by player i are attributed to j's

behavior. Let h_i^s denote the subhistory of h_i of length s, a_j^s the action of player in period s, and define

$$T_j = \left\{ s | d_{ji}(h_i^s) = 0 \text{ and } a_j^s = D \right\}$$

The probability of history $h^+(h_i)$ then satisfies

$$\begin{aligned} \theta^\varepsilon(G, h^+(h_i)) &= x(\varepsilon) y(\varepsilon) \prod_{j \in N_i} \prod_{s \in T_j} \varepsilon^{\alpha^s} \\ &= x(\varepsilon) y(\varepsilon) \, \varepsilon^{\sum_{j \in N_i} \sum_{s \in T_j} \alpha^s} \end{aligned}$$

since Lemma 4 applies, for $j \in N_i$, $d_{jk}(h^+(h_i)) = 0$ for any $k \in N_j \backslash \{i\}$. The term $x(\varepsilon)$ is a product that includes the prior and probabilities of "non-deviations", and $y(\varepsilon)$ a product of the probabilities of deviations to C by players in N_i directly observed by player i ($d_{ji}(h_i^s) > 0$ and $a_j^s = C$). Obviously,

$$\lim_{\varepsilon \to 0} x(\varepsilon) = f(G).$$

Now consider any other history such that $(G, h) \in \mathcal{I}(h_i)$. Suppose that such a history displays a deviation to C which is not directly observed by player i. Then, by construction

$$\theta^\varepsilon(G, h) \leq y(\varepsilon) \varepsilon^2.$$

Thus, $n \frac{\alpha}{1-\alpha} < 1$ implies that

$$\lim_{\varepsilon \to 0} \frac{\theta^\varepsilon(G, h)}{\theta^\varepsilon(G, h^+(h_i))} \leq \lim_{\varepsilon \to 0} \frac{1}{x(\varepsilon)} \varepsilon^{2 - \sum_{j \in N_i} \sum_{s \in T_j} \alpha^s} = 0,$$

since

$$\sum_{s \in T_j} \alpha^s < \sum_{s=0}^{\infty} \alpha^s < 2.$$

Consider now a history h' in which all deviations to C have been directly observed by player i. Let t denote the first period in which $d_{jk}(h'^t) > 0$ for some $k \in N_j \backslash i$. Then,

$$\theta^\varepsilon(G, h') \leq y(\varepsilon) \varepsilon^{\alpha^t} \prod_{j \in N_i} \prod_{s \in T_j | s \leq t} \varepsilon^{\alpha^s}$$

Now, $n \frac{\alpha}{1-\alpha} < 1$ implies that

$$\lim_{\varepsilon \to 0} \frac{\theta^\varepsilon(G, h')}{\theta^\varepsilon(G, h^+(h_i))} \leq \lim_{\varepsilon \to 0} \frac{1}{x(\varepsilon)} \varepsilon^{\alpha^t - \sum_{j \in N_i} \sum_{s \in T_j | s > t} \alpha^s} = 0$$

since

$$n \sum_{s \in T_j | s > t} \alpha^s < n \sum_{s=t+1}^{\infty} \alpha^s < \alpha^t.$$

Since there are only finitely many histories in $\mathcal{I}(h_i)$, it must be that $\lim_{\varepsilon\to 0}\beta^{\varepsilon}(G,h|h_i)>0$ only if $h=h^{+}(h_i)$. Therefore player i believes that $\mathcal{D}(G,h)\subseteq N_i\cup\{i\}$. ■

We now return to the proof of the Theorem.

Proof of Theorem 7. Property C is obvious. Tables are added as supplementary material to clarify the evolution of payoffs within a neighborhood after a defection. To prove Π^{NC}-IE, consider the system of beliefs β as in Lemma 9. Then, for any history $h_i\in H_{i,N_i}$ observed by player $i\in N$, if $\beta(G,h|h_i)>0$ for some $(G,h)\in\mathcal{I}(h_i)$, then $\mathcal{D}(G,h)\subseteq N_i\cup\{i\}$. Thus, since any relationship $ij\in G$ is a bridge, the conditions of Lemma 4 hold. Hence, for $j\in N_i$ and $k\in N_j\backslash\{i\}$, $d_{jk}(h')=0$ for any history h' which has h as a subhistory and $\mathcal{D}(G,h')\backslash\mathcal{D}(G,h)\subseteq\{i\}$. Thus, any player i believes that for any neighbor $j\in N_i$, $d_{jk}(h')=0$ for any $k\in N_j\backslash\{i\}$. Consequently, player i believes that the action of a neighbor $j\in N_i$ at any history h' is solely determined by $d_{ji}(h')$.

