INFORMATION TO USERS

THREE ESSAYS ON REPEATED GAMES WITHOUT PERFECT INFORMATION

Illtae Ahn

Supervisor of Dissertation

Graduate Group Chairperson

UMI Number: 9800840

Copyright 1997 by
Ahn, Illtae

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

To my father.

.

Acknowledgments

dissertation. My colleague Matti Suominen is the co-author of the first Chapter. It has been a great pleasure to work with him. I really enjoyed every bit of the discussions and, in some occasions, arguments we had and also learned a lot from him. Beyond the co-authorship and his continued academic advices, he has always been a good friend.

Words cannot describe my gratitude to my wife, Jeongwon. She made enormous efforts to make me happy and strong. Behind all her sacrifices and supports, there is her love. My parents are so special. They have always encouraged my study and all they care about is Jeongwon's and my well-being. Without their unconditional love and absolute supports, this small achievement would not be possible. So I am dedicating this dissertation to my parents.

ILLTAE AHN

University of Pennsylvania

August 1997

v

ABSTRACT

THREE ESSAYS ON REPEATED GAMES WITHOUT PERFECT INFORMATION

ILLTAE AHN

George Mailath

In repeated games in which some players do not observe other players' actions, effective information transmission among players is an essential element in supporting a nontrivial equilibrium. The purpose of this dissertation is to study the information transmission and understand how incentive problem that might arise due to the imperfect observability may restrict some equilibrium outcomes. Chapter 1 studies a repeated buyer-seller relationship in a random matching setting, where buyers privately "network" for information and sellers have a short term incentive to supply low quality. High-quality production equilibrium is provided even when each buyer periodically interacts with only a *small* number of other buyers: the number can grow only at a rate of square root of total population. When networking is costly, low quality has to be supplied with positive probability in any equilibrium. For this case, we characterize conditions for an equilibrium in which both high and low quality are supplied. However, the analysis here does not fully elucidate the incentive problem caused by imperfect observability. In general, a player's punishing behavior might not be distinguished from his own deviation by other players

and so the punisher might be punished as well. This potential confusion raises the incentive issue of why an observer of a deviation initiate punishments rather than conceal the information. Chapter 2 directly addresses this issue by studying repeated games in which at least one player observes all the other players' actions while the other players only observe actions of the perfect observer and possibly some other players. The restrictions on the stage game payoff are obtained for the Nash-threat Folk theorem. Chapter 3 considers a repeated game in which a single long-run player plays a fixed stage game against an infinite sequence of a different set of N short-run players. The stage game played by the $N + 1$ players is a finite game of perfect information. If short-run players only observe the plays of the last K stage games rather than all previous ones, for almost all discount factors the only pure strategy equilibrium of the repeated game is simply the repetition of the stage game equilibrium.

Contents

Chapter 1

Word-of-Mouth Communication

and Community Enforcement

1.1 Introduction

It has long been recognized that community enforcement can make sellers behave cooperatively even when they meet particular buyers only infrequently and have a short term incentive to cheat, e.g., to supply low quality or to shirk in a labor contract. For instance, Klein and Leffler (1981) study the problem of credibly committing to offer high quality in a model where a continuum of buyers are randomly matched with several sellers and each seller has a short term incentive to supply low quality at a lower cost. In their model community enforcement by the buyers, through a coordinated boycott after observing low quality, provides incentives for the seller to produce high quality. The results of Klein and Leffler (1981), along with most of the existing literature on community enforcement, depend upon the assumption that past quality choices of the seller are public information.

When the number of sellers and buyers is large and particular sellers and buyers meet only infrequently or only once, the assumption of public information seems rather demanding. Recently this observation has led to a number of articles looking at community enforcement with less stringent informational assumptions. These papers include Milgrom et. al., (1990), Okuno-Fujiwara and Postlewaite (1995), Kandori (1992) and Ellison (1994). However, partially due to the difficulties of dealing with private information, all of the above mentioned papers have made extreme informational assumptions: either complete anonymity of players together with the assumption that players observe only the actions chosen in their own games, or alternatively, locally complete information, which allows a player to perfectly observe the status of his current opponent, based on the opponents past behavior. See also Greif (1993), Harrington (1995), Greif (1994) and Greif et. al., (1994).

Kandori (1992) and Ellison (1994) study a repeated prisoners' dilemma in a large but finite-population random-matching setting, where players are unable to recognize their opponents. They show that even then there exist sequential equilibria where all players play cooperatively in every period. Cooperation is supported by community enforcement based on contagious strategies: all players who are cheated immediately start cheating their opponents, understanding that the whole society is in a process of switching into non-cooperative actions. In the equilibrium, players behave cooperatively to avoid initiating a general switch to non-cooperative actions. An important factor underlying Kandori's and Ellison's results is that defection is a dominant strategy of the stage game. Whether contagious strategies would work in a repeated random matching game that does not share this property, such as the buyer seller game we are about to study, is still unknown. In addition, at least under their informational assumptions, the cooperative equilibria based on conta-

2

gious strategies are unstable in the sense that a single "insane" player who does not cooperate can destroy the good equilibrium for all agents (Ellison, 1994).

Okuno-Fujiwara and Postlewaite (1995) and Kandori (1992) consider games with local information processing: 1) Each player carries a label observable to her opponent, 2) when two players are matched they observe each other's label before choosing their actions, 3) a player and his partner's actions and labels today determine their labels tomorrow. The information processing is "local" in the sense that the actions chosen by a pair of players are based only on their labels, not on the entire distribution of labels across the population, and the updating of each player's label depends only on the previous labels and the outcome of the stage game. When the population is large and players are randomly matched, the observability of the current trading partner's label and the updating of the labels require the existence of some efficient information transmission and processing mechanism. This could be a medieval law judge (Milgrom et. al., 1990) or institutions like credit bureaus which track the transactions of every agent.

In many real life situations, however, social norms and informal information transmission mechanisms can replace formal institutions and still facilitate cooperation. In this paper, we present a model of community enforcement that is based on word-of-mouth communication. The information transmission is highly imperfect in the sense that information about each defection spreads only to part of the player population and defectors can not always be immediately punished. Despite this, since players can be identified, *private* reputations evolve. This allows equilibria where only defectors are punished; making our equilibria more stable with respect to "insane" players, who do not cooperate, than those based on contagious strategies. In fact, when information is privately costly, only some of the sellers can

3

cooperate in any equilibrium. Nonetheless, word-of-mouth communication is shown to be surprisingly efficient in facilitating cooperation.

We assume there are M sellers and M buyers, where M is large but finite number, and that in each time period $t = 0, 1, 2, ...$, the sellers and buyers are randomly matched to play a stage game that contains an opportunity for a mutually beneficial trade. The sellers have a short-term incentive to supply low quality and will supply high quality only if the gains from maintaining good reputation outweigh the short-term loss. For simplicity, it is assumed that buyers would not knowingly purchase low quality at any price. We assume that buyers have networks of communication that, roughly speaking, work in the following manner: In each period each buyer observes N trades in addition to his own and N buyers, called spectators, observe his trade and send him signals regarding his current trading partner. We can think of these N spectators as friends of the buyer or just people who happen to pass by. We assume that the identities of the spectators can change from period to period.

Throughout the paper we consider two kinds of strategy profiles, which differ in the informativeness of the signals. The strategy profile with the less informative (actually totally uninformative) signals is equivalent to a model where signalling is not allowed. For these two strategy profiles we provide sufficient conditions on N and the discount factor for a sequential equilibrium where good quality is supplied by all sellers in every period. Assuming the existence of a public randomization device and high enough discount factor, these conditions can be stated as $N \geq N^*$ where N^* is a constant determined by the population size, discount factor and the payoff matrix. As one of our main results, we show that with informative signalling N^* is a diminishing fraction of the population size.

We then study a model where "networking" (i.e., setting up N connections) is costly. In this case, when M is large, we show that there must be a positive probability of sellers producing low quality goods in any equilibrium in order to give buyers an incentive to network. When the costs of networking are below a threshold value, we find a sequential equilibrium in which sellers initially randomize between high and low quality and continue to produce high quality if and only if they produced high quality in the first period. In this equilibrium the probability of buying low quality goods increases in M and the cost of networking. When the cost of networking reaches the threshold value, trade collapses because the probability of low quality goods that would provide agents with sufficient incentive to network is so large that each buyer no longer wishes to experiment with an unknown seller.

In the existing literature on quality provision, it is assumed that agents instantaneously learn about a seller's defection, see e.g. Klein and Leffler (1981) and Allen (1984). In this paper we show that word-of-mouth communication can spread information rather quickly and make such assumptions reasonable approximations in some settings. When the population is large and information privately costly, our results suggest, however, that both high and low quality would be produced in equilibrium. To further understand the role of institutions in transmitting information, as in Milgrom et. al. (1990), we feel it is useful to understand the workings of informal channels of information transmission. We hope that our formalization of word-of-mouth communication has interest on its own.

The rest of the paper is organized as follows: Section 2 gives the formal description of the model. Section 3 discusses players' strategies and presents sufficient conditions for sequential equilibria with informative and uninformative signalling, where high quality is produced by every seller in every period. Section 4 shows

that with informative signalling as M goes to infinity, trade can be sustained with buyers networking with a diminishing proportion of other buyers. Section 5 studies costly networking and section 6 concludes the paper.

1.2   The Model

There are two finite sets of players $M_k = \{1, 2, ..., M\}$, $k = S, B$. Denote by M_S the set of sellers and by M_B the set of buyers. We envisage the sellers as being positioned at fixed locations around a circle, where the locations are numbered clockwise from 1 to M. We refer to seller i as the seller at location i. It is assumed the buyers can identify the sellers by the number of their location, but the sellers can not recognize the identities of the buyers.[1]

Let Θ be the set of all permutations of M_B. In each period $t = 0, 1, 2, ...$, a permutation $\theta_t \in \Theta$ is chosen with uniform probability, independent of previous realizations. Buyer $\theta_t(i) \in M_B$ is placed at location i to play with seller $i \in M_S$ the following simultaneous move "trade" game:

$$Buyer\ \theta_t(i)$$

		B	NB
Seller i	H	$1, 1$	$0, 0$
	L	$1+g, -\ell$	$0, 0$

where both g and ℓ are taken to be strictly positive numbers. The first (second) number in each entry indicates the seller's (buyer's) payoff. Seller's actions H and

---

[1]The more general assumption that sellers can also recognize the identities of buyers does not change our results. The two strategy profiles that we consider would be equilibria of such a game under conditions slightly different from ours.

L refer to providing "*high-quality*" and "*low-quality*" while buyer's action B refers to "*buy*" and NB to "*not buy*". With g and ℓ strictly positive, L is the dominant strategy for the seller and (L, NB) is the only Nash equilibrium of the trade game. The sellers and buyers have a common discount factor $\delta \in (0, 1)$ and their overall payoffs are the discounted sum of payoffs from the trade games.

In each period $t = 0, 1, 2, ...$, there is preplay communication among neighboring buyers before the trade games. To be precise, the stage game proceeds as follows:

1. After θ_t is realized, buyer j observes θ_t, recognizes the identity of his opponent, $\theta_t^{-1}(j)$, as well as the identities of his "neighboring" sellers, $\theta_t^{-1}(j) + 1, \theta_t^{-1}(j) + 2, ..., \theta_t^{-1}(j) + N$, where $N \leq M/2 - 1$.[2,3,4] Let us denote by $S_j(\theta_t) = \{\theta_t^{-1}(j) + k\}_{1 \leq k \leq N}$ the subset of neighboring sellers, whom buyer j observes at period t, and by $N_j(\theta_t) = \{\theta_t(\theta_t^{-1}(j) + k)\}_{1 \leq k \leq N}$ their period t matches. We call $N_j(\theta_t)$ buyer j's neighboring buyers at period t. Also, let us denote by $N_j^s(\theta_t) = \{\theta_t(\theta_t^{-1}(j) - k)\}_{1 \leq k \leq N}$ the subset of buyers, who observe the interaction between j and $\theta_t^{-1}(j)$ at period t. We call $N_j^s(\theta_t)$ the spectators to buyer j's game at period t. Notice that the identities of the spectators, neighboring sellers and buyers depend on θ_t.

2. Buyer j sends a payoff-irrelevant signal to each of his neighboring buyers $n \in N_j(\theta_t)$ and receives a message from each spectator in $N_j^s(\theta_t)$. Let us

---

[2]The assumption that buyers observe θ_t is made to ease the notation. An alternative assumption that does not change our results would be that buyer j is able to recognize only the identities of his opponent $\theta_t^{-1}(j)$ as well as his neighboring sellers in period t.

We also make the assumption that $N \leq M/2 - 1$. The extension of our analysis to $M/2 \leq N \leq M - 1$ is trivial. The case $N = M - 1$ would then correspond to the game with perfect observability, while the case $N = 0$ to the game where each buyer observes only the outcome of his trade game and the identity of his opponent.

[3]Sellers are female and buyers are male.

[4]These locations are of Mod M.

7

introduce the following notation.

- $C = \{\gamma, \beta\}$: the set of possible signals. We can interpret a signal γ or β as meaning respectively "*Good*" or "*Bad*".

- $m_t^j(\ell) \in \{\gamma, \beta\}$: the signal from buyer j to buyer $\ell \in N_j(\theta_t)$.

- $m_t^j \in \{\gamma, \beta\}^N$: the N-tuple of the signals from j to each of his neighboring buyers in $N_j(\theta_t)$.

- $m_t(j) \in \{\gamma, \beta\}^N$: the N-tuple of the signals from j's spectators, $N_j^s(\theta_t)$, to buyer j.

3. Seller $\theta_t^{-1}(j)$ and buyer j play the 2×2 simultaneous move trade game described above. Denote the outcome (or the realized action profile) of that game by $(a_t^S(\theta_t^{-1}(j)), a_t^B(j))$, where $a_t^S(\theta_t^{-1}(j)) \in A_S = \{H, L\}$, $a_t^B(j) \in A_B = \{B, NB\}$.

4. In addition to his own outcome, buyer j observes the realized action profiles of the period t trade games played by the sellers $i \in S_j(\theta_t)$ and buyers $n \in N_j(\theta_t)$. Denote this observation by $o_t(j) = \left((a_t^S(i), a_t^B(\theta_t(i)))\right)_{i \in S_j(\theta_t) \cup \{\theta_t^{-1}(j)\}} \in (A_S \times A_B)^{N+1}$.

The information that buyer j receives in period t can now be written as $(\theta_t, m_t(j), o_t(j))$. We denote with $H^t(j)$ the set of all possible histories for buyer j up to but not including period t. By convention, let $H^0(j) = \emptyset$. An element $h_t(j) \in H^t(j)$ includes all past realizations of θ_s, all past messages to player j, $m_s(j)$, all past messages from player j, m_s^j, and all past observations of player j, $o_s(j)$, where $0 \le s < t$. Hence $h_t(j)$ is:

$$h_t(j) = (\theta_\tau, m_\tau(j), m_\tau^j, o_\tau(j))_{\tau=0}^{t-1}.$$

8

A pure strategy for buyer j is then a sequence $\{\widehat{m}_t^j, \widehat{b}_t^j\}_{t=0}^\infty$, where

$$\widehat{m}_t^j : \Theta \times H^t(j) \to \{\gamma, \beta\}^N$$

$$\widehat{b}_t^j : \Theta \times H^t(j) \times \{\gamma, \beta\}^N \to \{B, NB\}.$$

$\widehat{m}_t^j(\theta_t, h^t(j))$ specifies the N-tuple of signals that buyer j with private history $h^t(j)$ sends to his neighboring buyers $n \in N_j(\theta_t)$ in period t. $\widehat{b}_t^j(\theta_t, h^t(j), m_t(j))$ specifies the choice of action for buyer j in the period t trade game against seller $\theta_t^{-1}(j)$, when j has private history $h^t(j)$ and he receives signals $m_t(j)$ in the period t communication stage. Correspondingly, let $\{\widehat{\mu}_t^j, \widehat{\beta}_t^j\}_{t=0}^\infty$ denote a behavioral strategy for buyer j, where

$$\widehat{\mu}_t^j : \Theta \times H^t(j) \to \Delta\{\gamma, \beta\}^N$$

$$\widehat{\beta}_t^j : \Theta \times H^t(j) \times \{\gamma, \beta\}^N \to \Delta\{B, NB\}.$$

For seller i we define pure and behavioral strategies as sequences of maps $\{\widehat{s}_t^i\}_{t=0}^\infty$ and $\{\widehat{\sigma}_t^i\}_{t=0}^\infty$, where

$$\widehat{s}_t^i : (\{H, L\} \times \{B, NB\})^t \to \{H, L\}.$$

$$\widehat{\sigma}_t^i : (\{H, L\} \times \{B, NB\})^t \to \Delta\{H, L\}.$$

Note that the assumption that a seller does not recognize the identity of a buyer is implicit in this notation.

Because of the private histories that players have, the equilibrium concept that we apply is sequential equilibrium. Sequential equilibrium requires that after any history player's equilibrium strategy maximize his (her) expected payoff, taking as given all other player's strategies and his beliefs about the signals and actions taken by all other players in all previous periods. Furthermore, his beliefs should be "consistent" with the equilibrium strategy profile and private history, in the sense of

9

Kreps and Wilson (1982).[5] A trivial sequential equilibrium is one where sellers play L and buyers NB after any history: the repetition of the *only* Nash equilibrium of the trade game. We are interested, however, in sequential equilibria that support the efficient outcome where (H, B) is played by all players in every period.

For the most part we confine our analysis on a particular class of strategy profiles, which we call "unforgiving". These strategy profiles require sellers to sell high quality in period zero (in section 5 with some probability), and sell high quality thereafter if and only if 1) they have always done so, and 2) buyers have always purchased their goods. Under the unforgiving strategy profile buyers play B, except to punish a seller by playing NB when they are informed of her defection. The strategy profiles are unforgiving in the sense that informed buyers punish a defector whenever they meet her.

There are two reasons for focusing on these strategy profiles. First, they are simple: In fact, because of the private information that players have, it is difficult to imagine other strategies that could support the efficient outcome as a sequential equilibrium in this game. For instance, it is not obvious whether contagious strategy profiles, where a seller's defection affects how buyers treat other sellers, would be equilibria in this game. Checking the incentives of a buyer to follow such a strategy on off-the-equilibrium paths is very complicated, because his incentives depend on his belief about the previous plays, which in turn depend on his private history.[6] Strategies with less severe, finite punishments are also difficult to implement because buyers typically do not know the time of the first defection and therefore cannot

[5]In Kreps and Wilson (1982), the definition of sequential equilibrium requires the specification of beliefs system as well as a strategy profile. Because the beliefs system which is consistent with our strategy profiles is simple, we refer only to the strategy profile when describing a sequential equilibrium.

[6]See Kandori (1992) for a discussion on the difficulties of private information.

synchronize the last period of a punishment phase. Unforgiving strategies avoid these problems and the buyers incentives are easily shown to be satisfied. The second reason for focusing on these strategy profiles is that, in the class of non-contagious strategy profiles (i.e., where one sellers action does not affect how the other sellers are treated), these strategy profiles provide the maximum punishment for the seller. This is important because the conditions for the efficient outcome that we derive then characterize the minimum N that is necessary for the efficient outcome in any sequential equilibrium based on non-contagious strategy profiles.

1.3 Exogenous Connections and Trade

In this section we provide sufficient conditions in terms of N for a sequential equilibrium where (H, B) is played by all players in every period. In our model information about sellers' behavior may spread through two possible sources, the effectiveness of which depends on the number of spectators, N. First, by observing the outcomes of $N + 1$ trade games in each period, a buyer receives information about $N + 1$ sellers: he observes their current actions and may infer knowledge of their past defections from the actions of their opponents. Second, the information can be transmitted through direct communication among neighboring buyers. The effectiveness of direct communication depends, however, on the information content of the signals.

We now introduce two unforgiving strategy profiles that differ with respect to the informativeness of the buyer's signals. For obvious reasons we refer to the first as the *Uninformative Strategy Profile* and to the second as the *Informative Strategy Profile*. In all periods $t = 0, 1, 2....$, after θ_t is realized:

The strategy for seller i is same under both strategy profiles and is:

(I) In the first period play H. After that, if the outcome in seller i's past games was always (H, B), play H.

(II) Play L otherwise.

The *Uninformative Strategy* for buyer j is:

(III) Signal randomly γ or β with equal probabilities to all $n \in N_j(\theta_t)$, irrespective of the private history $h_t(j)$.

(IV) If j has ever observed $\theta_t^{-1}(j)$ play L or someone (including himself) play NB against her, play NB regardless of the messages $m_t(j)$.

(V) Play B otherwise.

The *Informative Strategy* for buyer j is:

(III)' If j has ever observed $\theta_t^{-1}(n)$, where $n \in N_j(\theta_t)$, play L or someone (including himself) play NB against her, signal β. Signal γ otherwise.

(IV)' If j previously observed $\theta_t^{-1}(j)$ play L or someone play NB against her, or if he received a message β from any of his current spectators, $h \in N_j^s(\theta_t)$, play NB.

(V)' Play B otherwise.

Under both strategies, in the beginning of each period t each buyer j categorizes sellers into two different status groups based on his private history $H^t(j)$. If he has observed a seller play L or someone (including himself) play NB against her by period $t - 1$, he gives her the *"Bad"* status β. Otherwise, he gives her the *"Good"*

12

status γ. During the communication stage in period t (phase 2 of the stage game), buyer j is given the opportunity to signal to his neighboring buyers $n \in N_j(\theta_t)$ the statuses that he has assigned to their opponents $\theta_t^{-1}(n)$, and to revise the status that he assigns to his current opponent by taking into account the signals that he receives from the period t spectators to his game $h \in N_j^s(\theta_t)$.

As can be seen from the conditions (III), (IV) and (V), under the *Uninformative Strategy Profile* buyers merely "babble", disregard their neighbors' signals and base their choices of action against their period t opponents on the statuses that they assigned to them after the period $t - 1$. So the communication stage is totally uninformative. Under the *Informative Strategy Profile*, on the other hand, signalling reveals all the relevant information (about receiver's opponent) of the senders, given the seller's strategy. Under this profile each buyer j sends a signal γ or β to each of his neighboring buyers $n \in N_j(\theta_t)$, depending on the statuses that he assigned to their opponents $\theta_t^{-1}(n)$ after period $t - 1$. He also fully respects the messages that his spectators $h \in N_j^s(\theta_t)$ send to him before the period t trade game, and revises the status that he has assigned to his current opponent $\theta_t^{-1}(j)$ accordingly, basing his period t trade game choice of action on the revised status of $\theta_t^{-1}(j)$. Both strategy profiles are unforgiving since once a buyer assigns a particular seller a status β, he never upgrades her status to γ.

Before proceeding it is convenient to introduce some additional notation. Take any two time periods t' and t, where $t' < t$. Under the *Uninformative Strategy Profile* denote by b the probability that $\theta_t(i)$ was among the $N + 1$ buyers who observed seller i at period t'. That is, let b denote the probability

$$\Pr\left\{ i \in S_{\theta_t(i)}(\theta_{t'}) \bigcup \{\theta_{t'}^{-1}(\theta_t(i))\} \right\}.$$

Correspondingly, under the *Informative Strategy Profile* denote by b the probability

that either $\theta_t(i)$ or some of the N $\theta_t(i)$'s time t spectators, $h \in N^s_{\theta_t(i)}(\theta_t)$, were among the $N + 1$ buyers who observed i at period t'. In this case, b is the probability

$$\Pr\left\{ i \in \bigcup_{j \in \{\theta_t(i)\} \cup N^s_{\theta_t(i)}(\theta_t)} [S_j(\theta_{t'}) \bigcup \{\theta_{t'}^{-1}(j)\}] \right\}.$$

It is straightforward to check that

$$b = \begin{cases} \dfrac{N+1}{M} & \text{with the } \textit{Uninformative Strategy Profile, and} \\[3em] 1 - \dfrac{\dbinom{M-N-1}{N+1}}{\dbinom{M}{N+1}} & \text{with the } \textit{Informative Strategy Profile} \end{cases}$$

For both strategies, note that since θ_t is i.i.d., b is time independent and does not depend on t and t'. Note also that for both strategies b increases in the number of spectators, N. In the proofs of Propositions 1 and 2 we need the unconditional probability of $\theta_t(i)$ assigning the seller i a status γ after the period t communication stage, given that seller i has defected in every period $t' \in \{ t_D, \dots t - 1 \}$. With our notation, this probability can be written as $(1 - b)^{t-t_D}$.

We are now ready to state our first two propositions. These propositions provide the conditions under which the *Uninformative* and the *Informative Strategy Profiles* are sequential equilibria of the random matching game. Notice that under both strategy profiles (H, B) is played in every period at each location along the equilibrium path.

Proposition 1: Define two constants δ^* and b^* as follows:

$$\delta^* = \left(1 + \frac{(\frac{M+g}{M})^2}{4(1+g)g(\frac{M-1}{M})} \right)^{-1},$$

14

$$b^* = \frac{g(1-\delta)}{\delta}.$$

i) If $\frac{g}{1+g} \leq \delta \leq \delta^*$, the *Uninformative Strategy Profile* is a sequential equilibrium of the above random matching game if $b \geq b^*$.

ii) If $\delta \geq \max[\frac{g}{1+g}, \delta^*]$, there exists constants b_L and b_H, where $b^* \leq b_L \leq b_H < 1$, such that the *Uninformative Strategy Profile* is a sequential equilibrium of the above random matching game if either $b^* \leq b \leq b_L$ or $b \geq b_H$. The constants b_L and b_H are given by

$$b_L = \frac{\frac{M+g}{M} - \sqrt{(\frac{M+g}{M})^2 - 4(1+g)\frac{g(1-\delta)}{\delta}(\frac{M-1}{M})}}{2(1+g)}$$

and

$$b_H = \frac{\frac{M+g}{M} + \sqrt{(\frac{M+g}{M})^2 - 4(1+g)\frac{g(1-\delta)}{\delta}(\frac{M-1}{M})}}{2(1+g)}.$$

The condition $\delta \geq g/(1+g)$ is necessary because with δ less than this, the efficient outcome could not be sustained by any equilibrium even with $M = 1$.[7]

Buyers' incentives to follow the unforgiving strategy profiles are easily satisfied: A buyer should expect his current opponent with status β to play L, regardless of his beliefs about the outcomes of her previous games, in which case NB is his best choice of action. If, on the other hand, he assigns her a status γ, playing B

---

[7]When $M = 1$, our random matching game is equivalent to a two-player standard repeated game with observable actions. In this case, the *Uninformative Strategy Profile*, which is identical with the *Informative Strategy Profile*, provides the maximum punishment for defection. This profile supports the efficient outcome as a Nash and a subgame perfect equilibrium if and only if $\delta \geq g/(1+g)$.

is optimal both on and off the equilibrium path given the consistent belief that she has never defected and will play H. Also the incentives for signalling are trivially satisfied.

Sellers' incentives to follow the uninformative strategy profile are characterized by two conditions: one preventing her from playing L on the equilibrium path and one that guarantees that sellers who have defected keep defecting irrespective of their private history. To give a seller an incentive to play H along the equilibrium path, the short-term gain from cheating, g, must be outweighed by the long-term loss resulting from the gradual loss of reputation among buyers. Given the buyers' strategies, this occurs if b (N) is sufficiently large that information about a seller's defection spreads quickly enough among the buyers. This results in the condition that $b \geq b^*$. On off-the-equilibrium paths, the strategy profile requires sellers to keep defecting rather than play H in an attempt to slow down the deterioration of her reputation. This off-the-equilibrium path constraint is satisfied when δ is small, $\delta \leq \delta^*$, since the short term gain g from selling low quality will then outweigh the future reward from trying to maintain a good reputation. When $\delta > \delta^*$, the condition is satisfied if either b is very large or b is small enough. If b is very large, $b \geq b_H$, it does not pay to slow down the deterioration of one's reputation, since with several informed buyers already playing NB against the seller, all buyers are soon likely to learn about seller's bad status anyway. On the contrary, if b is small enough, $b \leq b_L$, playing L is better than playing H simply because the information about her defections is not spreading very quickly. It can be shown that the off-the-equilibrium path constraint is always satisfied when $b = b^*$, implying that $b_L \geq b^*$.

This strategy profile is not a sequential equilibrium when $b \in (b_L, b_H)$. In

proposition 3, however, by using a public randomization device we construct a sequential equilibrium which supports the efficient outcome for any b greater than b^*.

Proof. When (II) holds a seller has incentive to follow (I) if and only if the following inequality holds:

$$\frac{1}{1-\delta} \geq \frac{1+g}{1-(1-b)\delta}. \tag{1.1}$$

The left-hand side is the payoff from playing H in every period whereas the right-hand side is the payoff from playing L in every period. Since b must be less than one, the inequality can hold only when $\delta \geq g/(1+g)$. In that case equation (3.1) can be written as:

$$b \geq b^*. \tag{1.2}$$

By the principle of dynamic programming to verify that (II) is optimal, it is enough to check that a one time switch to H is not profitable after any history in which the seller has obtained a bad status, i.e., she has played L or some buyer has played NB against her. We can show that the seller's incentives to follow (II) increase in the number of buyers who are aware of her bad status. Since consistency requires this number to be at least $N+1$ (in states that (II) is concerned with), it will be sufficient for us to show that a seller has incentive to follow (II) when exactly $N+1$ players assign her a bad status.

