Community Enforcement when Players Observe Partners' Past Play* (Job Market Paper)

Satoru Takahashi†

February 4, 2007

Abstract

I investigate whether a community can sustain cooperation in the repeated prisoner's dilemma by having cheaters sanctioned not by their victims but by third parties. Motivated by systems of credit history recording, online feedback systems, and some experimental settings, I assume that players can access information about their partners' past play for free, but that acquiring information about their partners' past partners' past play is prohibitively costly. In this setting, even though players cannot distinguish cheaters from those who punish cheaters, I show that any level of cooperation can be sustained by an equilibrium. The equilibrium I construct has the following two properties: every player chooses his actions independently of his own record of play, and he is indifferent between cooperation and defection at all histories. This equilibrium carries over to the finite-population setting and is robust to noise in the process of choosing actions or of recording past play. The technique of equilibrium construction is applied to more general stage games. I also analyze the possibility of cooperation either when players are required to have strict incentives to follow equilibrium strategies or when only summary statistics of records are stored in the community.

*I am grateful to Drew Fudenberg and Attila Ambrus for their guidance. I am also grateful to Bo Chen, Alex Gelber, Jerry Green, Fuhito Kojima, George Mailath, Mihai Manea, Parag Pathak, Tiantian Qiu, Tadashi Sekiguchi, Yuichi Yamamoto, and seminar participants at the Institute for Advanced Study for their helpful comments.

†Affiliation: Department of Economics, Harvard University. Address: 1875 Cambridge Street, Cambridge, MA 02138, USA. E-mail address: stakahas@fas.harvard.edu

1 Introduction

Cooperation can be sustained through long-term relationship. One way to sustain cooperation is by *personal enforcement*: A cooperator is rewarded by the recipient, and a cheater is revenged by the victim. Personal enforcement is a powerful mechanism especially in a small community where each person interacts with the same partner frequently. In a large community where people meet various partners over time, however, punishment by victims may not create an enough incentive to cooperate. In such a case, personal enforcement is replaced by another mechanism called *community enforcement*, which prescribes third parties to reward cooperators and punish cheaters. Since third parties are usually informationally inferior to parties concerned, community enforcement becomes effective only with extra information transmission within the community.

In the context of the infinitely repeated prisoner's dilemma in a continuum of population, I investigate whether first-order information is sufficient for sustaining cooperation by community enforcement. In this paper, first-order information means information about current partners' past play, whereas second-order information means information about current partners' past partners' past play, third-order information means information about current partners' past partners' past partners' past play, and so on. Without higher-order information, one cannot distinguish cheaters from those who punished cheaters. This makes it difficult to keep punishers' incentives. Indeed, the literature of community enforcement began with a negative result. Rosenthal (1979) provided an example of prisoner's dilemma in which, if each player only observes his partner's action at the previous period, then cooperation can be sustained by a pure-strategy equilibrium only when the discount factor takes a specific value.

Despite this difficulty, in Section 3, I show that first-order information is sufficient for sustaining cooperation if the discount factor is large enough. The equilibrium I construct satisfies two properties: players choose actions independently of their own records of play, and they are indifferent between cooperation and defection at all histories.[1] Due to the first property, a player's continuation payoff does not depend on who he meets or what record of play his partner has. This simplifies the analysis since I do not need to track distributions of records in the population. Moreover, since each player is indifferent between cooperation and defection, he has a weak incentive to reward or punish his partner based on the partner's record of play. I choose the

[1]These properties are closely related to the idea of "belief-free" equilibria in the literature on repeated games with private monitoring. See Subsections 3.2, 3,5, 3,6, and 5.3.

amount of reward/punishment so that the partner actually becomes indifferent between cooperation and defection. I do so by controlling the probability of cooperation carefully. In contrast, a pure strategy with bounded records as in Rosenthal (1979) can make players indifferent only under non-generic parameter values.

I analyze the robustness of this equilibrium in several ways. Subsection 3.3 introduces a one-time shock to the model, and investigates whether the community eventually goes back to the cooperation phase even after the shock. Subsection 3.4 checks the robustness to noise in the process of choosing actions or of recording past play. Subsection 3.5 shows that the equilibrium strategy, combined with any consistent belief system, forms a sequential equilibrium in the corresponding finite-population model. Moreover, the threshold of the discount factor sufficient for sustaining cooperation does not vary with the size of the population. This is in a clear contrast with Kandori (1992) and Ellison (1994). They analyzed the finite-population model when each player only observes the outcomes of matches in which he directly engaged before. As Kandori showed first and Ellison extended later, even in this minimum level of information transmission, there exist so-called "contagious" equilibria that sustain cooperation. See also Harrington (1995). For contagious equilibria to exist, the discount factor needs to be above a threshold. The threshold depends on the population size, and converges to 1 as the population size goes to the infinity. In the limit where the population is a continuum, there is no equilibrium that sustains cooperation. Thus my results show that an institution that transmits first-order information improves the possibility of cooperation even in a large finite population. Subsection 3.6 applies the technique of equilibrium construction to more general stage games.

The equilibrium I propose makes use of two somewhat extreme properties: heavy use of indifference conditions and dependency on the details of records. These issues are discussed in the following sections. Section 4 requires players to have strict incentives to follow equilibrium strategies. Then I show that, depending on payoff parameters, cooperation may or may not be sustained. However, if mixed strategies are allowed as long as mixing probabilities do not depend on payoff-irrelevant histories, then approximately full cooperation can be sustained for generic parameter values. Section 5 assumes that the community stores only summary statistics of records. Either when players observe the number of cooperation in their partners' records or when a finite bound is imposed on the length of records, I show that, under a restriction on payoff parameters, there exists an equilibrium that sustains cooperation. This parameter restriction can be dropped if not only first-order information but also second-order information is available. I also show that Rosenthal's (1979) non-genericity result is resolved if either mixed strategies or unboundedly long records are allowed.

Section 6 considers a version of the repeated prisoner's dilemma game with the possibility of exit. If the reservation payoff is low and the discount factor is high enough, then I can construct an equilibrium that uses the outside option as punishment.

There are many systems in the real world that rely only on first-order information. An example is credit histories of consumers recorded by credit bureaus (Klein (1992)). Online feedback systems are another example. At eBay, for example, after each transaction, both the seller and the buyer can post a rating (positive, negative, or neutral) and a short comment (Dellarocas (2003)). Customers are asked "to leave only fair and factual comments and ratings that relate to a specific transaction you have with your trading partner."[2] If customers actually follow this instruction, feedback scores contain only first-order information.[3] These systems use only first-order information partly because it is costly for a community to store and transmit higher-order information, and partly because it is cognitively demanding for community members to process such information. Milinski et al. (2001) reported that, in their experiment, subjects who were given first- and second-order information needed significantly more time to respond than those who were given only first-order information. Moreover, subjects with higher-order information often failed to distinguish simple cheating from punishment.[4]

My results are consistent with experimental findings. In many experiments on community enforcement, overall cooperation rates are higher in sessions with first-order information than in sessions with no information. Moreover, if first-order information is available, then subjects tend to use this information to discriminate between cooperators and defectors. See the survey article by Nowak and Sigmund (2005).

Related Literature

Belief-Free Equilibria In two-player repeated prisoner's dilemma games with perfect monitoring, Piccione (2002) and Ely and Välimäki (2002) con-

[2]http://pages.ebay.com/help/feedback/questions/feedback.html

[3]Higher-order information is, however, available through feedback comments.

[4]But see Panchanathan and Boyd (2003) and Bolton et al. (2005). Panchanathan and Boyd (2003) argued that people do not directly use higher-order information (sequences of actions) as in Milinski et al. (2001), but rely on its summary statistics such as "good" or "bad." Conducting an experiment similar to Milinski et al. (2001), Bolton et al. (2005) found that second-order information promoted cooperation more than first-order information did.

structed equilibria in which each player is indifferent between cooperation and defection at every history no matter what private signals the other player has received about his opponent's actions. Such equilibria are called *belief-free* equilibria. Due to their robustness to small noise in private signals, belief-free equilibria have played an important role in the literature of repeated games with private monitoring. See also Ely et al. (2004, 2005), Kandori and Obara (2006), Hörner and Olszewski (2006), and Yamamoto (2006).

In this paper, I argue that belief-free equilibria have another useful application in the literature on community enforcement. From the viewpoint of private-monitoring repeated games, what I do in this paper is to explain that a belief-free equilibrium can be interpreted as an equilibrium strategy in a community enforcement model (a repeated game with anonymous and random matching) if sufficiently rich information is transmitted among players. Specifically, my main result shows that first-order information is rich enough to re-interpret Piccione-type equilibria in the context of community enforcement. See Subsections 3.1, 3.2, 3.5, 3.6, and 5.3 for further discussions.

Community Enforcement Most of the theoretical studies on community enforcement are divided into four categories in terms of the level of information transmission available to the community. The first category is an extreme case in which every player knows the details of all matches in the population. In this situation, since sanction can be implemented by any member in the society, playing with varying partners is not an obstacle to enforcing cooperation (Kandori (1992)). Rather, community enforcement can expand the set of asymmetric equilibrium payoff profiles (Dal Bó (2004)).

The second category is the other extreme case: players are informed only about the outcomes of the matches in which they have been directly involved. As I explained before, in this setting, cooperation can be sustained only if the population is finite, and the threshold for the discount factor depends on the population size (Kandori (1992), Ellison (1994), and Harrington (1995)).