In order to check sequential rationality, we need to consider two separate cases. First assume that $l\geq 0$. Given any history, seven values of (d_{ij},d_{ji}) are possible, namely $(0,0)$, $(1,0)$, $(0,1)$, $(1,1)$, $(0,2)$, $(2,0)$, and $(2,2)$. First consider the case in which $\max_{j\in N_i}d_{ij}(h_i)=0$ and thus $\xi_i(h_i)=C$. If player i is sufficiently patient, he prefers to comply with the equilibrium strategy since the payoff differences between complying and a one shot deviation to D with any neighbor $j\in N_i$ are

$$\begin{array}{lll} (1+l)\left(\delta+\delta^2\right)-g & \text{if} & (d_{ij},d_{ji})=(0,0)\\ -l+\delta(1+l) & \text{if} & (d_{ij},d_{ji})=(0,1)\\ -l+\delta^2(1+l) & \text{if} & (d_{ij},d_{ji})=(0,2)\end{array}$$

which are positive by A1 and $l\geq 0$ when δ is sufficiently close to one.
If $\max_{j\in N_i}d_{ij}(h_i)=1$, then $\xi_i(h_i)=D$. A one shot deviation to C causes the maximum d_{ij} to remain equal to 1 in the next period for some $j\in N_i$. The payoff differences are

$$\begin{array}{lll} (1+g)(1-\delta)+\delta^3-1+l\left(\delta^3-\delta\right) & \text{if} & (d_{ij},d_{ji})=(0,0)\\ l+\left(\delta^2+\delta^3\right)(1+l)-\delta(1+g+l) & \text{if} & (d_{ij},d_{ji})=(0,1)\\ g+\delta & \text{if} & (d_{ij},d_{ji})=(1,0)\\ l+\delta & \text{if} & (d_{ij},d_{ji})=(1,1)\\ l(1-\delta) & \text{if} & (d_{ij},d_{ji})=(0,2)\end{array}$$

As $\delta\to 1$, the first and the last expression converge to zero, while the remaining three expressions become strictly positive. Since $\max_{j\in N_i}d_{ij}(h_i)=1$, a neighbor exists with whom player i strictly loses by deviating to C when δ is close to 1. Since $\eta_{ij}>0$ for any $j\in N_i$, a deviation to C strictly decreases payoffs for δ close to 1.
Finally, suppose that $\max d_{ij}(h_i)=2$. A one shot deviation to C causes the maximum

d_{ij} to remain equal to 2 in the next period for some $j \in N_i$. The payoff differences are

$$\begin{array}{ll} (1+g)(1-\delta) - \left(1-\delta^4\right) - l\left(\delta^2-\delta^4\right) & \text{if} \quad (d_{ij}, d_{ji}) = (0,0) \\ -\delta(1+g) + \delta^3 + \delta^4 + (1-\delta^2+\delta^3+\delta^4)l & \text{if} \quad (d_{ij}, d_{ji}) = (0,1) \\ (1+g)\left(1+\delta-\delta^2\right) - \left(1-\delta^4\right) - l\left(\delta^2-\delta^4\right) & \text{if} \quad (d_{ij}, d_{ji}) = (1,0) \\ (1+g)\left(\delta-\delta^2\right) + \delta^4 + (1-\delta^2+\delta^4)l & \text{if} \quad (d_{ij}, d_{ji}) = (1,1) \\ l(1-\delta^2) & \text{if} \quad (d_{ij}, d_{ji}) = (0,2) \\ (1+\delta)(1+g) + \delta^2 - 1 & \text{if} \quad (d_{ij}, d_{ji}) = (2,0) \\ l + \delta^2 & \text{if} \quad (d_{ij}, d_{ji}) = (2,2) \end{array}$$

As $\delta \to 1$ the first and the fifth expression converge to zero, while the remaining expressions become strictly positive. Since $\max_{j \in N_i} d_{ij}(h_i) = 2$, a neighbor exists with whom player i strictly loses by deviating to C when δ is close to 1. Since $\eta_{ij} > 0$ for any $j \in N_i$, a deviation to C strictly decreases payoffs for δ close to 1.

Next assume that $l \leq 0$. Given any history, five values of (d_{ij}, d_{ji}) are possible, namely $(0,0)$, $(1,0)$, $(0,1)$, $(1,1)$, and $(2,2)$. First consider the case in which $\max_{j \in N_i} d_{ij}(h_i) = 0$ and thus $\xi_i(h_i) = C$. If player i is sufficiently patient, he prefers to comply with the equilibrium strategy since the payoff differences between complying and a one shot deviation to D with any neighbor $j \in N_i$ are

$$\begin{array}{ll} -g + (1+l)\delta & \text{if} \quad (d_{ij}, d_{ji}) = (0,0) \\ -l & \text{if} \quad (d_{ij}, d_{ji}) = (0,1) \end{array}$$

As $\delta \to 1$, the first expression is strictly positive and the second weakly positive by A1 and $l \leq 0$.