Define $P_s(K)$ as the probability that $K+s$ buyers know about i's bad status after period t if i plays H in period t and K players knew about i's bad status after

17

the period $t - 1$. Let $\alpha = K/M$. It is straightforward to show that:

$$P_0(K) = (1 - \alpha) + \frac{\alpha \binom{K-1}{N}}{\binom{M-1}{N}}$$

and for $\alpha < 1$,

$$P_s(K) = \frac{\alpha \binom{K-1}{N-s}\binom{M-K}{s}}{\binom{M-1}{N}}, \forall\ 1 \leq s \leq \min[N, M - K].$$

Denote with $u_s(K) = 1 - (K + s)/M$ the associated conditional probability that seller i's period $t + 1$ match $\theta_{t+1}(i)$ does not know about i's bad status (after the period $t + 1$ communication stage), given that $K + s$ buyers know about i's bad status after the period t.

Then, assuming that K buyers are aware of a seller's bad reputation before the current period, a seller has incentive to follow (II) if and only if:

$$\frac{(1 + g)(1 - \alpha)}{1 - (1 - b)\delta} \geq (1 - \alpha) + \frac{\delta(1 + g)}{1 - (1 - b)\delta} \sum_{s=0}^{\min[N, M-K]} P_s(K)u_s(K). \qquad (1.3)$$

The left hand side is the payoff from playing L in each of the remaining periods, whereas the right hand side is the payoff for playing H in one period and then L thereafter. Realizing that s follows a hypergeometric distribution for $s = 1, 2, ...N$, it can be shown that

$$\sum_{s=0}^{\min[N, M-K]} P_s(K)u_s(K) = (1 - \alpha)^2 + \alpha(1 - \alpha)(1 - b)\left(\frac{M}{M - 1}\right).$$

It is now easy to see that equation (3.3) is relaxed as α is increased. The intuition for this result is quite simple: A seller who is matched with a buyer that is not aware of his bad status may by playing H keep his reputation among at most $N + 1$

18

players and benefit from this reputation later. When α is large many of her current spectators are likely to know about his defection already, which reduces the benefit to playing H. Since $\alpha \geq b$ by consistency, it is sufficient to check that equation (3.3) holds when $\alpha = b$. Setting $\alpha = b$ and rearranging, equation (3.3) can be written:

$$b^2 - \left(\frac{M+g}{(1+g)M}\right)b + \frac{(1-\delta)g}{(1+g)\delta}\frac{M-1}{M} \geq 0.$$

This quadratic inequality holds if either $\delta \leq \delta^*$ or if $\delta \geq \delta^*$ and either $0 \leq b \leq b_L$ or $b_H \leq b \leq 1$. Combined with equation (3.2) this result implies the equilibrium conditions stated in proposition 1.

(III) is trivially satisfied given (IV) and (V). It is also easy to see that (IV) and (V) are satisfied given (I),(II) and (III). If buyer j observed his current match, $\theta_t^{-1}(j)$, play L in the past or some buyer play NB against her, he should believe she will play L by (II); so playing NB is his best response. If he has never observed $\theta_t^{-1}(j)$ play L, nor some buyer play NB against her, then given (I), (II) and (III), he should believe she will play according to (I) regardless of the messages he has received. This being the case, B is his optimal choice of action. This establishes (IV) and (V) and completes the proof. ∎

As one would expect, the equilibrium conditions are very similar for the *Informative Strategy Profile*. In this case, however, it is not possible to state the off-the-equilibrium path conditions in terms of b as was true for the *Uninformative Strategy Profile*.

Proposition 2: For $\delta \geq g/(1 + g)$, the *Informative Strategy Profile is* a sequential equilibrium of the above random matching game if

$$1 - \frac{\dbinom{M-N-1}{N+1}}{\dbinom{M}{N+1}} \geq \frac{g(1-\delta)}{\delta} \qquad (1.4)$$

and

$$\frac{g}{\delta(1+g)} \geq \left(\frac{1+2g}{1+g}\right)\frac{\dbinom{M-N-1}{N+1}}{\dbinom{M}{N+1}} - I_{\{3(N+1)\leq M\}}\frac{\dbinom{M-2N-2}{N+1}}{\dbinom{M}{N+1}}, \qquad (1.5)$$

where I_A is a function that takes value 1 if A is true and 0 otherwise.

Equation (3.4) concerns on-the-equilibrium path behavior and can be written as $b \geq b^*$. Equation (3.5) is the constraint for off-the-equilibrium path behavior. As before, it is always satisfied when δ is small enough and when δ is large it is satisfied when b is either very large or b is small enough. Although the intuition for both conditions is exactly the same as under the *Uninformative Strategy Profile*, the second condition is not exactly the same in terms of b under the two strategy profiles. This constraint concerns a seller's possible deviation to H when playing on off-the-equilibrium paths, and is different for the two strategy profiles because the dissemination of a seller's bad reputation when playing H is different in terms of b under the two strategy profiles.

Proof. As was shown in the proof of proposition 1, assuming that (II) holds, a seller has incentive to follow (I) if and only if $b \geq b^*$. This is stated in equation (3.4).

Showing that she has incentive to follow (II) when equation (3.5) holds proceeds much as the proof of proposition 1. Denote with α the probability that seller i's period t opponent $\theta_t(i)$ is aware of i's bad status after the period t communication stage, if K buyers are aware of sellers i's bad status after the period $t-1$. Correspondingly, denote with ξ the probability that $\theta_t(i)$ is aware of i's bad status after the period t communication stage, if $K+N+1$ buyers are aware of sellers i's bad status after period $t-1$. Then

$$\alpha = 1 - I_{\{(N+1)\leq M-K\}}\frac{\binom{M-K}{N+1}}{\binom{M}{N+1}}$$

and

$$\xi = 1 - I_{\{2(N+1)\leq M-K\}}\frac{\binom{M-K-N-1}{N+1}}{\binom{M}{N+1}}.$$

Let P_s and u_s represent the same probabilities as in the proof of proposition 1.

Seller i has an incentive to follow (II) if and only if:

$$(1+g)\left((1-\alpha) + \frac{\delta((1-\alpha)(1-\xi) - (1-\alpha)^2 + \sum_{s=0}^{\min[N,M-K]} P_s u_s)}{1-(1-b)\delta}\right) \geq$$

$$(1-\alpha) + (1+g)\left(\frac{\delta \sum_{s=0}^{\min[N,M-K]} P_s u_s}{1-(1-b)\delta}\right) \tag{1.6}$$

The left hand side is the payoff from playing L in each of the remaining periods, whereas the right hand side is the payoff from playing H in one period and then L thereafter. Seller i's action makes a difference only when neither i's opponent $\theta_t(i)$ nor any of the spectators to i's game have assigned the bad status to i. This happens

with probability $(1 - \alpha)$. In that case, playing L results in a larger payoff by g, but $N+1$ new buyers learn about i's bad status, reducing the probability that i receives $(1+g)$ in the next period from $(1-\alpha)$ to $(1-\xi)$. If $\theta_t(i)$ or some of the spectators to i's game know about i's bad status, which happens with probability α, i receives nothing in that period and her reputation deteriorates similarly irrespective of the action that he takes.

This inequality can be written as:

$$\frac{g(1 - (1 - b)\delta)}{\delta(1 + g)} \geq \xi - \alpha. \tag{1.7}$$

It is now straightforward to check that equation (3.6) is relaxed as K increases. Since $K \geq N$ by consistency, it is sufficient to check that equation (3.6) holds when $K = N$ or $\alpha = b$. Setting $\alpha = b$ and rearranging gives equation (3.5).

If a buyer j has ever observed a seller $i \in S_j(\theta_t)$ play L or her opponent play NB against her, j is indifferent between signalling γ or β to $\theta_t(i)$, since given (II) he expects i to play L in all his future games, including those with j himself, irrespective of the signal that he sends. So we may assume that he sends a truthful signal β in this case. If j has never observed i's game or if j has observed i but the outcome in i's games were always (H, B) he strictly prefers to send the message γ instead of β. Sending the message β would result in $\theta_t(i)$ playing NB against seller i, giving her a bad status and making her play L in the future. This would reduce buyer j's future payoffs from games where he is matched with seller i. This establishes condition (III)'. Given (I), (II) and (III)' conditions (IV)' and (V)' are trivial. ∎

In order to create sequential equilibria that support trade for all $b \geq b^*$, let us now extend our basic model to include a public randomization device. The idea of using a public randomization device to adjust the severity of punishments is borrowed from Ellison (1994). In particular, we assume that before players choose their actions in period t, they observe a public random variable f_t which is drawn independently from a uniform distribution on $[0, 1]$. Let $f \in [0, 1]$ and consider adjusting the *Uninformative* and *Informative Strategy Profiles* as follows: In period t, sellers play according to the original strategies as long as $f_t \leq f$, but return immediately to the equilibrium path of the original strategies if $f_t > f$; buyers play according to the original strategies, except whenever $f_t > f$, at which point they forget all past actions of sellers and assign each seller status γ. Assuming that such a public randomization device is available, we can state the following proposition.

Proposition 3: For $\delta \geq g/(1 + g)$ there exists a function $f(\delta)$ such that the *adjusted Informative* and *Uninformative Strategy Profiles* with $f = f(\delta)$ are a sequential equilibrium of the random matching game if $b \geq b^*$, where $b^* \equiv g(1-\delta)/\delta$.

The idea in the proof is that whenever $b \geq b^*$, we can by an appropriate choice of f adjust the severity of punishment for a seller so that she becomes indifferent between playing H (following (I)) and deviating on the equilibrium path. Because at off-the-equilibrium paths a seller has less incentive to protect her reputation than on-the-equilibrium path, as her reputation is deteriorating anyway, this indifference can be shown to imply that the off-the-equilibrium path condition always holds.

Proof. For any f_t, a seller has incentive to follow (I) if and only if

$$\frac{1}{1-\delta} \geq \frac{1+g}{1-(1-b)f\delta} + \sum_{t=1}^{\infty} f^{t-1}(1-f)\frac{\delta^t}{1-\delta}$$

or

$$\frac{1}{1-\delta f} \geq \frac{1+g}{1-(1-b)f\delta} \qquad (1.8)$$

The left-hand side of the first inequality is the seller's payoff from following (I), while the right-hand side is her payoff from deviating and following (II), as long as $f_\tau \leq f$, and following (I) thereafter. For all $\delta \geq g/(1+g)$ and $b \geq b^*$, there exists $f(\delta) \in [0,1]$ such that equation (3.8) holds as an equality when $f = f(\delta)$. From now on, let us assume that $f = f(\delta)$.

A seller who is playing on the equilibrium path is now indifferent between playing H in every period and deviating. She is also indifferent between playing H in the current period and then deviating and deviating right away. By playing H in the current period she can keep her good reputation among $N+1$ buyers, until they, in some way, learn about her defections in the future. Now consider a seller who is following (II). If her opponent and all the N spectators to her game happen to assign her a good status, she also can keep her good reputation among $N+1$ buyers by playing H. This reputation, however, is worth less to her than if she were on the equilibrium path because, with some buyers already assigning her a bad status, these $N+1$ buyers are more likely to learn about her bad status before playing against her in the future. Since the short term gain from deviating is same in both cases, we conclude that playing L is optimal off the equilibrium path.

More formally, consider the sellers incentives to follow (II) in period t when $f_t \leq f(\delta)$. By the principle of dynamic programming, it is sufficient to show that

a single-period deviation to H is unprofitable. Let $\alpha, P_s(K)$, and $u_s(K)$ denote the same probabilities as in the proof of proposition 1. A seller has an incentive to follow (II) if and only if:

$$\frac{(1+g)(1-\alpha)}{1-(1-b)f(\delta)\delta} + \sum_{t=1}^{\infty} f(\delta)^{t-1}(1-f(\delta))\frac{\delta^t}{1-\delta} \geq$$

$$(1-\alpha) + \frac{\delta f(\delta)(1+g)\left(\sum_{s=0}^{\min[N,M-K]} P_s u_s\right)}{1-(1-b)f(\delta)\delta} + \sum_{t=1}^{\infty} f(\delta)^{t-1}(1-f(\delta))\frac{\delta^t}{1-\delta} \quad (1.9)$$

or

$$\frac{(1+g)(1-\alpha)}{1-(1-b)f(\delta)\delta} \geq (1-\alpha) + \frac{\delta f(\delta)(1+g)\left(\sum_{s=0}^{\min[N,M-K]} P_s u_s\right)}{1-(1-b)f(\delta)\delta}. \quad (1.10)$$

We can show that this inequality holds as follows:

$$\frac{(1+g)(1-\alpha)}{1-(1-b)f(\delta)\delta} = \frac{(1-\alpha)}{1-\delta f(\delta)} =$$

$$(1-\alpha) + \frac{(1+g)\delta f(\delta)(1-\alpha)}{1-(1-b)f(\delta)\delta} \geq$$

$$(1-\alpha) + \frac{(1+g)\delta f(\delta)\left(\sum_{s=0}^{\min[N,M-K]} P_s u_s\right)}{1-(1-b)f(\delta)\delta}.$$

The first two equalities come from the fact that (3.8) holds as an equality with $f = f(\delta)$. The inequality follows since

$$\sum_{s=0}^{\min[N,M-K]} P_s u_s \leq \sum_{s=0}^{\min[N,M-K]} P_s u_0 = u_0 = 1-\alpha.$$

When $f_t \leq f(\delta)$ a buyer's problem is similar to that in propositions 1 and 2 so he is better off following the original strategies. On the other hand, given that a past defector plays H after $f_t > f(\delta)$ it is optimal for the buyer to assign her a status γ and treat her like the seller who never defected. ∎

1.4 Large Population Results

In this section we study how fast N must grow in relation to M in order to sustain (H, B) as the outcome of the trade game for our strategies. If we denote $N^*(M)$ as the smallest integer N that satisfies the constraint $b \geq g(1 - \delta)/\delta$ (i.e., $b \geq b^*$), then given propositions 1,2 and 3, the question can be reformulated as how fast $N^*(M)$ grows in relation to M.

If seller i has defected in the previous period, the $N + 1$ buyers who observed the defection are the only ones who are informed of the defection before the current period's trade game starts. Then b, the probability that seller i meets a buyer who is informed of her previous defection, is simply $(N + 1)/M$ under the *Uninformative Strategy Profile*. With this strategy profile it is clear that N^* and M must grow at the same rate in the limit.

This is not true, however, with the *Informative Strategy Profile*. In this strategy profile i's current opponent $\theta_t(i)$ assigns her a bad status if he observed this defection or he received an informative signal based on this defection during the current period's communication stage. With N spectators to $\theta_t(i)$'s game each of whom observes $N + 1$ games, $\theta_t(i)$ obtains information from $N^2 + 2N + 1$ possibly overlapping games and assigns i a bad status if any of these games was played at location i. This suggests that N^* may grow more slowly than M. Below we show

that in order to sustain trade, N^* must grow only at a rate $\sqrt{M}$. To prove this result formally we need the following lemma.

Lemma 1: Under the *Informative Strategy Profile* $\lim_{M \to \infty} \dfrac{N^*(M)}{M} = 0$.

Proof. By the definition of $N^*(M)$ we have the following inequalities

$$\frac{\dbinom{M - N^*(M) - 1}{N^*(M) + 1}}{\dbinom{M}{N^*(M) + 1}} \leq (1 - b^*) < \frac{\dbinom{M - N^*(M)}{N^*(M)}}{\dbinom{M}{N^*(M)}} \tag{1.11}$$

First note $\limsup_{M \to \infty} N^*(M)/M < 1/2$. For the subsequences of $N^*(M)/M$ whose limits are $1/2$, the numerator of the right-hand side of the strict inequality in (4.1) approaches 1 and the denominator goes to the infinity, leading to a contradiction since b^* is assumed to be strictly less than one.

Using Stirling's formula

$$\sqrt{2\pi n}\, n^n e^{-n} \leq n! \leq \sqrt{2\pi n}\, n^n e^{-n} e^{\frac{1}{12n}}$$

the second inequality in (4.1) implies that

$$(1 - b^*) <$$

$$\frac{(M - N^*)^{2(M-N^*)+1}}{M^{M+\frac{1}{2}}(M - 2N^*)^{M-2N^*+\frac{1}{2}}} e^{\frac{1}{6(M-N^*)}} \tag{1.12}$$

Given that $\limsup_{M \to \infty} N^*/M < 1/2$, it is easy to see that the exponential term that appears at the right-hand side of (4.2) goes to 1 as M approaches infinity.

The nonexponential term then has to be bounded away from zero for large M. In what follows we show this implies $\lim_{M \to \infty} N^*/M = 0$.

Let $q = N^*/M$. The nonexponential terms in the right hand side of (4.2) can now be written as:

$$\left(\frac{(1-q)^{2-2q}}{(1-2q)^{1-2q}} \right)^M \frac{(1-q)}{(1-2q)^{1/2}}.$$

Since $0 < q < 1/2$ and $\lim \sup_{M \to \infty} q < 1/2$ the second term of the above expression is bounded. Furthermore, we can show the first term is strictly decreasing with respect to q for $0 \le q \le 1/2$ and for that range of q it is one if and only if $q = 0$. Thus if $\lim_{M \to \infty} q \ne 0$, the first term is very close to zero for large M, which is a contradiction. Hence it must be that $\lim_{M \to \infty} N^*/M = 0$. ■

Proposition 4: Under the *Informative Strategy Profile* $0 < \lim_{M \to \infty} \dfrac{N^*(M)^2}{M} < \infty$.

Proof. If we rewrite the inequalities (4.1) using Stirling's formula we get

$$\frac{(M - N^* - 1)^{2(M-N^*)}}{M^M (M - 2N^* - 2)^{M-2N^*}} \left(\frac{(M - 2N^* - 2)^3}{M(M - N^* - 1)^2} \right)^{\frac{1}{2}} e^{-\frac{1}{12(M-2N^*-2)} - \frac{1}{12M}}$$

$$\le (1 - b^*) <$$

$$\frac{(M - N^*)^{2(M-N^*)}}{M^M (M - 2N^*)^{M-2N^*}} \left(\frac{(M - N^*)^2}{M(M - 2N^*)} \right)^{\frac{1}{2}} e^{\frac{1}{6(M-N^*)}} \tag{1.13}$$

28

Since the last two terms of both sides of the inequalities (4.3) approach one as $M \to \infty$ by lemma 1 and both of the first terms exhibit the same behavior in the limit, we must have

$$\frac{(M - N^*)^{2(M-N^*)}}{M^M (M - 2N^*)^{M-2N^*}} \to (1 - b^*) \ as \ M \to \infty. \tag{1.14}$$

Next we show that $0 < \lim_{M \to \infty} N^{*2}/M < \infty$ in order for (4.4) to hold.

First of all, note that we can write

$$\frac{(M - N^*)^{2(M-N^*)}}{M^M (M - 2N^*)^{M-2N^*}} =$$

$$\left[\left(1 - \left(\frac{N^*}{M - N^*} \right)^2 \right)^{-\left(\frac{M-N^*}{N^*} \right)^2} \right]^{\frac{N^{*2}}{M-N^*}} \left[\left(1 - \frac{2N^*}{M} \right)^{\frac{-M}{2N^*}} \right]^{\frac{-2N^{*2}}{M}} =$$

$$(e + \alpha_M)^{\frac{N^{*2}}{M-N^*}} (e + \beta_M)^{\frac{-2N^{*2}}{M}}$$

where

$$\alpha_M = \left(1 - \left(\frac{N^*}{M - N^*} \right)^2 \right)^{-\left(\frac{M-N^*}{N^*} \right)^2} - e,$$

$$\beta_M = \left(1 - \frac{2N^*}{M} \right)^{\frac{-M}{2N^*}} - e.$$

Since $\lim_{M \to \infty} N^*/M = 0$, both sequences $\{\alpha_M\}, \{\beta_M\}$ converge to zero from above. Now define new sequences $\{a_M\}, \{b_M\}$ such that $e^{a_M} = e + \alpha_M$, $e^{b_M} = e + \beta_M$. We can then write

$$\frac{(M - N^*)^{2(M-N^*)}}{M^M (M - 2N^*)^{M-2N^*}} = e^{a_M \frac{N^{*2}}{M-N^*} - b_M \frac{2N^{*2}}{M}}. \tag{1.15}$$

29

Given (4.4) and (4.5), all that remains to show is that $\{N^{*2}/M\}$ must converge to some positive number in order for $\{a_M N^{*2}/(M - N^*) - b_M 2N^{*2}/M\}$ to converge to $\ln(1 - b^*)$. If $\lim_{M \to \infty} N^{*2}/M = \infty$, $a_M N^{*2}/(M - N^*) - b_M 2N^{*2}/M$ approaches minus infinity as M goes to infinity. This can be easily shown using the fact that $\{a_M\}$ and $\{b_M\}$ converge to one from above. If $\lim_{M \to \infty} N^{*2}/M = 0$, we obtain another contradiction since $a_M N^{*2}/(M - N^*) - b_M 2N^{*2}/M \to 0$ as $M \to \infty$. Since $\{N^{*2}/M\}$ is bounded, it is straightforward to show $\lim_{M \to \infty} N^{*2}/M = -\ln(1 - b^*)$. ∎

1.5 Endogenous Connections

In this section we extend our model by assuming that networking is costly. Let us say that strategies are non-contagious when only the sellers who have produced low-quality are punished. When either inviting spectators, $N_j^s(\theta_t)$, observing neighboring buyers, $N_j(\theta_t)$, or both are costly to buyer j, the following result holds:

Proposition 5: In any Nash equilibrium of the random matching game with costly networking that uses non-contagious strategies, low quality is produced with positive probability when $M > \delta/[g(1 - \delta)]$.[8]

Proof. The proposition is proved by contradiction. Assume that there is a Nash equilibrium with non-contagious strategies in which every seller produces high quality with probability one in every period. Buyers then do not have any incentive to network and the optimal N_j must be zero for all buyers j. If, however, N_j is zero for all buyers and $M > \delta/[g(1 - \delta)]$ all sellers have incentive to unilaterally deviate

---

[8]A similar proposition could be stated allowing for contagious strategies for $M > \overline{M}(\delta, g, \ell)$, where $\overline{M} < \infty$.

and produce low-quality. Contradiction. ■

There clearly exists the equilibrium where $N_j = 0$ for all buyers and every seller produces low quality. More interestingly, we show that if the costs of networking are small enough, there exist sequential equilibria with strictly positive probability of trade.

Let us concentrate on the case where observing neighboring buyers $n \in N_j(\theta_t)$ is costly to buyer j and where buyers are unable to affect the number of spectators to their game. This assumption corresponds to the idea that buyers network to gather information about their trading environment. An alternative - that leads to similar results - would be that buyers invited other buyers to their games in an attempt to obtain information regarding their current opponents. Clearly the second alternative would make sense only under the *Informative Strategy Profile*. More specifically, let us extend our basic model by assuming that in period zero, before θ_0 is realized, all buyers j can invest in N_j connections that allow them to observe the games in N_j consecutive locations to their own in every future period. To obtain N_j connections j must pay $N_j c$, where $c > 0$, in period zero.

With these assumptions, whenever the costs of networking are less than some threshold value $\bar{c}(M)$, we can find sequential equilibria with slightly modified *Informative and Uninformative Strategy Profiles* such that the sellers initially randomize between high and low quality and produce that level of quality in the future. A positive probability of low quality goods is necessary to provide buyers with an incentive to network. This probability tends to increase with M. When the costs of networking exceed the threshold value $\bar{c}(M)$, trade collapses because the probability of low quality goods that would provide buyers sufficient incentives to network is

so high that buyers are unwilling to buy from unknown sellers. For simplicity, let us confine our analysis to the *Uninformative Strategy Profile*.

Consider the following *modified Uninformative Strategies*:

For the seller i, in all periods $t = 0, 1, 2....$, after θ_t is realized:

(I) In the first period play H with probability $1 - p$ and L with probability p. After that, if the outcome in all the trade games where seller i played was (H, B), play H.

(II) Play L otherwise.

For the buyer j:

(III) In period 0, before θ_0 is realized, invest in $N^* - 1$ connections with probability r and in N^* connections with probability $1 - r$,

and in all periods $t = 0, 1, 2....$, after θ_t is realized:

(IV) Signal randomly γ or β with equal probabilities to all $n \in N_j(\theta_t)$, irrespective of the private history $h_t(j)$.

(V) If j has ever observed $\theta_t^{-1}(j)$ play L or someone (including himself) play NB against her, play NB regardless of the messages $m_t(j)$.

(VI) Play B otherwise.

Before proceeding, we need some additional notation. If buyer j invested in N_j connections, the probability that he observed $\theta_t^{-1}(j)$ at period t', where $t' < t$, is $(N_j + 1)/M$. Let us denote this probability by $b^e(N_j)$.[9] Then the probability that buyer j has never observed $\theta_t^{-1}(j)$ until period t is simply $(1 - b^e(N_j))^t$.

---

[9]More precisely, $b^e(N_j) = \Pr[\theta_t^{-1}(j) \in S_j(\theta_{t'}) \cup \{\theta_{t'}^{-1}(j)\}]$

32

Proposition 6: The *modified Uninformative Strategy Profile* is a sequential equilibrium of the above random matching game with

$$p = cM \left(\frac{\left(1 - \delta(1 - b^e(N^* - 1))\right)\left(1 - \delta(1 - b^e(N^*))\right)}{\ell \delta} \right),$$

where

$$c \leq \bar{c} \equiv \frac{1}{M} \left(\frac{\ell \delta \left(M(1 - \delta) + \delta\right)}{\left(1 - \delta(1 - b^e(N^* - 1))\right)\left(1 - \delta(1 - b^e(N^*))\right)\left((\ell + 1)M(1 - \delta) + \delta\right)} \right).$$

Proof. To have a sequential equilibrium in this extended game for our strategies, the following three equations must hold:

$$\frac{1}{1 - \delta} = \frac{r(1 + g)}{1 - (1 - b^e(N^* - 1))\delta} + \frac{(1 - r)(1 + g)}{1 - (1 - b^e(N^*))\delta} \tag{1.16}$$

$$N^*, N^* - 1 \in \arg\max_{N_j} \frac{1 - p}{1 - \delta} - \frac{p\ell}{1 - \delta(1 - b^e(N_j))} - N_j c, \tag{1.17}$$

$$(1 - p) + \frac{\delta(1 - p)}{M(1 - \delta)} \geq p\ell. \tag{1.18}$$

A seller is willing to randomize in the initial period between providing high quality goods forever and providing low quality goods forever if and only if equation (5.1) holds. The left-hand side of (5.1) is the payoff from providing high quality goods forever and the right-hand side is the expected payoff from providing low quality goods forever. To see this, note that the probability that $\theta_t(i)$ has never observed i is just $(1 - b^e(N_{\theta_t(i)}))^t$. Given this indifference in the initial period, the seller who once provided low quality good can be shown to keep on providing low

33

quality goods; the proof is exactly the same as that of proposition 3. So (I) and (II) are established.

Equation (5.2) requires that buyers are willing to randomize between N^* and $N^* - 1$ connections. We can show that the right hand side of equation (5.2) has a single peak if N_j is treated as a positive real number. Hence if the expected payoff to buyer j is the same with N^* and $N^* - 1$, then N^* and $N^* - 1$ both solve j's maximization problem. Equation (5.2) is therefore satisfied if

$$\frac{p\ell}{(1 - \delta(1 - b^e(N^*)))} = \frac{p\ell}{(1 - \delta(1 - b^e(N^* - 1)))} + c,$$

or

$$p = cM \left(\frac{(1 - \delta(1 - b^e(N^* - 1))) (1 - \delta(1 - b^e(N^*)))}{\ell\delta} \right) \qquad (1.19)$$

Given that other buyer's actions do not depend on the messages, (IV) is obvious. If a buyer has observed his current opponent play L or someone play NB against her, he should play NB against her given (II). And if a buyer has observed his current opponent and the outcomes have always been (H, B), he should believe she will play H and he should play B. If the buyer has never observed his current opponent, he should believe that she will play H with probability $(1 - p)$. For a buyer to play B against her rather than give her a bad status by playing NB, equation (5.3) has to be satisfied. This equation can be rewritten as

$$p \leq \frac{M(1 - \delta) + \delta}{(\ell + 1)M(1 - \delta) + \delta}. \qquad (1.20)$$

Therefore the *modified Uninformative Strategy Profile* with p defined by equation (5.4) is a sequential equilibrium if p satisfies equation (5.5). This happens when $c \leq \bar{c}.$ ∎

34

Several interesting results now follow: First, for $c > 0$ equation (5.4) requires that p is strictly positive as was shown in proposition 5. L has to be played with positive probability to provide buyers with the incentive to network. Secondly, this probability is increasing in M. Under the *Uninformative Strategy Profile* approximately proportionally and with *Informative Strategy Profile* (it can be shown) less than proportionally. More striking result is the knife edge property of our equilibria: If even one more buyer invested in one more connection there would be no low quality at all (increasing the utility of all buyers and sellers discontinuously). But networking to reduce production of low quality is a public good and, as usual, everyone wants to free ride in its production. Because of this, the economy is stuck in an inefficient equilibrium.

Another interesting observation is that when equation (5.5) fails, trade collapses even when it might be beneficial for buyers to keep trading. This occurs because a buyer, who considers whether to trade with an unknown seller or to give her a bad status by playing NB, does not take into account the future trading opportunities of other buyers with her. With the *Informative Strategy Profile* there would still be another externality because of the informative signalling to neighbors. When choosing the number of locations to observe the buyers would not take into account the learning by their neighbors, but would only be interested in their own learning.

1.6 Conclusion

In many real life situations particular sellers and buyers trade with each other infrequently, or only once. In such instances, when one or both parties have short-term incentives to cheat, community enforcement may be needed to facilitate cooperation and trade. This paper studied community enforcement in the absence of institutions to transmit information.