The third category assumes that each player is labeled with a status, which is observable to his partner, and that a player can condition his action on his partner's status. Also, the transition of a player's status over time depends not only on the realized action profile of the stage game and his own status at the previous period, but also on his partner's status. For example, each player in the population is labeled "good" or "bad". A good player remains being good either if he cooperates when his partner has a good status or if he defects when his partner has a bad status (Example 1 in Okuno-Fujiwara and Postlewaite (1995)). Since a player's status at the next period depends on his current partner's status, which, in turn, depends on the partner's previous partner's status, and so on, a status is a summary statistics that contains higher-order information. In this setting, the folk theorem is proved in many games including the prisoner's dilemma (Kandori (1992) and Okuno-Fujiwara and Postlewaite (1995)). See also Dixit (2003a) for his analyses on self-governance and on information intermediaries.

The last category is to use only first-order information. My model and Rosenthal (1979) fall into this category. Rosenthal and Landau (1979) studied the effect of reputation in bargaining. In their model, a player's reputation is a summary statistic of his past play. Unlike in my model, the major role of the reputation is not sanction of deviators, but a coordination device that determines which party should obtain the larger proportion of a pie. Klein (1992) studied the role of credit bureaus in the setting of repeated games between consumers and firms. A credit bureau is an institute that records whether each consumer has ever defaulted or not. Klein analyzed a one-sided incentive problem, whereas the prisoner's dilemma is a two-sided incentive problem. Greif (1993, 1994) and Tirole (1996) also analyzed one-sided incentive problems. Dixit (2003b) investigated the effect of trade expansion in a model where the geographic, economic, or social distance affects the probability of matching, gain from trade, and information transmission. The basic part of his model is the twice repeated prisoner's dilemma with the option to exit. I will discuss the effect of the exit option in details in Section 6.

Besides these papers, Milgrom et al. (1990) assumed the existence of a judge (law merchant). A judge in their model plays two roles. One is to store information. The other is to ask a cheater to pay a certain amount of money to the partner. The payment is voluntary, but the information about remaining unpaid judgments is recorded and available to other players. See also Dixit (2003a) for his analysis on enforcement intermediaries.

Evolutionary biologists have built a model of indirect reciprocity based on "image scoring" (Nowak and Sigmund (1998a, b)). A person's image score is a publicly observable number that counts how many times he has cooperated minus how many times he has defected. Thus image scores are a summary statistics of first-order information. Since many works by evolutionary biologists, including these papers, use evolutionarily stable strategies (ESS) or dynamical stability as an equilibrium refinement, their analysis differs from economists' analysis that emphasizes credibility of threat and sequential rationality.

2 Repeated Prisoner's Dilemma

Random matching I consider a continuum of players who repeatedly play a stage game with varying partners.[5] A matching in this population is a function $m\colon [0,1] \to [0,1]$ such that $m(m(i)) = i \neq m(i)$ for all $i \in [0,1]$. At each period $t = 1, 2, \dots$, matching m_t is randomly drawn independently across time, and every player $i \in [0,1]$ plays a stage game with his partner at this period, $m_t(i)$.[6]

Stage game This paper uses prisoner's dilemma as a stage game except in Subsection 3.6 and Section 6. Subsection 3.6 analyzes a general stage game; Section 6 allows players to opt out of prisoner's dilemma. In the prisoner's dilemma game, players can cooperate (C) or defect (D). Payoffs are given by

	C	D
C	$1, 1$	$-l, 1+g$
D	$1+g, -l$	$0, 0$

$g,\, l > 0.$

[5]Although this paper uses a single-population model, it is straightforward to modify it to a setting where a player and his partner at each period are chosen from different populations.

[6]There have been well-known difficulties (so-called "measurability problem") in providing a probability-theoretic foundation to the idea of the "law of large numbers" for a continuum of random variables (Judd (1985) and Feldman and Gilles (1985)). As a consequence of the measurability problem, in their footnote 4, McLennan and Sonnenschein (1991) proved an impossibility theorem for the existence of random matching if the population is $[0,1]$ with the Borel σ-algebra and the Lebesgue measure. These difficulties were recently resolved by Sun (2006) and Duffie and Sun (2004, 2006). More specifically, Duffie and Sun (2004, 2006) showed the existence of random matching m on an atomless probability space $([0,1], \mathcal{I}, \lambda)$ such that

(1) (i) for any realization of m, $\lambda(m(E)) = \lambda(E)$ holds for any $E \in \mathcal{I}$, (ii) $\mathrm{Prob}(m(i) \in E) = \lambda(E)$ holds for any $i \in [0,1]$ and $E \in \mathcal{I}$, and (iii) for any E, $F \in \mathcal{I}$, $\lambda(E \cap m^{-1}(F)) = \lambda(E)\lambda(F)$ holds almost surely, and

(2) for any $\mathcal{I}$-measurable type function α from $[0,1]$ to a finite set of types, for λ-almost every $i \in [0,1]$, $\alpha(m(i))$ and $\alpha(m(j))$ are stochastically independent for λ-almost every $j \in [0,1]$.

Especially, a player is matched with a specific partner with probability 0. See also Alós-Ferrer (1999) for another construction of random matching. His construction is simpler but depends on type functions.

In this paper, I do not use a population with countably many players since there is no atomless probability distribution on a countably infinite set.

In the case of $g > l$, the more likely a player is to cooperate, the more willing his partner is to defect, and vise versa for $g < l$.[7]

Let $A = \{C, D\}$ denote the action set, and $u\colon A^2 \to \mathbb{R}$ denote the payoff function, where $u(a, \bar{a})$ is the payoff of a player who chooses action $a \in A$ when his partner chooses action $\bar{a} \in A$. For a finite set X, let $\Delta(X)$ be the set of probability distributions on X. With a slight abuse of notation, a point $x \in X$ is identified with the distribution with unit mass on the point x. The domain of u is extended to $(\Delta(A))^2$ by the expected utility hypothesis.

Payoffs Player i's action at period t is denoted by a_{it}. For any realization $\{m_t\}$ of a sequence of random matchings, player i's payoff in the repeated game is given by

$$(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}u(a_{it}, a_{m_t(i),t})$$

with a common discount factor $\delta \in (0, 1)$.

Information Except in Subsection 5.3, information about past play is transmitted through two channels: by direct observation and through a "criminal history repository." The former channel means that each player i observes the outcomes of matches in which he directly engaged before period t, $(a_{is}, a_{m_s(i),s})$ for $s \leq t-1$. The latter channel means that the criminal history repository honestly keeps track of players' actions, and that each player can access his current partner's record of play for free.[8] More precisely, let $a_i^t = (a_{i1}, a_{i2}, \dots, a_{it})$ denote the sequence of player i's actions up to period t. After the matching m_t for period t is determined and before the stage game is played, player i receives a report about his current partner's record of play up to period $t-1$, $a_{m_t(i)}^{t-1}$, from the criminal history repository. I assume that a player's record does not contain information about his identity.[9]

[7]Dixit (2003a) called $g > l$ the *offensive* case and $g < l$ the *defensive* case. The case of $g = l$ is non-generic, but this case arises naturally as a superposition of two "gift-giving games." In a gift-giving game, one player decides to give a gift to the other player or not. If the gift is transferred, the donor pays cost $c > 0$, and the recipient receives benefit $b > c$. The gift-giving game has been used in many experiments. If two players simultaneously decide whether to give a gift or not, the payoff structure of this game is equivalent to prisoner's dilemma with $g = l = c/(b-c)$.

[8]It would be difficult to analyze my model if players had to pay subscription fees or the criminal history repository could cheat. These issues were investigated in other models such as Klein (1992), Milgrom et al. (1990), and Dixit (2003a).

[9]Even if a player's record contained his identity, it would not affect my analysis since I will soon assume *ex ante* symmetry of strategies.

Subsection 5.3 assumes that, at period $t \geq 2$, each player i receives a report about the outcome of his current partner's previous match, $(a_{m_t(i),t-1}, a_{m_{t-1}(m_t(i)),t-1})$, from the criminal history repository.

Strategies Now I can define strategies. In general, player i's action at period t can depend on all information that he has gathered so far. This information includes his own record of play, a_i^{t-1}, his current partner's record of play, $a_{m_t(i)}^{t-1}$, and his past partners' records of play, $a_{m_s(i)}^s$ for $s \leq t-1$. (Record $a_{m_s(i)}^s$ consists of two parts: report $a_{m_s(i)}^{s-1}$ from the criminal history repository and direct observation $a_{m_s(i),s}$.) However, I argue that, in the continuum-population model, player i uses only his and his current partner's records of play, $(a_i^{t-1}, a_{m_t(i)}^{t-1})$, without loss of generality. The other aspects of the information are player i's private information that is not shared by his current or future partners. Although such private information can be correlated with other players' information if they have been matched with player i directly or indirectly (e.g., being matched with a player who was matched with player i before), the number of such players is finite and hence negligible when the population is a continuum. Even if a player chooses actions based on his private information about his past partners' records of play, such information is, from all but finitely many players' points of view, no more than an independent and private randomization device.[10]

Given the above simplification, I define player i's behavior strategy by sequence $\sigma_i = \{\sigma_{it}\}$ with $\sigma_{it} \colon A^{2t-2} \to \Delta(A)$, where $\sigma_{it}(a_i^{t-1}, a_{m_t(i)}^{t-1})$ denotes player i's mixed action at period t. In this paper, I focus on *ex ante* symmetric strategies for simplicity: $\sigma_{it}(a_i^{t-1}, a_{m_t(i)}^{t-1}) = \sigma_{jt}(a_j^{t-1}, a_{m_t(j)}^{t-1})$ whenever $a_i^{t-1} = a_j^{t-1}$ and $a_{m_t(i)}^{t-1} = a_{m_t(j)}^{t-1}$.[11] Thus I write $\sigma_t(a^{t-1}, \bar{a}^{t-1})$ for a player's mixed action at period t when his own record of play is $a^{t-1} = (a_1, \ldots, a_{t-1})$ and his current partner's record of play is $\bar{a}^{t-1} = (\bar{a}_1, \ldots, \bar{a}_{t-1})$.