If $\max_{j \in N_i} d_{ij}(h_i) = 1$, then $\xi_i(h_i) = D$. A one shot deviation to C causes the maximum d_{ij} to increase to 2 in the next period for some $j \in N_i$. The payoff differences are

$$\begin{array}{ll} g - (1+g+l)\delta + \delta^2 & \text{if} \quad (d_{ij}, d_{ji}) = (0,0) \\ l - \delta g + \delta^2 & \text{if} \quad (d_{ij}, d_{ji}) = (0,1) \\ g + \delta + \delta^2 & \text{if} \quad (d_{ij}, d_{ji}) = (1,0) \\ l + \delta + \delta^2 & \text{if} \quad (d_{ij}, d_{ji}) = (1,1) \end{array}$$

As $\delta \to 1$, the first expression is weakly positive and the remaining expressions become strictly positive, since $1 > g - l$ by A1. Since $\max_{j \in N_i} d_{ij}(h_i) = 1$, a neighbor exists with whom player i strictly loses by deviating to C when δ is close to 1. Since $\eta_{ij} > 0$ for any $j \in N_i$, a deviation to C strictly decreases payoffs for δ close to 1.

Finally, suppose that $\max d_{ij}(h_i) = 2$. A one shot deviation to C causes the maximum

d_{ij} to remain equal to 2 in the next period for some $j \in N_i$. The payoff differences are

$$\begin{array}{cl} g-(1+g)\,\delta+\delta^2 & \text{if} \;\; (d_{ij},d_{ji})=(0,0) \\ l(1-\delta^2) & \text{if} \;\; (d_{ij},d_{ji})=(0,1) \\ g+(1+g)\,\delta-l\delta^2 & \text{if} \;\; (d_{ij},d_{ji})=(1,0) \\ l(1-\delta^2)+(1+g)\,\delta & \text{if} \;\; (d_{ij},d_{ji})=(1,1) \\ l+\delta^2 & \text{if} \;\; (d_{ij},d_{ji})=(2,2) \end{array}$$

As $\delta \to 1$, the first and the second expression converge to zero, while the remaining expressions become strictly positive. Since $\max_{j\in N_i} d_{ij}(h_i) = 2$, a neighbor exists with whom player i strictly loses by deviating to C when δ is close to 1. Since $\eta_{ij} > 0$ for any $j \in N_i$, a deviation to C strictly decreases payoffs for δ close to 1.

Since the incentives to conform to ξ_N are not affected by the beliefs about the graph, the proof is complete. ■

Supplementary Notes

The following tables clarify the incentive constraints in the proof of theorem 7. Each entry shows the payoff in periods following either no deviation or a one shot deviation by player i from the strategy ξ_i when the relationship with player j was in state (d_{ij}, d_{ji}). Payoffs are omitted after a relationship returns to the state $(0,0)$. If $l \geq 0$ and $\max_{j\in N_i} d_{ij}(h_i) = 0$:

	Equilibrium: C			Deviation: D		
(d_{ij}, d_{ji})	t	$t+1$	$t+2$	t	$t+1$	$t+2$
$(0,0)$	1	1	1	$1+g$	$-l$	$-l$
$(0,1)$	$-l$	1	1	0	$-l$	1
$(0,2)$	$-l$	$-l$	1	0	$-l$	$-l$

If $l \geq 0$ and $\max_{j\in N_i} d_{ij}(h_i) = 1$:

	Equilibrium: D				Deviation: C			
(d_{ij}, d_{ji})	t	$t+1$	$t+2$	$t+3$	t	$t+1$	$t+2$	$t+3$
$(0,0)$	$1+g$	$-l$	$-l$	1	1	$1+g$	$-l$	$-l$
$(0,1)$	0	$-l$	1	1	$-l$	$1+g$	$-l$	$-l$
$(1,0)$	$1+g$	1	1	1	1	0	1	1
$(1,1)$	0	1	1	1	$-l$	0	1	1
$(0,2)$	0	$-l$	$-l$	1	$-l$	0	$-l$	1

If $l \geq 0$ and $\max_{j \in N_i} d_{ij}(h_i) = 2$:

	Equilibrium: D					Deviation: C				
(d_{ij}, d_{ji})	t	$t+1$	$t+2$	$t+3$	$t+4$	t	$t+1$	$t+2$	$t+3$	$t+4$
$(0,0)$	$1+g$	0	$-l$	$-l$	1	1	$1+g$	0	$-l$	$-l$
$(0,1)$	0	0	$-l$	1	1	$-l$	$1+g$	0	$-l$	$-l$
$(1,0)$	$1+g$	$1+g$	$-l$	$-l$	1	1	0	$1+g$	$-l$	$-l$
$(1,1)$	0	$1+g$	$-l$	$-l$	1	$-l$	0	$1+g$	$-l$	$-l$
$(0,2)$	0	0	$-l$	$-l$	1	$-l$	0	0	$-l$	1
$(2,0)$	$1+g$	$1+g$	1	1	1	1	0	0	1	1
$(2,2)$	0	0	1	1	1	$-l$	0	0	1	1