We studied a large population, random matching game between buyers and sellers, where the sellers have a short-term incentive to cheat and supply low quality. We studied informal networks of communication as the mechanism that spreads information about sellers' behavior and facilitates trade. We looked at both informative and uninformative signalling and for the latter we showed that high quality can be sold in a sequential equilibrium with population M where each buyer networks with only $N^*(M)$ players with $\lim_{M \to \infty} N^*/M = 0$.

We studied the case of costly networking and showed that in this case, when M is large, low-quality goods must be supplied with positive probability in any equilibrium to provide buyers with an incentive to network. When the costs of networking were below a threshold value, we found a sequential equilibrium in which sellers initially randomize between high and low quality with probabilities $(1 - p)$ and p respectively and then continue to produce high quality if and only if they did so in the first period. In this equilibrium p is strictly positive and increasing in both M and the costs of networking.

Chapter 2

Imperfect Information Repeated Games with a Single Perfect Observer

2.1 Introduction

A central result in the theory of repeated games is that non-Nash equilibrium outcomes of the stage game are consistent with equilibrium play of the repeated game. The key in proving this result is the construction of punishments that are to be imposed upon deviators to offset short run gains. The crucial assumption in the construction is that each player observes the other players' actions, so that a deviation is identified by all the players.

The assumption of perfect observability seems rather demanding in many economic situations, especially involving a large number of players. For example, consider a joint-project with a large number of participants. To achieve the goal

of the project, each member has to exert a certain level of individual effort. If the project, because of its character, requires members to work in different locations or in specific fields with which other members are not familiar, the informational assumption that each member does not monitor all the other members' effort levels is quite natural. Each member may observe only the effort levels of the members with whom he works closely. Also, in models of social norms in which in each period each player is involved in a trade with an opponent who is randomly selected among a large population of players, it is reasonable to assume that players observe only the outcomes of their own matches and possibly their neighbors' matches.

Imperfect observability causes a nontrivial problem if we are to support an equilibrium that specifies players to play a non-Nash equilibrium of the stage game. If some player's deviation is not observed by all the players, such an equilibrium requires at least one observer, if any, of the deviation to choose an action different from the action to be played on the equilibrium path. While it depends on the structure of the stage game payoffs whether the required behavior of observers is by itself an effective punishment against the deviation, the behavior acts as a signal to other players who will in turn punish the deviator. However, because of the informational constraint, the punishing or signalling behavior may be regarded as a deviation itself by other players who are not aware of the initial deviation. As a result, the punisher or signaller might be punished as well. Then an observer of a deviation might be better off by continuing to play the action to specified on the equilibrium path rather than initiate punishments (or to delay punishments). The difficulty is that incentives for a player must simultaneously ensure that she initiate punishments when she observes some other player's deviation, while not doing so when there are no deviations. This is a major concern of Kandori (1992) and Ellison

(1994) when they construct the "contagious" strategy profile to support cooperation in a random matching model of the prisoners' dilemma.[1] Ahn and Suominen (1996) also face a similar problem when they study a random matching version of buyer-seller game.

To investigate the problem more closely, we consider particular types of imperfect observability. We assume that only one player, denoted by player 0, observes the actions of all the other players. The other players only observe the perfect observer's action and possibly some other players' actions.[2] For instance, in the joint-project example, the existence of the perfect observer would mean that there is a project leader among the members who monitors the other members' effort levels and whose effort level is also monitored by the other members. Since the leader's effort level is perfectly monitored by the other members, her deviation from the required effort level can be easily prevented (if she is sufficiently patient). On the other hand, to prevent a deviation of the other members, if one occurs, the leader must choose a different effort from the required effort either to punish the deviator or signal the deviation to the other members, who can then punish the deviator. The question is : what makes the leader signal a deviation to trigger punishments despite the potential losses from the ensuing punishments after the signal?

Given the information structure specified above, our objective is to find a sequential equilibrium of the repeated game in which any action profile that strictly Pareto-dominates a stage game Nash equilibrium is played in every period. As a first step, we consider a simple case where each player only observes player 0's action at

---

[1]The success of the "contagious" strategy profile depends on the specific payoff structure of the prisoners' dilemma in which defection is a dominant strategy.

[2]After the first draft of this paper is written, the author is informed of the work by Verboben (1994), which studies a similar problem in the context of a triopoly model. In terms of the payoff structures and the information structure, the setting of this paper includes the triopoly model as a special case.

the end of each period and provide an equilibrium by imposing a restriction on the stage game payoffs. The restriction is that each player i has an action m_i such that player 0's stage game payoff from the action profile in which player i plays m_i and the other players (including player 0) play the actions specified on the equilibrium path is lower than her stage game payoff along the equilibrium path. Essentially, player i can independently punish player 0. Without the restriction, player 0 can guarantee herself at least the equilibrium payoff even if she does not signal at all (by playing the action to be played on the equilibrium path) after observing a deviation. The equilibrium we construct requires player i to play m_i after his deviation until player 0 signals. By doing so, player i can effectively punish player 0 if she does not signal after his deviation. We examine more general information structures under which each player can possibly observe some players other than player 0 and show the same result with stronger restrictions on the stage game payoffs. The equilibria constructed in these cases generate the same outcome paths as the equilibrium in the simple case except when simultaneous deviations occur.

We also study finitely repeated games without discounting. We assume the stage game has at least two Pareto-ranked Nash equilibrium. We then show that under a certain restriction on the stage game payoffs, there is a sequential equilibrium in which any action profile that strictly Pareto-dominates the Pareto-inferior Nash equilibrium is played in every period except for the last few periods of Nash equilibrium phase. In the Nash equilibrium phase along the equilibrium path, the Pareto-superior Nash equilibrium is played as rewards. The restriction on the stage game payoffs is as follows.: For each player i, consider all the Nash equilibria of the game played by player i, player 0 and other possible player i's observers, fixing the other players' plays as the ones specified by the Pareto-superior Nash equilib-

rium. The restriction is that among all the Nash equilibria of the modified game, there exists a Nash equilibrium such that (i) player i's stage game payoff when player i, player 0 and the player i's observers play the Nash equilibrium and the other players play the Pareto-superior equilibrium is lower than his payoff from the Pareto-superior equilibrium and (ii) player 0's stage game payoffs between the former profile and the latter are not the same. Given (i), a deviation at the period just before the Nash equilibrium phase is prevented. If player i deviates, the players who observed his deviation will credibly punish him during the Nash equilibrium phase. The second restriction makes it possible either to punish player 0 or not to reward her in case she does not properly signal a deviation that occurred during the late periods near the Nash equilibrium phase. For a deviation in the early periods, the equilibrium induces player 0 to signal by making her indifferent between signalling the deviation and delaying it. Here we use no discounting assumption, which replaces the restrictions for the infinite repeated games.

In recent years, repeated games without perfect observability have been the focus of a number of papers. Most successful works among thses are on repeated games with imperfect public information. They analyze games where each player does not observe other players' actions, but at the end of each period observes a public outcome which depends the action profile played in that period. Fudenberg, Levine and Maskin (1994) consider public equilibria in which the players base their actions only on the public outcomes and give sufficient conditions on the informativeness of the public outcomes to obtain the folk theorem with the equilibria. In our framework, since each player observes player 0's action at the end of each period, some type of public information is also available. The difference is that a priori, the public information is not informative at all. Our task is how to endogenize the

public information to be informative.

For the literatures on games without public information, we would like to mention Fudenberg and Levine (1991) and Ben-Porath and Kahneman (1996). In the setting where players only observe private signals at the end of each period, Fudenberg and Levine provides a partial folk theorem using an epsilon-sequential equilibrium. However, the solution concept allows nonoptimal behaviors at the off-equilibrium paths. For instance, after a deviation occurs, the observers of the deviation initiate punishments even though it is not optimal as long as the punishments last finite periods. Ben-Porath and Kahneman set up a model where a player only observes some other players in each period but there is a public announcement stage after the period. In the announcement stage, each player announces who deviated among his neighbors and he also observes all the other players' announcements. In this case, if a mechanism where the players do not lie during the stage is available, the players can identify a deviator and coordinate the punishments against him. They provide such a mechanism for the case where each player is observed by at least two other players and show the folk theorem. In our model, there is no announcement stage. Players communicate only by their actions.

The rest of the paper is organized as follows. The model is introduced in section 2. In section 3, we study examples which illustrate the main difficulties that arise due to the imperfect observability. In section 4 and 5, a Nash folk theorem is introduced for the infinitely repeated games and for finitely repeated games without discounting, respectively. In section 6 we discuss a "robust" equilibrium which is a sequential equilibrium under any information structure among the ones we look at. We conclude the paper in section 7.

2.2 The Model

In the stage game, each player i, $i = 0, 1, ..., n$, simultaneously chooses an action a_i from a finite set A_i. We denote the set of players by $N \cup \{0\}$ where $N = \{1, 2, ..., n\}$. Let $u_i(a_0, a_1, ..., a_n)$ be player i's stage game payoff from the action profile $(a_0, a_1, ..., a_n)$.

Consider a finite or an infinite repetition of the stage game. As for the information structures of the repeated game, we assume that each player observes the actions of only a subset (not necessarily proper) of the players at the end of each period. Specifically, at the end of each period, while player 0 observes the actions of all the players, player i, $i \in N$, observes player 0's and possibly some other players' actions besides his own. We are therefore interested in certain types of imperfect observability that we call the case of *player 0 being a perfect observer*. We do not, however, exclude the possibility that there are other perfect observers who are observed by all the other players and also observe all the other players. For $i \in N$, denote by N_i the set of players in N who are observed by player i. These are player i's *neighbors*. Similarly, denote by S_i the set of players in N who observes player i. We call them player i's *spectators*. We assume the information structures are fixed across the periods.

It is convenient to describe the information structures in terms of graphs. Envisage the players as $n + 1$ nodes and the observability from player i to player j as an arrow directed from node i to node j on a graph. Then the information structures we are looking at are depicted as the graphs with at least one node (node 0) connected by n undirected arrows to the other n nodes in N.[3,4] We call them

---

[3] An arrow between two nodes is undirected if it is directed from each node to the other.

[4] For $i \in N$, N_i is then the set of nodes in N to which the arrows from node i are directed and S_i is the set of nodes in N from which the arrows are directed to node i.

$P - 0$ *graphs*. Two extreme kinds of $P - 0$ *graphs* are the *complete graph* where each node is connected to the other n nodes and the *star graph* where each node is connected to only node 0 by an undirected arrow. The *complete graph* represents the perfect observability case (where for all $i \in N$, $N_i = S_i = N - \{i\}$), while the *star graph* describes the other extreme case where $N_i = S_i = \emptyset$ for all $i \in N$. We also denote by the *symmetric graphs* the information structures where for all $i \in N$, $N_i = S_i$.

We denote by $H_i(t)$ the set of histories available at period t to player i, $i \in N \cup \{0\}$, by letting:[5]

for $t = 2, 3, ..., T \leq \infty$.

$$H_0(t) = (\times_{i \in N \cup \{0\}} A_i)^{t-1} \quad \text{and} \quad \forall i \in N, \ H_i(t) = (\times_{j \in N_i \cup \{0, i\}} A_j)^{t-1}.$$

$$H_i(1) = \emptyset, \ \forall i \in N \cup \{0\}.$$

We denote a typical element in $H_i(t)$ by $h_i(t)$. Note $h_i(t)$ is one of the player i's information sets which can be reached at period t.

A pure strategy for player i, σ_i, is a sequence of maps $\sigma_i(t)$ where

$$\sigma_i(t) : H_i(t) \to A_i.$$

A behavior strategy for player i is a sequence of maps from $H_i(t)$ to the set of the probability distribution over A_i.

As the repeated game payoff, we use the average undiscounted sum of the stage game payoffs for the finitely repeated game and the average discounted sum of the stage game payoffs for the infinitely repeated game. In other words, if $(a(t))$

---

[5]We assume players do not observe the realized stage game payoffs after each stage game.

is the action profile played by all the players at period t, player i's repeated game payoff is given by

$$(1/T) \sum_{t=1}^{T} u_i(\mathbf{a}(t)) \quad \text{for } T < \infty,$$

$$(1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} u_i(\mathbf{a}(t)) \quad \text{for some } \delta \text{ with } 0 < \delta < 1.$$

We apply Kreps and Wilson (1982)'s *sequential equilibrium* as a solution concept for the finitely repeated game. A strategy profile $\sigma = (\sigma_i)_{i \in N \cup \{0\}}$ is a sequential equilibrium if (i) for each $i \in N \cup \{0\}$ and each of his information sets $h_i(t)$, player i has a belief $\mu_i(h_i(t))$ about which node he is at such that σ_i is optimal given the other players' strategies and the belief and (ii) the beliefs system $\mu = (\mu_i)_{i \in N \cup \{0\}}$ is consistent with σ. Remember a beliefs system μ is consistent with a strategy profile σ if there is a sequence $\{(\sigma^v, \mu^v)\}_{v=1}^{\infty}$ which converges to (σ, μ) in Euclidean space where each σ^v is totally mixed and each μ^v is derived from σ^v by Bayes' rule.[6]

While Kreps and Wilson (1982) explicitly defines sequential equilibrium only for finite games, the definition can be extended for the infinitely repeated games in a straightforward way, as suggested by Fudenberg and Levine (1994). The only change from Kreps and Wilson in the definition of sequential equilibrium is for the consistency of a beliefs system: A beliefs system μ is consistent with a strategy profile σ if there is a sequence $\{(\sigma^v, \mu^v)\}_{n=1}^{\infty}$ which converges to (σ, μ) at every information set with each σ^v being totally mixed and each μ^v being derived from σ^v by Bayes' rule.

Since the stage game has at least one Nash equilibrium, the repeated game always has a sequential equilibrium where the Nash equilibrium is played in every

---

[6]A (behavior) strategy profile is totally mixed if it puts a positive probability on every action at every information set.

45

period. Our interests are however beyond the trivial equilibrium. In particular, we try to provide more efficient equilibria supporting an outcome that strictly Pareto-dominates the stage game Nash equilibrium.

2.3 Examples and Discussion

2.3.1 Finitely repeated games

Consider the following three-person stage game, where player 0 chooses a row, player 1 chooses a column, and player 2 chooses a matrix:

		1 (I)	1 (N)				1 (I)	1 (N)
0	I	30,30,20	0,0,-5		0	I	14,14,21	0,0,0
	N	0,0,-5	10,10,10			N	0,0,0	$\sqrt{3}, \sqrt{3}, 15$

$$H \qquad\qquad L$$

$$2$$

The information structure of the repeated game is the *star graph*: At the end of each period, player 0 observes the actions that player 1 and player 2 chose and player 1 and player 2 also observe player 0's action. However, player 1 and player 2 cannot observe each other.

Suppose now the stage game is played for two periods. In the perfect observability case, $\{(I, I, H), (I, I, L)\}$ is an outcome path supported by a subgame perfect equilibrium. The reversion to the less favorable Nash equilibrium profile, (N, N, L), prevents player 2 from deviating to the myopic best response, L, in the first period.

However, the outcome path is not attainable by a sequential equilibrium under the *star graph*. There is no credible threat to player 2 if he deviates in the first period. After player 2's deviation, the only sequentially rational outcome is for player 0 and player 2 still to play I and L given that player 1, who does not know the deviation, will play I. The point is that fixing player 3's action to I, there is no other Nash equilibrium of the modified stage game between player 0 and player 2 than (I, L). Further, for the three-periods repeated game, we can make the following claim.

Claim 1. Under the *star graph*, if the stage game is played for three periods, there is no pure strategy sequential equilibrium in which player 2 ever plays H.

Proof. Consider a sequential equilibrium σ^*. In the 3rd period, player 2 plays L regardless of his private history $h_2(3)$, since L is the dominant action of the stage game and this is the last period. Furthermore, in each period, player 0 knows player 1's private history $h_1(t)$ and so his action prescribed by σ^*, $\sigma_1^*(h_1(t))$. So, in the last period, given her history $h_0(3)$ (which includes player 1's history $h_1(3)$), she will play the best response to $\sigma_1^*(h_1(3))$ and L. This implies that there is no player 2's history for which player 2 plays H in the second period under σ^*. It is because neither player 1, who does not know player 2's second period's action, nor player 0, who will play the best response in the last period, change their actions in the last period after player 2 deviated to L in the second period.

Suppose player 2 plays H in the first period under σ^*. Let $(a_0^*(t), a_1^*(t), a_2^*(t))_{t=1}^3$ be the equilibrium path of σ^*. Notice that $a_2^*(1) = H$, $a_2^*(2) = a_2^*(3) = L$ and $(a_0^*(3), a_1^*(3), a_2^*(3))$ is a Nash equilibrium of the stage game. Furthermore, especially, in the second period after $(a_0^*(1), a_1^*(1), a_2^*(1))$ is played, player 0 should not have an incentive to deviate to an action $\hat{a}_0(2)$, different from $a_0^*(2)$. In other words,

47

if we let $(\hat{a}_0(3), \hat{a}_1(3), L)$ is the action profile to be prescribed by σ^* after player 0 plays $\hat{a}_0(2)$ in the second period, the following inequality has to be satisfied:

$$\sum_{t=2}^{3} u_0(a_0^*(t), a_1^*(t), L) \geq u_0(\hat{a}_0(2), a_1^*(2), L) + u_0(\hat{a}_0(3), \hat{a}_1(3), L). \qquad (1)$$

Notice that $(\hat{a}_0(3), \hat{a}_1(3), L)$ is a Nash equilibrium of the stage game.

On the other hand, to deter player 2's deviation to L in the first period, player 0 has to play an action different from $a_0^*(2)$, $\hat{a}_0(2)$, in the second period if he does. If player 0 plays $a_0^*(2)$, the second period and the third period outcome is $(a_0^*(t), a_1^*(t), L)_{t=2}^{3}$, making player 2 better off by playing L in the first period. This contradicts the fact that σ^* is an equilibrium, so we require:

$$\sum_{t=2}^{3} u_0(a_0^*(t), a_1^*(t), L) \leq u_0(\hat{a}_0(2), a_1^*(2), L) + u_0(\hat{a}_0(3), \hat{a}_1(3), L). \qquad (2)$$

Hence, in order for σ^* to be a sequential equilibrium, both inequalities (1) and (2) must hold and so are satisfied with equality. A careful inspection shows this is impossible: If $(\hat{a}_0(3), \hat{a}_1(3), L)$ is identical to $(a_0^*(3), a_1^*(3), L)$, the inequality (in (1) and (2)) cannot be satisfied by equality since $u_0(a_0^*(2), a_1^*(2), L) \neq u_0(\hat{a}_0(2), a_1^*(2), L)$ given $a_0^*(2) \neq \hat{a}_0(2)$. Suppose now $(a_0^*(3), a_1^*(3), L)$ and $(\hat{a}_0(3), \hat{a}_1(3), L)$ are the two distinct Nash equilibria. To obtain the equality in (1) and (2), the absolute value of the difference between $u_0(a_0^*(2), a_1^*(2), L)$ and $u_0(\hat{a}_0(2), a_1^*(2), L)$ has to be exactly $14 - \sqrt{3}$, which is impossible given the payoffs. $\blacksquare$

The problem demonstrated in the three-periods repeated game does not disappear in longer, but still finite, repeated game. By extending the previous argument along the lines of the proof of Proposition 6.2, we can show that given any

T, there is no pure strategy sequential equilibrium of T-period repeated game in which player 2 ever plays H.

If we allow mixed strategies, the result is no longer true. For example, when $T = 2$, (I, I, H) can be played in the first period, after which the mixed strategy Nash equilibrium is played unless player 2 played L in the first period.[7] In that case she plays I instead of randomizing, making player 2 strictly worse off. This threat is credible since player 0 is indifferent between playing I and playing other strategies given that player 1, without knowing player 2's deviation, plays the mixed strategy Nash equilibrium.

The negative result of the previous example also heavily depends on the payoff of the stage game, especially in that L is the strictly dominant strategy of the stage game for player 2. If we let $u_2(N, N, H) = 15$, $u_2(N, N, L) = 10$ instead of 10 and 15, respectively, without changing any other payoffs, there is a sequential equilibrium for sufficiently long periods of repetition where player 2 does not always play L. If $T = 4$, for example, $\{(I, I, H), (I, N, L), (I, I, L), (I, I, L)\}$ is an outcome path of the following sequential equilibrium; If player 2 played L in the first period, player 0 plays N and player 2 H in the second period. If player 0 played N in the first or the second period, player 2 plays H and player 0 and 1 play N everafter. If player 0 played I in the first two periods, player 2 plays L in the last two periods. If player 0 played I in the first two periods and player 1 played I in the second period, player 0 and 1 play N in the last two periods. Other than those cases, each player plays the action prescribed in the outcome path. For this modified

---

[7]It is interesting to note that this sequential equilibrium path is not supported by a subgame perfect equilibrium in perfect observability case. Under perfect observability, one of the stage game Nash equilibria has to be played in the second period. However, in the example, the mixed strategy gives player 2 the lowest payoff among the stage game Nash equilibria, and then he is better off by deviating to L in the first period.

payoff, furthermore, we can show there is a $\underline{T}$ such that for all $T \geq \underline{T}$, the T period repetition of the stage game has a sequential equilibrium where (I, I, H) is played for the first $T - \underline{T}$ periods. In section 5, we will discuss sufficient conditions for a Nash-threat folk theorem in more general payoff structures of the stage games under $P - 0$ *graphs*.

2.3.2 Infinitely repeated games

If the stage game in the previous subsection is infinitely repeated, there is a sequential equilibrium in which (I, I, H) is played in every period. However, the infinite repetition itself cannot get rid of the problem arises due to the imperfect observability. To see this, let us consider the following stage game. This stage game has a unique Nash equilibrium (N, N, L) and a unique individually rational and efficient outcome (I, I, H).

		1	
		I	N
0	I	$30, 30, 30$	$0, 0, 0$
	N	$0, 0, 0$	$10, 10, 10$

H

		1	
		I	N
0	I	$40, -10, 40$	$0, 0, 0$
	N	$0, 0, 0$	$15, 15, 15$

L

2

In the infinite repetition of the stage game with perfect observability, (I, I, H) can be played in every period by using the grim trigger strategy; player 0 and player 1 play I and player 2 plays H unless player 2 played L before, in which case the players play (N, N, L) everafter. Under the *star graph*, however, there is no

sequential equilibrium in which (I, I, H) is played in every period.[8] In order to deter player 2's deviation to L, his average continuation payoff after the deviation should be strictly less than 30. This requires player 0 to play N with positive probability at some period after player 2's deviation. If she does not, player 2's continuation payoff would be at least 30 since player 1 keeps on playing I. On the other hand, to induce player 0 to do so, her continuation payoff of playing N after player 2's deviation has to be at least 30 since she can guarantee 30 by keeping on playing I. Apparently, it is impossible to make her continuation payoff more than or equal to 30 with holding player 2's continuation payoff less than 30.

Claim 2. Under the *star graph*, there is no sequential equilibrium in which (I, I, H) is played in every period.

Proof. Suppose π^* is a sequential equilibrium where (I, I, H) is played in every period. Now suppose that player 2 for the first time deviated to L at period τ. Let $\alpha_t(i, j, k), 1 \leq i, j, k \leq 2$ be the probability, specified by π^*, that player 0 plays the i'th action, player 1 the j'th action and player 2 the k'th action at period $t \geq \tau + 1$ if player 3 deviated to L at period τ. Let $\alpha(i, j, k) = (1 - \delta) \sum_{t=\tau+1}^{\infty} \delta^{t-\tau-1} \alpha_t(i, j, k)$ be the average discounted probability of $\alpha_t(i, j, k)$.

Since player 2's continuation payoff after his deviation should be less than 30, we can write

$$30\alpha(1, 1, 1) + 10\alpha(2, 2, 1) + 40\alpha(1, 1, 2) + 15\alpha(2, 2, 2) < 30,$$

$$\text{where } \alpha(2, 1, 1) + \alpha(2, 1, 2) > 0.$$

On the other hand, to induce player 0 to play N with positive probability, her

---

[8]We note this negative result depends on the fact that $u_0(I, I, L) \leq u_2(I, I, L)$. For example, if we let $u_0(I, I, L) = 100$ and keep the other payoffs unchanged, there is a sequential equilibrium in which (I, I, H) is played in every period

continuation payoff of playing N after player 2's deviation has to be at least 30. In other words,

$$30\alpha(1,1,1) + 10\alpha(2,2,1) + 40\alpha(1,1,2) + 15\alpha(2,2,2) \geq 30,$$

$$\text{where } \alpha(2,1,1) + \alpha(2,1,2) > 0.$$

So, we have a contradiction. ∎

If $u_0(I,I,L) < 30$, we can find a sequential equilibrium where (I,I,H) is played in every period for a discount factor sufficiently close to 1. The equilibrium strategy for more general stage game payoffs will be introduced in next section.

2.3.3 Discussion

If an equilibrium path specifies a non-Nash equilibrium outcome of the stage game in some periods, then at least one of the players has a myopic incentive to deviate from the path. To prevent his deviation, the equilibrium has to provide a punishment severe enough that his short run gains from the deviation are washed out. Also, the punishment should be credible in the sense that the players are better off by participating in the punishment. If all the players observe the other players's actions, it is easy to specify such punishment.[9] However, if some players do not observe all the other players' actions and so a deviation may not be observed by all the players, we encounter a nontrivial problem. Punishing a deviation might be regarded as a deviation itself by the players who do not know about the initial deviation. As a result, the punisher may be punished. Then she might be better off by keeping on playing the action to be played on the equilibrium path not to initiate punishments (or to delay punishments). To study this problem more

[9]See Fudenberg and Maskin (1986) or Benoit and Krisna (1985).

closely, we consider particular types of imperfect observability that we called $P - 0$ *graphs*. Under these information structures, because her action is observed by all the players, player 0's deviation from the equilibrium path can be easily prevented once the punishments against her deviation are severe enough. The key is then to induce player 0 to initiate punishments after some other player's deviation despite the potential loss from the ensuing punishments. However, as the previous examples illustrate, there are stage games for which it is not possible. In those examples, it is not possible to punish player 0 in case she did not *signal* properly nor to reward her for *signalling* to compensate her loss during the punishments phase. In section 4 and 5, we provide sufficient conditions in terms of the stage game payoffs for a Nash-threat folk theorem.

Imperfect observability, including $P - 0$ *graphs*, causes other complications. Consider a strategy profile $\hat{\sigma}$ that is sequentially rational on the equilibrium path and after any sequence of unilateral deviations from the equilibrium path. Under perfect observability, we can easily construct a subgame perfect equilibrium that supports the outcome path of $\hat{\sigma}$ by ignoring simultaneous deviations.[10] However, under imperfect observability, we cannot simply let the play return to the ongoing path after simultaneous deviations. The players, who knew only a single, possibly different, part of the deviations, would play the actions that a strategy profile prescribes after the single deviation. Then other players who knew more than a single part of the deviations might want take advantage of the situations.[11] So

---

[10]The sequential rationality of the subgame perfect equilibrium on the equilibrium path and after any sequence of unilateral deviations from the equilibrium path is immediate by the construction. Ignoring simultaneous deviations is also sequentially rational given the noncooperative solution concept.

[11]The question is : given a strategy profile that is sequentially rational on the equilibrium path and after any sequence of unilateral deviations from the equilibrium path, does there exist another strategy profile which exhibits the same behaviors after those histories and is also sequentially rational after the other histories? If existence is guaranteed, we can ignore simultaneous deviations

finding a sequential equilibrium here is potentially more complicated than under perfect observability.

For finitely repeated games, we construct an equilibrium which does not face the complication caused by simultaneous deviations. However, since the construction is based on the assumption that players do not discount the future, the equilibrium cannot be applied to the infinitely repeated games with discounting. For the infinitely repeated game, we first consider the *star graph*, a special case of $P - 0$ *graphs* where $N_i = \emptyset$ for all $i \in N$. In the equilibrium we construct in this case, the difficulty after simultaneous deviations is easily resolved as player 0's decision making problem. Here player 0 takes advantage of the beliefs system that each player does not suspect any deviation he does not observe. For more general $P - 0$ *graphs*, however, the strategy profile that generates the same outcome paths as the equilibrium under the *star graph* for all the histories may fail to be an equilibrium. Actions that the strategy profile specifies after some simultaneous deviations may not be optimal under $P - 0$ *graphs* although they are so after the other histories. We take two approaches to deal with this problem. The first one is to restrict the stage game payoffs so that the strategy profile is an equilibrium under $P - 0$ *graphs*. Secondly, we consider symmetric stage games under the *symmetric graphs*, $P - 0$ *graphs* where all $i \in N$, $N_i = S_i$. Using the symmetry, we can construct a strategy profile that exhibits sequentially rational behaviors after simultaneous deviations, while generating the same behaviors as the equilibrium under the *star graph* except after simultaneous deviations.

---

without loss of generality and have only to find a strategy profile which is sequentially rational on the equilibrium path and after any sequence of unilateral deviations from the equilibrium path as we could under perfect observability. Unfortunately, we are not able to provide the answer to this existence problem.

2.4 Infinitely Repeated Games

2.4.1 A Nash-threat folk theorem under the star graph

Our objective, throughout this section, is to find a sequential equilibrium in which any action profile $a^* = (a_i^*)_{i \in N \cup \{0\}}$ that strictly Pareto dominates a stage game Nash equilibrium is played in each period. As an important benchmark, we first consider the *star graph*. Before stating the proposition in this case, we impose the following restriction on the action profile a^*.