Distributions of records If I followed Kreps and Wilson (1982) to define sequential equilibria, the next step would be to define beliefs. A player's belief is a probability distribution about who met whom and who played

[10]Thus my continuum-population model cannot explain the tendency that people whose recent partners cooperated are more likely to cooperate than those whose recent partners defected. This tendency was reported in several experimental studies (Engelmann and Fischbacher (2004) and Bolton et al. (2005)).

[11]If different players adopted different strategies, then a player could make inferences about his current partner's identity from the record of play. This information would change the former player's expectation about the partner's action at this period. In this case, unlike in footnote 9, it would affect analysis whether reports from the criminal history repository are anonymous or not.

which action in the past conditional on what he knows. This is a complicated mathematical object. What is relevant in my setting is, however, only the aggregate data: the distribution of records of play in the population. For $t \geq 1$ and $a^t \in A^t$, let $\mu_t(a^t)$ be the fraction of players in the population whose record of play up to period t is a^t, and $\mu_t \in \Delta(A^t)$ be the distribution of records of play up to period t in the population.

The sequence of such distributions, $\mu = \{\mu_t\}$, is determined as follows. At period 1, all players choose action a_1 with probability $\sigma_1(\emptyset)(a_1)$ independently. Since there are a continuum of players, by the "law of large numbers," the fraction of players who choose action a_1 at the first period is $\sigma_1(\emptyset)(a_1)$. Thus I have

$$\mu_1(a_1) = \sigma_1(\emptyset)(a_1). \tag{1}$$

For $t \geq 2$, μ_t can be computed from σ_t and μ_{t-1}. At the beginning of period t, there are $\mu_{t-1}(a^{t-1})$ mass of players with record a^{t-1}. Among them, the fraction of players who meet players with record $\bar{a}^{t-1}$ is $\mu_{t-1}(\bar{a}^{t-1})$, and such players choose action a_t with probability $\sigma_t(a^{t-1}, \bar{a}^{t-1})(a_t)$ independently. Thus I have

$$\mu_t(a^t) = \mu_{t-1}(a^{t-1}) \sum_{\bar{a}^{t-1} \in A^{t-1}} \mu_{t-1}(\bar{a}^{t-1}) \sigma_t(a^{t-1}, \bar{a}^{t-1})(a_t), \tag{2}$$

where $a^t = (a^{t-1}, a_t)$. I say that μ is *generated by* σ if (1) and (2) are satisfied.

By the "law of large numbers", players believe with certainty that the distribution of records of play at the end of period t is equal to μ_t. I assume that this is true even after an off-path history. In other words, no information with finitely many observations is convincing enough for a player to change his belief about the distribution of records of play in the population. This assumption follows the spirit of sequential equilibria for finite games, although my model is not a finite game. For example, consider an equilibrium that sustains cooperation. If a player is matched with another player whose record of play shows defection in the past, the former player resorts to the "trembling-hand theory" to explain his observation. Since trembles by finitely many players are infinitely more likely than trembles by a positive mass of players, he concludes that this off-path history was caused by finitely many players' mistakes, which will not affect his future matches with positive probability.

Since each player's partner is chosen uniformly from the population, he believes that his partner at period t has record of play a^{t-1} with probability $\mu_{t-1}(a^{t-1})$.

Continuation payoffs Let $U_t(\sigma, \bar{\sigma} \mid a^{t-1}, \bar{a}^{t-1}, \mu)$ be the continuation payoff of a player who follows strategy σ when all other players follow strategy $\bar{\sigma}$, given that the former player has record a^{t-1}, his partner at period t has record $\bar{a}^{t-1}$, and the distributions of records in the population are given by μ. By definition, it holds that

$$
\begin{aligned}
&U_t(\sigma, \bar{\sigma} \mid a^{t-1}, \bar{a}^{t-1}, \mu) \\
&= (1-\delta) \sum_{a_t \in A} \sigma_t(a^{t-1}, \bar{a}^{t-1})(a_t) \Bigg(u(a_t, \bar{\sigma}_t(\bar{a}^{t-1}, a^{t-1})) \\
&\quad + \delta \sum_{a_{t+1} \in A, \bar{b}^t \in A^t} \sigma_{t+1}(a^t, \bar{b}^t)(a_{t+1}) \Bigg(u(a_{t+1}, \bar{\sigma}_{t+1}(\bar{b}^t, a^t)) \mu_t(\bar{b}^t) \\
&\quad + \delta \sum_{a_{t+2} \in A, \bar{c}^{t+1} \in A^{t+1}} \sigma_{t+2}(a^{t+1}, \bar{c}^{t+1})(a_{t+2}) \Bigg(u(a_{t+2}, \bar{\sigma}_{t+2}(\bar{c}^{t+1}, a^{t+1})) \mu_{t+1}(\bar{c}^{t+1}) + \cdots \Bigg)\Bigg)\Bigg) \\
&= \sum_{a_t \in A} \sigma_t(a^{t-1}, \bar{a}^{t-1})(a_t) \left((1-\delta) u(a_t, \bar{\sigma}_t(\bar{a}^{t-1}, a^{t-1})) + \delta \sum_{\bar{b}^t \in A^t} U_{t+1}(\sigma, \bar{\sigma} \mid a^t, \bar{b}^t, \mu) \mu_t(\bar{b}^t) \right).
\end{aligned}
$$

The last line is a recursive representation. Consider a player with record a^{t-1} who faces his partner with record $\bar{a}^{t-1}$ at period t. He chooses action a_t with probability $\sigma_t(a^{t-1}, \bar{a}^{t-1})(a_t)$ and obtains short-run payoff $u(a_t, \bar{\sigma}_t(\bar{a}^{t-1}, a^{t-1}))$. At the end of period t, his record is updated to $a^t = (a^{t-1}, a_t)$. Since his next partner is randomly chosen from the population, his continuation payoff is given by the expectation of $U_{t+1}(\sigma, \bar{\sigma} \mid a^t, \bar{b}^t, \mu)$, where $\bar{b}^t$ is distributed according to μ_t.

Equilibria Now I define the equilibrium concept.

Definition 1. A strategy σ^* is a *continuum-population perfect equilibrium* (or *equilibrium*) if, for every $t \geq 1$, every $a^{t-1}, \bar{a}^{t-1} \in A^{t-1}$, and every strategy σ, it holds that

$$U_t(\sigma^*, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu^*) \geq U_t(\sigma, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu^*), \tag{3}$$

where μ^* is generated by σ^*.

In this definition, the same μ^* is used in both sides of (3). It means that each player recognizes that he is negligible in the continuum of population and believes that no change in his own strategy can affect the distribution of records of play in the population.

Note that my definition requires sequential rationality. If I required optimizing behavior only at the initial period, which would correspond to Nash equilibria rather than perfect equilibria, I could easily construct an equilibrium that sustains cooperation on the equilibrium path. For example, consider the pairwise grim-trigger strategy defined as follows.

Definition 2. The *pairwise grim-trigger strategy* is a strategy σ that prescribes each player to cooperate if neither he nor his current partner has ever defected, and to defect otherwise, i.e., $\sigma_t(a^{t-1}, \bar{a}^{t-1}) = C$ if $t = 1$ or ($t \geq 2$ and $a^{t-1} = \bar{a}^{t-1} = (C, \dots, C)$); $\sigma_t(a^{t-1}, \bar{a}^{t-1}) = D$ otherwise.

I add the adjective *pairwise* because a player's action depends only on his and his current partner's records of play.[12] The pairwise grim-trigger strategy forms a Nash equilibrium (not necessarily perfect) if and only if $\delta \geq g/(1+g)$. In Subsection 4.1, I will check under what condition the pairwise grim-trigger strategy forms a perfect equilibrium.

It follows from the standard argument in dynamic programming that sequential rationality is equivalent to the nonexistence of one-shot profitable deviations in pure actions at any history. Namely, σ^* is a continuum-population perfect equilibrium if and only if, for every $t \geq 1$, every a^{t-1}, $\bar{a}^{t-1} \in A^{t-1}$, and every $a_t \in A$, it holds that

$$\begin{aligned} &U_t(\sigma^*, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu^*) \\ &\quad \geq (1-\delta)u(a_t, \sigma_t^*(a^{t-1}, \bar{a}^{t-1})) + \delta \sum_{\bar{b}^t \in A^t} U_{t+1}(\sigma^*, \sigma^* \mid a^t, \bar{b}^t, \mu^*)\mu_t^*(\bar{b}^t), \end{aligned} \tag{4}$$

where μ^* is generated by σ^*.