If $l \leq 0$ and $\max_{j \in N_i} d_{ij}(h_i) = 0$:

	Equilibrium: C			Deviation: D		
(d_{ij}, d_{ji})	t	$t+1$	$t+2$	t	$t+1$	$t+2$
$(0,0)$	1	1	1	$1+g$	$-l$	1
$(0,1)$	$-l$	1	1	0	1	1

If $l \leq 0$ and $\max_{j \in N_i} d_{ij}(h_i) = 1$:

	Equilibrium: D				Deviation: C			
(d_{ij}, d_{ji})	t	$t+1$	$t+2$	$t+3$	t	$t+1$	$t+2$	$t+3$
$(0,0)$	$1+g$	$-l$	1	1	1	$1+g$	0	1
$(0,1)$	0	1	1	1	$-l$	$1+g$	0	1
$(1,0)$	$1+g$	1	1	1	1	0	0	1
$(1,1)$	0	1	1	1	$-l$	0	0	1

If $l \leq 0$ and $\max_{j \in N_i} d_{ij}(h_i) = 2$:

	Equilibrium: D				Deviation: C			
(d_{ij}, d_{ji})	t	$t+1$	$t+2$	$t+3$	t	$t+1$	$t+2$	$t+3$
$(0,0)$	$1+g$	0	1	1	1	$1+g$	0	1
$(0,1)$	0	$1+g$	$-l$	1	$-l$	$1+g$	0	1
$(1,0)$	$1+g$	$1+g$	$-l$	1	1	0	0	1
$(1,1)$	0	$1+g$	$-l$	1	$-l$	0	0	1
$(2,2)$	0	0	1	1	$-l$	0	0	1

5.2 Omitted Proofs

Proof of Lemma 2. The proof first establishes (1) and then proceeds by induction to prove (2) and (3). Consider a history (h, a). Notice that, by definition,

$$e_{ij}(h,a) = e_{ij}(h) + \mathbb{I}(a_i \neq a_j)\left[\mathbb{I}(a_i = C) - \mathbb{I}(a_i = D)\right]$$

Hence, for any path $\pi = (j_1, .., , j_m) \in P_{if}$:

$$\begin{aligned} E_\pi(h,a) &= E_\pi(h) + \sum_{k=1}^{m-1} \mathbb{I}\left(a_{j_k} \neq a_{j_{k+1}}\right)\left[\mathbb{I}(a_{j_k} = C) - \mathbb{I}(a_{j_k} = D)\right] = \\ &= E_\pi(h) + \mathbb{I}(a_i \neq a_f)\left[\mathbb{I}(a_i = C) - \mathbb{I}(a_i = D)\right] \end{aligned}$$

The last equality holds by a simple counting argument. Consider the sequence of action pairs $\{(a_{j_k}, a_{j_{k+1}})\}_{k=1}^{m-1}$. First remove all the pairs of actions $(a_{j_k}, a_{j_{k+1}})$ for which $a_{j_k} = a_{j_{k+1}}$ since $\mathbb{I}(a_{j_k} \neq a_{j_{k+1}}) = 0$. Since the stage game has only two actions, if the actions played at the beginning and at the end of the path coincide ($a_i = a_f$), we are left an even number of alternating pairs. If actions played at the beginning and at the end do not coincide ($a_i \neq a_f$), we are left an odd number of alternating pairs. The desired equality then follows. Figure 3 below presents a visual intuition for the claim.

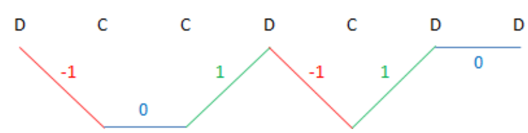

Figure 3: Changes in excess defections are reported on any given link for a particular action profile chosen by the players on a path.

Notice that (1) and a simple induction argument imply (2). When h is empty, (2) holds trivially. If (2) holds for any history h, it will also hold for a history (h, a) since $a_i = a_f$ in a cycle. A similar induction argument also establishes (3).

Claim (4) is also proved by induction. When h is the empty history, $d_{ij}(h) = 0$ for any $ij \in G$, and (4) holds trivially since $S(h) = N$. Suppose that (4) holds for a history h. Consider the history $h' = (h, a)$ and a player $i \in S(h)$. If $i \in S(h')$, the claim holds. Suppose then that $i \notin S(h')$. Since $i \in S(h)$, by (1) there exists at least one path $\pi \in P_{ij}$ such that $E_\pi(h') = 1$. We will show that this implies that $j \in S(h')$. Consider any path $\pi' \in P_{jf}$ and any path $\pi'' \in P_{if}$ for any $f \in N$. Note that, by (1), $E_{\pi''}(h') \leq 1$ and, by (3):

$$\begin{aligned} E_{\pi'}(h') &= E_{\pi''}(h') - E_\pi(h') = \\ &= E_{\pi''}(h') - 1 \leq E_{\pi''}(h) \leq 0 \end{aligned}$$

which establishes (4). ■

Proof of Lemma 3. Fix an information network G. Consider any history $h \in H$ of length t. Following any history, the players' actions for the remainder of the game are determined by ζ_N. Thus, in any relationship $ij \in G$, the state transitions take place according to the following table:

d_{ij}	0	0	0	0	0	0	+
d_{ji}	0	0	0	0	+	+	+
a_i	D	D	C	C	D	C	D
a_j	D	C	D	C	D	D	D
Δd_{ij}	0	0	1	0	0	0	-1
Δd_{ji}	0	1	0	0	0	-1	-1