(IP) For each player $i \in N$, there exists an action $m_i \in A_i$ s.t.

$$u_0(\ a^*) > u_0(a_0^*, m_i, (a_j^*)_{j \in N - \{i\}}).$$

We call the restriction "Independent Punishments" or (IP) because under the restriction, player i can independently punish player 0. In the equilibria we construct in this section, player 0 is punished by player i through m_i if she does not *signal* after player i's deviation from a_i^*. Without (IP), there is no way he can punish player 0 for not *signalling* since she can guarantee herself at least the equilibrium payoff, $u_0(\ a^*)$, by keeping on to play a_0^*.[12]

Proposition 4.1. *Consider an action profile* $a^* = (a_i^*)_{i \in N \cup \{0\}}$ *that strictly Pareto dominates a stage game Nash equilibrium* f.[13] *Suppose also that* a^* *satisfies (IP). Then under the star graph, there exists* $\bar{\delta} < 1$ *such that* a^* *is played in every period as a sequential equilibrium outcome for* $\delta \in (\bar{\delta}, 1)$.

The equilibrium which we call the *modified finite periods Nash reversion* strategy profile- hereafter M.F.N.R.- is the *star graph* version of the finite periods

---

[12](IP) is not a necessary condition as footnote 6 in section 3.2 suggests.
[13]We do not exclude possibility that the stage game Nash equilibrium is a mixed strategy.

Nash reversion strategy profile under the *complete graph*. In the finite periods Nash reversion strategy profile, a deviation from a* is immediately followed by punishment phase where the stage game Nash equilibrium f is played for finite periods, after which the play returns to a*. The punishment is not player-specific in the sense that the punishment following a deviation is the same regardless of the identity of the deviator. M.F.N.R. has a similar structure. However, it has one more step between the play on the equilibrium path and the punishment phase if a player other than player 0 deviates from a*. In this case, player 0 *signals* by playing a predetermined action different from a_0^*, say s_0, to trigger the punishment phase.

More precisely, M.F.N.R. is as follows. It has four phases, normal phase, signalling phase, finite punishment phase and the infinite punishment phase. The play starts from normal phase where a* is played. If player 0 plays s_0, an action different from a_0^*, finite punishment phase follows where the stage game Nash equilibrium f is played for K periods. K is an integer satisfying $2z < Ky$ where

$$z = \max_{i \in \{0\} \cup N} \max_{a,b} u_i(a) - u_i(b) \text{ and}$$

$$y = \min_{i \in \{0\} \cup N} \min_{a,b} u_i(a) - u_i(b) \ s.t. \ u_i(a) - u_i(b) > 0.$$

If she played an action different from a_0^* or s_0, the punishment will continue forever. On the other hand, if some player(s) other than player 0 deviated from normal phase, his (or their) signalling phase starts from the next period of the deviation(s). In player i's, $i \in N$, signalling phase, player i is required to play m_i. Player i's signalling phase ends if either player 0 plays an action different from a_0^* or player 0 and he simultaneously play a_0^* and a_i^*. After the former case the finite or the infinite punishment phase follows while normal phase resumes in the latter case.

56

After the final period of finite punishment phase, the play returns to normal phase unless some player(s) deviated at the end of the last signalling phase.[14] In this case his (or their) signalling phase starts. Now let us prescribe player 0's action in signalling phase. In a single player's signalling phase, player 0 *signals* by playing s_0. To see player 0's action in multiplayers' signalling phase, let Q be the set of players in N who are in their signalling phase. If $u_0(\mathbf{a}^*)$ is strictly larger than $u_0(a_0^*, (m_i)_{i \in Q}, (a_j^*)_{j \in N-Q})$, player 0 plays s_0 to trigger finite punishment phase. If $u_0(a_0^*, (m_i)_{i \in Q}, (a_j^*)_{j \in N-Q})$ is strictly larger than or equal to $u_0(\mathbf{a}^*)$, on the other hand, she plays a_0^*.

The beliefs system underlying M.F.N.R. is that player i never believes the other players in N (i.e., who he does not observe) ever deviated before. So if he observes player 0 play an action other than a_0^* in normal phase, he believes it is her own deviation, not *signalling* other player's deviation. Another implication of this beliefs system is that player i never suspects other deviation(s) if he has been playing m_i, but player 0 has unexpectedly continued to play a_0^*. He believes that player 0's not *signalling* is solely due to her own mistake. Checking the consistency of this beliefs system is straightforward.[15]

No player i, $i \in N$, has the incentive to deviate from normal phase or his signalling phase since K periods of punishment phase will be triggered and wash out the short run gains from the deviation. Nor does he have in punishment phase

---

[14]Throughout the section, by saying player i deviated at the end of the last signalling phase, we mean two possibilities: The first is when he, in nomal phase, did not play a_i^* at the period when the last punishment phase is triggered (i.e., the last period when player 0 started to play s_0.) The second case is when he did not play m_i at the period the last punishment phase is triggered even if he was in his signalling phase.

[15]For example, consider the following sequence of totally mixed strategy profiles $(\sigma^v)_{v=1}^\infty$: For player 0, for each t and each of her information sets $h_0(t)$, assign a probability $(1/v)^{1/t^2}$ to each action other than the one that player 0 chooses in M.F.N.R.. As for player $i \in N$, assign the probability $(1/v)^n$ at every information set.

where the stage game Nash equilibrium is played. It is also obvious player 0 is also worse off by deviating from normal phase or punishment phase. Now consider player 0's incentive to *signal* in a single player's, say player i's, signalling phase. Since player i plays m_i in his signalling phase and the signalling phase continues as long as player 0 delays *signalling*, player 0's delaying the *signal* will result in $(a_0^*, m_i, (a_j^*)_{j \in N-\{i\}})$ during the delayed periods, followed by the same outcome path that *signalling* without delay would generate. So, the effect from the delay for one period is to get $u_0(a_0^*, m_i, (a_j^*)_{j \in N-\{i\}})$ during the delayed period and to resume normal phase one period later. (IP) guarantees she is better off by *signalling* without delay if the discount factor is close to one.

Suppose now several players played actions other than a^* simultaneously in normal phase. As we discussed before, we cannot simply ignore the simultaneous deviations and let the play to return to normal phase unlike in the perfect observability case. If player 0 continues to play a_0^* instead of *signalling*, the deviators, each of whom believes player 0's not *signalling* is solely due to her mistake without knowing the simultaneous deviations, would keep on playing m. Player 0 then chooses her action by solving a decision making problem: Let the current period be the signalling phase for the players in the set Q. If $u_0(a_0^*, (m_i)_{i \in Q}, (a_j^*)_{j \in N-Q})$ is strictly larger than or equal to $u_0(a^*)$, player 0 finds it better to maintain Q's signalling phase by playing a_0^* and get $u_0(a_0^*, (m_i)_{i \in Q}, (a_j^*)_{j \in N-Q})$ in every future period rather than to *signal* the deviations and to get $u_0(a^*)$ after the punishment phase initiated by her *signalling*. If $u_0(a^*)$ is strictly larger than $u_0(a_0^*, (m_i)_{i \in Q}, (a_j^*)_{j \in N-Q})$, however, it is optimal for player 0 to *signal*. The same argument that shows the optimality of *signalling* in a single player's signalling phase applies.

2.4.2 A Nash-threat folk theorem under $P - 0$ graphs

Let us now extend M.F.N.R. to $P - 0$ *graphs* such that given any previous plays, the extension generates the same outcome paths that M.F.N.R. does. The *extended* M.F.N.R. can be then described as follows.

The *extended* M.F.N.R. has normal phase, signalling phase, finite punishment phase and infinite punishment phase as M.F.N.R..[16] Consider a player i in N. Suppose period t was in normal phase or the signalling phase for some players in $N_i \cup \{i\}$. Suppose player 0 played a_0^* at period t. Let $D_i(t)$ be the set of players in $N_i \cup \{i\}$ who did not play a^* at period t. If $D_i(t)$ is empty, normal phase continues at period $t + 1$. If $D_i(t)$ is not empty, the signalling phase for the players in $D_i(t)$ starts at period $t + 1$. Player i plays m_i at period $t + 1$ if he belongs to $D_i(t)$. If he does not, he plays a_i^*. If player 0 plays a_0^* at period $t + 1$ and the set of players in $N_i \cup \{i\}$ who did not play a^* at period $t + 1$ is $D_i(t+1)$, the signalling phase for the players in $D_i(t + 1)$ starts at period $t + 2$ unless $D_i(t + 1)$ is empty, in which case period $t + 2$ is in normal phase. On the other hand, if player 0 plays s_0 at period $t + 1$, K periods of punishment phase will start from period $t + 2$. To describe the play after the punishment phase, we define $R_i(t+1)$, a subset of $N_i \cup \{i\}$, as follows.

$$R_i(t + 1) = \{j \in N_i \cup \{i\} |\ (i)\ a_j(t + 1) \neq a_j^* \text{ for } j \notin D_i(t) \text{ and}$$

$$(ii)\ a_j(t + 1) \neq m_j \text{ for } j \in D_i(t)\}.$$

Notice that $R_i(t + 1)$ is the set of the players in $N_i \cup \{i\}$ who did not follow the specified actions at period $t + 1$. If $R_i(t + 1)$ is empty, normal phase resumes after the final period of the punishment phase. If it is not empty, the signalling phase for

---

[16] Here, because of the same behaviors the *extended* M.F.N.R. and M.F.N.R. exhibits, we use the same names to the histories which we called nomal, signalling finite punishment and inifinite punishment phase in M.F.N.R..

the players in $R_i(t+1)$ starts after the punishment. The infinite punishment phase is triggered after player 0 played an action other than a_0^* and s_0 at period $t+1$. The specification of player 0's strategy is precisely the same as under M.F.N.R..

The beliefs system under the *extended* M.F.N.R. is also the extension of the beliefs system underlying M.F.N.R.. That is, player i believes player l who is not one of his neighbors has always played the equilibrium action a_l^* except for punishment phases where the stage game Nash equilibrium is played.

The *extended* M.F.N.R. prescribes the optimal behavior to player i in normal phase, his own signalling phase and punishment phase since his deviation in those phases will result in the same outcome paths that it does under M.F.N.R.. Player i is also better off by playing a_i^* when $D_i(t) = \{j\}$, player j's signalling phase. In this case he is expecting that player 0 plays s_0 at period $t+1$ to punish player j . So, if he deviates at period $t+1$, $R_i(t+1)$ will be $\{i\}$ and a new punishment phase against him will start from $t+K+3$. Given the other players' strategies, the optimality of player 0's actions is immediate.

However, some players who observed simultaneous deviations by several players might not find it optimal to follow the actions specified above. Suppose player 0 played a_0^* at period t and $D_i(t) = \{i, j\}$. Suppose also the stage game payoffs are such that $u_0(a_0^*, m_i, m_j, (a_k^*)_{k \in N-\{i,j\}})$ is greater than $u_0(\mathbf{a}^*)$ and $u_i(a_0^*, m_i, m_j, (a_k^*)_{k \in N-\{i,j\}})$ is the worst stage game payoff to player i. Then player i is better off by deviating from the prescribed action m_i since he expects player 0 will continue to play a_0^* as long as he plays m_i, which gives him the lowest continuation payoff. How about the incentive of player k who observed the player i and j's simultaneous deviations? The difficulty in this case is that the problem after simultaneous deviations can no longer be solved as player 0's decision making problem since players other than

player 0 may notice (parts of) the simultaneous deviations.

These complications disappear if we give a stronger restriction on the stage game payoffs as follows. We refer to this restriction as "General Punishments" or simply "GP".

(GP) Given the action profile $(m_j)_{j \in N}$ as defined in (IP), suppose that for all subset $Q \subset N$,

$$u_0(\mathbf{a^*}) > u_0(a_0^*, (a_i^*)_{i \in N-Q}, (m_j)_{j \in Q}) \ .$$

With (GP), player 0 *signals* in any signaling phases. If this is case, player i's deviation in any signalling phases will result in another punishment phase after the punishment phase to be triggered at the next period. So, the *extended* M.F.N.R. is an equilibrium and we have Proposition 4.2.

Proposition 4.2. *Consider an action profile* $\mathbf{a^*} = (a_i^*)_{i \in N \cup \{0\}}$ *that strictly Pareto dominates a stage game Nash equilibrium* f. *Suppose (GP) holds. Then there is* $\delta' < 1$ *such that* $\mathbf{a^*}$ *is played in every period as a sequential equilibrium outcome for* $\delta \in (\delta', 1)$.

Remark 1. In proposition 4.2., the *extended* M.F.N.R. is sequentially rational given any beliefs system. In other words, in each of his information set, it is optimal for player i to follow the extended M.F.N.R. regardless of his belief about which node he is at.

Remark 2. If there are at least two perfect observers, the same result holds without (IP) or (GP). It is obvious to see the following strategy profile is a sequential equilibrium for a high discount factor.: It has also four phases, normal phase, signalling phase, K periods of punishment phase, the infinite punishment

phase. Take two perfect observers, say player 0 and player 1. For each of them, fix an action, $s_0 \neq a_0^*$ and $s_1 \neq a_1^*$, respectively. As in (the *extended*) M.F.N.R., the play starts at normal phase. If any player(s) other than the two plays an action other than a^* in normal or the last signalling phase, his (their) signalling phase begins where the two play s_0 and s_1 and all the other players play a^*. If the two play s_0 and s_1 in this phase, K periods of punishment phase starts after which the play return to normal phase unless some player(s) other than the two deviated in the last signalling phase. In that case, the deviator(s)' signalling phase starts. If the actions of the two perfect observers are different from (a_0^*, a_1^*) or (s_0, s_1), punishment phase continues forever. Here, a perfect observer's incentive to *signal* in signalling phase is given by the threat of the infinite punishment that would follow if she did not *signal.*

2.4.3 A Nash-threat folk theorem for symmetric stage games under the symmetric graphs

Even though (GP) fits many interesting economic situations, we introduce another way to deal with the difficulties after simultaneous deviations without (GP). We assume the stage game with symmetric payoffs and confine the information structure to the *symmetric graphs.*[17]

Proposition 4.3. *Consider a stage game with symmetric payoffs. Let an action profile* $\mathbf{a}^* = (a_i^*)_{i \in N \cup \{0\}}$ *that strictly Pareto dominates a stage game Nash equilibrium* f. *Suppose that (IP) holds and also that for each player* $i \in N$,

---

[17]We say the stage game has symmetric payoffs if $\forall i, j \in N \cup \{0\}$, $\forall a_i \in A_i, \forall a_j \in A_j$ $u_i(a_i, a_j, a_{-i-j}) = u_j(a_i', a_j', a_{-i-j})$ $\forall a_{-i-j} \in \times_{k \neq i,j} A_k$ where $a_i' = a_j$, $a_j' = a_i$.

$u_i(\ \mathbf{a}^*) \neq u_i(a_0^*, a_i^*, (a_j^*)_{j \in N - E_i - \{i\}}, (m_k)_{k \in E_i})$ *for any subset* $E_i \subset N_i$.[18] *Then under the symmetric graphs, there is* $\delta'' < 1$ *such that* $\mathbf{a}^*$ *is played in every period as a sequential equilibrium outcome for* $\delta \in (\delta'', 1)$.

The equilibrium we construct here prescribes the same actions to the players as the *extended* M.F.N.R. with some exceptions. The differences are the behavior of a deviator who observes another deviation(s) and player 0's *signalling* behavior after some simultaneous deviations. In particular, after several players in N simultaneously deviated and each deviator observes the other deviator(s), the equilibrium prescribes player 0 to play a_0^* rather than to *signal*. Deviator i who observes another deviator(s) is required to play a_i^* in stead of playing m_i, expecting the other deviator(s) to play $\mathbf{a}^*$ as well. Player i's expectation makes sense since the other deviator(s) also observed player i's own deviation due to the symmetry of the observability. So, if there occurred simultaneous deviations where each deviator observes the other deviators, the play returns to normal phase.

Proof. See appendices. ∎

2.5 Finitely Repeated Games without Discounting

In this section, we consider a finite repetition of a stage game which has at least two Pareto-ranked Nash equilibria. The information structure of the repeated game is *P-0 graphs*. Our question is whether there is a sequential equilibrium of the repeated game in which any action profile strictly Pareto-dominating the Pareto-inferior Nash equilibrium is played in every period, except for the last few periods.

---

[18](A3) can be dispensed with.

As the example in subsection 3.1 illustrates, the answer may not be true for some stage games. However, under a certain restriction of the stage game payoffs, we can construct such an equilibrium.

The restrictions we impose on the stage game payoffs are as follows.

(M) There exist two Pareto-ranked stage game Nash equilibria e and f with $u_i(e) > u_i(f)$ $\forall i \in \{0\} \cup N$.

(SP) Given the Pareto-superior stage game Nash equilibrium e, for each $i \in N$, there exists $\alpha(i) = (\alpha_j(i))_{j \in \{0\} \cup \{i\} \cup S_i}$, a Nash equilibrium of the *modified stage game* among 0, i, and S_i, fixing the other players' actions to the ones specified by e, such that

(i) $u_i(e) > u_i(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})$ and

(ii) $u_0(e) \neq u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})$.

(M) is simply the requirement that the stage game has "multiple" Pareto-ranked Nash equilibria. (SP) refers to as "Spectators' Punishments". While (M) guarantees collective punishments by all the players against each player, (SP-i) says that player 0 and player i's spectators can punish player i independent of the other players.

Although it has structures similar to the *extended* M.F.N.R. for the infinitely repeated games, the equilibrium we construct differs from the *extended* M.F.N.R. in terms of the incentive for player 0 to *signal* after observing a deviation. In the *extended* M.F.N.R., we needed (IP) or (GP) to induce player 0 to *signal* a single deviation or simultaneous deviations. In the equilibrium here, we use no discounting assumption and (SP-ii) to induce player 0 to do so instead of (IP) or (GP).

Proposition 5.1. *Suppose the stage game satisfies (M) and (SP).*[19] *Let* $\mathbf{a}^* =$

---

[19] As we discussed through an example in section 3.1., (SP) is not a necessary condition.

$(a_i^*)_{i \in N \cup \{0\}}$ *be an action profile that is strictly Pareto-dominating the Pareto-inferior stage game Nash equilibrium* f. *Then there is a* $\underline{T}$ *such that for all* $T \geq \underline{T}$, *the* T *period repetition of the stage game has a sequential equilibrium where* $\mathbf{a}^*$ *is played for the first* $T - \underline{T}$ *periods.*

Along the equilibrium path, $\mathbf{a}^*$ is played for the first $T - \underline{T}$ periods and e is played for the last $\underline{T}$ periods. If player 0 plays an action other than a_0^* at the next to the last $\underline{T}$ periods, f is played for the next K periods. Given (M), she does not have the incentive to deviate at that period. (SP-i) makes it possible to punish player i in N at some points during the last $\underline{T}$ periods if he deviates at the next to the last $\underline{T}$ periods.

For the first $T - \underline{T} + 1$ periods, the equilibrium has similar structures to the *extended* M.F.N.R.. It has normal phase, signalling phase and punishment phase. The difference of the equilibrium from the *extended* M.F.N.R. is the behavior of a player in N after his deviation (i.e., in his signalling phase or in the signalling phase for multiple players including himself). If player i deviates, whether he is the only deviator or one of the simultaneous deviators, he is required to play a_i^* until player 0 *signals* while he is to play m_i under the *extended* M.F.N.R.. Player 0's *signalling* behavior is the same as the *extended* M.F.N.R.. Player 0 *signals* in any signalling phase whether it is for a single player or multiple players. In the equilibrium here, in a signalling phase of the early period when there are more than $2 + K + \underline{T}$ periods left, player 0's *signalling* after one period of delay will result in the same path as the one that would be generated if she *signalled* without delay. The effect of the delay is then to get $\mathbf{a}^*$ played during the delayed period and to resume normal phase where $\mathbf{a}^*$ is played one period later. Given no discounting assumption, player 0 is indifferent between the two paths. This is why (IP) or (GP) can be dispensed with

in proposition 5.1. On the other hand, if she does not *signal* in a signalling phase when there are at most $1 + K + \underline{T}$ periods left, player 0 will be strictly worse off. (SP-ii) guarantees that it is possible to punish player 0 or not to reward her, for not properly *signalling*, at some points during the last $\underline{T}$ periods.

Proof. See appendices. ∎

2.6 Robust Equilibrium

Until now, we discussed equilibria which support efficient behaviors under the general information structure given that player 0 is a perfect observer - $P - 0$ *graphs*. Another important assumption was that the information structure does not change from period to period and each player knows the structure. If a player observes possibly different players across periods and he does not know who is observing his action, then some of the equilibria we constructed may not be ones in this case even when player 0 is a perfect observer. In this section, we maintain our basic assumption that player 0 is a perfect observer and discuss a "robust" equilibrium which is an equilibrium under any of $P - 0$ *graphs*, including the *star graph* and the *complete graph*.[20]

Consider an arbitrary graph g among $P - 0$ *graphs*. For an arbitrary outcome path $h(t) = (a_0(s), (a_j(s))_{j \in N})_{s=1}^{t-1}$, let $h_i^g(t)$ be the previous actions that player 0, player i and his neighbors chose. In other words, $h_i^g(t) = (a_0(s), (a_j(s))_{j \in N_i^g \cup \{i\}})_{s=1}^{t-1} |$ $h(t)$ where N_i^g is player i's neighbors defined by the graph g. Now we define a "robust equilibrium" as follows.

---

[20]The extended M.F.N.R., without (GP), is not a "robust" strategy profile since it may not be an equilibrium under an arbitrary $P - 0$ *graph* as we discussed in section 4.2.. And, for example, the grim trigger strategy which is an equilibrium in the *complete graph* is not "robust" because it cannot be implemented in *star graph* for the obvious reason.

Definition 6.1. Consider two arbitrary $P - 0$ graphs, g and g'. Let σ^g and $\sigma^{g'}$ be two strategy profiles for the graphs g and g' such that for an arbitrary outcome path $h(t)$, (i) $\sigma_0^g(h(t)) = \sigma_0^{g'}(h(t))$ and (ii) for each $i \in N$, $\sigma_i^g(h_i^g(t)) = \sigma_i^{g'}(h_i^{g'}(t))$. σ^g is a robust equilibrium if $\sigma^{g'}$ is a sequential equilibrium for the graph g' whenever σ^g is a sequential equilibrium for the graph g.

In order for a strategy profile to be "robust", it has to be implementable in the *star graph* and so the action that it prescribes to each player other than player 0 should be based on only player 0's and his previous actions. We call a strategy profile with this property a restricted strategy profile.

Definition 6.2. σ is a restricted strategy profile if $\forall i \in N$, $\forall t$, $\sigma_i(h_i(t)) = \sigma_i(\widetilde{h_i}(t))$ whenever $(a_0(s), a_i(s))_{s=1}^{t-1} | h_i(t) = (a_0(s), a_i(s))_{s=1}^{t-1} | \widetilde{h_i}(t)$.

Also, a "robust" strategy profile has to be a subgame perfect equilibrium because it has to be an equilibrium under the *complete graph*.

Definition 6.3. *A restricted subgame perfect equilibrium is a restricted strategy profile that is a subgame perfect equilibrium under the complete graph.*

While a "robust" strategy profile has to be a restricted subgame perfect equilibrium, the next proposition shows the opposite is also true.

Proposition 6.1. *A restricted subgame perfect equilibrium is a sequential equilibrium under any of $P - 0$ graphs.*

Proof. Consider a restricted subgame perfect equilibrium (of the *complete graph*) σ^c. Take an arbitrary graph g among $P - 0$ *graphs*. For an arbitrary outcome path $h(t) = (a_0(s), (a_j(s))_{j \in N})_{s=1}^{t-1}$, which is nothing but an information set under the

complete graph, and for each $i \in N$, let $h_i^g(t) = (a_0(s), (a_j(s))_{j \in N_i \cup \{i\}})_{s=1}^{t-1} | h(t)$. In other words, $h_i^g(t)$ is the outcome path that player 0, player i and his neighbors chose, or player i's private information. Now let σ^g be a strategy profile under the graph g such that for an outcome path $h(t)$, (i) for each $i \in N$, $\sigma_i^g(h_i^g(t)) = \sigma_i^c(h(t))$ and (ii) $\sigma_0^g(h_0^g(t)) = \sigma_0^c(h(t))$ (note $h_0^g(t) = h(t)$). We want to show σ^g is a sequential equilibrium.

First of all, it is obvious to see that σ_0^g, by the construction, specifies to player 0 the optimal behavior at each of her information set $h_0^g(t)$.

Now, fix a typical player in N, say player 1 and his information set $h_1^g(t)$. Let $x(h_1^g(t))$ be a node in the information set $h_1^g(t)$ and $h(t)(x)$ be the outcome path that the node x ($= x(h_1^g(t))$ represents. To conclude the proof, it suffices to show that for all $x \in h_1^g(t)$, playing $\sigma_1^g(h_1^g(t))$ is optimal for player 1. However, the optimality is immediate from the observations that a deviation from $\sigma_1^g(h_1^g(t))$ will lead to the same outcome path that the deviation from $\sigma_1^c(h(t)(x))$ at the *information set* $h(t)(x)$ under the *complete graph* and that $\sigma_1^c(h(t)(x))$ is optimal for player 1 at the *information set* $h(t)(x)$ under the *complete graph*. So, we are done. ∎

In other words, we can completely characterize "robust" equilibrium by restricted subgame perfect equilibrium. A trivial restricted subgame perfect equilibrium is to play a stage game Nash equilibrium in each period, regardless of the previous plays. Now we examine a "robust" equilibrium which supports a stage game outcome that strictly Pareto-dominates the stage game Nash equilibrium.

The extended M.F.N.R. we constructed in Proposition 4.2 is a restricted subgame perfect equilibrium of the infinitely repeated game under the same conditions given in 4.2. So, the following proposition is true.

Corollary 6.2. *Consider an action profile* $\mathbf{a}^* = (a_i^*)_{i \in N \cup \{0\}}$ *which strictly Pareto dominates a stage game Nash equilibrium* f. *Suppose that (GP) holds. Then, for the infinitely repeated game, there is* $\bar{\delta} < 1$ *such that* $\mathbf{a}^*$ *is played in every period as a restricted subgame perfect equilibrium outcome for* $\delta \in (\bar{\delta}, 1)$.

For the finitely repeated games, we can get a result corresponding to Proposition 5.1 if we strengthen (SP) to (DPN). (DPN) refers to "Deviator's Punishments by Nash equilibrium".

(DPN) Given the Pareto-superior stage game Nash equilibrium e, for each $i \in N$, there exists another Nash equilibrium $(\beta_0, \beta_i, (e_j)_{j \in N - \{i\}})$ such that

(i) $u_i(e) > u_i(\beta_0, \beta_i, (e_j)_{j \in N - \{i\}})$ and

(ii) $u_0(e) \neq u_0(\beta_0, \beta_i, (e_j)_{j \in N - \{i\}})$.

Corollary 6.3. *Suppose the stage game satisfies (M) and (DPN). Let* $\mathbf{a}^* = (a_h^*)_{h \in N \cup \{0\}}$ *be an action profile that is strictly Pareto-dominating the Pareto-inferior stage game Nash equilibrium* f. *Then there is a* $\underline{T}$ *such that for all* $T \geq \underline{T}$, *the* T *period repetition of the stage game has a restricted subgame perfect equilibrium where* $\mathbf{a}^*$ *is played for the first* $T - \underline{T}$ *periods.*

Proof. Given (DPN), the strategy profile, constructed in Proposition 5.1, is a restricted subgame perfect equilibrium. ∎

(DPN) is quite a strong restriction. For some other stage games which do not satisfy (DPN), we present a negative result.

Consider a stage game satisfying the following restriction, which we call "player 0's action is payoff-irrelevant" and denote by (P-0 PI).

(P-0 PI) For each $h \in N \cup \{0\}$ and for all $(a_i)_{i \in N} \in \times_{i \in N} A_i$,

$$u_h(c_0, (a_i)_{i \in N}) = u_h(d_0, (a_i)_{i \in N}) \qquad \text{for all } c_0, d_0 \in A_0.$$

For a stage game that satisfies (P-0 PI), player 0's action is payoff-irrelevant in the sense that it does not affect the other players' payoffs as well as her own payoff. It is of course possible that player 0 has a strict preference over the actions played by the other players. For example, consider a joint-project consisting of n workers and 1 manager. Suppose total output of the project depends on the effort levels of the n workers. The manager is unproductive in the sense that her action does not affect the total output. However, the total output is equally distributed among the $n + 1$ members including the manager. In this joint-project example, the manager's action is payoff-irrelevant.

Suppose the stage game is played for T periods and the information structure of the repeated game is the *star graph*. If an equilibrium path specifies a non-Nash equilibrium outcome of the stage game in some periods, then at least one of the players has a myopic incentive to deviate from the path. To prevent the deviation, the equilibrium should provide punishments against the deviator. However, since other players did not observe the deviation, player 0 must initiate punishments if there is a deviation. So, the equilibrium must ensure that player 0 be better off by initiating punishments when she observes a deviation. On the other hand, player 0 should not do so when there are no deviations. Proposition 6.4 shows that for a large class of stage games that satisfy (P-0 PI), it is impossible to satisfy both constraints for player 0 simultaneously.

A payoff matrix for a game satisfying (P-0 PI) can be viewed as an element of $\Re^{(n+1)\prod_{i=1}^{n}|A_i|}$.