Unlike the previous literature such as Rosenthal (1979), Rosenthal and Landau (1979), and Okuno-Fujiwara and Postlewaite (1995), I do not impose stationarity on σ^* or μ^* in the definition of equilibria.

3 Independent and Indifferent Equilibria

This section studies a class of equilibria with a simple structure, and shows that this class is large enough to sustain cooperation on the equilibrium path.

[12] The pairwise grim-trigger strategy is different from the "full" grim-trigger strategy, which prescribes players to defect if and only if some player in the population has ever defected. The full grim-trigger strategy is not feasible under the restriction on information transmission in my setting.

3.1 Definitions

I consider equilibrium σ^* with the following properties

independence of own play Strategy σ^* prescribes each player to choose actions with probability independent of his own record of play, i.e., $\sigma_t^*(a^{t-1}, \bar{a}^{t-1}) = \sigma_t^*(b^{t-1}, \bar{a}^{t-1})$ for all $t \geq 1$ and a^{t-1}, b^{t-1}, $\bar{a}^{t-1} \in A^{t-1}$.

indifference at all histories Each player is indifferent between actions C and D at all histories if the other players follow σ^*, i.e., (4) is satisfied with equality for all $t \geq 1$, a^{t-1}, $\bar{a}^{t-1} \in A^{t-1}$, and $a_t \in A$, where μ^* is generated by σ^*.

Note that any strategy that satisfies indifference at all histories is an equilibrium.

Independence of own play simplifies analysis. Under this assumption, if a player has record of play a^{t-1}, his partner chooses an action based only on a^{t-1} no matter what record of play the partner has. Therefore, a player's belief about other players' records of play does not affect his payoffs, i.e., $U_t(\sigma, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu)$ is independent of $\bar{a}^{t-1}$ and μ. Thus equilibrium conditions can be verified without computing distributions of records in the population.

Independence of own play, however, restricts players' behavior. Since a player does not meet the same partner twice or more, he does not have a strict incentive to vary his action non-trivially in response to the information about his current partner's record of play when the partner himself does not use this information. Thus, in order to construct any equilibrium other than the repetition of D, players have to be indifferent between C and D at *some* histories. Here I simply assume indifference at *all* histories. Since players are indifferent between two actions, they have weak incentives to change their mixing probabilities to punish partners. This is how flexibility of players' behavior is restored under the restriction of independence of own play.

A similar technique has been used in the literature of repeated games with private monitoring. In the repeated prisoner's dilemma with private monitoring, Piccione (2002) and Ely and Välimäki (2002) constructed equilibria in which each player is indifferent between cooperation and defection at all histories no matter what private signals the other player has received. As in my case, this simplifies the analysis since there is no need to compute a player's belief about the other player's private histories. Such equilibria are called *belief-free* equilibria (Ely et al. (2005)).

In my setting, indifference at all histories implies independence of own play if $g \neq l$. The reason is simple: if $g \neq l$, then any change in a player's mixed action based on his own record of play causes a non-zero influence on his partner's trade-off between C and D, which violates at least one of the two indifference conditions.

Proposition 1. *If $g \neq l$, then any strategy with indifference at all histories satisfies independence of own play.*

Proof. See Appendix A.1. □

Definition 3. A strategy σ^* is called an *independent and indifferent equilibrium* if it satisfies both independence of own play and indifference at all histories.

By Proposition 1, requiring independence of own play is redundant in the above definition if $g \neq l$.

3.2 Equilibrium Construction

I will show that independent and indifferent equilibria can sustain any level of cooperation for a sufficiently large discount factor.

Proposition 2. *Suppose that $\delta \geq \max(g/(1+g), l/(1+l))$. Then there is an independent and indifferent equilibrium with symmetric payoff x if and only if $x \in [0,1]$.*

If $g - l \leq 1$, then $[0,1]$ is the set of feasible payoffs under symmetric strategies. Thus, in this case, Proposition 2 characterizes the set of symmetric (not necessarily independent and indifferent) equilibrium payoffs for any large discount factor. For the case that $g - l > 1$, I leave it open whether there exists a symmetric equilibrium that achieves a payoff higher than 1 by using *ex post* asymmetric outcomes (C, D) and (D, C).

For any $x \in [0,1]$, I will construct an independent and indifferent equilibrium with symmetric payoff x as follows:

- At period t, players with record a^{t-1} are assigned with "target payoff" $V_t(a^{t-1})$. Set $V_1(\emptyset) = x$. I will construct an equilibrium under which all players obtain their target payoffs as their continuation payoffs.
- A player who meets another player with record $\bar{a}^{t-1}$ chooses action C with probability $p_t(\bar{a}^{t-1})$ and action D with probability $1 - p_t(\bar{a}^{t-1})$.
- The target payoff in the next period, $V_{t+1}(a^t)$, is computed recursively from the current target payoff $V_t(a^{t-1})$ and the current action a_t.

To accomplish the construction, I need to specify parameters $\{p_t\}$ and the transition rule of target payoffs.

Suppose that a player has record a^{t-1} at the beginning of period t. His partner at period t cooperates with probability $p_t(a^{t-1})$ and defects with probability $1 - p_t(a^{t-1})$. If he cooperates at period t, then he obtains continuation payoff $V_{t+1}(a^{t-1}, C)$ from period $t+1$ on. Thus his total payoff is

$$(1-\delta)[p_t(a^{t-1}) - (1-p_t(a^{t-1}))l] + \delta V_{t+1}(a^{t-1}, C).$$

Similarly, if he defects at period t, then his total payoff is

$$(1-\delta)p_t(a^{t-1})(1+g) + \delta V_{t+1}(a^{t-1}, D).$$

Since the equilibrium satisfies indifference at all histories, the total payoffs in the two situations are the same and equal to $V_t(a^{t-1})$. Thus I obtain

$$\begin{aligned} V_t(a^{t-1}) &= (1-\delta)[p_t(a^{t-1}) - (1-p_t(a^{t-1}))l] + \delta V_{t+1}(a^{t-1}, C) \\ &= (1-\delta)p_t(a^{t-1})(1+g) + \delta V_{t+1}(a^{t-1}, D). \end{aligned} \tag{5}$$

The next lemma shows that, given any target payoff at each period, I can find a mixing probability at this period and target payoffs at the next period such that (5) is satisfied.

Lemma 1. *If $\delta \geq \max(g/(1+g), l/(1+l))$, then, for every $t \geq 1$, every $a^{t-1} \in A^{t-1}$, and every $V_t(a^{t-1}) \in [0,1]$, there exist a probability $p_t(a^{t-1}) \in [0,1]$ and target payoffs $V_{t+1}(a^{t-1}, C)$, $V_{t+1}(a^{t-1}, D) \in [0,1]$ such that (5) is satisfied.*

Proof. Let

$$\begin{aligned} p_t(a^{t-1}) &= V_t(a^{t-1}), \\ V_{t+1}(a^{t-1}, C) &= \frac{V_t(a^{t-1})}{\delta} - \frac{1-\delta}{\delta}[p_t(a^{t-1}) - (1-p_t(a^{t-1}))l], \\ V_{t+1}(a^{t-1}, D) &= \frac{V_t(a^{t-1})}{\delta} - \frac{1-\delta}{\delta}p_t(a^{t-1})(1+g). \end{aligned}$$

By the construction, (5) is satisfied. Since $V_t(a^{t-1}) \in [0,1]$, I have $p_t(a^{t-1}) \in [0,1]$. Also, by the assumption on δ, I have $V_{t+1}(a^{t-1}, C)$, $V_{t+1}(a^{t-1}, D) \in [0,1]$. □

Proof of Proposition 2. "If" part: Repeatedly applying Lemma 1, I can construct $\{p_t\}$ and $\{V_t\}$ that satisfy (5) with initial condition $V_1(\emptyset) = x$. Now I define strategy σ^* by $\sigma_t^*(a^{t-1}, \bar{a}^{t-1})(C) = p_t(\bar{a}^{t-1})$, which satisfies independence of own play. Since target payoffs $\{V_t\}$ are uniformly bounded, $\{V_t\}$ are equal to the value functions under σ^*. By (5), σ^* satisfies indifference at all histories.

"Only if" part: For any independent and indifferent equilibrium σ^*, let $p_t(a^{t-1}) = \sigma_t^*(\bar{a}^{t-1}, a^{t-1})$ for an arbitrary $\bar{a}^{t-1}$. Since σ^* satisfies independence of own play, this definition does not depend on the choice of $\bar{a}^{t-1}$. Similarly, let $V_t(a^{t-1}) = U_t(\sigma^*, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu^*) = \sup_\sigma U_t(\sigma, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu^*)$, where μ^* is generated by σ^*, which does not depend on $\bar{a}^{t-1}$ as well. By indifference at all histories, $\{p_t\}$ and $\{V_t\}$ satisfy (5).

Let $\underline{V} = \inf_{t,a^{t-1}} V_t(a^{t-1})$, which is finite since stage-game payoffs are bounded from below by $-l$. It follows from (5) that

$$V_t(a^{t-1}) = (1-\delta)p_t(a^{t-1})(1+g) + \delta V_{t+1}(a^{t-1}, D) \geq \delta \underline{V}.$$

The inequality holds since $p_t(a^{t-1}) \geq 0$ and $V_{t+1}(a^{t-1}, D) \geq \underline{V}$. Since this is true for any t and any a^{t-1}, $\delta\underline{V}$ is one of the lower bounds for $V_t(a^{t-1})$. Since $\underline{V}$ is the largest lower bound for $V_t(a^{t-1})$, I have $\underline{V} \geq \delta\underline{V}$. Thus $V_1(\emptyset) \geq \underline{V} \geq 0$.