(2)

Let

$$T(h) = \max_{ij\in G} \{\min \{d_{ij}(h), d_{ji}(h)\}\}.$$

and h_+^s denote the history s periods longer than h that is generated by ζ_N after history h. If all players play according to ζ_N after history h, for any $z > T(h)$ all the relationships ij will satisfy $\min\{d_{ij}(h_+^z), d_{ji}(h_+^z)\} = 0$, that is, either $d_{ij}(h_+^z)$ or $d_{ji}(h_+^z)$ is equal to zero. To show that the strategy satisfies Π^A-stability, it will be sufficient to prove that, for any history $h \in H$ and for any $z > T(h)$,

(A) $S(h_+^z) \subseteq S(h_+^{z+1})$
(B) If $S(h_+^z) \neq N$, $S(h_+^z) \neq S(h_+^{z+k})$ for some $k > 0$

Indeed, if both statements were to hold, Π^A-S would follow trivially as $S(h_+^z) = N$ for z sufficiently large, and $S(h_+^z) = N$ if and only if $\max_{ij\in G}\{d_{ij}(h_+^z)\} = 0$. We establish (A) by contradiction. Consider a player i such that $i \in S(h_+^z)$ for $z > T(h)$ and $i \notin S(h_+^{z+1})$. Then, there exists a path $\pi \in P_{if}$ such that

$$E_\pi(h_+^z) = 0 \text{ and } E_\pi(h_+^{z+1}) = 1$$

Since $i \in S(h_+^z)$, by (1) of Lemma 2, $\zeta_f\left(h_+^z\right) = D$. For player f to choose D along the equilibrium path it must be that $d_{fk}(h_+^z) > 0$ for some $k \in N_f$. Since $z > T(h)$, by definition it must be that $d_{kf}(h_+^z) = 0$ and thus, for $\pi' \in P_{ik}$,

$$E_{\pi'}(h_+^z) = E_\pi(h_+^z) + e_{fk}(h_+^z) = e_{fk}(h_+^z) > 0$$

which contradicts that $i \in S(h_+^z)$. Hence, (A) must hold.

For the proof of (B), take $j \in N_i$ such that $i \in S(h_+^z)$ and $j \notin S(h_+^z)$ for $z > T(h)$. Notice that such player i must exist by (4) of Lemma 2. By (A), $d_{ij}(h_+^{z+z'}) = 0$ for any $z' \geq 0$. Since

$$d_{ji}(h_+^{z+z'+1}) = \max\left\{d_{ji}(h_+^{z+z'}) - 1, 0\right\}$$

for any $z' \geq 0$, it follows that $d_{ji}(h_+^{z+z'}) = 0$ for any $z' > d_{ji}(h_+^z)$. The claim follows noting that, for any history h, if $e_{ij}(h) = 0$ and $i \in S(h)$, then $j \in S(h)$. ■

Proof of Lemma 4. First consider any player $j \in \mathcal{D}(G, h)$ such that $j \neq i$. Let $(N(G_j), G_j)$ denote the component of the graph $G \backslash \{ij\}$ to which player j belongs. By condition (ii), such component cannot include player i and players in $N_i \backslash \{j\}$, or else relationship ij would not be a bridge. We want to establish that $d_{jk}(h) = 0$ for $k \in N_j$, where $k \neq i$. Partition players in the $N(G_j)$ based on their distance from j. In particular, let N_j^z denote the set of players in $N(G_j)$ whose shortest path to player j contains z relationships and let $N_j^0 = \{j\}$. Clearly, $N_j^1 = N_j \backslash \{i\}$.