Proposition 6.4. *For any T, there exists a full measure open subset of $\Re^{(n+1)\prod_{i=1}^{n}|A_i|}$, G_T, such that if the stage game payoff satisfying (P-0 PI) is in G_T, there is no pure strategy restricted subgame perfect equilibrium of the T period repetition of the stage game in which an action profile other than a Nash equilibrium is ever played.*[21]

In a stage game that satisfies (P-0 PI), player 0's action can be regarded as a cheap-talk since her action is payoff-irrelevant. Player 0's only role is sending a public signal about the behaviors of the other players. The result of the proposition can be then interpreted as nonexistence of a pure strategy equilibrium in which player 0 tells the truth, unless the equilibrium path is playing a Nash equilibrium of the stage game. As a matter of fact, the result does not change if we explicitly allow player 0's cheap-talk after each period. Consider the following setup. Suppose players other than player 0 play a n–person game in each period. Suppose the actions they choose determine player 0's payoff as well as theirs although player 0 is not directly involved in the stage game. After the stage game, player 0 announces who deviated in the stage game. In this setup, we can show following. Given T, for generic payoffs of the stage game played by the n players, there is no pure strategy restricted subgame perfect equilibrium of the T period repetition of the stage game in which an action profile other than a Nash equilibrium is ever played. This negative result presents a striking contrast to Ben-Porath and Kahneman (1996) who

---

[21](P-0 PI) itself is not necessary for the negative result. The crucial property we need is that if an action profile $(c_0, b_i, (b_j)_{j \in N-\{i\}})$ is a pure strategy Nash equilibrium of the stage game, for each $i \in N$ and for any action $d_0 \in A_0$, there is no $a_i' \neq b_i$ such that $(d_0, a_i', (b_j)_{j \in N-\{i\}})$ is a Nash equilibrium. As a matter of fact, we can show a similar negative result for a stage game which satisfies this property instead of (P-0 P1).

provide the folk-theorem in a slightly different setting. In their papers, each player is observed by two other players and all the players make public announcements about the actions they observed. In the equilibrium they build, players do not lie during the announcement stage since both of two monitors are punished in case their announcement are incompatible. In our case, we cannot deter player 0 from lying to her advantage.

Proof. Consider a restricted subgame perfect equilibrium σ^* of the repeated game. Notice first that for any history, a Nash equilibrium has to be played in the last period. Suppose we have now established that for each period t, for an arbitrary history available at period t, $h(t)$, σ^* prescribes a Nash equilibrium for $h(t)$ and for any history following $h(t)$. We want to show that for an arbitrary history available at period $t - 1$, $h(t - 1)$, $\sigma^*(h(t - 1))$ is a Nash equilibrium. Suppose not. That is, there is a history $h(t - 1)$ and player $i \in N$ such that $\sigma_i^*(h(t - 1))$ is not the best response for $(\sigma_0^*(h(t - 1)), (\sigma_j^*(h(t - 1)))_{j \in N - \{i\}})$.

Let $(a_0^*(\tau), a_i^*(\tau), (a_j^*(\tau))_{j \in N - \{i\}})_{t-1}^T$ be the outcome path that σ^* generates given the history $h(t - 1)$. Notice that $a_i^*(t - 1) \neq Br_i(a_0^*(t - 1), (a_j^*(t - 1))_{j \in N - \{i\}})$ by the assumption and that for each τ, $t \leq \tau \leq T$, $(a_0^*(\tau), a_i^*(\tau), (a_j^*(\tau))_{j \in N - \{i\}})$ is a Nash equilibrium by the inductive hypothesis.

Now, let us explore a restriction that σ^* impose to deter player i's deviation to $b_i = Br_i(a_0^*(t - 1), (a_j^*(t - 1))_{j \in N - \{i\}})$ at period $t - 1$.

First of all, notice that if player i deviates to b_i at period $t - 1$, there has to be a period t', $t \leq t' \leq T$, where player 0 plays an action different from $a_0^*(t')$, say $\hat{a}_0(t')$. To see this, suppose there is no such period under σ^*. Then the actions that

σ^* prescribes to player $j \neq i$ must be $(a_j^*(\tau))_{\tau=t}^T$ since σ^* is a restricted strategy. Moreover, player i should play $a_i^*(\tau)$ at period τ, $t \leq \tau \leq T$. It is because the action profile that σ^* prescribes from period t is a Nash equilibrium by the inductive hypothesis and (P-0 PI) and the genericity imply that given a Nash equilibrium $(a_0^*(\tau), a_i^*(\tau), (a_j^*(\tau))_{j \in N-\{i\}})$, there is no action $\hat{a}_i(\tau) \neq a_i^*(\tau)$ for player i such that $(a_0^*(\tau), \hat{a}_i(\tau), (a_j^*(\tau))_{j \in N-\{i\}})$ is another Nash equilibrium. So, the outcome path that σ^* generates after player i's deviation will be $(a_0^*(\tau), a_i^*(\tau), (a_j^*(\tau))_{j \in N-\{i\}})_{\tau=t}^T$, the same path to be followed after player i did not deviated from $a_i^*(t-1)$. Then, player i is strictly better off by deviating to b_i at period $t-1$, which contradicts the fact that σ^* is a subgame perfect equilibrium.

So we assume, without loss of generality, that σ^* prescribes player 0 to play an action $\hat{a}_0(t) \neq a_0^*(t)$ at period t if player i deviates to b_i at period $t-1$. Therefore, the following inequality has to be satisfied.

$$u_0(\hat{a}_0(t), a_i^*(t), (a_j^*(t))_{j \in N-\{i\}}) + \sum_{\tau=t+1}^T u_0(\hat{a}_0(\tau), \hat{a}_i(\tau), (\hat{a}_j(\tau))_{j \in N-\{i\}})$$

$$\geq \sum_{\tau=t}^T u_0(a_0^*(\tau), a_i^*(\tau), (a_j^*(\tau))_{j \in N-\{i\}}) \tag{1}$$

Here, $(\hat{a}_0(\tau), \hat{a}_i(\tau), (\hat{a}_j(\tau))_{j \in N-\{i\}})_{\tau=t+1}^T$ is the outcome path to be generated by σ^* after the history $\hat{h}(t+1)$

where $\hat{h}(t+1) = [h(t-1), (a_0^*(t-1), b_i, (a_j^*(t-1))_{j \in N-\{i\}}), (\hat{a}_0(t), a_i^*(t), (a_j^*(t))_{j \in N-\{i\}})]$.

While the left-hand side of the inequality is player 0's payoff from the outcome path that σ^* generates after player i deviates at period $t-1$, the right hand side is player 0's payoff when she keeps on playing $a_0^*(\tau)$ at each period τ. Notice that player i's action at period t after his deviation at period $t-1$ should be $a_i^*(t)$. It is because the

73

action profile that σ^* prescribes at period t is a Nash equilibrium by the inductive hypothesis and (P-0 PI) and the genericity imply that given a Nash equilibrium $(a_0^*(t), a_i^*(t), (a_j^*(t))_{j \in N-\{i\}})$, there is no action $a_i' \neq a_i^*(t)$ for player i such that $(\hat{a}_0(t), a_i', (a_j^*(t))_{j \in N-\{i\}})$ is another Nash equilibrium. For the same reason, player i plays $a_i^*(\tau)$ at each period τ, $\tau \geq t+1$, as long as player 0 plays $a_0^*(\tau-1)$ at period $\tau - 1$.

On the other hand, in order for σ^* to be a subgame perfect equilibrium, given the history $h(t)$ where $h(t) = [h(t-1), (a_0^*(t-1), a_i^*(t-1), (a_j^*(t-1))_{j \in N-\{i\}})]$, player 0 should not have an incentive to deviate to an action other than $a_0^*(t)$ at period t. In particular, she should be worse off by playing $\hat{a}_0(\tau)$ in any period τ, $t \leq \tau \leq T$. In other words,

$$\sum_{\tau=t}^{T} u_0(a_0^*(\tau), a_i^*(\tau), (a_j^*(\tau))_{j \in N-\{i\}})$$

$$\geq u_0(\hat{a}_0(t), a_i^*(t), (a_j^*(t))_{j \in N-\{i\}}) + \sum_{\tau=t+1}^{T} u_0(\hat{a}_0(\tau), a_i'(\tau), (a_j'(\tau))_{j \in N-\{i\}}). \quad (2)$$

Here, for each τ, $t+1 \leq \tau \leq T$,

$$(a_i'(\tau), (a_j'(\tau))_{j \in N-\{i\}}) = (\sigma_i^*(h'(\tau)), (\sigma_j^*(h'(\tau)))_{j \in N-\{i\}}) \text{ where}$$

$$h'(t+1) = [h(t), (\hat{a}_0(t), a_i^*(t), (a_j^*(t))_{j \in N-\{i\}})] \text{ and}$$

$$h'(\tau+1) = [h'(t+1), (\hat{a}_0(s), a_i'(s), (a_j'(s))_{j \in N-\{i\}})_{s=t+1}^{\tau}].$$

Notice that for each τ, $t+1 \leq \tau \leq T$, $(\sigma_j^*(h'(\tau)))_{j \in N-\{i\}} = (\hat{a}_j(\tau))_{j \in N-\{i\}}$ since σ^* is a restricted strategy and so the actions that it prescribes to player $j \in N - \{i\}$ only depends on player 0's and player j's own previous actions. Given $(a_j'(\tau))_{j \in N-\{i\}} = (\hat{a}_j(\tau))_{j \in N-\{i\}}$, $a_i'(\tau)$ has also to be same as $\hat{a}_i(\tau)$ by the inductive hypothesis that

$(\tilde{a}_0(\tau), a'_i(\tau), (a'_j(\tau))_{j \in N-\{i\}})$ is a Nash equilibrium and by (P-0 PI) and the genericity. That is, the right-hand side of the inequality (2) is identical to the left-hand side of the inequality (1).

This implies that the inequality (1) and (2) must be satisfied with equality. However, we notice that $(a_0^*(\tau), a_i^*(\tau), (a_j^*(\tau))_{j \in N-\{i\}})_{\tau=t}^T$ is not just the rearrangement of $((\tilde{a}_0(t), a_i^*(t), (a_j^*(t))_{j \in N-\{i\}}), (\tilde{a}_0(\tau), \tilde{a}_i(\tau), (\tilde{a}_j(\tau))_{j \in N-\{i\}}))$. If those two paths are different only in terms of the order, then player 0 is strictly better off by deviating to b_i at period $t-1$. With this observation, it is straightforward to see that given T, the set of payoffs of the stage game which satisfies the inequality (1) and (2) is closed and of measure zero. ∎

2.7 Concluding Remarks

In this paper, we consider repeated games with a particular type of imperfect observability where while there is at least one player who observes the actions of all the players, other players only observe the perfect observer's action and possibly some other players' actions. Given the information structure, we provide Nash-threat folk theorems under certain restrictions of the stage game payoffs. In each of the equilibria we constructed, after a deviation occurs along the equilibrium path, the perfect observer plays an action other than the equilibrium-path action. This acts as a *signal* to the players after which they play the stage game Nash equilibrium for finite, say K, periods. A major concern is given to ensure that the perfect observer initiate the punishment when she observed some other player's deviation, while not doing so when there are no deviations.

Given our setup, we doubt the possibility of the general folk theorem that

any individually rational payoff vector is supportable as a sequential equilibrium outcome. Since the perfect observer can signal the identity of a deviator by playing some specific action or a sequence of actions, it might not be impossible to implement a player-specific punishment which is essential for the folk theorem-type results. The difficulty is to prevent the perfect observer's incentive to lie that some player has deviated if she is better off by minmaxing the player. However, we believe that the folk theorem can be obtained if there are two perfect observers. The argument in Ben-Porath and Kahneman (1996) can be applied.

The existence of the perfect observer is crucial for our results. The perfect observer's *signalling* action serves as a coordination device according to which the players synchronize the timing of the punishment. Without a perfect observer, however, the players do not know when to start to execute the punishment (for example, to play the stage game Nash equilibrium) since they may not be able to get the "feedback" that all the *signalling* process is properly done.

Regarding to the informational requirement for the players in repeated games, our results suggest an important observation. The *star graph* provides sufficient informations to the players to sustain an outcome that Pareto-dominate a stage game Nash equilibrium if the stage game satisfy a certain restriction on the payoffs. Compared with $n(n-1)$ observability for the n players under the *complete graph*, the *star graph* requires only $2n$ observability.

2.8 Appendices

The proof of proposition 4.3.

To fix the ideas, let us first give a sketch of the strategy profile.

Suppose player 0 neither deviated from a_0^* at period t nor before. Consider a typical player in N, say player 1. Let $D_1(t)$ be the set of the players among $N_1 \cup \{1\}$ who did not play a^* at period t.

If $D_1(t)$ is empty or a singleton, the actions player 1 should follow after period t are exactly same as in the *extended* M.F.N.R.. That is, he plays m_1 if $D_1(t) = \{1\}$ and a_1^* if $D_1(t)$ is empty or a singleton other than $\{1\}$. If $D_1(t)$ consists of at least two players, player 1 is required to play a_1^* at period $t+1$ whether he is one of them or not. If player 0 plays a_0^* at period $t+1$ and the set of the players in $N_1 \cup \{1\}$ who did not play a^* at period $t+1$ is $D_1(t+1)$, player 1 does the same at period $t+2$ as he would do at period $t+1$ if $D_1(t+1)$ were $D_1(t)$. On the other hand, if player 0 plays s_0 at period $t+1$, K periods of punishment phase where f is played starts from the next period. At period $t+K+2$, after the final period of the punishment phase, the action player 1 plays depends on $R_1(t+1)$, the set of the players whose period $t+1$'s actions are different from the ones as specified above. More formally, $R_1(t+1)$ is defined as follows:

For all $i \in N$, let $CN_{1i} = [N_1 \cup \{1\}] \cup [N_i \cup \{i\}]$.

$R_1(t+1) = \{i \in N_1 \cup \{1\} \mid$ (i) $a_i(t+1) \neq m_i$ for $\{i\} = D_1(t) \cap CN_{1i}$ and

$\qquad$ (ii) $a_i(t+1) \neq a_i^*$ for $\{i\} \neq D_1(t) \cap CN_{1i}$ $\}$

Player 1 plays a_1^* at period $t+K+2$ unless $R_1(t+1) = \{1\}$, in which case he plays m_1. That is, he plays exactly as he would do if player 0 played a_0^* at period

77

$t + K + 1$ and $D_1(t + K + 1)$ were $R_1(t + 1)$. If player 0 plays an action other than a_0^* and s_0, the infinite punishment phase is triggered.

Player 0, after observing all the players who did not play a^* at period t, first calculates the set of the players each of whom observed only his own deviation (from a^*). If we let the set $\widetilde{D}_0(t)$ (i.e., $\widetilde{D}_0(t) = \{i \in N \mid D_i(t) = \{i\}\}$), only those in the set will play m among N in the next period since the other deviators, who observed another deviation(s) besides their own, will play a^* as well as nondeviators. Given the other players' strategies, player 0 will *signal* at period $t+1$ if and only if $u_0(\ a^*)$ is strictly greater than $u_0(a_0^*, (m_j)_{j \in \widetilde{D}_0(t)}, (a_k^*)_{k \in N - \widetilde{D}_0(t)})$. If either $\widetilde{D}_0(t)$ is empty or the inequality is in the opposite direction, she plays a_0^*.

The beliefs system underlying the strategy profile is that player 1 believes player l who is not one of his neighbors has always played the equilibrium action a_l^* except for punishment phases where the stage game Nash equilibrium is played. It is quite straightforward to show this beliefs system is consistent with the strategy profile introduced in the above.

An important implication of this beliefs system is that if period t was in nonpunishment phase and player 0 played a_0^* at the period, $D_1(t)$ is the set of all the players who player 1 believes did not play a^* at period t. Also he believes that for each player $i \in N$, $D_1(t) \cap CN_{1i}$ is the set of all the players in N who player i believes did not play a^* at period t. Then $D_1(t)$ along with the strategy profile reflects player 1's prediction for the future plays[22]: Let $\widetilde{D}_1(t)$ be the subset of $D_1(t)$ such that a deviator in $\widetilde{D}_1(t)$ does not observe other deviators in $D_1(t)$.

---

[22]Note also that if $D_1(t) = \{i\}$ but player 0 and player i played a_0^* and a_i^* at period $t + 1$ instead of m_0 and m_i, player 1 believes the unexpected actions are due to their simultaneous mistakes rather than due to another deviation by one of player i's neighbors at period t. Another important implication of this beliefs system is that player 1, if he has been the only deviator among his nighbors and himself, never suspects another deviation which he does not observe even though player 0 unexpectedly continues to play a_0^* just as in M.F.N.R. under the *star graph*.

The deviators in $\widetilde{D}_1(t)$ are only those who are expected to play m since the other deviators in $D_1(t) - \widetilde{D}_1(t)$ observe another deviation(s) besides their own and will play a^*. So he expects the play to return to the equilibrium path if $\widetilde{D}_1(t)$ is empty. For the case where $\widetilde{D}_1(t)$ is not empty, he predicts player 0's *signalling* if $u_0(\mathbf{a^*})$ is strictly greater than $u_0(a_0^*, (m_j)_{j \in \widetilde{D}_1(t)}, (a_k^*)_{k \in N - \widetilde{D}_1(t)})$. If the inequality is in the opposite direction, he expects player 0 to continue to play a_0^*.

Checking the optimality of player 1's actions in this case is quite apparent because he can predict the future plays, whether he followed or did not followed the prescribed actions, only by $D_1(t)$.

However, if period t was the final period of K periods of the punishment phase triggered by player 0's *signalling* at period $t - K$, checking player 1's incentive at period $t + 1$ is more subtle than in the former case. In this case, $R_1(t - K)$, the set of the players in $N_1 \cup \{1\}$ who do not follow the prescribed actions at period $t - K$, may not be sufficient as player 1's predictor for the play of period $t + 1$. To see this, assume $N = \{1, 2, 3\}, N_1 = \{2, 3\}, N_2 = \{1\}$ and $N_3 = \{1\}$. Suppose period $t - K - 1$ was in nonpunishment phase and all the players played a^* at the period. If the action profile played at period $t - K$ is (s_0, a_1^*, m_2, a_3^*), $R_1(t - K) = \{2\}$ and player 1 believes $R_2(t - K) = \{2\}$. So, player 1's prediction for the play of period $t + 1$ will be (s_0, a_1^*, m_2, a_3^*). Now consider another situation where player 1 and player 3 played m_1 and m_3 at period $t - K - 1$. (so, $D_1(t - K - 1) = \{1, 3\}$.) If the action profile played at period $t - K$ is (s_0, a_1^*, m_2, a_3^*) as before, $R_1(t - K)$ will be also $\{2\}$. In this situation, however, player 1 expects the play to return to the equilibrium path from period $t + 1$. It is because he believes that $R_2(t - K) = \{1, 2\}$ since player 2, who deviated at period $t - K$, also regards player 1's unexpected action a_1^* as his mistake rather than as the "right" action and that player 2 and

player 0 will play a_2^* and a_0^* at period $t+1$. The point is that for each player $i \in N$, $R_1(t-K) \cap CN_{1i}$ may not be player 1's beliefs about the set of all the players in N who player i believes deviated at period $t-K$.

.

This problem motivates us to introduce the following machines to describe the strategy profile.

First of all, consider a typical player in N, say player 1. Player 1's machine is defined as follows. The set of states for player 1 is $[Z^+ \times (2^N)^n]$ where Z^+ is the set of nonnegative integers. A typical state is $(k, (Q_1^i)_{i \in N})$ where $k = 0, 1, ..., K, K+1$ and $Q_1^i \subset N$ and the initial state is $(0, (\emptyset))$. The output function for player 1 prescribes an action to player 1 as a function of a state for player 1 and is as follows:

$$\sigma_1(k, (Q_1^i)_{i \in N}) = \begin{cases} m_1 & \text{if } k = 0, \ Q_1^1 = \{1\} \\ a_1^* & \text{if } k = 0, \ Q_1^1 \neq \{1\} \\ f_1 & \text{if } k \neq 0 \end{cases}$$

Notice that player 1's action depends on only k and Q_1^1. The transition function for player 1 specifies the state for player 1 at the next period as a function of the state for player 1 of the current period and the actions player 1 observes at the current period. Before specifying the transition function, we introduce the following notations.

<u>Notation</u>

- $D_1(\mathbf{a}_1) = \{i \in N_1 \cup \{1\} \mid a_i \neq a_i^*\}$ where $\mathbf{a}_1 = (a_i)_{i \in N_1 \cup \{1\}}$.

- We define "common neighbors" between player 1 and player i as follows:

$\forall i \in N$,

$$CN_{1i} = [N_1 \cup \{1\}] \cap [N_i \cup \{i\}].$$

- For each $i \in N$, $\widetilde{Q_1^i} = \{j \in Q_1^i \mid Q_1^i \cap CN_{1j} = \{j\}\}$.

- For each $i \in N$, $(a_j(Q_1^i))_{j \in N} \in \times_{j \in N} A_j$ where

$$a_j(Q_1^i) = \begin{cases} m_j \text{ if } j \in \widetilde{Q_1^i} \\ a_j^* \text{ if } j \notin \widetilde{Q_1^i} \end{cases}$$

- For each $i \in N$, $R_1(Q_1^i, \mathbf{a}_1) = \{j \in CN_{1i} \mid a_j \neq a_j(Q_1^i)\}$

The transition function is as follows.

- $[(0, (Q_1^i)_{i \in N}), (a_0^*, \mathbf{a}_1)] \rightarrow [0, (D_1(\mathbf{a}_1) \cap CN_{1i})_{i \in N}]$

    - If the state at period t is $(0, (Q_1^i)_{i \in N})$ and player 0 played a_0^* at period t, player 1's state at period $t + 1$ would be $(0, (D_1(\mathbf{a}_1) \cap CN_{1i})_{i \in N})$ where $D_1(\mathbf{a}_1) \cap CN_{1i}$ is the set of players in CN_{1i} who did not play a^* at period t.

- $[(0, (Q_1^i)_{i \in N}), (s_0, \mathbf{a}_1)] \rightarrow [1, (R_1(Q_1^i, \mathbf{a}_1))_{i \in N}]$

- $[(k, (Q_1^i)_{i \in N}), (.,.)] \rightarrow [k + 1, (Q_1^i)_{i \in N}]$ for $k = 1, 2, ..., K - 1$

- $[(K, (Q_1^i)_{i \in N}), (.,.)] \rightarrow [0, (Q_1^i)_{i \in N}]$

    - If the state at period t is $(0, (Q_1^i)_{i \in N})$ and player 0 played s_0 at period t, K periods of punishment phase starts from the next period. After the final period of the punishment phase, player 1's state will be $(0, (R_1(Q_1^i, \mathbf{a}_1))_{i \in N})$ where $R_1(Q_1^i, \mathbf{a}_1)$ is the set of players in N whose actions at period t are different from the ones given by $(a_j(Q_1^i))_{j \in N}$.

81

- $[(0, (Q_1^i)_{i \in N}), (\hat{a}_0, \ a_1)] \to [K+1, \emptyset]$ where $\hat{a}_0 \neq a_0^*, s_0$

- $[(K+1, \emptyset), (.,.)] \to [K+1, \emptyset]$

  - If the state at period t is $(0, (Q_1^i)_{i \in N})$ and player 0 plays an action other than a_0^* or s_0 at period t, the next state will be $(K+1, \emptyset)$ that indicates the infinite punishment phase. The infinite punishment phase is absorbing.

  For player 0, define

  $$Q_0(((Q_j^i)_{i \in N})_{j \in N}) = \{i \in N \mid Q_i^i = \{i\}\}.$$

  Then states and transition function for player 0 are also well-defined. Her output function is given by

  - $\sigma_0(0, Q_0) = \begin{cases} s_0 & \text{if } u_0(\ \mathbf{a}^*) > u_0(a_0^*, (m_i)_{i \in Q_0}, (a_j^*)_{j \in Q_0}) \\ \\ a_0^* & \text{if } u_0(\ \mathbf{a}^*) \leq u_0(a_0^*, (m_i)_{i \in Q_0}, (a_j^*)_{j \in Q_0}) \text{ or} \\ & \qquad Q_0 = \emptyset \end{cases}$

  - $\sigma_0(k, Q_0) = f_0$ for $k \neq 0$.

  In each period t, the states for player 1, $(k, (Q_1^i)_{i \in N})$, reflects player 1's prediction for period t's play. Unless $k = 0$, player 1 expects f to be played. The k indicates that the current state is the kth period of K periods of punishment phase if $1 \leq k \leq K$ and that for $k = K + 1$, it is in the infinite punishment phase. If $k = 0$, given that each player i in N bases his action on Q_i^i and player 0 bases her action on $(Q_i^i)_{i \in N}$, player 1's prediction for period t's plays requires his beliefs about

82

$(Q_t^i)_{i \in N}$. However, lemma 1 shows that the beliefs system described above implies that for all $i \in N$, player 1's belief about Q_t^i is nothing but Q_1^i.

Lemma 1. For all $i \in N$, $Q_1^i = Q_\tau^i$ if $\forall l \in N - (N_1 \cup \{1\}), \forall \tau \geq 2$,

    (i) $a_l(\tau) = f$ if there is a period $s, 1 \leq s \leq K$, such that $a_0(\tau - s) \neq a_0^*$.

    (ii) $a_l(\tau) = a_l^*$ if $a_0(\tau - s) = a_0^*$ for all $s, 1 \leq s \leq K$.

To see the meaning of Q_1^i, suppose player 1's state at period t is $(0, (Q_1^i)_{i \in N})$. Given the interpretation of $(Q_1^i)_{i \in N}$, player 1 expects player i in N to play $a_i(Q_1^i)$. That is, player 1 calculates a set defined as

$$I_1((Q_1^i)_{i \in N})) = \{i \in N \mid Q_1^i = \{i\}\}.$$

I_1 is the set of players in N who are expected to m at period t. Also, player 1 expects player 0 to play s_0 if $u_0(\mathbf{a}^*) > u_0(a_0^*, (m_i)_{i \in I_1}, (a_j)_{j \in N - I_1})$. If I_1 is empty or the inequality is the opposite direction, she is expected to play a_0^*.

If player 0 played a_0^* at period t, player 1's state at period $t + 1$ would be $(0, (Q_1'^i)_{i \in N})$ where $Q_1'^i = D_1(\mathbf{a}_1) \cap CN_{1i}$ and $D_1(\mathbf{a}_1) \cap CN_{1i}$ is the set of players in CN_{1i} who did not play a^* at period t. Given his belief that those in $D_1(\mathbf{a}_1)$ are all the players in N who did not play a^* at period t, player 1 believes that for each player $i \in N$, $D_1(\mathbf{a}_1) \cap CN_{1i}$ is the set of all the players in N who player i believes did not play a^* at period t.

Now suppose player 0 played s_0 at period t. Then player 1's state at the next k period, $1 \leq k \leq K$, is $(k, (R_1^i(Q_1^i, \mathbf{a}_1))_{i \in N})$, which indicates the kth period of K periods of punishment phase. Here, $R_1(Q_1^i, \mathbf{a}_1)$ is the set of players in N whose actions at period t are different from the ones given by $(a_j(Q_1^i))_{j \in N}$.

Before checking the optimality of the output functions, we introduce useful lemmas.

Lemma 2. **(i)** For all $l \notin N_1 \cup \{1\}$, $Q_1^l \neq \{l\}$.

For all $i \in N_1 \cup \{1\}$,

(ii) if $Q_1^i = \{i\}$, $Q_1^j \neq \{j\}$ for all $j \in CN_{1i} - \{i\}$ and

(iii) if $Q_1^i = \{j\}$ for some $j \in CN_{1i} - \{i\}$, then $Q_1^j \neq \{i\}$.

Lemma 3. For all $i \in N$,

(i) if $Q_i^i = \{i\}$, $Q_j^j \neq \{j\}$ for all $j \in N_i$ and

(ii) if $j \in \widetilde{Q_i^i}$ for some $j \in N_i$, then $i \notin \widetilde{Q_j^j}$.

Lemma 2 shows some restrictions between player 1's states. According to (i) and (ii) of lemma 2., player 1 expects the other players in $N - \{1\}$ to play a^* if his state is $(0, (Q_1^i)_{i \in N})$ with $Q_1^1 = \{1\}$. Also, if the state is $(0, (Q_1^i)_{i \in N})$ where there is some $i \in N_1$ such that $Q_1^i = \{i\}$, all the player i's neighbors are expected to play a^* and player i to play m_i. Lemma 3, on the other hand, shows some restrictions on the states between the players in N. Lemma 3.(i) implies that there are no two neighbors who are to play m if the current state is not in punishment phase. In other words, if $i \in Q_0$ and $j \in Q_0$, they are not neighbors. The proofs of the lemmas will be provided at the end.

The optimality of player 1's output function

(1) $(0, (Q_1^i))$ with $Q_1^1 = \{1\}$.

Since there is no $i \in N - \{1\}$ such that $Q_1^i = \{i\}$ by (i) and (ii) of lemma 2., he expects player 0 to play s_0 and the other players in $N - \{1\}$ to play a^*. So if he plays m_1 as the output function specifies, he gets

$$(1 - \delta)u_1(s_0, m_1, (a_i^*)_{i \in N - \{1\}}) + \delta(1 - \delta^K)u_1(f) + \delta^{K+1}u_1(\mathbf{a}^*). \qquad (8.1)$$

On the other hand, if he does not play m_1, he expects that the state after the final period of the punishment phase will be $(0, (Q_1'^i))$ where $Q_1'^1 = \{1\}$, which leads to another punishment phase. So he will get

$$(1 - \delta) \max_{a_1 \neq m_1} u_1(s_0, a_1, (a_i^*)_{i \in N - \{1\}}) + \delta(1 - \delta^K)u_1(f)$$

$$+\delta^{K+1}(1 - \delta)u_1(s_0, m_1, (a_i^*)_{i \in N - \{1\}})$$

$$+\delta^{K+2}(1 - \delta^K)u_1(f) + \delta^{2K+2}u_1(\mathbf{a}^*). \qquad (8.2)$$

Given the construction of K, it is easy to see (8.1) is strictly larger than (8.2) for δ sufficiently close to 1.