Similarly, let $\overline{V} = \sup_{t,a^{t-1}} V_t(a^{t-1})$, which is finite since stage-game payoffs are bounded from above by $1+g$. It follows from (5) that

$$V_t(a^{t-1}) = (1-\delta)[p_t(a^{t-1}) - (1-p_t(a^{t-1}))l] + \delta V_{t+1}(a^{t-1}, C) \leq (1-\delta) + \delta\overline{V}$$

The inequality holds since $p_t(a^{t-1}) \leq 1$ and $V_{t+1}(a^{t-1}, C) \leq \overline{V}$. Since this is true for any t and any a^{t-1}, $(1-\delta)+\delta\overline{V}$ is one of the upper bounds for $V_t(a^{t-1})$. Since $\overline{V}$ is the smallest upper bound for $V_t(a^{t-1})$, I have $\overline{V} \leq (1-\delta) + \delta\overline{V}$. Thus $V_1(\emptyset) \leq \overline{V} \leq 1$. □

In the repeated prisoner's dilemma with private monitoring, Piccione (2002) constructed a belief-free equilibrium in which each player's mixing probability is independent of his own play; it depends only on the sequence of signals about his opponent's actions. In Proposition 2, each player follows a Piccione-type equilibrium based on the report he receives from the criminal history repository.[13]

Although the basic idea is the same, there are two technical differences between Piccione's and my constructions. First, Piccione focused on belief-free equilibria that not only satisfy independence of own play, but also have a flavor of the tit-for-tat strategy: each of his equilibrium strategies generates the periodic repetition of finitely many actions if each player observes his

[13] Unlike Piccione, Ely and Välimäki (2002) constructed a belief-free equilibrium in which each player chooses an action based only on what he observed at the previous period, i.e., his own action and the signal about his opponent's action at the previous period. An Ely and Välimäki-type equilibrium would have a counterpart in the community enforcement model if second-order information were available. See Subsection 5.3.

opponent's actions correctly, and any past deviation is immediately forgiven once this cycle is observed. Since this class of strategies covers at most countably many payoff profiles under a fixed discount factor, he proved only an approximate folk theorem: for any target payoff profile, there exists a sequence of equilibrium payoffs that converges to the target as the discount factor goes to 1. Second, in some cases, Piccione used asymmetric strategies to sustain even symmetric payoff profiles. For example, when $g-l>1$, due to his restriction to tit-for-tat-like strategies, for any $x \in [0,1]$, the equilibrium he constructed in order to (approximately) achieve symmetric payoff profile (x,x) is asymmetric. In the context of community enforcement, however, it is not clear whether and how such an asymmetric equilibrium can be translated to a single-population model, where players are not assigned with the role of "player 1" or "player 2" at each match. In contrast, for any $x \in [0,1]$ and any large but fixed discount factor, the proof of Proposition 2 constructed a nonperiodic and symmetric equilibrium that exactly achieves symmetric payoff x.

I call the recursively constructed equilibrium in the above proof the *linear* independent and indifferent equilibrium with symmetric payoff x since $p_t(a^{t-1})$ depends on $V_t(a^{t-1})$ linearly.

In the case of $g=l$, I can provide explicit formulae of the linear independent and indifferent equilibrium with symmetric payoff x. Let $\lambda = [(1-\delta)/\delta]g$, which belongs to $(0,1]$ as long as $\delta \geq g/(1+g)$. For $a^{t-1} = (a_1, a_2, \ldots, a_{t-1})$, it holds that

$$p_t(a^{t-1}) = \lambda \sum_{\substack{1 \leq s \leq t-1, \\ a_s = C}} (1-\lambda)^{t-s-1} + x(1-\lambda)^{t-1}, \tag{6}$$

$$1 - p_t(a^{t-1}) = \lambda \sum_{\substack{1 \leq s \leq t-1, \\ a_s = D}} (1-\lambda)^{t-s-1} + (1-x)(1-\lambda)^{t-1}. \tag{7}$$

(6) is interpreted as the discounted sum of rewards. Every time a player cooperates, this choice is recorded and other players will reward him in the future by cooperating with higher probabilities. The impact of cooperation at a given period remains positive forever (except for the knife-edge case that $\delta = g/(1+g)$), but converges to zero at an exponential rate as time goes by. Also, the probability of cooperation at a given period responds more sensitively to cooperation in the recent past than to cooperation in the distant past. These qualitative features remain valid for $g \neq l$. In the same vein, (7) is interpreted as the discounted sum of punishments.

Note that there may be nonlinear independent and indifferent equilibria. For example, if $\delta \geq (1+g+l)/(2+g+l)$, then there is another construction of $p_t(a^{t-1})$, $V_{t+1}(a^{t-1}, C)$, and $V_{t+1}(a^{t-1}, D)$ to prove Lemma 1. To see this, pick any $v^* \in [(1-\delta)(1+g), -(1-\delta)l+\delta]$. For any $V_t(a^{t-1}) \in [0,1]$, let

$$\begin{aligned} p_t(a^{t-1}) &= 1, \\ V_{t+1}(a^{t-1}, C) &= \frac{V_t(a^{t-1})}{\delta} - \frac{1-\delta}{\delta}, \\ V_{t+1}(a^{t-1}, D) &= \frac{V_t(a^{t-1})}{\delta} - \frac{1-\delta}{\delta}(1+g) \end{aligned}$$

if $V_t(a^{t-1}) \geq v^*$, and

$$\begin{aligned} p_t(a^{t-1}) &= 0, \\ V_{t+1}(a^{t-1}, C) &= \frac{V_t(a^{t-1})}{\delta} + \frac{1-\delta}{\delta}l, \\ V_{t+1}(a^{t-1}, D) &= \frac{V_t(a^{t-1})}{\delta} \end{aligned}$$

if $V_t(a^{t-1}) < v^*$. Since $p_t(a^{t-1})$ is either 0 or 1, this construction defines a pure-strategy independent and indifferent equilibrium.

3.3 Long-Run Stability

Consider the linear independent and indifferent equilibrium with symmetric payoff 1 under discount factor $\delta \geq \max(g/(1+g), l/(1+l))$. In this equilibrium, every player cooperates forever on the equilibrium path. This subsection investigates whether this long-run prediction is robust to one-time shocks. Suppose that a small group of players mistakenly defect at some period. This defection induces their future partners to punish with a positive probability, leading to the second defection. The second defection leads to the third defection and so on. Does this chain of defection spread over the community or remain in a small part of the population in the long run?

Before answering this question, I have to make three clarifications. First, a finite number of mistakes do not change the distribution of records in the continuum population. Thus, in order to check robustness of continuum-population equilibria, I use a shock that causes simultaneous mistakes by a small but positive mass of players. Second, I assume that the shock occurs once and unexpectedly. Players keep cooperating until the shock realizes. Third, players do not notice the shock after it realizes. Players, including those who made mistakes because of the shock, believe the same distributions of records in the population as if there had not been a shock, and follow the original equilibrium strategy. If a player observes defection in his partner's record, he does not forgive it even though his partner did not deliberately defect.

Note that long-run stability is a criterion different from maximizing the average of the players' equilibrium payoffs. There are two motivations for this criterion. The first motivation comes from normative analysis. Although it may seem odd to use a welfare criterion based on undiscounted payoffs, I can justify this criterion if each player in the model is not an individual but a dynasty that consists of infinitely many agents. Under this interpretation, each agent cares his offspring's payoffs with discount factor δ, but the social planner has no reason to care older generations more than younger generations. In this situation, long-run outcomes are a desirable social welfare criterion that treats all generations equally. The second motivation comes from positive analysis. If I seriously believe my model as an abstraction of the real world, and also if there are many communities in reality that successfully sustain cooperation for long time, then this phenomenon should be explained by an equilibrium with long-run stability.[14]

In my model, long-run stability is analyzed in the following way. Let σ^* be any equilibrium. Let

$$P_t(\sigma_t^*, \mu_{t-1}) = \sum_{a^{t-1}, \bar{a}^{t-1} \in A_{t-1}} \mu_{t-1}(a^{t-1}) \mu_{t-1}(\bar{a}^{t-1}) \sigma_t^*(a^{t-1}, \bar{a}^{t-1})(C)$$

be the fraction of players in the population who play C at period t. I add a shock at the end of period T so that μ_T differs from μ_T^* generated by σ^*. From period $T+1$ on, all players follow σ^*. The sequence $\{\mu_t\}_{t \geq T}$ of distributions of records in the population is defined recursively by (2) with initial condition μ_T and strategy σ^*. I say that σ^* *sustains cooperation in the long run from* μ_T if $P_t(\sigma_t^*, \mu_{t-1}) \to 1$ as $t \to \infty$.