By induction on the history length, we will first prove that, if $\mathcal{D}(G, h) \cap N(G_j) = \{j\}$, then for any distance $z \geq 0$, any player $r \in N_j^z$, and any relationship $rk \in G_j$:

$$d_{rk}(h) = \begin{cases} 0 & \text{if} \quad k \in N_r \backslash N_j^{z-1} \\ b_z(h) & \text{if} \quad k \in N_j^{z-1} \end{cases} \tag{3}$$

where the second condition holds only for $z > 0$ and $b_z(h)$ depends only on z and h, and is independent of the identity of the two players. Observe that the claim holds the empty history, as $d_{rk}(\emptyset) = 0$ for any $rk \in G_j$. Further observe that for $m \in N_j^z$ and $z > 0$, $N_m \subset N_j^{z-1} \cup N_j^z \cup N_j^{z+1}$ and $N_m \cap N_j^{z-1} \neq \emptyset$. Now assume that the claim holds for any history of length up to T. We will show that it holds for length $T+1$. Let (h^T, a) denote a history of length $T+1$, where a denotes the profile of actions chosen in period $T+1$. Observe that, for any distance $z > 0$ and any player $r \in N_j^z$,

$$a_r = D \Leftrightarrow d_{rk}(h^T) > 0 \text{ for } k \in N_j^{z-1} \tag{4}$$

since $r \notin \mathcal{D}\left(G, h^T\right)$ and since, by the induction hypothesis, $d_{rk}(h^T) = 0$ for any $k \in N_r \backslash N_j^{z-1}$. Thus, for any $z > 0$, all players in N_j^z must choose the same action since $d_{rk}(h^T) = b_z(h^T)$ for any $r \in N_i^z$ and $k \in N_j^{z-1} \cap N_r$, and since $N_j^{z-1} \cap N_r \neq \emptyset$ given that a path exists connecting player r to player j (r belongs to component G_j). Thus, for any distance $z > 0$, any player $r \in N_j^z$, and any relationship $rk \in G_j$,

$$d_{rk}(h^T, a) = 0 \ \text{ if } k \in N_i^z$$

since $d_{rk}(h^T) = d_{kr}(h^T) = 0$, and since $a_r = a_k$. Similarly, observe that for any distance

$z \geq 0$, any player $r \in N_j^z$, and any relationship $rk \in G$,

$$d_{rk}(h^T, a) = 0 \text{ if } k \in N_j^{z+1}$$

since $d_{rk}(h^T) = 0$ if $k \in N_j^{z+1}$, and because (4) immediately implies that $d_{rk}(h^T, a) = 0$, by the transition rules. Finally note that for any distance $z > 0$, any player $r \in N_j^z$, and any relationship $rk \in G$,

$$d_{rk}(h^T, a) = b_z(h^T, a) \text{ if } k \in N_j^{z-1}$$

since $d_{rk}(h^T) = b_z(h^T)$ if $k \in N_j^{z-1}$, and because $a_l = a_m$ for any two players $l, m \in N_j^s$ for any $s \geq 0$. Thus, condition (3), must hold for a history of arbitrary length in which only player j has deviated in component G_j. This establishes that for any history $h \in H$, if conditions (i) and (ii) in the lemma hold, $d_{jk}(h) = 0$, for any $j \in \mathcal{D}(G, h) \backslash \{i\}$ and any one of his neighbors $k \in N_j \backslash \{i\}$.

To conclude the proof consider the neighbors of player i in $N_i \backslash \mathcal{D}(G, h)$. In particular, consider the component of the network G to which player i belongs when all the relationships between player i and players in $\mathcal{D}(G, h)$ have been removed from the network G. Label such network $(N(G_i), G_i)$. Clearly, $N_i \backslash \mathcal{D}(G, h) \subset N(G_i)$. Furthermore, $N(G_i) \cap \mathcal{D}(G, h) = \{i\}$ by construction. Hence, since by condition (ii) in the lemma $N(G_i) \cap G_j = \emptyset$ for any $j \in \mathcal{D}(G, h) \backslash \{i\}$, the previous induction argument can still be used to establish that for any distance $z \geq 0$, any player $r \in N_i^z$, and any relationship $rk \in G_i$,

$$d_{rk}(h) = \begin{cases} 0 & \text{if } \; k \in N_r \backslash N_i^{z-1} \\ b_z(h) & \text{if } \; k \in N_i^{z-1} \end{cases}$$

where N_i^z denotes the set of player at distance $z \geq 0$ from i in G_i, as in the previous part of the proof. Therefore, $d_{jk}(h) = 0$, for any $j \in N_i \backslash \mathcal{D}(G, h)$ and any one of his neighbors $k \in N_j \backslash \{i\}$, which with the previous part of the argument establishes the result. ■

Proof of Lemma 5

We begin with a preliminary result. For any history $h \in H$, let h^t denote the sub-history of length $t < T$. The next lemma relates the sets of defecting players $\mathcal{D}(G_i^*, h^*(h_i), t)$ and $\mathcal{D}(G, h, t)$ for two nodes $(G_i^*, h^*(h_i)), (G, h) \in \mathcal{I}(h_i)$.

Lemma 10 *Consider a node $(G, h) \in \mathcal{I}(h_i)$ where history h is of length T. If*

(i) $\mathcal{D}(G_i^*, h^*(h_i), t) = \mathcal{D}(G, h, t)$ for any $t < T$, and

(ii) $N_j = \{i\}$ for any $j \in \mathcal{D}\left(G, h^{T-1}\right) \backslash \{i\}$,

then $\mathcal{D}(G_i^*, h^*(h_i), T) \subseteq \mathcal{D}(G, h, T)$.