(2) $(0, (Q_1^i))$ where there is some $i \in N_1$ such that $Q_1^i = \{i\}$.

Let I be the set of players in N_1 such that $Q_1^i = \{i\}$. Then, notice that there are no i and j in I who observe each other by (ii) of lemma 2. Player 1 expects those in I will keep on playing m if player 0 plays a_0^*, given the other players playing a_0^* as specified by the output functions.

(2.1) Suppose first that $u_0(\mathbf{a}^*) > u_0(a_0^*, (m_i)_{i \in I}, (a_j^*)_{j \in N - I})$.

In this case, player 1 expects player 0 to play s_0, players in I to play m and the others to play a^*. So if he plays a_1^* following the output function, he gets

85

$$(1-\delta)u_1(s_0, a_1^*, (m_i)_{i\in I}, (a_j^*)_{j\in N-I-\{1\}}) + \delta(1-\delta^K)u_1(f) + \delta^{K+1}u_1(\mathbf{a}^*).$$

As in (1), if δ is sufficiently close to 1, it is strictly larger than the payoff player 1 will get if he does not play a_1^*, which is

$$(1-\delta)\max_{a_1 \neq a_1^*} u_1(s_0, a_1, (m_i)_{i\in I}, (a_j^*)_{j\in N-I-\{1\}}) + \delta(1-\delta^K)u_1(f)$$

$$+\delta^{K+1}(1-\delta)u_1(s_0, m_1, (a_i^*)_{i\in N-\{1\}}) + \delta^{K+2}(1-\delta^K)u_1(f) + \delta^{2K+2}u_1(\mathbf{a}^*).$$

(2.2) Suppose now that $u_0(\mathbf{a}^*) \leq u_0(a_0^*, (m_i)_{i\in I}, (a_j^*)_{j\in N-I})$.

If this is case, player 1 expects player 0 to keep on playing a_0^*. If he plays a_1^* as required, his continuation payoff will be $u_1(a_0^*, (m_i)_{i\in I}, (a_j^*)_{j\in N-I})$. On the other hand, if he does not play a_1^*, all those in I will observe his deviation as well as their own and the state will be $(0, (Q_1'^i))$ where $Q_1'^1 = \{1, i\}$ for all $i \in I$, the play returning to the equilibrium path. So the payoff will be

$$(1-\delta)\max_{a_1 \neq a_1^*} u_1(s_0, a_1, (m_i)_{i\in I}, (a_j^*)_{j\in N-I-\{1\}}) + \delta u_1(\mathbf{a}^*).$$

Since $u_1(\mathbf{a}^*) \leq u_1(a_0^*, (m_i)_{i\in I}, (a_j^*)_{j\in N-I})$ by the symmetry of the payoffs and $u_1(\mathbf{a}^*) \neq u_1(a_0^*, (m_i)_{i\in I}, (a_j^*)_{j\in N-I})$ by (A3), player 1 will be better off in the former case for δ close to 1.

(3) $(0, (Q_1^i))$ where there is no $i \in N_1 \cup \{1\}$ such that $Q_1^i = \{i\}$.

In this case, player 1 expects the equilibrium path's play $\mathbf{a}^*$. Since his deviation will lead to K periods of punishment phase, he will be better off by playing a_1^*.

(4) $(k, (Q_1^i))$ where $k = 1, ..., K+1$

Since a Nash equilibrium f is played when $k \neq 0$ and the action he chooses in this state does not influence the future, player 1 will play f_1, the myopic best response.

The optimality of player 0's output function

(5) $(0, Q_0)$ with $Q_0 = \emptyset$

Since her action other than a_0^* will trigger the finite punishment phase or the infinite punishment, player 0 is better off by playing a_0^* and getting $u_0(\mathbf{a}^*)$ in each period.

(6) $(0, Q_0)$ with $Q_0 \neq \emptyset$

Those in Q_0 are to play m in the current period. And lemma 3.(i) implies that they will continue to play m unless player 0 plays an action other than a_0^*.

(6.1) Suppose first that $u_0(\mathbf{a}^*) \leq u_0(a_0^*, (m_i)_{i \in Q_0}, (a_j^*)_{j \in Q_0})$.

Then, player 0 is better off by keeping on playing a_0^* rather than playing s_0 for high discount factor since

$$u_0(a_0^*, (m_i)_{i \in Q_0}, (a_j^*)_{j \in Q_0})$$
$$> \quad (1 - \delta)u_0(s_0, (m_i)_{i \in Q_0}, (a_j^*)_{j \in Q_0}) + \delta(1 - \delta^K)u_0(f) + \delta^{K+1}u_0(\mathbf{a}^*).$$

The inequality is guaranteed by the construction of K and the restriction that $u_0(\mathbf{a}^*) \leq u_0(a_0^*, (m_i)_{i \in Q_0}, (a_j^*)_{j \in Q_0})$.

(6.2) Suppose now that $u_0(\mathbf{a}^*) > u_0(a_0^*, (m_i)_{i \in Q_0}, (a_j^*)_{j \in Q_0})$.

87

The following inequality shows the optimality of playing m_0 in this case.

$$(1 - \delta)u_0(s_0, (m_i)_{i \in Q_0}, (a_j^*)_{j \in Q_0}) + \delta(1 - \delta^K)u_1(f) + \delta^{K+1}u_1(\mathbf{a}^*)$$

$$> \quad (1 - \delta)u_0(a_0^*, (m_i)_{i \in Q_0}, (a_j^*)_{j \in Q_0}) + \delta(1 - \delta)u_0(s_0, (m_i)_{i \in Q_0}, (a_j^*)_{j \in Q_0})$$

$$+ \delta^2(1 - \delta^K)u_0(f) + \delta^{K+2}u_0(\mathbf{a}^*).$$

The left-hand side is player 0's continuation payoff from playing s_0 while the right hand side is her continuation payoff from playing a_0^* in the current period and playing s_0 in the next period. Notice that in both cases, the play returns to the equilibrium path after K periods of punishment phase. The restriction that $u_0(\mathbf{a}^*) > u_0(a_0^*, (m_i)_{i \in Q_0}, (a_j^*)_{j \in Q_0})$ guarantees the strict inequality if δ is sufficiently close to 1.

Proofs of lemma 1, 2, and 3

Lemma 1. For all $i \in N$, $Q_1^i = Q_i^i$ if $\forall l \in N - (N_1 \cup \{1\}), \forall \tau \geq 2$,

(i) $a_l(\tau) = f_l$ if $a_0(\tau - s) \neq a_0^*$ for $1 \leq s \leq K$

(ii) $a_l(\tau - 1) = a_l^*$, otherwise.

Proof. If $CN_{1i} = \emptyset$, $Q_1^i = \emptyset$ since $Q_1^i \subset CN_{1i}$. Q_i^i is also empty given the assumptions (i) and (ii). From now on, we prove the lemma for $i \in N$ such that CN_{1i} is not empty. We prove by induction. In the initial period, notice $Q_1^i = Q_i^i = \emptyset$. Let the states of player 1 and player i at period t be $(k, (Q_1^j)_{j \in N})$, $(k, (Q_i^j)_{j \in N})$, respectively, and suppose $Q_1^i = Q_i^i$. If $k = 1, 2, ..., K$, the states at period $t + 1$ will be simply $(k+1, (Q_1^j)_{j \in N})$, $(k+1, (Q_i^j)_{j \in N})$ for $k = 1, 2, ..., K - 1$ and $(0, (Q_1^j)_{j \in N})$, $(0, (Q_i^j)_{j \in N})$

for $k = K$. So the claim is obviously true. Suppose now the states at period t be

$$(0, (Q_1^j(t))_{j \in N}), (0, (Q_i^j(t))_{j \in N}).$$

First, consider the case where player 0's action at period t was a_0^*. Then,

$$Q_1^i(t+1) = D_1(\mathbf{a}_1) \cap CN_{1i} = \{j \in CN_{1i} | \ a_j(t) \neq a_j^*\}.$$

On the other hand,

$$Q_i^i(t+1) = D_i(\mathbf{a}_i) \cap [N_i \cup \{i\}] = \{j \in N_i \cup \{i\} | \ a_j(t) \neq a_j^*\}$$

$$= \{j \in CN_{1i} | \ a_j \neq a_j^*\}.$$

The last equality comes from the assumption that $a_l(t) = a_l^*$ for all $l \in N - (N_1 \cup \{1\})$. So, $Q_1^i(t+1) = Q_i^i(t+1)$.

Now, consider the case where player 0's action at period t was s_0. Suppose $j \in Q_1^i(t+1)$. For the case that $j \notin \widetilde{Q_1^i}(t)$ (or $j \in \widetilde{Q_1^i}(t)$), it has to be that $a_j(t) \neq a_j^*$ (or $a_j(t) \neq m_j$). Since $j \notin \widetilde{Q_i^i}(t)$ (or $j \in \widetilde{Q_i^i}(t)$) by the inductive hypothesis and $j \in N_i \cup \{i\}$, $j \in Q_i^i(t+1)$.[23] The same argument will show that if $j \in Q_i^i(t+1)$, $j \in Q_1^i(t+1)$. So we are done. ∎

Lemma 2. (i) For all $l \notin N_1 \cup \{1\}$, $Q_1^l \neq \{l\}$.

For all $i \in N_1 \cup \{1\}$,

(ii) if $Q_1^i = \{i\}$, $Q_1^j \neq \{j\}$ for all $j \in CN_{1i} - \{i\}$ and

(iii) if $j \in \widetilde{Q_1^i}$ for some $j \in CN_{1i} - \{i\}$, then $i \notin \widetilde{Q_1^j}$.

Proof. Since $Q_1^i \subseteq CN_{1i}$ for all $i \in N$, (i) is obvious.

For (ii) and (iii), we prove by induction. In the initial period, (ii) and (iii) are trivially true since $Q_1^i = \emptyset$ for all $i \in N$. Now, let the state of player 1 at period t

---

[23]It is obvious to see that that $Q_1^i = Q_i^i$ implies that $\widetilde{Q_1^i} = \widetilde{Q_i^i}$.

be $(k, (Q_1^i)_{i \in N})$ and suppose (ii) and (iii) are satisfied. Notice for $k = 1, 2, ..., K$, (ii) and (iii) are also satisfied at period $t+1$ since the state will be simply $(k+1, (Q_1^i)_{i \in N})$ and $(0, (Q_1^i)_{i \in N})$, respectively. We now suppose the state of player 1 at period t is $(0, (Q_1^i(t))_{i \in N})$ where $(Q_1^i(t))_{i \in N}$ satisfies (ii) and (iii).

First, consider the case where player 0's action at period t was a_0^*. Fix a player $i \in N_1 \cup \{1\}$. Since $Q_1^h(t+1) = D_1(a_1(t)) \cap CN_{1h}$ for all $h \in N$ in this case, $Q_1^i(t+1) = \{i\}$ holds only if $a_i(t) \neq a_i^*$ and $a_j(t) = a_j^*$ for all $j \in CN_{1i} - \{i\}$. So it must be the case that $i \in Q_1^j(t+1)$. On the other hand, if $j \in \widetilde{Q_1^i}(t+1)$ for some $j \in CN_{1i} - \{i\}$, it must be true that $a_j(t) \neq a_j^*$ and $a_i(t) = a_i^*$, which implies $i \notin \widetilde{Q_1^j}(t+1)$. Hence (ii) and (iii) are obviously satisfied.

Suppose player 0 played s_0 at period t and so the state of player 1 at period $t+1$ is $(1, (Q_1^i(t+1))_{i \in N})$.

We show first $(Q_1^i(t+1))_{i \in N}$ satisfies (ii). Fix a player $i \in N_1 \cup \{1\}$.

(ii-1) Suppose $Q_1^i(t) = \{i\}$. In this case, $Q_1^i(t+1) = \{i\}$ only if $a_i(t) \neq m_i$ and $a_j(t) = a_j^*$ for all $j \in CN_{1i} - \{i\}$. Then $Q_1^j(t+1) \neq \{j\}$ since $Q_1^j(t) \neq \{j\}$ by the inductive hypothesis (ii) and $a_j(t) = a_j^*$.

Suppose now $Q_1^i(t) \neq \{i\}$ and $Q_1^i(t+1) = \{i\}$. Let J be the set of $j \in CN_{1i} - \{i\}$ such that $.j \in \widetilde{Q_1^i}(t)$. ($J$ may be empty.) Notice then that in order for $Q_1^i(t+1) = \{i\}$ to be true, it has to be that $a_i(t) \neq a_i^*$, $a_j(t) = m_j$ for all $j \in J$ and $a_k(t) = a_k^*$ for all $k \in CN_{1i} - J - \{i\}$.

(ii-2) First of all, it is easy to see that for all $j \in J$, $i \in Q_1^j(t+1)$ (so, $Q_1^j(t+1) \neq \{j\}$) since $i \notin \widetilde{Q_1^j}(t)$ by the inductive hypothesis (iii) and $a_i(t) \neq a_i^*$.

(ii-3) Furthermore, for all $k \in CN_{1i} - J - \{i\}$, $Q_1^k(t+1) \neq \{k\}$. If there is $k \in CN_{1i} - J - \{i\}$ such that $Q_1^k(t+1) = \{k\}$, it must be that $k \in \widetilde{Q_1^k}(t)$ given $a_k(t) = a_k^*$. Notice that $k \in \widetilde{Q_1^k}(t)$ implies $Q_1^k(t) = \{k\}$. It is because if $k \in Q_1^k(t)$ and there is $q \in CN_{1k}$ such that $q \in Q_1^k(t)$, $k \notin \widetilde{Q_1^k}(t)$. But if $Q_1^k(t) = \{k\}$, $i \in Q_1^k(t+1)$ since $a_i(t) \neq a_i^*$. That is a contradiction to the fact that $Q_1^k(t+1) = \{k\}$.

Now we show $(Q_1^i(t+1))_{i \in N}$ also satisfies (iii). Fix a player $i \in N_1 \cup \{1\}$.

(iii-1) Suppose $Q_1^i(t) = \{i\}$ and there is some $j \in CN_{1i} - \{i\}$ such that $j \in \widetilde{Q_1^i}(t+1)$. Given $Q_1^i(t) = \{i\}$ (and so, $\widetilde{Q_1^i}(t) = \{1\}$), it must be the case that $a_j(t) \neq a_j^*$ in order for $j \in \widetilde{Q_1^i}(t+1)$ to be true. Notice $a_j(t) \neq a_j^*$ implies $j \in Q_1^j(t+1)$ since $Q_1^j(t) \neq \{j\}$ by the inductive hypothesis (ii). However, if $j \in Q_1^j(t+1)$, $\widetilde{Q_1^j}(t+1)$ is either $\{j\}$ or empty by the definition of $\widetilde{Q_1^j}(t+1)$. So, $i \notin \widetilde{Q_1^j}(t+1)$.

Suppose $Q_1^i(t) \neq \{i\}$ and there is some $q \in CN_{1i} - \{i\}$ such that $q \in \widetilde{Q_1^i}(t+1)$. In this case, notice first that $a_i(t) = a_i^*$ has to be true since $i \notin Q_1^i(t+1)$. Let J be the set of $j \in CN_{1i} - \{i\}$ such that $j \in \widetilde{Q_1^i}(t)$. ($J$ may be empty.)

(iii-2) First, if $q \in J$, it is obvious that $i \notin Q_1^q(t+1)$ since $i \notin \widetilde{Q_1^q}(t)$ by the inductive hypothesis (iii) and $a_i(t) = a_i^*$.

(iii-3) Now suppose $q \notin \widetilde{Q_1^i}(t)$ (i.e., $q \notin J$.) and $i \in \widetilde{Q_1^q}(t+1)$. Since $q \notin \widetilde{Q_1^i}(t)$, first of all, it must be that $a_q(t) \neq a_q^*$ in order for $q \in Q_1^i(t+1)$ to be true. On the other hand, in order for $i \in Q_1^q(t+1)$ to be true, it has to be that $i \in \widetilde{Q_1^q}(t)$ given $a_i(t) = a_i^*$. However, if $i \in \widetilde{Q_1^q}(t)$ (and so, $q \notin Q_1^q(t)$) and $a_q(t) \neq a_q^*$, $q \in Q_1^q(t+1)$. It is a contradiction to the fact that $i \in Q_1^q(t+1)$ since $\widetilde{Q_1^q}(t+1)$ is either $\{q\}$ or empty if $q \in Q_1^q(t+1)$. $\blacksquare$

Lemma 3. For all $i \in N$,

91

(i) if $Q_i^i = \{i\}$, $Q_j^j \neq \{j\}$ for all $j \in N_i$ and

(ii) if $j \in \widetilde{Q_i^i}$ for some $j \in N_i$, then $i \notin \widetilde{Q_j^j}$.

Proof. The proof is similar to the one in lemma 2.. We prove by induction. In the initial period, (i) and (ii) are trivially true since $Q_i^i = \emptyset$ for all $i \in N$. If player i's state is $(k, (Q_i^j)_{j \in N})$ where $k \neq 0$, the induction argument also trivially holds since player i's next state will simply be $(k+1, (Q_i^j)_{j \in N})$. We now suppose the states of the players in N at period t are $[(0, (Q_i^j(t))_{j \in N})]_{i \in N}$ where $(Q_i^i(t))_{i \in N}$ satisfies (i) and (ii).

First, consider the case where player 0's action at period t was a_0^*. Fix a player $i \in N$. Since $Q_i^i(t+1) = D_i(\ a_i) = \{j \in N_i \cup \{i\} \mid a_j(t) \neq a_j^*\}$ in this case, $Q_i^i(t+1) = \{i\}$ implies that $a_j(t) = a_j^*$ for all $j \in N_i$. So, $j \notin Q_j^j(t+1)$. Also, if $j \in \widetilde{Q_i^i}(t+1)$ for some $j \in N_i$, it must be the case that $a_i(t) = a_i^*$, which implies $i \notin Q_j^j(t+1)$. Hence, (i) and (ii) are satisfied.

Suppose player 0 played s_0 at period t and so the state of player i at period $t+1$ is $(1, (Q_i^j(t+1))_{j \in N})$.

We show first $(Q_i^i(t+1))_{i \in N}$ satisfies (i). Fix a player $i \in N$.

(i-1) Suppose $Q_i^i(t) = \{i\}$. In this case, $Q_i^i(t+1) = \{i\}$ only if $a_i(t) \neq m_i$ and $a_j(t) = a_j^*$ for all $j \in N_i$. Then $Q_j^j(t+1) \neq \{j\}$ since $Q_j^j(t) \neq \{j\}$ by the inductive hypothesis (i) and $a_j(t) = a_j^*$.

Suppose now $Q_i^i(t) \neq \{i\}$ and $Q_i^i(t+1) = \{i\}$. Let J ($\subset N_i$) be the set of $j \in \widetilde{Q_i^i}(t)$. ($J$ may be empty.) Notice that in order for $Q_i^i(t+1) = \{i\}$ to be true, it has to be that $a_i(t) \neq a_i^*$, $a_j(t) = m_j$ for all $j \in J$ and $a_k(t) = a_k^*$ for all $k \in N_i - J$.

92

(i-2) First of all, it is easy to see that for all $j \in J$, $i \in Q_j^j(t+1)$ (so, $Q_j^j(t+1) \neq \{j\}$) since $i \notin \widetilde{Q_j^j}(t)$ by the inductive hypothesis (ii) and $a_i(t) \neq a_i^*$.

(i-3) Furthermore, for all $k \in N_i - J$, $Q_k^k(t+1) \neq \{k\}$. If there is $k \in N_i - J$ such that $Q_k^k(t+1) = \{k\}$, it must be that $k \in \widetilde{Q_k^k}(t)$ given $a_k(t) = a_k^*$. Notice that $k \in \widetilde{Q_k^k}(t)$ implies $Q_k^k(t) = \{k\}$. It is because if $k \in Q_k^k(t)$ and there is $q \in N_k$ such that $q \in Q_k^k(t)$, $k \notin \widetilde{Q_k^k}(t)$. But if $Q_k^k(t) = \{k\}$, $i \in Q_k^k(t+1)$ since $a_i(t) \neq a_i^*$. That is a contradiction to the fact that $Q_k^k(t+1) = \{k\}$.

Now we show $(Q_i^i(t+1))_{i \in N}$ also satisfies (ii). Fix a player $i \in N_1 \cup \{1\}$.

(ii-1) Suppose $Q_i^i(t) = \{i\}$ and there is some $j \in N_i$ such that $j \in \widetilde{Q_i^i}(t+1)$. Given $Q_i^i(t) = \{i\}$ (and so, $\widetilde{Q_i^i}(t) = \{i\}$), it must be the case that $a_j(t) \neq a_j^*$ in order for $j \in \widetilde{Q_i^i}(t+1)$ to be true. Notice $a_j(t) \neq a_j^*$ implies $j \in Q_j^j(t+1)$ since $Q_j^j(t) \neq \{j\}$ by the inductive hypothesis (i). However, if $j \in Q_j^j(t+1)$, $\widetilde{Q_j^j}(t+1)$ is either $\{j\}$ or empty by the definition of $\widetilde{Q_j^j}(t+1)$. So, $i \notin \widetilde{Q_j^j}(t+1)$.

Suppose $Q_i^i(t) \neq \{i\}$ and there is some $q \in N_i$ such that $q \in \widetilde{Q_i^i}(t+1)$. In this case, notice first that $a_i(t) = a_i^*$ has to be true since $i \notin Q_i^i(t+1)$. Let J be the set of j in N_i such that $j \in \widetilde{Q_i^i}(t)$. ($J$ may be empty.)

(ii-2) First, if $q \in J$, it is obvious that $i \notin Q_q^q(t+1)$ since $i \notin \widetilde{Q_q^q}(t)$ by the inductive hypothesis (ii) and $a_i(t) = a_i^*$.

(ii-3) Now suppose $q \notin \widetilde{Q_i^i}(t)$ (i.e., $q \notin J$.) and $i \in \widetilde{Q_q^q}(t+1)$. Since $q \notin \widetilde{Q_i^i}(t)$, first of all, it must be that $a_q(t) \neq a_q^*$ in order for $q \in Q_i^i(t+1)$ to be true. On the other hand, in order for $i \in Q_q^q(t+1)$ to be true, it has to be that $i \in \widetilde{Q_q^q}(t)$ given $a_i(t) = a_i^*$. However, if $i \in \widetilde{Q_q^q}(t)$ (and so, $q \notin Q_q^q(t)$) and $a_q(t) \neq a_q^*$, $q \in Q_q^q(t+1)$. It is a contradiction to the fact that $i \in \widetilde{Q_q^q}(t+1)$ since $\widetilde{Q_q^q}(t+1)$ is either $\{q\}$ or empty if $q \in Q_q^q(t+1)$. ∎

The proof of Proposition 5.1

For convenience, we denote by period t the point of time when there are t periods left to end the game so that the game starts at period T and ends at period 1.

First of all, let K be an integer satisfying $2z < Ky$ where

$$z = \max_{i \in \{0\} \cup N} \max_{a,b} u_i(a) - u_i(b) \text{ and}$$

$$y = \min_{i \in \{0\} \cup N} \min_{a,b} u_i(a) - u_i(b) \text{ s.t. } u_i(a) - u_i(b) > 0.$$

We also let $\underline{T} = K + n(K + 2K^2)$.

Along the equilibrium path, $\mathbf{a}^*$ is played for the first $T - \underline{T}$ periods and e is played for the last $\underline{T}$ periods. If player 0 plays an action other than a_0^* at $1 + \underline{T}$ period, f is played for the next K periods. For the first $T - \underline{T} + 1$ periods, the equilibrium has three phases, normal phase, signalling phase and punishment phase. The play starts from normal phase where $\mathbf{a}^*$ is played unless some player(s) deviates. If player i, $i \neq 0$, is the only player who deviates from normal phase or deviated at the end of the last signalling phase, signalling phase for player i starts. If he is one of the deviators, signalling phase for the deviators starts. In the signalling phase, whether it is for player i or for multiplayers including player i, player i is required to play a_i^* while player 0 is required to play an action different from a_0^*.[24] The signalling phase continues unless player 0 *signals*, which triggers punishment phase where f is played for K periods. After the final period of punishment phase, the play returns to normal phase unless there are deviations at the end of the last signalling phase, in which case a new signalling phase follows.

---

[24] In this equilibrium, saying that player i deviated at the end of the last signalling phase has a slightly different meaning from saying in M.F.N.R.. Here, we mean the case where he did not play a_i at the period where the last punishment phase is triggered(i.e., at the last period when player 0 started to play an action other than a_0^*) whether he is in normal phase or his or other player(s)' signalling phase.

The last $\underline{T}$ periods is divided into $n+1$ blocks. We call K periods of the first block "adjustment phase". In this phase, punishment phase will be completed if not finished at period $1+\underline{T}$, and then e is played for the remaining periods of this phase. Each block i of the next n blocks consists of K periods of the "evaluation phase for player i" and the next $2K^2$ periods of the "player i's rewarding phase for player 0", denoted by $E(i)$ and $R(i)$, respectively. If period $\underline{T}+1$ was not during punishment phase, e is played during $E(i)$ unless player i deviated at period $\underline{T}+1$, in which case player h, $h \in S_i \cup \{i,0\}$, plays $\alpha_h(i)$ instead. For the case where period $\underline{T}+1$ was during punishment phase, e is played unless player i deviated at the period when the punishment is triggered. In this case, player h, $h \in S_i \cup \{i,0\}$, plays $\alpha_h(i)$ during this phase. Finally, in $R(i)$, player 0 would be either punished if she did not *signal* properly in the signalling phase for player i or players including player i or would be rewarded if she properly *signals*. More specifically, the play in $R(i)$ goes as follows:

Suppose first $u_0(e) > u_0(\alpha(i), (e_l)_{l \in N-(\{i\}\cup S_i)})$.

(i) Play e for the whole $R(i)$ phase if $\tau(i)$, the period at which ends the last signalling phase of player i (or players including him), is greater than or equal to $1+ K +\underline{T}$ or if he was never unpunished since period $1 + K +\underline{T}$.[25] Also play e if he never deviated since the initial period.

(ii) If $\tau(i)$, the period at which ends the last signalling phase of player i (or players including him) who was unpunished at period $\tau(i)+1$, is smaller than $1+K+\underline{T}$, $\alpha_h(i)$ is played for $[K + 1 + \underline{T} - \tau(i)] \cdot K$ periods. Play e for the remaining

---

[25]We say that player i stayed unpunished at period t if it was in the signalling phase of player i(or players including him) in period t but player 0 did not *signal* at that period. We also say that player i was never unpunished if player 0 *signalled* as soon as signalling phase of player i(or players including him) starts.

periods.

(iii) If player i stayed unpunished at period $1 + \underline{T}$, $\alpha_h(i)$ is played for the whole $R(i)$ phase(i.e., for $2K^2$).

Now consider the case $u_0(e) < u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})$.

(i)' Play $\alpha_h(i)$ for the whole $R(i)$ phase if $\tau(i)$, the period at which ends the last signalling phase of player i (or players including him), is greater than or equal to $1 + K + \underline{T}$ or if he was never unpunished since period $1 + K + \underline{T}$.

(ii)' If $\tau(i)$, the period at which ends the last signalling phase of player i (or players including him) who was unpunished at period $\tau(i)+1$, is smaller than $1 + K + \underline{T}$, e is played for $[K + 1 + \underline{T} - \tau(i)] \cdot K$ periods. Play $\alpha_h(i)$ for the remaining periods.

(iii)' If player i either stayed unpunished at period $1 + \underline{T}$ or he never deviated since the initial period, e is played for the whole $R(i)$ phase.

The beliefs system underlying the equilibrium is that player k believes player i who he does not observe has always played the equilibrium action a_i^* except for punishments phases where he played f_i. The implication of this belief is that the unexpected actions of player j, who is not only his neighbor but also a player i's spectator, during $E(i)$ or $R(i)$ are due to his own mistakes during those phases, not due to player i's deviation or player 0's not *signalling* at some periods before adjustment phase. It is straightforward to show this beliefs system is consistent.

Let us start with noticing the sequential rationality of the (claimed) equilibrium for the last $\underline{T}$ periods(i.e., $1 \leq t \leq \underline{T}$) given the beliefs system of the players: Each player plays Nash equilibrium of the stage game in adjustment phase. In $E(i)$ and $R(i)$, player j, $j \in S_i \cup \{0, i\}$, plays either e_j, Nash equilibrium of the stage game, or $\alpha_j(i)$, the Nash equilibrium of the modified stage game. On the other hand, player l plays e_l if he does not observe player i. This is optimal for player l given his belief that player i never deviated and so the other players would play e.

By the construction of K, it is also apparent that player 0 is better off playing a_0^* in normal phase rather than playing other action, which triggers K periods of f

.

The incentive of player i, $i \in N$, who is in normal phase or in the signalling phase of Q_i where $Q_i \subseteq N_i \cup \{i\}$, at period $t \geq 1 + \underline{T}$ is easily seen to be satisfied. While player i's gain from the deviation on early normal phase $t \geq 2 + \underline{T}$ or signalling phase $t \geq 2 + K + \underline{T}$, would be washed out by the ensuing K periods of punishment phase[26](which can be extended to adjustment phase), the incentive to deviate from the last normal phase $t = 1 + \underline{T}$ or the signalling phase in period $t \leq 1 + K + \underline{T}$ will be deterred by the threat through $E(i)$.