Now I focus on any linear independent and indifferent equilibrium. Then it holds that

$$P_{t+1}(\sigma_{t+1}^*, \mu_t) = P_t(\sigma_t^*, \mu_{t-1}) - \frac{1-\delta}{\delta}(g-l) P_t(\sigma_t^*, \mu_{t-1})(1 - P_t(\sigma_t^*, \mu_{t-1})) \tag{8}$$

for $t \geq T+1$. See Appendix A.2 for the derivation of (8). I can approximate (8) by the linear difference equation

$$1 - P_{t+1}(\sigma_{t+1}^*, \mu_t) \approx \left(1 + \frac{1-\delta}{\delta}(g-l)\right)(1 - P_t(\sigma_t^*, \mu_{t-1}))$$

[14] The concept of long-run stability was introduced by Kandori (1992). Ellison (1994) clarified the difference between long-run stability and robustness to noises analyzed in the next subsection.

if $P_t(\sigma_t^*, \mu_{t-1})$ is close to 1.

If $g < l$, then the fraction of players who choose D declines exponentially. In fact, σ^* sustains cooperation in the long run from any μ_T. See Appendix A.3 for the proof. In contrast, if $g > l$, then defection spreads in the whole population at an exponential rate. Thus even a small fraction of errors leads to the corruption of cooperation in the long run. If $g = l$, then the fraction of D in the population remains constant.

3.4 Noise in Actions or in Records

The real world contains persistent noise. A player may not be able to choose an action as he wants. Even if he chooses an action correctly, it may still be recorded wrong or interpreted with biases. For example, people often forget to pay their bills on time. Even if they make the payment, credit history bureaus occasionally make mistakes in recording. In an online transaction, a seller sometimes fails to deliver a good to the buyer on time due to lost or delayed package. Even if the seller delivers on time, the buyer may nevertheless submit a negative rating.

In this subsection, I add noise to the process of choosing actions, and show that independent and indifferent equilibria are robust to such noise. (A similar result is obtained when noise is added to the recording process. The two kinds of noise are theoretically equivalent under a certain transformation of stage-game payoffs.) At each period, if a player intends to cooperate, then he mistakenly defects with probability ε_C; if a player intends to defect, then he mistakenly cooperates with probability ε_D. So the probability of cooperation is restricted between ε_D and $1 - \varepsilon_C$. Note that, unlike in the previous subsection, shocks can occur repeatedly over time. These shocks are expected, and players take into account the possibility that they make mistakes. Also I use players' discounted payoffs as the welfare criterion.

The next proposition shows that the set of independent and indifferent equilibrium payoffs changes continuously with respect to noise levels. Although very low or very high payoffs are not sustained by independent and indifferent equilibria if players make mistakes with positive probabilities, the set of independent and indifferent equilibrium payoffs converges to $[0, 1]$ as the probabilities of mistakes go to zero.

Proposition 3. *In the model with noise in actions, suppose that* $\varepsilon_C(1+l) + \varepsilon_D(1+g) < 1$ *and*

$$\delta \geq \frac{1}{1-\varepsilon_C-\varepsilon_D} \max\left(\frac{(1-\varepsilon_C)g+\varepsilon_C l}{1+g}, \frac{(1-\varepsilon_D)l+\varepsilon_D g}{1+l}\right).$$

Then there is an independent and indifferent equilibrium with symmetric payoff x if and only if $x \in [\varepsilon_D(1+g), 1-\varepsilon_C(1+l)]$.

Proof. Proved similarly to Proposition 2. See Appendix A.4. □

3.5 Finite Population

In a large community, people may ignore the possibility that a person's private observation can alter his belief about aggregated variables in the community. My continuum-population model is an idealization that captures this aspect of the real world. In a finite-population model, in contrast, I can explicitly analyze the effect of private information on beliefs about other players' histories. Since the two models are different, there is no *a priori* reason to believe that equilibria in the continuum-population model carry over to the finite-population model or the vice versa.

As the next proposition shows, however, independent and indifferent equilibria carry over to the corresponding finite-population model. The reason is simple. If an equilibrium satisfies independence of own play, then, to compute a player's continuation payoff, it does not matter who he meets or what record of play his partner has. Therefore, his belief about other players' histories does not matter, either. Note that the only difference between finite- and continuum-population models is the way of updating beliefs, but this difference does not affect players' incentives.

Proposition 4. *If a continuum-population perfect equilibrium strategy satisfies independence of own play, then, the strategy combined with any consistent belief system forms a sequential equilibrium in the corresponding finite-population model.*

Proof. For any equilibrium σ^* that satisfies independence of own play, continuation payoff $U_t(\sigma, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu)$ is independent of μ. Even though a player's belief about other players' records of play depends on his private history in a finite-population model, this does not alter his continuation payoff. Therefore, in the finite-population model, σ^* is sequentially rational under any belief system. □

Note that Proposition 4 holds even if matchings are not chosen uniformly. Moreover, the threshold of δ sufficient for the existence of independent and indifferent equilibria does not vary with the size of the population. This is not the case of contagious equilibria. In the finite-population model without information transmission through the criminal history repository, Kandori (1992) and Ellison (1994) showed that cooperation can be sustained by a

popul. size	contagious eqm	indep. & indiff. eqm		
		$l \leq 1$	$l = 2$	$l = 10$
2	0.50	0.50	0.67	0.91
4	0.68	0.50	0.67	0.91
10	0.79	0.50	0.67	0.91
100	0.89	0.50	0.67	0.91
1000	0.93	0.50	0.67	0.91

Table 1: Discount factor sufficient to sustain cooperation: $g = 1$. The column for contagious equilibrium is taken from Table 1 in Ellison (1994).

contagious equilibrium if δ is above a threshold. The threshold increases as the population becomes larger, and converges to 1 in the limit. Therefore, to sustain cooperation by a contagious equilibrium in a large community, interactions need to be frequent. Table 1 gives the threshold of the discount factor for several population sizes and several values of l with $g = 1$. Note that the threshold for Ellison's contagious equilibrium does not depend on l. The table shows that first-order information improves the possibility of cooperation when the population size is large and l is small.

If the population size is 2, then the finite-population model collapses to the standard two-player repeated prisoner's dilemma. Thus, by Propositions 2 and 4, I can construct an equilibrium that sustains cooperation in the two-player repeated prisoner's dilemma. The constructed equilibrium is one of the belief-free equilibria in Ely et al.'s (2005) terminology.

The next subsection exploits this connection in the other way. Namely, I will use a technique that was originally developed for two-player repeated games to construct equilibria in the community enforcement model.

3.6 General Stage Games

Ely et al. (2004, 2005) analyzed belief-free equilibria for two-player repeated games. In Proposition 5 together with Lemma 3 of their 2005 paper, they characterized the limit set of belief-free equilibrium payoffs as δ goes to 1 if public randomization devices are available. Their 2004 paper showed that the same characterization holds without public randomization devices. Here I mimic their proofs to characterize the limit set of equilibrium payoffs with independence of own play as δ goes to 1.

Suppose that the stage game is given by a finite symmetric two-player game $G = (A, u)$, where A is a nonempty finite set of actions and $u\colon A^2 \to \mathbb{R}$ is a player's payoff function. For each stage game G, I follow Section 2 verbatim to build the infinitely repeated game of G with random matching in a continuum of population. I do not allow players to observe public randomization devices. I keep the definition of independence of own play as in Subsection 3.1. Let $\mathcal{P}$ be the set of all nonempty subsets of A. For each $t \geq 1$ and $B_t \in \mathcal{P}$, I say that σ^* satisfies *indifference within* B_t *at period* t if (4) is satisfied with equality for all a^{t-1}, $\bar{a}^{t-1} \in A^{t-1}$ and $a_t \in B_t$, where μ^* is generated by σ^*. For each $B \in \mathcal{P}$, I define

$$\underline{v}(B) = \min_{\alpha \in \Delta(B)} \max_{a \in A} u(a, \alpha), \tag{9}$$

$$\overline{v}(B) = \max_{\alpha \in \Delta(B)} \min_{a \in B} u(a, \alpha). \tag{10}$$

If $\underline{v}(B) < \overline{v}(B)$, then, as in Proposition 2, I can show that, for a sufficiently large δ, there exists an equilibrium with symmetric payoff x that satisfies independence of own play and indifference within $B_t = B$ at all periods if and only if $x \in [\underline{v}(B), \overline{v}(B)]$. I omit the proof of this claim. Instead, I will construct equilibria that cover a wider range of payoffs by changing B_t over time. Namely, I will characterize the limit set of equilibrium payoffs with independence of own play. The limit set can be strictly larger than the union of $[\underline{v}(B), \overline{v}(B)]$ for $B \in \mathcal{P}$ such that $\underline{v}(B) < \overline{v}(B)$.

The stage game G belongs to the *positive case* if $\underline{v}(B) < \overline{v}(B)$ for some $B \in \mathcal{P}$; G belongs to the *negative case* if $\underline{v}(B) \geq \overline{v}(B)$ for all $B \in \mathcal{P}$. Let

$$W = \left\{ w \in \Delta(\mathcal{P}) \mid \sum_{B \in \mathcal{P}} w(B)\underline{v}(B) \leq \sum_{B \in \mathcal{P}} w(B)\overline{v}(B) \right\}$$

be the set of weights w on $\mathcal{P}$ such that the weighted average of $\underline{v}(B)$ is less than or equal to the weighted average of $\overline{v}(B)$, and

$$\underline{v} = \min_{w \in W} \sum_{B \in \mathcal{P}} w(B)\underline{v}(B), \qquad \overline{v} = \max_{w \in W} \sum_{B \in \mathcal{P}} w(B)\overline{v}(B).$$

be the minimum and the maximum of such weighted averages, respectively, when the weight w is taken from W.