Proof. Suppose that the (i) and (ii) hold. Observe that by definition of $h^*(h_i)$,

$$\mathcal{D}\left(G_i^*, h^*(h_i), t\right) \subseteq N_i \cup \{i\}.$$

Moreover, note that Lemma 4 can be applied to establish that for any sub-history h^t of length $t < T$ and for any player $j \in N_i$,

$$d_{jk}\left(h^t\right) = 0 \text{ for } k \in N_j \backslash \{i\}.$$

Now observe that, since $\left(G_i^*, h^*(h_i)\right), (G, h) \in \mathcal{I}(h_i)$, we must have that for any sub-history h^t of length $t < T$ and for any player $j \in N_i$,

$$d_{ji}\left(h^t\right) = d_{ji}\left(h^*(h_i)^t\right) \text{ and } d_{ij}\left(h^t\right) = d_{ij}\left(h^*(h_i)^t\right).$$

The latter observation immediately implies that if $i \in \mathcal{D}\left(G_i^*, h^*(h_i), T\right)$, then $i \in \mathcal{D}(G, h, T)$. Now consider a player $j \in \mathcal{D}\left(G_i^*, h^*(h_i), T\right) \backslash \{i\}$. If player j plays C at T, then $d_{ji}\left(h^*(h_i)^{T-1}\right) > 0$, and thus $j \in \mathcal{D}(G, h, T)$ since $d_{ji}\left(h^{T-1}\right) > 0$ as well. If player j plays D at T, then $d_{ji}\left(h^*(h_i)^{T-1}\right) = 0$, and thus $j \in \mathcal{D}(G, h, T)$ since $d_{jk}\left(h^{T-1}\right) = 0$ for $k \in N_j$. ■

We now return to the proof of Lemma 5.

Proof of Lemma 5. For any player i, consider trembles such that:
(i) If $n_i = 1$, a deviation in period t from ξ_N occurs with probability ε^{α^t}, where $\frac{\alpha}{1-\alpha} n < 1$
(ii) If $n_i > 1$, a deviation in period t from ξ_N occurs with probability ε^2.

Note that, for any $t > 1$, such trembles imply that, as ε vanishes, a single deviation of type (i) at time $t < T$ is infinitely less likely than deviations of type (i) by all the players in periods $t+1, t+2, ..., T$ since $\alpha^t > n \sum_{s=t+1}^{\infty} \alpha^s$. Given the sequence of completely mixed behavior strategy profiles ξ_N^ε obtained by adding the above trembles to the profile ξ_N, let $\theta^\varepsilon(G, h)$ be the probability of node (G, h). The strategy ξ_N^ε is such that, for every information set $\mathcal{I}(h_i)$ of player i, the conditional belief of node $(G, h) \in \mathcal{I}(h_i)$

$$\beta^\varepsilon(G, h | h_i) = \frac{\theta^\varepsilon(G, h)}{\sum_{(G', h') \in \mathcal{I}(h_i)} \theta^\varepsilon(G', h')}$$

converges as $\varepsilon \to 0$, since each $\theta^\varepsilon(G, h)$ is a polynomial of the form

$$x \prod_{k=1}^{W} \left(1 - \varepsilon^{y_k}\right) \prod_{k=1}^{V} \varepsilon^{z_k}, \tag{5}$$

for some parameters $W, V \leq nT$, $x \in (0, 1)$, and $y_k, z_k \in \mathbb{R}_+$ for k in the appropriate range. For any node $(G,h) \in \mathcal{I}(h_i)$ define

$$\beta(G,h|h_i) = \lim_{\varepsilon\to 0} \beta^{\varepsilon}(G,h|h_i).$$

We first establish (a). Consider $(G,h) \in \mathcal{I}(h_i)$. Recall that the history $h^*(h_i)$ is such that $(G_i^*, h^*(h_i)) \in \mathcal{I}(h_i)$ and every player $j \notin N_i \cup \{i\}$, plays C in every period. Obviously, for any $j \in N_i$,

$$h_i(j) = h^*(h_i, j) = h(j)$$

where $h_i(j)$, $h^*(h_i, j)$, and $h(j)$ denote player j's play in histories h_i, $h^*(h_i)$, and h.