We show now *signalling* is optimal for player 0 in any player(s) signalling phase. Let Q, $Q \subseteq N$, be the nonempty set of the players who are in their signalling phase at the beginning of period $t \geq 1 + \underline{T}$. Let Q^n and Q^{up} be the two subset of Q where Q^n is the set of the players who are in their *new* signalling phase at the beginning of period $t \geq 1 + \underline{T}$ [27] and Q^{up} is the set of players who are unpunished at period $t + 1$. Partition Q^a, $a = n, up$, into two subsets Q'^a and

---

[26]Also by $2K^2$ periods of $\alpha_h(i)$ in $R(i)$ if $u_0(e) < u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})$ and he never deviated before.

[27]So, player i, $i \in Q^n$, deviated either at period $t + 1$ which is in his normal phase or at period $t + K + 1$ when player 0 triggered the last punishmentand.

Q''^a where

$$Q'^a = \{i \in Q^a | \ u_0(e) > u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})\}.$$

$$Q''^a = \{i \in Q^a | \ u_0(e) < u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})\}.$$

We also partition $N - Q$, the set of the players who are *not* in their signalling phase at the beginning of period $t \geq 1 + \underline{T}$, into three subsets, ND, P' and P'': ND is the set of the players who never deviated. $P' \cup P''$ is the set of the players who deviated before but are already punished at period t where

$$P' = \{i \in N - (ND \cup Q) \ | \ u_0(e) > u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})\}.$$

$$P'' = \{i \in N - (ND \cup Q) \ | \ u_0(e) < u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})\}.$$

Suppose, first, $t \geq 2 + K + \underline{T}$. Then player 0 is indifferent between *signalling* at period t and playing a_0^* at period t and then *signalling* at period $t - 1$. To see this, note that the payoffs for the last $\underline{T}$ periods do not change by this one-shot deviation and that the payoffs for the first $t - \underline{T}$ periods from *signalling* are the left-hand side of the following equation and the left-hand side is the payoffs from the one-shot deviation for these periods.

$$\arg\max_{a_0 \neq a_0^*} u_0(a_0, (a_i^*)_{i \in N}) + K \cdot u_0(f) + [t - (1 + K + \underline{T})] \cdot u_0(a^*)$$

$$= \ u_0(a^*) + u_0(m_0, (a_i^*)_{i \in N}) + K \cdot u_0(f) + [t - 1 - (1 + K + \underline{T})] \cdot u_0(a^*)$$

Consider now period t, where $2 + \underline{T} \leq t \leq 1 + K + \underline{T}$.

$$\arg\max_{a_0 \neq a_0^*} u_0(a_0, (a_i^*)_{i \in N}) + K \cdot u_0(f) + [t - (1 + \underline{T})] \cdot u_0(e) + nK \cdot u_0(e)$$

$$+ (\#ND + \#P') \cdot 2K^2 \cdot u_0(e) + 2K^2 \cdot \sum_{i \in P''} u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})$$

$$+2K^2 \cdot \#Q'^n \cdot u_0(e) + 2K^2 \cdot \sum_{i \in Q''^n} u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})$$

$$+(K + 1 + \underline{T} - t) \cdot K \cdot \sum_{i \in Q'^{up}} u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})$$

$$+[K^2 + (t - (1 + \underline{T})) \cdot K] \cdot \#Q'^{up} \cdot u_0(e)$$

$$+(K + 1 + \underline{T} - t) \cdot K \cdot \#Q''^{up} \cdot u_0(e)$$

$$+[K^2 + (t - (1 + \underline{T})) \cdot K] \cdot \sum_{i \in Q''^{up}} u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})$$

$$> \ u_0(\mathbf{a}^*) + \arg\max_{a_0 \neq a_0^*} u_0(a_0, (a_i^*)_{i \in N}) + K \cdot u_0(f)$$

$$+[t - 1 - (1 + \underline{T})] \cdot u_0(e) + nK \cdot u_0(e)$$

$$+(\#ND + \#P') \cdot 2K^2 \cdot u_0(e) + 2K^2 \cdot \sum_{i \in P''} u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})$$

$$+[K + 1 + \underline{T} - (t - 1)] \cdot K \cdot \sum_{i \in Q'^{up} \cup Q'^n} u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})$$

$$+[K^2 + ((t - 1) - (1 + \underline{T})) \cdot K] \cdot (\#Q'^{up} + \#Q'^n) \cdot u_0(e)$$

$$+(K + 1 + \underline{T} - (t - 1)) \cdot K \cdot (\#Q''^{up} + \#Q''^n) \cdot u_0(e)$$

$$+[K^2 + ((t - 1) - (1 + \underline{T})) \cdot K] \cdot \sum_{i \in Q''^{up} \cup Q''^n} u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})$$

The left-hand side of the inequality is the payoff from *signalling* at period t which is in Q's signalling phase while the right-hand side is the payoff from the one-shot deviation(i.e., playing a_0^* and period t and *signalling* at period $t+1$). The inequality is established by the construction of K.

The optimality of *signalling* at period $t = 1 + \underline{T}$ can be shown by the following inequality.

$$\arg\max_{a_0 \neq a_0^*} u_0(a_0, (a_i^*)_{i \in N}) + K \cdot u_0(f) + nK \cdot u_0(e)$$

$$+(\#ND + \#P') \cdot 2K^2 \cdot u_0(e) + 2K^2 \cdot \sum_{i \in P''} u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})$$

$$+2K^2 \cdot \#Q'^n \cdot u_0(e) + 2K^2 \cdot \sum_{i \in Q''^n} u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})$$

$$+K^2 \cdot \sum_{i \in Q'^{up}} u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)}) + K^2 \cdot \#Q'^{up} \cdot u_0(e)$$

$$+K^2 \cdot \#Q''^{up} \cdot u_0(e) + K^2 \cdot \sum_{i \in Q''^{up}} u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})$$

$$> \quad u_0(\mathbf{a}^*) + K \cdot u_0(e) + nK \cdot u_0(e)$$

$$+(\#ND + \#P') \cdot 2K^2 \cdot u_0(e) + 2K^2 \cdot \sum_{i \in P''} u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)})$$

$$+2K^2 \cdot \sum_{i \in Q'^{up} \cup Q'^{n}} u_0(\alpha(i), (e_l)_{l \in N - (\{i\} \cup S_i)}) + 2K^2 \cdot (\#Q''^{up} + \#Q''^{n}) \cdot u_0(e).$$

Chapter 3

Repeated Games with A Single Long-lived Player and Short-lived Players with Bounded Memory

3.1 Introduction

Classic repeated games refer to a situation where the same set of players play a fixed stage game in every period. A central result in the theory of these repeated games is that non-equilibrium outcomes of the stage game are consistent with equilibrium play of the repeated game. However, there are many interesting economic situations where not all of the players are involved in everlasting relationship. Some players play the stage game infinitely often, but others play only once and stay out of the game. For example, while a firm is expected to remain in business for substantially long periods, buyers usually leave the market after the trade with the firm. Even though the situations where some players are only short-lived do not exactly fit

102

to the classic repeated games, similar results have been expected. Consider, for instance, the example that appears Fudenberg, Kreps and Maskin (1990). One long-run player plays a fixed stage game against an infinite sequence of short-run players, each of whom plays the stage game only once. In the stage game, the short-run player moves first, and then the long-run player chooses an action available to her. If both the long-run and the short-run player observe all previous plays before the stage game starts, the repeated game has an equilibrium in which an outcome other than stage game equilibrium can be played in every period. The short-run player will not deviate from the desired play if the long-run player punishes him immediately after his deviation, while the long-run player also can abstrain from the myopic best response because her deviation will bring punishments from the subsequent short-run players. This is the basic intuition underlying many papers that explain why a firm can be trustworthy in the employment relationship or it produces high-quality goods despite short-run costs. (See, for example, Kreps (1986) and Klein and Leffler (1981).)

In this paper, we show that the intuition given in the above example is not true in general unless the short-run players observe all the previous plays. In particular, we consider a single long-run player who plays a fixed stage game against an infinite sequence of a different set of N short-run players. The stage game played by the $N + 1$ players is a standard game of perfect information. The main result of this note is that if each short-run player only observes the plays of a given finite number of previous stage games rather than all previous ones, for almost all discount factors the only pure strategy equilibrium outcome of the repeated game is simply the repetition of the stage game equilibrium.

This is a striking result considering the seemingly innocuous informational

constraint for the short-run players. While the long-run player is informed of all previous plays by its nature, it is rather demanding to expect the same for the short-run players. Due to their one-shot or short-run interaction with the long-run player, the short-run players are more likely to obtain the relevant information from only some of their predecessors than from all of them. In the above example, suppose the short-run players are an infinite sequence of buyers, each of whom has to decide whether to purchase a good or not, and the long-run player is a single firm who has to make a quality choice against a buyer who has decided to purchase. It is then reasonable to assume that while the firm usually has the records of its own quality choice and the buyer's purchasing behavior for all previous trades, buyers observe the outcomes of only a finite number of previous trades. Unfortunately, the implication of the note is that if this is the case, low-quality production and no-purchase is the only pure strategy equilibrium outcome of the repeated game for almost all discount factors.

The idea that bounded memory of a player may restrict the set of equilibria is not new. In a closely related paper, Tiffany (1988) study a model of intertemporal economic transfer in a setting of overlapping generations. In his model, each period a single agent is born and lives two periods. Young agents have the option of transferring a portion of their endowment to the old agents, who have no endowment. The main result of the model is that if agents cannot observe all transfers made before, the only sequential equilibrium is for each agent to transfer nothing. The literature on reactive equilibria in standard repeated games also concerns the role of bounded memory. In infinitely repeated duopoly games, Stanford (1986) consider a particular type of equilibria in which the quantity choice for each firm in the current period only depends on the other firm's quantity choice in the last period and show

that among those equilibria the only subgame perfect equilibrium is the repetition of Cournot-Nash equilibrium. Kalai, Samet and Stanford (1988) extends this result in the context of the prisoners' dilemma game by allowing both players to use more distant past, not only the previous period. They show that if at least one player bases the current action only on all previous actions of his opponent, ignoring his own actions, the non-cooperative outcome is the only reasonable subgame perfect equilibrium outcome of the repeated game. Even though the analysis on reactive strategy profiles may be interesting for its own sake, it is not very convincing that a player disregards the history of his own actions.

In next section, we provide an example which captures the main idea of this note. Section 3 introduces the notations for the model we consider. Section 4 presents the main result. As the conclusion, section 5 discusses certain issues to be addressed.

3.2   A Leading Example

Consider the sequential-move version of the buyer-seller game. The buyer begins by choosing whether or not to purchase a good from the seller. If he chooses not to buy, both players receive zero. If he buys, the seller can produce high quality or low quality. High quality gives both players a payoff of 1; low quality gives the seller 2 and the buyer -1. While trade, purchase and high-quality production, is the efficient outcome, the seller produces low quality and so the buyer refuses to buy in the unique subgame perfect equilibrium of this game.

To examine the plausibility of the efficient trade outcome, consider the same repeated game as in Fudenberg and Tirole (1991). Suppose the seller faces an

infinite sequence of buyers, each of whom plays the stage game only once. If the buyers and the seller observe all previous plays, this repeated game has a subgame perfect equilibrium in which the trade outcome is sustained in every period for high discount factors: The seller produces high quality in the first period and continues to do so if all previous buyers have purchased and she has always produced high quality. She produces low quality if she ever did so or there is at least one period when a buyer did not purchase. The first buyer starts out buying the good from the seller, and the other buyers continue to do so as long as all previous buyers have purchased and the seller has always produced high quality. If the seller ever produced low quality or there is at least one period when a buyer did not purchase, buyers refuse to buy.

If she is patient, the seller will obviously produce high quality along the equilibrium path for fear of the subsequent buyers' boycotts. The seller who ever produced low quality will provide low quality in case the current buyer unexpectedly purchases, because the subsequent buyers would not purchase regardless of her current quality choice. The buyer's strategy is also optimal given the seller's strategy. In particular, the buyer's punishment against the seller who ever produced low quality is indeed credible because she would produce low quality in case he purchases despite her cheating in the past.

This result, however, crucially depends on the perfect observability assumption that buyers observe all previous outcomes. If each buyer observes the outcomes of only a finite number of previous stage games, for almost all discount factors the only pure strategy equilibrium outcome is the repetition of the subgame perfect equilibrium of the stage game, the non-trade outcome. In particular, if each buyer observes only the outcome of the previous stage game, the only pure strategy equi-

librium outcome for a discount factor other than $\delta = 1/2$ is the repetition of the non-trade outcome.

To motivate the intuition, let us consider the case when each buyer can observe the outcome of only the previous stage game. Notice first that to induce the seller to produce high quality along the equilibrium path, the buyer should not purchase unless the seller produced high quality in the last period in a pure strategy equilibrium. If the seller is sufficiently patient, say $\delta > 1/2$, the seller will strictly prefer high-quality production to low-quality along the equilibrium path, given the buyers' subsequent punishments. However, the seller who produced low quality in the last period has no reason to produce low quality in case the current buyer unexpectedly purchases. In this case, she can be strictly better off by producing high quality because the next buyer will not observe the original low-quality production and thus will purchase as long as she produces high quality in the current period. Then the buyer, expecting the seller to always produce high quality, will not boycott her despite the low-quality production in the last period. This forgiveness of the buyers leads the seller to produce low quality, repudiating any pure strategy equilibrium supporting the trade outcome for $\delta > 1/2$.

To see this more formally, let a pure strategy profile for the seller and the buyers, (σ_S, σ_B), be an equilibrium of the repeated game where

$$\sigma_S \; : \; \{(B, H), (B, L), N\}^{t-1} \times \{B\} \rightarrow \{H, L\}$$

and

$$\sigma_B \; : \; \{(B, H), (B, L), N\} \rightarrow \{B, N\}.$$

Denote by $o_1 \in \{(B, H), (B, L), N\}$ a typical outcome of the last period' stage

game. Denote a typical outcome path until period t by $h_t \in \{(B,H),(B,L),N\}^t$. Let $V(h_{t-1},a_B)$ be the seller's value function induced by the equilibrium when the outcome of all the previous stage games was h_{t-1} and the current buyer's action was $a_B \in \{B,N\}$. Notice first that $V(h_{t-2},o_1,a_B) = V(h'_{t-2},o_1,a_B)$ for $h_{t-2} \neq h'_{t-2}$, for all $o_1 \in \{(B,H),(B,L),N\}$ and $a_B \in \{B,N\}$. Furthermore, $V(h_{t-2},o_1,a_B) = V(h'_{t-2},o'_1,a_B)$ for $o_1 \neq o'_1$. So the seller's value function only depends on the current buyer's action and we denote it by $V(a_B)$.

In order to support an outcome other than non-trade, the equilibrium has to satisfy

$$u_S(B,H) + \delta V(\sigma_B(B,H)) \geq u_S(B,L) + \delta V(\sigma_B(B,L)),$$

$$\text{where } V(N) = u_S(N) + \delta V(\sigma_B(N)).$$

Otherwise, the seller produces low quality at every opportunity and buyers therefore do not buy. On the other hand, if the above inequality holds with the strict one, the seller will produce high quality whenever the current buyer purchases. Given the seller's quality choice, buyers will buy regardless of the previous period's outcome. That leads to a contradiction since the seller will be then strictly better off by producing low quality. So the only possible case for (σ_S,σ_B) to support an outcome other than non-trade is when the above inequality holds with equality. It is quite straightforward to check the above inequality holds with equality if and only if $\delta = 1/2$, $\sigma_B(B,H) = B$ and $\sigma_B(B,L) = \sigma_B(N) = N$.

In fact, if $\delta = 1/2$, there is an equilibrium in which the trade outcome is sustained in every period: The seller produces high quality in the first period and continues to do so if the last period's outcome was purchase and high-quality

production. She produces low quality if the last period's outcome was non-purchase or low-quality production. The first buyer starts out buying the good from the seller, and the other buyers continue to do so as long as the last period's outcome was purchase and high-quality production. If the seller produced low quality in the last period or the buyer in the last period did not purchase, buyers refuse to buy. In this equilibrium, the seller is indifferent between high-quality production and low-quality production. So it is trivially optimal to produce high quality if she did so in the last period and to produce low quality if she produced low quality or the buyer did not purchase in the last period. Given the seller's strategy, the buyer's strategy is also optimal because each buyer cares only about the current period's payoff and thus should buy if and only if the seller is expected to produce high quality.

For the case when each buyer observes the outcomes of a finite number of the previous stage games rather than only one, we obtain a similar result:

Claim *If each buyer observes the outcomes of the last K stage games, there are only finite numbers of discount factors for which the repeated game has a pure strategy equilibrium outcome other than the repetition of the subgame perfect equilibrium of the stage game, the non-trade outcome.*

The proof of the claim basically follows the same argument as for $K = 1$ and consists of three steps.

As a first step, we realize that the seller's value function only depends on the outcomes of the last $K - 1$ stage games and the current buyer's action.

In the second step, we consider an equilibrium with the property that for each of the last $K - 1$ stage games' outcomes, the seller is not indifferent between high-quality production and low-quality production. Using backward-induction,

we develop the argument that in the end, the seller's value function induced by this equilibrium only depends on the last period's outcome and the current buyer's action, returning to the case of $K = 1$. This establishes that in the equilibrium satisfying the specified property, the seller always produces low quality and thus buyers never buy.

In the last step, we examine existence of a pure strategy equilibrium in which the seller is indifferent between high-quality and low-quality production for one of the last $K - 1$ stage games' outcomes. We show that such an equilibrium may exist only for finitely many discount factors.

Proof. Let a pure strategy profile of the seller and the buyers, (σ_S, σ_B), be an equilibrium where

$$\sigma_S \ : \ \{(B,H),(B,L),N\}^{t-1} \times \{B\} \to \{H,L\}$$

and

$$\sigma_B \ : \ \{(B,H),(B,L),N\}^K \to \{B,N\}.$$

Let $V : \{(B,H),(B,L),N\}^{t-1} \times \{B,N\} \to \Re$ denote the seller's value function induced by (σ_S, σ_B). Let $o_k \in \{(B,H),(B,L),N\}$ denote the realized outcome of the stage game of k period ago.[1] Of course, for all t and for all $(o_{t-\tau})_{\tau=1}^{t-1}$, $\sigma_S((o_{t-\tau})_{\tau=1}^{t-1}, B)$ solves

$$V((o_{t-\tau})_{\tau=1}^{t-1}, B) \tag{2.0}$$

$$= \max_{a_S \in \{H,L\}} u_S(B, a_S) + \delta V((o_{t-\tau})_{\tau=1}^{t-1}, (B, a_S), \sigma_B((o_{K-\tau})_{\tau=1}^{K-1}, (B, a_S))),$$

---

[1]In period t, we understand o_{t-k} to be empty if $t \leq k$.

where $\quad V((o_{t-\tau})_{\tau=1}^{t-1}, N)$

$$= u_S(N) + \delta V((o_{t-\tau})_{\tau=1}^{t-1}, (N), \sigma_B((o_{K-\tau})_{\tau=1}^{K-1}, (N))).$$

Notice that for all t,

$$V((o_{t-\tau})_{\tau=1}^{t-K-1}, o_K, o_{K-1}, ..., o_1, a_B) = V((o'_{t-\tau})_{\tau=1}^{t-K-1}, o'_K, o_{K-1}, ..., o_1, a_B)$$

for $(o_{t-\tau})_{\tau=1}^{t-K-1} \neq (o'_{t-\tau})_{\tau=1}^{t-K-1}, o_K \neq o'_K$.

So, the seller's value function only depends on the outcomes of the last $K-1$ stage games and the current buyer's action. We denote it by $V((o_{K-\tau})_{\tau=1}^{K-1}, a_B)$, where for all t,

$$V((o_{K-\tau})_{\tau=1}^{K-1}, a_B) = V((o_{t-\tau})_{\tau=1}^{t-K}, (o_{K-\tau})_{\tau=1}^{K-1}, a_B) \qquad (2.1)$$

for all $(o_{t-\tau})_{\tau=1}^{t-K} \in \{(B, H), (B, L), N\}^{t-K}$.

Then (2.0) can be written as the following. For all t and $(o_{t-\tau})_{\tau=1}^{t-1}$, $\sigma_S((o_{t-\tau})_{\tau=1}^{t-1}, B)$ solves:

$$V((o_{K-\tau})_{\tau=1}^{K-1}, B) =$$

$$\max_{a_S \in \{H,L\}} u_S(B, a_S) + \delta V((o_{K-\tau})_{\tau=2}^{K-1}, (B, a_S), \sigma_B((o_{K-\tau})_{\tau=1}^{K-1}, (B, a_S))), \qquad (2.2)$$

where $\quad V((o_{K-\tau})_{\tau=1}^{K-1}, N)$

$$= u_S(N) + \delta V((o_{K-\tau})_{\tau=2}^{K-1}, N, \sigma_B((o_{K-\tau})_{\tau=1}^{K-1}, (N))).$$

111

Case 1. Suppose first that for all $(o_{K-\tau})_{\tau=1}^{K-1} \in \{(B,H),(B,L),N\}^{K-1}$, (2.2) has a unique solution, i.e.,

$$u_S(B,H) + \delta V((o_{K-\tau})_{\tau=2}^{K-1}, (B,H), \sigma_B((o_{K-\tau})_{\tau=1}^{K-1}, (B,H))) \quad (2.3)$$

$$\neq u_S(B,L) + \delta V((o_{K-\tau})_{\tau=2}^{K-1}, (B,L), \sigma_B((o_{K-\tau})_{\tau=1}^{K-1}, (B,L))).$$

This means that the equilibrium does not allow two different outcome paths to generate the same payoff to the seller. In this case, we show that the repetition of non-trade is the only equilibrium outcome.

Notice first the seller's quality after the current buyer's purchase is only dependent upon the last $K-1$ periods' outcomes. Given a particular path of the last $K-1$ periods' outcomes, $(o_{K-\tau})_{\tau=1}^{K-1}$, the seller will choose the quality which gives higher payoffs. We denote it by $\sigma_S((o_{K-\tau})_{\tau=1}^{K-1}, B)$, where

$$\sigma_S((o_{K-\tau})_{\tau=1}^{K-1}, B) = \sigma_S((o_{t-\tau})_{\tau=1}^{t-K}, (o_{K-\tau})_{\tau=1}^{K-1}, B)$$

for all t and $(o_{t-\tau})_{\tau=1}^{t-K}$.

Then the current buyer also bases his decision only on the outcomes of the last $K-1$ stage games, $(o_{K-\tau})_{\tau=1}^{K-1}$, regardless of the outcome of the stage game of K period ago, o_K. Given $(o_{K-\tau})_{\tau=1}^{K-1}$, the buyer will buy if $\sigma_S((o_{K-\tau})_{\tau=1}^{K-1}, B) = H$ and will not buy if $\sigma_S((o_{K-\tau})_{\tau=1}^{K-1}, B) = L$. Denote the buyer's decision by $\sigma_B((o_{K-\tau})_{\tau=1}^{K-1})$, where

$$\sigma_B((o_{K-\tau})_{\tau=1}^{K-1}) = \sigma_B(o_K, (o_{K-\tau})_{\tau=1}^{K-1}) \quad (2.4)$$

for all $o_K \in \{(B,H),(B,L),N\}$.

The seller's value function then only depends on the last $K - 2$ periods' outcomes and the current buyer's action and will be denoted by $V((o_{K-\tau})_{\tau=1}^{K-2}, a_B)$, where

$$V((o_{K-\tau})_{\tau=2}^{K-1}, a_B) = V(o_{K-1}, (o_{K-\tau})_{\tau=2}^{K-1}, a_B) \qquad (2.5)$$

for all $o_{K-1} \in \{(B, H), (B, L), N\}$.

Notice that (2.3), along with (2.4) and (2.5), implies that all $(o_{K-\tau})_{\tau=2}^{K-1}$,

$$u_S(B, H) + \delta V((o_{K-\tau})_{\tau=3}^{K-1}, (B, H), \sigma_B((o_{K-\tau})_{\tau=2}^{K-1}, (B, H)))$$

$$\neq u_S(B, L) + \delta V((o_{K-\tau})_{\tau=3}^{K-1}, (B, L), \sigma_B((o_{K-\tau})_{\tau=2}^{K-1}, (B, L))).$$

Then the seller's quality choice after the current buyer's purchase turns out to depend only on the last $K - 2$ periods' outcomes, and so does the current buyer's action, regardless of the outcomes of the first two stage games, o_K or o_{K-1}.

Continuing the inductive argument leads to the conclusion that the seller's quality choice after the purchase does not depend on any of the previous periods' outcomes and thus nor does the buyer's purchase decision. Then the seller will produce low quality, making it optimal for buyers not to purchase. So in order for (σ_S, σ_B) to support an outcome other than the repetition of non-trade outcome, it must be the case that there is at least one $(o_{K-\tau})_{\tau=1}^{K-1}$ for which (2.2) does not have a unique solution.

Case 2. Suppose that under (σ_S, σ_B) there is at least one $(o_{K-\tau})_{\tau=1}^{K-1}$ for which (2.2) does not have a unique solution.

In this case, we will show that the number of δ's for which such (σ_S, σ_B) exists is at most finite.

First of all, to each $((o_{K-\tau})_{\tau=1}^{K-1}, a_B) \in \{(B, H), (B, L), N\}^{K-1} \times \{B, N\}$, let us assign a number $i \in I = \{1, 2, ..., 2 \times 3^{K-1}\}$. Consider now a correspondence, $\hat{\sigma} : I \Rightarrow \{u_S(B, H), u_S(B, L), u_s(N)\} \times I$, that generates a simultaneous equation system, with $2 \times 3^{K-1}$ unknown variables $\{V(i)\}_{i \in I}$, of the following form by specifying combinations of u and j for each i : For all $i \in I$,

$$V(i) = u + \delta V(j)$$

for some $u \in \{u_S(B, H), u_S(B, L), u_S(N)\}$,

for some $j \in I$.

While an arbitrary correspondence does not necessarily correspond to a pure strategy profile, any pure strategy equilibrium generates a particular correspondence, or a simultaneous equation system, because for each i, it determines particular combinations (or a combination) of (u, j) by solving (2.2). Furthermore, the simultaneous equation system generated by a pure strategy equilibrium has a solution, $(V(i))_i$, the same one as defined in (2.1).[2]

Notice that the simultaneous equation system, say σ^*, generated by any pure strategy equilibrium satisfying Case 2, has more equations than variables since such an equilibrium allows at least one $i \in I$ that has more than one specification of

[2]Of course, the converse is not true because the simultaneous equation system does not take the incentive compatibility conditions into account.

114

(u, j).[3] Notice also σ^* has the property that if $\sigma^*(i) = \{(u_S(B, H), j), (u_S(B, L), j')\}$, then $j \neq j'$. If $j = j'$, this means that for the outcome path i, the next period's buyer's purchasing decision does not depend on the seller's current quality choice and then the seller will be strictly better off by producing low quality in the current period. Then it cannot be the case that $\sigma^*(i) = \{(u_S(B, H), j), (u_S(B, L), j')\}$.

Let us call a simultaneous equation system with more equations than variables that satisfies the above property simply a candidate system. Obviously, a candidate system generated by any pure strategy equilibrium must have a solution. However, it can be shown that for an arbitrary candidate system, the number of δ's for which that candidate system has a solution is at most finite.[4] This means that the number of δ's for which that candidate system can be generated by a pure strategy equilibrium is at most finite. Since for each i there are only finitely many possible ways of specifying combinations of u and j, the number of all the candidate systems is finite and thus so is the number of δ's for which there exists a candidate system that has a solution. This implies that the number of δ's for which there exists a pure strategy equilibrium that allows Case 2 is at most finite. ∎

In the next section, we extend this result to a more general class of stage games, finite games of perfect information.

---

[3] The number of equations are $(3^{K-1} \times 2)$ +(the number of i's, $(o_{K-\tau})_{\tau=1}^{K-1}$, for which (2.1) does not have a unique solution).

[4] If, for all i, (2.1) has a unique solution, so that the simultaneous equation system has the same number of equations and variables, we can show the simutaneous equation system indeed has a solution for all $\delta \in [0, 1)$.

3.3 The Model

Suppose that a single long-run player, player l, plays a fixed stage game against an infinite sequence of a different set of N short-run players. The stage game is a game of perfect information. The stage game has a finite number of nodes and for each node, there is a finite number of actions available. The description of the stage game is as follows:

1. Denote by X the finite set of non-terminal nodes and the finite set of terminal nodes by Z.

   - For all $x \in X$, let $S(x), \subset X \cup Z$, denote the successors of x, the set of nodes that can be reached from node x.

   - For all $y \in X \cup Z$, let $P(y), \subset X$, denote the predecessors of y, the set of nodes that precedes node y.

   - For convenience, we partition the set of non-terminal nodes, X, into $N+1$ subsets of X_l and $\{X_i\}_{i=1,2,\ldots,N}$. X_l represents the set of nodes where the long-run player moves and X_i the set of nodes for short-run player i.

2. For each $x \in X$, let $A(x)$ denote the finite set of actions available at x.

   - For each $x \in X$, for each $a \in A(x)$, let $\vec{a}$ denote the immediate successor of x connected by a.

3. For all $y \in X \cup Z$, let $o(y) \in \times_{x \in P(y)} A(x)$ denote the preceding outcome path (a string of actions) reaching y.

   - If y is the initial node, we understand $o(y)$ to be empty.

116

- Let $O(Z) = \cup_{z \in Z}\{o(z)\}$ be the set of all possible outcome paths (strings of actions) reaching terminal nodes in Z. We denote $O(Z)$ and $o(z)$ simply by O and o, respectively.

4. The stage game payoff for player i is a function $u_i : O \to \Re$ with $u_i(o)$ being player i's stage game payoff if o is the realized outcome.