Let *NE* be the set of symmetric Nash equilibrium payoffs of the stage game.

Proposition 5. (1) *Suppose that the stage game belongs to the positive case. Then, the set of symmetric equilibrium payoffs with independence of own play is a subset of* $[\underline{v}, \overline{v}]$. *Conversely, for any* $x \in (\underline{v}, \overline{v})$, *there exists* $\bar{\delta} < 1$ *such that, for any* $\delta > \bar{\delta}$, *there exists an equilibrium with symmetric payoff* x *that satisfies independence of own play.*

(2) *Suppose that the stage game belongs to the negative case. Then, the set of symmetric equilibrium payoffs with independence of own play is a subset of* $[\min NE, \max NE]$. *Conversely, for any* $x \in [\min NE, \max NE]$ *and any* $\delta \geq 1/2$, *there exists an equilibrium with symmetric payoff* x *that satisfies independence of own play.*

Sketch of the Proof. The proof is based on Ely et al. (2004, 2005). I will sketch the proof of the second statement of Part (1). If the stage game belongs to the positive case, then, for any $x \in (\underline{v}, \overline{v})$, there exists a sequence $\{B_t\}$ such that, for sufficiently large δ, it holds that

$$\underline{v}_t \equiv \sum_{B \in \mathcal{P}} w_t(B)\underline{v}(B) < x < \sum_{B \in \mathcal{P}} w_t(B)\overline{v}(B) \equiv \overline{v}_t$$

for every $t \geq 1$, where

$$w_t(B) = (1-\delta) \sum_{s \geq t,\, B_s = B} \delta^{s-t}$$

is the discounted fraction of periods in which $B_s = B$ after period t. I will construct an equilibrium that satisfies independence of own play and indifference within B_t at each period t. The construction is recursive as in Lemma 1. For a player's target payoff in $[\underline{v}_t, \overline{v}_t]$, I can control his partner's action and target payoffs at period $t+1$ so that he is indifferent within B_t but any action outside B_t is severely punished by the worst continuation payoff $\underline{v}_{t+1}$. Also, by having the partner choose a mixture of a solution to the minmax problem (9) and a solution to the maxmin problem (10), I can keep the next-period target payoffs within $[\underline{v}_{t+1}, \overline{v}_{t+1}]$. See Appendix A.5 for the details.

There are two main differences between Proposition 5 and Ely et al. (2004). One is that I use symmetric strategies in the single-population model. This simplifies the definitions of positive and negative cases, and makes the "abnormal case" empty. The other is that a player in the two-player repeated game recalls his own past actions whereas a player in the community enforcement model does not observe the realizations of his partner's past partners' mixed actions. Therefore, a player's target payoff at the next period cannot depend on whether his current partner chooses a solution to the minmax problem (9) or a solution to the maxmin problem (10). This is why an equilibrium strategy with independence of own play is not described by a two-state automaton as in Ely and Välimäki (2002). To fix this problem, I modify Ely and Välimäki-type equilibria constructed in Ely et al. (2004) into Piccione-type equilibria. Compare Lemma 6 in Appendix A.5 with step (iv) in Ely et al. (2004). □

B	$\{C\}$	$\{D\}$	$\{E\}$	$\{C,D\}$	$\{C,E\}$	$\{D,E\}$	$A=\{C,D,E\}$
$\underline{v}(B)$	5	0	4	0	5/2	0	0
$\overline{v}(B)$	1	0	2	1	2	0	10/7

Table 2: $\underline{v}(B)$ and $\overline{v}(B)$ for Example 2.

Note that the threshold $\bar{\delta}$ depends on the target payoff x in Proposition 5 (1). Also note that Proposition 5 (1) does not say anything about whether the boundary of $[\underline{v},\overline{v}]$ can be sustained as an equilibrium payoff.

Example 1. For prisoner's dilemma, I have

$$\begin{aligned} &\underline{v}(\{C\}) = 1+g, &&\underline{v}(\{D\}) = 0, &&\underline{v}(\{C,D\}) = 0,\\ &\overline{v}(\{C\}) = 1, &&\overline{v}(\{D\}) = 0, &&\overline{v}(\{C,D\}) = 1. \end{aligned}$$

The prisoner's dilemma game belongs to the positive case, and, by Proposition 5 (1), the limit set of equilibrium payoffs with independence of own play is given by $[\underline{v},\overline{v}] = [0,1]$. Thus, in the case of prisoner's dilemma, it does not change the set of equilibrium payoffs whether players are indifferent between C and D at all histories or not.

Example 2. Consider the following 3×3 game with $A=\{C,D,E\}$:

	C	D	E
C	$1,1$	$-1,5$	$4,0$
D	$5,-1$	$0,0$	$0,0$
E	$0,4$	$0,0$	$2,2$

See Table 2 for the list of $\underline{v}(B)$ and $\overline{v}(B)$ for this game. The game belongs to the positive case. Thus, by Proposition 5 (1), the limit set of equilibrium payoffs with independence of own play is given by $[\underline{v},\overline{v}] = [0, 50/27]$. Here the upper bound 50/27 is the weighted average of $\overline{v}(\{C,E\})$ and $\overline{v}(\{C,D,E\})$ with weights 20/27 and 7/27, respectively. Note that $[0,50/27]$ is strictly larger than $[0,10/7]$, the union of $[\underline{v}(B),\overline{v}(B)]$ such that $\underline{v}(B) < \overline{v}(B)$. Thus, varying sets of indifferent actions over time expands the set of equilibrium payoffs in this game.

Example 3. Consider the following 3×3 game with $A=\{a,a',a''\}$:

	a	a'	a''
a	$0,0$	$2,0$	$1,2$
a'	$0,2$	$1,1$	$0,0$
a''	$2,1$	$0,0$	$0,0$

B	$\{a\}$	$\{a'\}$	$\{a''\}$	$\{a,a'\}$	$\{a,a''\}$	$\{a',a''\}$	$A=\{a,a',a''\}$
$\underline{v}(B)$	2	2	1	1	2/3	1	2/3
$\overline{v}(B)$	0	1	0	1	2/3	0	2/3

Table 3: $\underline{v}(B)$ and $\overline{v}(B)$ for Example 3.

See Table 3 for the list of $\underline{v}(B)$ and $\overline{v}(B)$ for this game. The game belongs to the negative case and $NE = \{2/3\}$. Thus, by Proposition 5 (2), 2/3 is the unique repeated-game equilibrium payoff with independence of own play. Note that $[\underline{v}, \overline{v}] = [2/3, 1]$ is different from $\{2/3\}$. Thus, this example shows that the analysis on the negative case cannot be absorbed to the positive case.

4 Strict Incentives

Every independent and indifferent equilibrium is subject to the criticism that players lack strict incentives to follow the equilibrium strategy. Motivated by this criticism, in this section, I explore equilibria that do not use indifference conditions as heavily as independent and indifferent equilibria.

4.1 Strict Equilibria

I call an equilibrium *strict* if, at any history, each player strictly prefers the action prescribed by the equilibrium to any one-shot deviation.

Definition 4. A strategy σ^* is a *continuum-population strict perfect equilibrium* (or *strict equilibrium*) if (4) holds with strict inequality whenever $a_t \neq \sigma_t^*(a^{t-1}, \bar{a}^{t-1})$.

This definition is weaker than the standard definition of strict equilibria for normal-form games. At each period, this definition ignores deviations at histories that are not reachable from the current history under the equilibrium play. A strict equilibrium is always in pure strategies: $\sigma_t^*(a^{t-1}, \bar{a}^{t-1})$ is either C or D for every $t \geq 1$ and $a^{t-1}, \bar{a}^{t-1} \in A^{t-1}$.

The next proposition shows that cooperation can be sustained by a strict equilibrium for a sufficiently large discount factor if and only if $g < l$.

Proposition 6. (1) *If $g < l$ and $\delta > g(1+l)/[(1+g)l]$, then there exists a strict equilibrium with symmetric payoff* 1*.*

(2) *If $g \geq l$, then there is no strict equilibrium with a positive symmetric payoff.*

To prove Proposition 6 (1), consider the pairwise grim-trigger strategy defined in Section 2 as a candidate for strict equilibria. Recall that the pairwise grim-trigger strategy prescribes a player to cooperate if and only if neither he nor his current partner has ever defected. Since both cheaters and those who punished cheaters are punished equally, punishment in the pairwise grim-trigger strategy has two effects that conflict with each other. One effect is to encourage a player to cooperate when neither his nor his current partner's record contains D. The other effect is to discourage a player to defect when his own record consists of C only but his partner's record contains D. In this situation, he may prefer keeping a clean record to punishing the partner in the current match if the punishment is too severe. Therefore, punishment needs to be severer than the short-term gain g in the cooperation phase, but milder than the short-term loss l in the punishment phase. This is why $g < l$ is assumed in Proposition 6 (1). The next lemma formalizes this argument.