Now consider a player $j \in N_i$ that i-deviates from ξ_N at the observed history h_i. That is, $j \in \mathcal{D}(G_i^*, h^*(h_i))$. Since at node $(G_i^*, h^*(h_i))$ all deviations are of type (i),

$$\theta^{\varepsilon}(G_i^*, h^*(h_i)) \geq f(G_i^*)(1-\varepsilon)^{nT}\varepsilon,$$

where the lower bound is obtained by setting W to be equal to nT, $y_k = 1$ in (5) and noting that

$$\sum_{k=1}^{V} z_k \leq \sum_{t=1}^{T} n\alpha^t < 1$$

since $\frac{\alpha}{1-\alpha}n < 1$. Thus, for ε sufficiently close to zero, there exists a constant $q > 0$ such that

$$\theta^{\varepsilon}(G_i^*, h^*(h_i)) \geq q\varepsilon.$$

The constant q is positive since, by hypothesis, $f(G_i^*) > 0$.

Now consider a node $(G', h') \in \mathcal{I}(h_i)$ such that $N_j' \neq \{i\}$, where N_j' is neighborhood of player j in G'. Consider two separate cases:

1. First suppose that $j \in \mathcal{D}(G', h')$. As the deviation of player j at period t is of type (ii), $\theta^{\varepsilon}(G', h') \leq \varepsilon^2$. Thus,

$$\beta^{\varepsilon}(G', h'|h_i) \leq \frac{\theta^{\varepsilon}(G', h')}{\theta^{\varepsilon}(G_i^*, h^*(h_i))} \leq \frac{\varepsilon}{q}$$

   which implies that $\beta(G', h'|h_i) = 0$. Thus, the claim holds.

2. Then suppose that $j \notin \mathcal{D}(G', h')$. Let t^* denote the earliest period t in which

$$\mathcal{D}(G_i^*, h^*(h_i), t) \neq \mathcal{D}(G', h', t).$$

   By the previous argument, we can assume that if $r \in \mathcal{D}(G', h') \cap N_i$, then $N_r' = \{i\}$, as otherwise the node would have a null probability. Lemma 10 then yields

$$\mathcal{D}\left(G_i^*, h^*\left(h_i\right), t^*\right) \subseteq \mathcal{D}\left(G', h', t^*\right),$$

which implies that

$$\mathcal{D}\left(G_i^*, h^*\left(h_i\right), t^*\right) \subset \mathcal{D}\left(G', h', t^*\right).$$

For any $t \leq T$, let $K(t)$ denote the number of player in $\mathcal{D}(G', h', t)$. Then

$$\theta^{\varepsilon}\left(G', h'\right) \leq \varepsilon^{\sum_{t=1}^{t^*} K(t)\alpha^t}$$

$$\theta^{\varepsilon}\left(G_i^*, h^*\left(h_i\right)\right) \geq f\left(G_i^*\right)(1-\varepsilon)^{nT}\, \varepsilon^{-\left(1-n\frac{\alpha}{1-\alpha}\right)\alpha^{t^*}+\sum_{t=1}^{t^*} K(t)\alpha^t}$$

where the upper-bound in the first inequality is obtained setting $y_k = \infty$, $k = 1, ..., W$, and $x = 1$ in (5), and the lower-bound in the second inequality is obtained by setting $W = nT$ and $y_k = 1$ in (5), and noting that

$$\sum_{k=1}^{V} z_k \leq \sum_{t=1}^{t^*-1} K(t)\,\alpha^t + \left(K\left(t^*\right)-1\right)\alpha^{t^*} + \sum_{t=t^*+1}^{\infty} n\alpha^t$$

Hence, for some constant $q' > 0$, when ε is close to zero,

$$\theta^{\varepsilon}\left(G_i^*, h^*\left(h_i\right)\right) \geq q'\varepsilon^{-\left(1-n\frac{\alpha}{1-\alpha}\right)\alpha^{t^*}+\sum_{t=1}^{t^*-1} K(t)\alpha^t}$$

Then

$$\beta^{\varepsilon}\left(G', h'|h_i\right) \leq \frac{\theta^{\varepsilon}(G', h')}{\theta^{\varepsilon}(G_i^*, h^*\left(h_i\right))} \leq \frac{\varepsilon^{\left(1-n\frac{\alpha}{1-\alpha}\right)\alpha^{t^*}}}{q'}$$

and thus, $\beta\left(G', h'|h_i\right) = 0$ since $\frac{\alpha}{1-\alpha}n < 1$.

This establishes part (a) and implies that, if $\beta(G, h|h_i) > 0$, player i believes that $\mathcal{D}(G, h) \subseteq N_i \cup \{i\}$.

To prove (b), observe that (a) implies that we can restrict attention to networks G such that $N_j = \{i\}$ for any $j \in \mathcal{D}\left(G_i^*, h^*\left(h_i\right)\right) \setminus \{i\}$. We prove the claim by contradiction. Let t^* be the earliest period t such that

$$\mathcal{D}\left(G_i^*, h^*\left(h_i\right), t\right) \neq \mathcal{D}\left(G, h, t\right).$$

Observe that the same argument as in (a) shows that

$$\mathcal{D}\left(G_i^*, h^*\left(h_i\right), t^*\right) \subset \mathcal{D}\left(G, h, t^*\right)$$

and the claim is proved analogously. ■