As for the information structure of the repeated game, the long-run player observes the outcome of all the previous stage games as well as the current period's outcome path while a short-run player observes the outcome of only the last K stage games and the current period's outcome path. Let $o_k \in O$ denote the realized outcome of the stage game of k period ago. In period t, we understand o_{t-k} to be empty if $t \leq k$. A pure strategy for the long-run player is a map σ_l, where for all period t,

$$\sigma_l : O^{t-1} \times X_l \to \times_{x \in X_l} A(x) .$$

A pure strategy for short-run player $i \in \{1, 2, ..., N\}$ is a map σ_i, where

$$\sigma_i \quad : \quad O^K \times X_i \to \times_{x \in X_i} A(x) \text{ if } t > K ,$$

$$\sigma_i \quad : \quad O^{t-1} \times X_i \to \times_{x \in X_i} A(x) \text{ if } t \leq K.$$

While the long run player's repeated game payoff is the discounted sum of the stage game payoffs for a discount factor $\delta \in [0, 1)$, a short-run player's payoff is simply given by the stage game payoff.

We apply Perfect Bayesian as an equilibrium concept for the repeated game. A strategy profile $(\sigma_l, (\sigma_i)_{i=1,...,N})$ is a Perfect Bayesian equilibrium if in each period t, for each $(o_{t-\tau})_{\tau=1}^{t-1}$ and for each node $x \in X_l$, σ_l is optimal given $(\sigma_l, (\sigma_i)_{i=1,...,N})$ and

for each $(o_{K-\tau})_{\tau=0}^{K-1}$ and a node $x \in X_i, \sigma_i$ is optimal given a belief about $(o_{t-\tau})_{\tau=0}^{t-K-1}$ and $(\sigma_l, (\sigma_i)_{i=1,...,N})$.

3.4 Pure strategy equilibria

Since a finite game of perfect information has always a pure strategy subgame perfect equilibrium, the repeated game has also a pure strategy equilibrium, the repetition of the stage game equilibrium. Our interest is of course beyond the trivial equilibrium and finding another pure strategy equilibria. The main result of this paper, however, suggests that it is almost (i.e., except for a finite number of discount factors) impossible.

Proposition 1 *For generic payoffs of the stage game, there are at most finitely many δ's for which the repeated game has a pure strategy equilibrium other than the repetition of the subgame perfect equilibrium of the stage game.*

The only difference in the proof of this general case from the example in section 2 is due to the fact that the long-run player in the stage game may have more than one decision node. So we define the long-run player's value function on the outcomes of the previous $K - 1$ stage games and all her decision nodes of the current stage game. In the first step, we consider an equilibrium with the property that for each of the last $K - 1$ stage games' outcomes and for each of her decision nodes, the long-run player *strictly* prefers one action to the other available actions. Using backward-induction, we then develop the argument that in this equilibrium, the long-run player's decision depends on none of the previous periods' outcomes, thus nor do the short-run players' decisions. This history-independence immediately establishes that the equilibrium satisfying the property is the repetition of

118

the subgame perfect equilibrium of the stage game. Secondly, we examine existence of a pure strategy equilibrium that has an outcome path after which the long-run player is indifferent between choosing at least two distinct actions. We find out that the number of discount factors for which such an equilibrium may exist is at most finite.

Proof. Here we assume that $K = 1$, so that each short-run player observes only the last period's outcome and the current period's outcome path. This proof can be easily extended to more general K by using the argument given in the Claim.

Since short-run players only care about the payoff of the current stage game and the stage game is of perfect information, we can treat short-run players as one player and denote them by player s, as opposed to player l for the long-run player, instead of player $i \in \{1, 2, ..., N\}$.

Consider a pure strategy equilibrium (σ_l, σ_s) of the repeated game. Then for all t, for all $(o_{t-\tau})_{\tau=1}^{t-1} \in O^{t-1}$, and for all $x \in X$, we can define the outcome path of the stage game after node x that (σ_l, σ_s) generates. We will denote it by $q^e((o_{t-\tau})_{\tau=1}^{t-1}; x)$. On the other hand, let $q^*(x)$ denote the outcome path of the stage game after node x that is determined by the subgame perfect equilibrium of the stage game. Notice that genericity of payoffs, i.e., distinct payoffs for different terminal nodes, guarantees uniqueness of $q^*(x)$ for each x.

To appropriately introduce the long-run player's value function induced by a pure strategy equilibrium, let $\widehat{X_l} = \{x \in X_l | P(x) \cap X_l = \emptyset\}$ and $\widehat{Z} = \{z \in Z | P(z) \cap X_l = \emptyset\}$. Obviously, if the initial node belongs to the long-run player, $\widehat{X_l}$ is the singleton set consisting of the initial node and $\widehat{Z} = \emptyset$.

119

Now let $V : O^{t-1} \times (X_l \cup \widehat{Z}) \to \Re$ denote the long-run player's value function determined by (σ_l, σ_s). Of course, for all t, for all $(o_{t-\tau})_{\tau=1}^{t-1} \in O^{t-1}$, and for all $x \in X_l$, $\sigma_l((o_{t-\tau})_{\tau=1}^{t-1}, x)$ solves

$$V((o_{t-\tau})_{\tau=1}^{t-1}; x) = \max_{a_l \in A(x)} u_S(o(x), a_l, q^e((o_{t-\tau})_{\tau=1}^{t-1}; \overrightarrow{a_l}))$$

$$+ \delta V((o_{t-\tau})_{\tau=1}^{t-1}, (o(x), a_l, q^e((o_{t-\tau})_{\tau=1}^{t-1}; \overrightarrow{a_l})); y^{\sigma_s}),$$

(i) where $\overrightarrow{a_l}$ is the immediate successor of x connected by $a_l \in A(x)$, and

(ii) $(o(x), a_l, q^e((o_{t-\tau})_{\tau=1}^{t-1}; \overrightarrow{a_l}))$ is the outcome path of period t's stage game that (σ_l, σ_s) generates when the long-run player chooses a_l at x.

(iii) $y^{\sigma_s} \in \widehat{X_l} \cup \widehat{Z}$ is the node to be uniquely determined by σ_s given $(o(x), a_l, q^e((o_{t-\tau})_{\tau=1}^{t-1}; \overrightarrow{a_l})).$[5]

Notice that y^{σ_s} only depends on $(o(x), a_l, q^e((o_{t-\tau})_{\tau=1}^{t-1}; \overrightarrow{a_l}))$ since y^{σ_s} will reached only by the short-run player's move at period $t+1$ and the short-player's move at period $t+1$ only depends on period t's outcome path and period $t+1$'s ongoing path.

Another interpretation of the seller's equilibrium strategy of the repeated game is that for all t, for all $(o_{t-\tau})_{\tau=1}^{t-1} \in O^{t-1}$, and for all $x \in X_l$, she chooses an outcome path of the stage game after x among the ones that are available given the short-run player's strategy. To formalize this idea, for all $o_1 \in O$ and $x \in X_l$, let $Q^{\sigma_s}(o_1; x)$ denote the set of all the outcome paths (of the stage game after x) that are available given σ_s. Notice that $Q^{\sigma_s}(o_1; x)$ does not depend on $(o_{t-\tau})_{\tau=1}^{t-2}$ since the short-run player's move at period t only depends on period $t-1$'s outcome path and

---

[5]If the initial node belongs to the long-run player, $y(\sigma_s)$ is the initial node for all σ_s and $(o(x), a_l, q^e((o_{t-\tau})_{\tau=1}^{t-1}; \overrightarrow{a_l}))..$

period t's ongoing path. Then an equivalent way of writing the problem in the last paragraph is : for all t, for all $(o_{t-\tau})_{\tau=1}^{t-1} \in O^{t-1}$, and for all $x \in X_t$, $q^e((o_{t-\tau})_{\tau=1}^{t-1}; x)$ solves

$$V((o_{t-\tau})_{\tau=1}^{t-1}; x) \tag{4.0}$$

$$= \max_{q(x) \in Q^{\sigma_*}(o_1; x)} u_s(o(x), q(x)) + \delta V((o_{t-\tau})_{\tau=1}^{t-1}, (o(x), q(x)); y^{\sigma_*}).$$

Since $Q^{\sigma_*}(o_1 : x)$ only depends on o_1 and y^{σ_*} only depends on $(o(x), q^{\sigma_*}(x))$, notice that

$$V((o_{t-\tau})_{\tau=1}^{t-2}, o_1; x) = V((o'_{t-\tau})_{\tau=1}^{t-2}, o_1; x)$$

$$\text{for } (o_{t-\tau})_{\tau=1}^{t-2} \neq (o'_{t-\tau})_{\tau=1}^{t-2}, \text{ for all } o_1 \in O \text{ and } x \in X_t \cup \hat{Z}.$$

We redefine $V : O \times (X_t \cup \hat{Z})$ by letting $V(o_1; x) = V((o_{t-\tau})_{\tau=1}^{t-2}, o_1; x)$ for all $(o_{t-\tau})_{\tau=1}^{t-2} \in O^{t-2}$. Then (4.0) can be rewritten as : for each $o_1 \in O$ and for each $x \in X_t$, $q^e((o_{t-\tau})_{\tau=1}^{t-2}, o_1; x)$ solves

$$V(o_1; x) \tag{4.1}$$

$$= \max_{q(x) \in Q^{\sigma_*}(o_1; x)} u_s(o(x), q(x)) + \delta V((o(x), q(x)); y^{\sigma_*}).$$

Case 1. Suppose for all $o_1 \in O$ and for all $x \in X_t$, (4.1) has a unique solution.

In this case, we show that the repetition of the subgame perfect equilibrium is the only equilibrium outcome. That is, for all t and $(o_{t-\tau})_{\tau=1}^{t-1} \in O^{t-1}$, $q^e((o_{t-\tau})_{\tau=1}^{t-1}; x) = q^*(x)$ for all $x \in X$.

Before preceding, let us partition X_l, the set of nodes for the long-run player, into $\{X_l^j\}_{j=0,1,...,J}$ for some integer J in the following way[6]:

$$X_l^0 = \{x \in X_l | S(x) \cap X_l = \emptyset\},$$

$$X_l^1 = \{x \in X_l - X_l^0 | S(x) \cap X_l \subset X_l^0\},$$

$$\vdots$$

$$X_l^j = \{x \in X_l - (\cup_{i=0,1,...,j-1} X_l^i) | S(x) \cap X_l \subset (\cup_{i=0,1,...,j-1} X_l^i)\},$$

$$\text{for } j = 2, 3, ..., J.$$

This partitioning gives a partial order on X_l by specifying the order of the long-run player's move. If $x \in X_l^j$ and $x' \in X_l^k$ with $J \geq k > j$, either x' or other node $x'' \in X_l^k$ precedes x. X_l^0 is the set of nodes for the long-run player whose successors are either terminal nodes or nodes for the short-run players. So, the long-run player at $x \in X_l^0$, whatever her decision is, will not move again in the stage game. If $x \in X_l^j$, $j \neq 0$, there are succeeding nodes belong to the long-run players depending on her move at x and the following short-run player's move, but there is always at least a succeeding node in X_l^{j-1}. If the initial node belongs to the long-run player, X_l^J is the singleton set consisting of the initial node. If $x \in X_l^J$, $P(x) \cap X_l = \emptyset$.

---

[6]It is easy to check that $\{X_l^j\}_{j=0,1,2,...,J}$ is indeed a partition. In particular, for all $j = 0, 1, 2, ..., J$, $X_l^j \neq \emptyset$. Also, if $X_l^i = \emptyset$, then $X_l^{i+1} = \emptyset$. Hence J is well-defined.

Given $\{X_l^j\}_{j=0,1,...,J}$, we can also partition X_s, the set of nodes for the short-run player, into partition, $\{X_s^j\}_{j=0,1,...,J+1}$, as follows[7]:

$$X_s^0 = \{x \in X_s | S(x) \cap X_l = \emptyset\},$$

$$X_s^1 = \{x \in X_s - X_s^0 | S(x) \cap X_l \subset X_l^0\},$$

$$\vdots$$

$$X_s^j = \{x \in X_s - (\cup_{i=0,1,...,j-1}X_s^i) | S(x) \cap X_l \subset (\cup_{i=0,1,...,j-1}X_l^i)\},$$

$$\text{for } j = 2, 3, ..., J + 1.$$

The short-run player's move at $x \in X_s^0$, whatever it is, will not reach a node for the long-run player. If $x \in X_s^j$, $j \neq 0$, there is always at least one succeeding node for the long-run player in X_l^{j-1}. For $x \in X_s^{J+1}$, there are no preceding nodes for the long-run players. Especially, if the initial node belongs to the long-run player, $X_s^{J+1} = \emptyset$.

(At X_s^0) For all $o_1 \in O$ and for all $x \in X_s^0$, we can use the typical backward induction argument to establish that the short-run player chooses an action that maximizes his stage game payoff given his decision at the succeeding nodes, regardless of the outcome of the last stage game. So, for all $(o_{t-\tau})_{\tau=1}^{t-1} \in O^{t-1}$ and for all $x \in X_s^0$, $q^e((o_{t-\tau})_{\tau=1}^{t-1}; x)$ only depends on x and will be denoted by $q^e(x)$. Note $q^e(x) = q^*(x)$.

(At X_l^0) Now consider a node $x \in X_l^0$. Notice that $V(o_1; x) = V(o_1'; x)$ for all $o_1 \neq o_1'$, since

---

[7]For some j, $X_s^j = \emptyset$.

123

$$V(o_1; x) = V(o_1'; x) =$$

$$\max_{a_l \in A(x)} u_l(o(x), a_l, q^e(\overrightarrow{a_l})) + \delta V((o(x), a_l, q^e(\overrightarrow{a_l})); y(\sigma_s)). \tag{4.2}$$

Define $V(x) = V(o_1; x)$ for all $o_1 \in O$. Then (4.2) can be rewritten as[8] :

$$V(x) = \max_{a_l \in A(x)} u_l(o(x), a_l, q^e(\overrightarrow{a_l})) + \delta V(y(\sigma_s)). \tag{4.3}$$

Since (4.1) has a unique solution, (4.3) also has a unique solution, a_l^*. Notice this solution only depends on the outcome of the current period's stage game, $o(x)$, because $q^e(\overrightarrow{a_l})$ does not depend on o_1 and $y(\sigma_s)$ only depends on $(o(x), a_l, q^e(\overrightarrow{a_l}))$. If we denote it by $\sigma_l(x)$, for all $(o_{t-\tau})_{\tau=1}^{t-1} \in O^{t-1}$ and for all $x \in X_l^0$, $q^e((o_{t-\tau})_{\tau=1}^{t-1}; x) = (\sigma_l(x), q^e(\overrightarrow{\sigma_l(x)})) \in X_l^0 \times (\times_{y \in X_s^0} A(y))$. We write $q^e((o_{t-\tau})_{\tau=1}^{t-1}; x)$ by $q^e(x)$.

(At X_s^1) Consider a node $x \in X_s^1$ that does not have a succeeding node X_s^1. The short-run player at node $x \in X_s^1$ chooses an action $a_s \in A(x)$ that maximizes $u_s(o(x), a_s, q^e(\overrightarrow{a_s}))$.[9] By the genericity of the stage game payoff, the action is unique and only depends on the outcome of the current period's stage game. Denote it by $\sigma_s(x)$. Let $q^e(x) = (\sigma_s(x), q^e(\overrightarrow{\sigma_s(x)}))$ denote the uniquely determined equilibrium path after x. For $x \in X_s^1$ that has a succeeding node X_s^1, we can again continue the argument to establish that for all $x \in X_s^1$, $\sigma_s(o_1, x)$ is the same action for all o_1. Let $\sigma_s(x)$ denote the action. Then, for all $x \in X_s^1$, we can define by $q^e(x) = (\sigma_s(x), q^e(\overrightarrow{\sigma_s(x)}))$ the uniquely determined equilibrium path after x.

---

[8] By the definition of X_l^0, for all $x \in X_l^0$ and for all $a_l \in A(x)$, either $\overrightarrow{a_l} \in Z$ or $\overrightarrow{a_l} \in X_s^0$. Hence $q^e(\overrightarrow{a_l})$ is well-defined.

[9] Note either $\overrightarrow{a_s} \in X_l^0$ or $\overrightarrow{a_s} \in X_s^0$. Thus $q^e(\overrightarrow{a_s})$ is well-defined

(At X_l^1) Consider a node $x \in X_l^1$. Given $q^*(\overrightarrow{a_l})$ for each $a_l \in A(x)$, defined in (At X_s^1) and possibly in (At X_s^0) and (At X_l^0), we can give the exact same argument as in (At X_l^0) to establish that uniqueness of the solution for (4.1) guarantees the unique solution for (4.3). Denote the solution by $\sigma_l(x)$ and conclude that for all $(o_{t-\tau})_{\tau=1}^{t-1} \in O^{t-1}$ and for all $x \in X_l^0$, $q^e((o_{t-\tau})_{\tau=1}^{t-1}; x) = (\sigma_l(x), q^e(\overrightarrow{\sigma_l(x)})) \in X_l^0 \times (\times_{y \in X_s^0} A(y))$. We write $q^e((o_{t-\tau})_{\tau=1}^{t-1}; x)$ by $q^e(x)$.

$$\vdots$$

(At X_l^J) By continuing the previous argument in the previous steps, conclude that for all $(o_{t-\tau})_{\tau=1}^{t-1} \in O^{t-1}$ and for all $x \in X_l^0$, $q^e((o_{t-\tau})_{\tau=1}^{t-1}; x)$ only depends on x and denote it by $q^e(x)$.

Suppose first X_l^J is the singleton set consisting of the initial node, i.e. the initial node belongs to the long-run player. Then (4.3) immediately implies that for all $x \in X_l^0$, $\sigma_l(x) = \max_{a_l \in A(x)} u_l(o(x), a_l, q^e(\overrightarrow{a_l}))$ since $y(\sigma_s)$ in (4.3) is the initial node for all $(o(x), a_l, q^e(\overrightarrow{a_l}))$. Since $q^e(\overrightarrow{a_l}) = q^*(\overrightarrow{a_l})$, $\sigma_l(x)$ is the same as in the subgame perfect equilibrium of the stage game. Hence, for all $x \in X_l^0$ $q^e(x) = q^*(x)$. Given the equilibrium paths, $q^e(x)$, defined in (At X_i^j) for all $x \in X_i^j, i = l, s, j = 0, 1, ..., J$, continuing the backward induction argument establishes that for all $x \in X_i^j$, $q^e(x) = q^*(x)$.

If the initial node does not belongs to the long-run player, so that $X_s^{J+1} \neq \emptyset$, follow the argument in (At X_s^1) to argue that for all $x \in X_s^{J+1}$, $\sigma_s(x) = \max_{a_s \in A(x)} u_s(o(x), a_s, q^e(\overrightarrow{a_s}))$. Then for all $x \in X_l^0, \sigma_l(x) = \max_{a_l \in A(x)} u_l(o(x), a_l, q^e(\overrightarrow{a_l}))$ since we have now established that for all $x \in X_s, \sigma_s$ does not depend on the out-

come of the previous period's stage game and thus $y(\sigma_s)$ in (4.3) does not depend on a_l. We apply the same argument as in the last paragraph to conclude that the only equilibrium outcome of the repeated game is just the repetition of the subgame perfect equilibrium of the stage game.

Case 2. Suppose now there exists at least one $(o_1; x) \in O \times X_l$ for which (4.1) has a multiple solution.

For all $x \in X_l \cup \widehat{Z}$, let $U(x) = \{u_l(o(x), q(x))\}_{q(x) \in Q^{\sigma_s}(x)}$ and $U = \cup_{x \in X_l \cup \widehat{Z}} U(x)$. Consider a correspondence $\widehat{\sigma}$, where

$$\widehat{\sigma} : O \times (X_l \cup \widehat{Z}) \Rightarrow U \times O,$$

that generates a simultaneous equation system of the following form by specifying combinations of u and j for each i : For all $o \in O$ and for all $x \in X_l \cup \widehat{Z}$,

$$V(o; x) = u + \delta V(o', x')$$

$$\text{for some } u \in U(x),$$

$$\text{for some } (o', x') \in O \times (X_l \cup \widehat{Z}).$$

By using the same argument on this correspondence as in Case 2 of Claim, we can show there are at most finitely many δ's for which there exists a pure strategy equilibrium that allows Case 2. ∎

3.5 Discussion

3.5.1 Generic Result

The proof of the proposition basically consists of two steps. We first consider an arbitrary equilibrium with the property that for each outcome path the long-run player strictly prefers one action to the other available actions. We could then develop the argument that in this equilibrium, the long-run player's decision depends on none of the previous periods' outcomes, thus nor do the short-run players' decisions. This history-independence immediately establishes that the equilibrium satisfying the property is the repetition of the subgame perfect equilibrium of the stage game. As a second step, we examine existence of a pure strategy equilibrium that has an outcome path after which the long-run player is indifferent between choosing at least two distinct actions. We find out that the number of discount factors for which such an equilibrium may exist is at most finite.

The proposition is only a generic result because we cannot establish the history-independence for the finite number of discount factors. If the long-run player is indifferent between two distinct actions after each of two different outcome paths, we can freely specify any of those two actions after the two different paths, so that the long-run player's action choice can be different after the two different paths. This dependence of the long-run player's action choice on the previous outcome paths generates the short-run players' dependence as well, and thus may allow us to construct a non-trivial equilibrium by resolving the indifference in an appropriate way.

For instance, in the buyer-seller game introduced in Section 2, suppose the buyers only observe the last period's outcome. Notice that in any non-trivial equi-

librium, the buyer has to purchase if and only if the seller produced high quality in the last period. Given the buyer's strategy, the seller is indifferent between high-quality production and low-quality production if and only if $\delta = 1/2$. Using this indifference, we could construct an equilibrium in which the seller produces high-quality in every period, by specifying high-quality production after she produced high quality in the last period, but low-quality production in case she produced low quality in the last period or the last period's buyer did not purchase.

3.5.2 Mixed strategy equilibria

The generic result of the proposition does not apply to mixed strategy equilibria. By allowing the short-run players to use mixed strategies that depend on the discount factor, we can construct an equilibrium that allows certain outcome paths after which the long-run player is indifferent between choosing two distinct actions for generic discount factors. Then the possibility of existence of a non-trivial equilibrium is left open for those generic discount factors, again by breaking the indifference in an appropriate way.

In the example of the buyer-seller game, a mixed strategy equilibrium supporting the trade outcome exists for all $\delta > 1/2$: The seller starts to produce high quality in the first period and continues to do so if she did so in the last period. She randomizes between high-quality and low-quality production with probability $1/2$ if the buyer in the last period did not purchase and the current buyer purchased. She produces low quality if she produced low quality in the last period but the current buyer purchased. On the other hand, the buyers purchase if the seller produced high quality in the last period, and do not purchase if the seller produced low quality in the last period. They randomize the purchasing decision after the

last buyer did not purchase. They purchase with probability $(2\delta - 1)/\delta$ and refuse to buy with the other probability. Given the buyers' randomization, the seller is indifferent between high-quality and low-quality production after all the outcome paths, making the seller's strategy trivially optimal. The buyers' strategy is also optimal given the seller's strategy. In particular, their randomization is optimal given the seller's randomization because they get the payoff of 0, regardless of their purchasing decision, after the last buyer refused to buy.

3.5.3 General extensive-form games

The argument in the proposition crucially depends on the structure of the stage game, games of perfect information. In the example of the buyer-seller game, for instance, suppose the stage game is the following simultaneous-move version instead of the sequential one. So the seller is not informed of the current buyer's purchasing behavior when she makes her quality choice.

buyer

		B	N
seller	H	1, 1	0, 0
	L	2, −1	0, 0

Then we can easily find an equilibrium supporting the trade outcome by imposing a typical punishment scheme against the seller who produced low quality in the last period: The buyer purchases and the seller produces high quality, in the first period and only after the last period's outcome is that the buyer purchased and the seller produced high quality. After the other last period's outcomes, the seller produces low quality and the buyers refuse to buy.

129

In this equilibrium, it is optimal for the seller to produce low quality after she has done so, in which case she expects the current buyer not to purchase and thus her high-quality production cannot stop the subsequent buyers' boycotts. Hence the buyers' punishment is also optimal given the expectation that the seller who produced low quality in the last period will do so again. In the sequential-move version, this kind of punishment is not credible. The seller will of course produce high quality along the equilibrium path if she expects severe punishments after her deviation. However, the seller who produced low quality in the last period has no reason to produce low-quality if the current buyer purchases, because the succeeding buyer will not observe the original low-quality production and thus will purchase if she produces high quality in the current period. Then the buyer, despite the seller's low-quality production in the last period, will not boycott her since she is expected to produce high quality. This forgiveness of the buyers eventually leads the seller to produce low quality, repudiating any non-trivial equilibrium.

Another interesting counterexample of a general extensive-form stage game is the two-buyer version of the original sequential-move buyer-seller game. Suppose first two buyers independently and simultaneously have to decide whether purchase a good from the seller or not. If at least one buyer purchases, the seller can provide high quality or low quality. The same quality applies to both buyers. If none of them purchases, the game ends and all three players receive zero. Each buyer's payoffs do not depend on the other buyer's purchasing decision and are the same as in the original sequential-move game while the seller stage game payoffs are the sum of the payoffs from the original sequential-move game. So the seller gets a payoff of 4 if both buyers purchase and she produces low quality, 2 if both purchase and she produces high quality, 2 if only one purchases and she produces low quality, 1

130

if only one purchases and she produces high quality, and 0 if none purchases.

As in the previous example, suppose each buyer observes the outcome of the last period's stage game while the seller observes all previous outcomes. Then we can also construct an equilibrium in which each buyer purchases and the seller produces high quality. The seller produces high quality in the first period and continues to do so only after both purchased and she produced high quality in the last period. If she produced low quality or at least one buyer did not purchase in the last period, she produce low quality. The buyers purchase in the first period and only after both buyers in the last period purchased and the seller produced low quality. If the seller produced low quality or at least one buyer did not purchase in the last period.

In this equilibrium, the buyer will punish the seller who observed the seller produce low quality in the last period by refusing to purchase from her because the other buyer will not purchase and so she is expected to produce low quality in case only he purchases. On the other hand, the seller who produced low quality in the last period will do again if only one buyer unexpectedly purchases, in which case both buyers in the next period refuse to purchase regardless of her quality choice and so in the subsequent future periods. Hence the buyers' punishments against low-quality production is indeed credible, contrary to the single buyer case. However, the buyers' purchasing decisions are made sequentially, the result of the proposition is still effective since the stage game is then a game of perfect information.

Bibliography

[1] Ahn, I. and M. Suominen (1996). Word-of-Mouth Communication and Community Enforcement. Caress Working Paper #96-02, University of Pennsylvania.

[2] Allen F. (1984). Reputation and Product Quality. *Rand Journal of Economics* 15: 311-327.

[3] Benoit, J. P. and V. Krishna (1985). Finitely Repeated Games. *Econometrica* 53: 890-904.

[4] Ben-Porath, E. and M. Kahnemam (1996). Communication in Repeated Games with Private Information. *Journal of Economic Theory*, forthcoming.

[5] Ellison G. (1994). Cooperation in The Prisoners Dilemma with Anonymous Random Matching. *Review of Economic Studies* 61: 567-588.

[6] Fudenberg, D., D. Kreps and E. Maskin (1990). Repeated Games with Long-run and Short-run Players. *Review of Economic Studies* 57: 555-574.

[7] Fudenberg, D. and D. Levine (1991). An Approximate Folk Theorem with Imperfect Private Information. *Journal of Economic Theory* 54: 26-47.

[8] Fudenberg, D. and D. Levine (1994). Efficiency and Observability with Long-run and Short-run players. *Journal of Economic Theory* 62: 103-135.

[9] Fudenberg, D., D. Levine. and E. Maskin (1994). The Folk Theorem with Imperfect Public Information. *Econometrica* 62: 997-1040.

[10] Fudenberg, D. and E. Maskin (1986). Folk Theorems for Repeated Games with Discounting and Incomplete Information. *Econometrica* 54: 533-554.

[11] Fudenberg, D. and J. Tirole (1991). *Game Theory.* Cambridge: M.I.T. Press.

[12] Greif, A. (1993). Contract Enforceability and Economic Institutions in Early Trade: The Magrabi Traders' Coalition. *The American Economic Review* 83: 525-548.

[13] Greif, A. (1994). Cultural Beliefs and The Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies. *Journal of Political Economy* 102: 912-950.

[14] Greif, A., P. Milgrom and Weingast, B (1994). Coordination, Commitment, and Enforcement: The Case of The Merchant Quild. *Journal of Political Economy* 102: 745-776.

[15] Harrington, J. (1995). Cooperation in a One-shot Prisoners' Dilemma. *Games and Economic Behavior* 8: 364-377.

[16] Kalai, E., D. Samet and W. Stanford (1988). A Note on Reactive Equilibria in the Discounted Prisoner's Dilemma and Associate Games. *International Journal of Game Theory* 17: 177-186.

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

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

[19] Kreps, D. (1990). "Corporate Culture and Economic Theory," in *Rational Perspectives on Political Science*, ed. J. Alt and K. Shepsle: Harvard University Press.

[20] Kreps, D. and R. Wilson (1982). Sequential Equilibria. *Econometrica* 50: 863-894.

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

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

[23] Standford, W. (1986). Subgame Perfect Reaction Function Equilibria in Discounted Duopoly Supergames Are Trivial. *Journal of Economic Theory* 39: 226-232.

[24] Tiffany, F. (1988). "The Role of Memory in Dynamic Games of Intertemporal Transfers," in *Ph.D. Dissertation*, University of Pennsylvania.

[25] Verboven, F. (1994). Localized Competition, Multimarket Operation and Collusive Behavior. mimeo.