Lemma 2. *The pairwise grim-trigger strategy is an equilibrium if and only if $g/(1+g) \leq \delta \leq l/(1+l)$, and a strict equilibrium if and only if $g/(1+g) < \delta < l/(1+l)$.*

Proof. Consider a match at period t between a player with record a^{t-1} and another player with record $\bar{a}^{t-1}$. First, suppose that neither a^{t-1} nor $\bar{a}^{t-1}$ contains D. Let μ^* be generated by the pairwise grim-trigger strategy. Under μ^*, the player with record a^{t-1} believes that none of his future partners has a record that contains D. Thus he believes that his future partners will cooperate if he follows the equilibrium strategy, but defect otherwise. In this situation, C is a best response for the player with record a^{t-1} if and only if

$$1 \geq (1-\delta)(1+g). \tag{11}$$

Second, suppose that a^{t-1} does not contain D but $\bar{a}^{t-1}$ does. The only difference from the previous situation is that the player with record a^{t-1} faces the partner who will play D at period t. In this situation, D is a best response for the player with record a^{t-1} if and only if

$$0 \geq -(1-\delta)l + \delta. \tag{12}$$

Lastly, suppose that a^{t-1} contains D. Then, no matter what the player with record a^{t-1} does, other players will defect to him. Thus D is always a strict best response for him.

Therefore, it follows from (11) and (12) that the pairwise grim-trigger strategy is an equilibrium if and only if $g/(1+g) \le \delta \le l/(1+l)$.

Similarly, the pairwise grim-trigger strategy is a strict equilibrium if and only if (11) and (12) are satisfied with strict inequalities, i.e., $g/(1+g) < \delta < l/(1+l)$. □

If $g < l$, then there is an open interval of δ under which the pairwise grim-trigger strategy is a strict equilibrium. Then I can apply a well-known trick by Ellison (1994) to modify the pairwise grim-trigger strategy into an equilibrium for any sufficiently large discount factor.

Proof of Proposition 6. (1) For any $\delta > g(1+l)/[(1+g)l]$, there exists an integer T such that $g/(1+g) < \delta^T < l/(1+l)$. For this T, I define a strategy σ^* as follows: $\sigma_t^*(a^{t-1}, \bar{a}^{t-1}) = C$ if $a_{t-kT} = \bar{a}_{t-kT} = C$ for every positive integer k such that $kT < t$; $\sigma_t^*(a^{t-1}, \bar{a}^{t-1}) = D$ otherwise. In other words, the whole repeated game is divided into T mini-games such that the stage game at period t belongs to the t'-th mini-game if and only if $t - t'$ is divisible by T, and players follow the pairwise grim-trigger strategy in each mini-game, ignoring information about the other mini-games. Since the effective discount factor in each mini-game is δ^T, by Lemma 2, σ^* is a strict equilibrium.

(2) Suppose that σ^* is a strict equilibrium with a positive symmetric payoff. Since σ^* is a pure strategy, I can define $(a_1, a_2, \dots)$ as the sequence of pure actions on the equilibrium play. Since σ^* yields a positive payoff, there exists a period t at which $a_t = C$. If a player deviates at period t by choosing D, he has to be punished in the future. Thus there exists another period $t' > t$ at which the equilibrium play $a_{t'}$ is C but a player chooses D if he encounters a player who deviated at period t. Let $b^{t'-1}$ be the deviator's record of play before period t'.

Here I consider two incentive compatibility constraints of a non-deviator whose record of play is $a^{t'-1} = (a_1, \dots, a_{t'-1})$. If his partner has followed the equilibrium play (i.e., the partner's record of play is also $a^{t'-1}$), then he strictly prefers C to D:

$$\begin{aligned}&(1-\delta)u(C,C) + \delta U_{t'+1}(\sigma^*, \sigma^* \mid (a^{t'-1}, C), a^{t'}, \mu^*)\\ &\quad > (1-\delta)u(D,C) + \delta U_{t'+1}(\sigma^*, \sigma^* \mid (a^{t'-1}, D), a^{t'}, \mu^*), \qquad (13)\end{aligned}$$

where μ_s^* puts weight 1 on a^s for any s. Also, he strictly prefers D to C if the partner's record of play is $b^{t'-1}$:

$$\begin{aligned}&(1-\delta)u(D,\bar{a}) + \delta U_{t'+1}(\sigma^*, \sigma^* \mid (a^{t'-1}, D), a^{t'}, \mu^*)\\ &\quad > (1-\delta)u(C,\bar{a}) + \delta U_{t'+1}(\sigma^*, \sigma^* \mid (a^{t'-1}, C), a^{t'}, \mu^*), \qquad (14)\end{aligned}$$

where $\bar{a} = \sigma^*_{t'}(b^{t'-1}, a^{t'-1})$ is the deviator's action at period t'. Adding up (13) and (14), I can cancel out the future payoff terms because the future payoffs do not depend on the current partner's record of play. Thus I have

$$u(C, C) + u(D, \bar{a}) > u(D, C) + u(C, \bar{a}),$$

which does not hold no matter whether $\bar{a} = C$ or D since $g \geq l$. □

4.2 Equilibria without History-Dependent Mixing

One might think that using mixed strategies in independent and indifferent equilibria can be justified by Harsanyi's (1973) purification theorem. He showed that, for a generic normal-form game, all mixed-strategy equilibria can be approximated by a sequence of pure-strategy equilibria in incomplete-information games with payoff perturbations. Moreover, players in incomplete-information games have strict best responses with probability 1.

The purification theorem, however, does not apply to extensive-form games. Moreover, Bhaskar (1998) provided an example of overlapping generation game with a mixed-strategy equilibrium that cannot be purified if payoff perturbations enter in the additively separable form. He argued that non-purifiability does not stem from mixed equilibrium *per se*, but from having mixing probabilities depend on payoff-irrelevant histories. See Bhaskar and van Damme (2002) for a similar argument in the context of two-player repeated games.[15]

No independent and indifferent equilibrium passes Bhaskar's critique. This subsection considers whether cooperation can be sustained by an equilibrium that does not use history-dependent mixing.

Definition 5. An equilibrium σ^* is *without history-dependent mixing* if the following condition is satisfied: for any $t \geq 1$, it holds that $\sigma^*_t(a^{t-1}, \bar{a}^{t-1}) = \sigma^*_t(a^{t-1}, \bar{b}^{t-1})$ if a player with record a^{t-1} is indifferent between C and D both when his partner has record $\bar{a}^{t-1}$ and when his partner has record $\bar{b}^{t-1}$.

This definition is motivated by Bhaskar's argument. Consider two situations: a player with record a^{t-1} faces a partner with record $\bar{a}^{t-1}$ and a partner with record $\bar{b}^{t-1}$. In both situations, he is indifferent between cooperation and defection. Then he has to decide his action based on a payoff perturbation to his current payoff. No matter how the perturbation is distributed, it must induce the same mixing probabilities. Note that I do not

[15] But see Bhaskar et al. (2006). In the repeated prisoner's dilemma game, they showed that Ely and Välimäki's (2002) equilibrium can be purified by a sequence of equilibria with unboundedly long memory even though it uses history-dependent mixing.

require that two players' mixing probabilities be the same when their own records of play are different. If records are different, then their current decisions can be affected not only by the current payoff perturbation but also by future payoff perturbations, and the two mixing probabilities do not need to coincide with each other. I conjecture that an equilibrium being without history-dependent mixing is sufficient (may not be necessary) for the equilibrium to be purifiable when payoffs are perturbed in an additively separable way.

No strict equilibrium uses history-dependent mixing. Since I have already shown in Proposition 6 (1) that strict equilibria can sustain cooperation when $g < l$ for sufficiently large δ, I will focus on $g \geq l$ in this subsection.

The next proposition shows that, if $g > l$ and players are patient enough, then cooperation can be approximately sustained by equilibria without history-dependent mixing, although full cooperation cannot be sustained exactly. It is also shown that no positive payoff can be sustained by equilibria without history-dependent mixing if $g = l$.

Proposition 7. *Suppose that* $g \geq l$*.*

(1) *If* $g > l$ *and* $\delta > (1 + g + l)/(2 + g + l)$*, then, for any* $x \in [0, 1)$*, there exists an equilibrium without history-dependent mixing that sustains symmetric payoff* x*.*

(2) *There is no equilibrium without history-dependent mixing that sustains full cooperation on the equilibrium path.*

(3) *If* $g = l$*, then there is no equilibrium without history-dependent mixing that sustains a positive symmetric payoff.*

For Part (1), since always choosing D is a strict equilibrium with symmetric payoff $x = 0$, I can focus on achieving $x \in (0, 1)$ without loss of generality.

For each $x \in (0, 1)$, I will construct an equilibrium with the following properties.

- At the beginning of period t (before the matching at period t is determined), players with record a^{t-1} are assigned with target payoff $V_t(a^{t-1})$. I will construct an equilibrium under which all players obtain their target payoffs as their continuation payoffs. Set $V_1(\emptyset) = x$.
- Players are divided into two groups based on their target payoffs. If a player's target payoff is greater than or equal to a certain threshold v^*, he is categorized to "high reputation" group H, and to "low reputation" group L otherwise.

- If two players who are matched are from the same level of reputation $r = H$ or L, then they cooperate with probability p_t^r. If one player is from the low-reputation group and the other from the high-reputation group, then the low-reputation player cooperates whereas the high-reputation player defects.

- The target payoff in the next period, $V_{t+1}(a^t)$, is computed recursively from the current target payoff $V_t(a^{t-1})$ and the current action a_t.

To accomplish the construction, I need to specify v^*, $\{p_t^H\}$, $\{p_t^L\}$, and the transition rule of target payoffs.

Fix any $x \in (0, 1)$. Since $\delta > (1+g+l)/(2+g+l)$, there exists a threshold level v^* such that

$$(1 - \delta)(1 + g) < v^* < -(1 - \delta)l + \delta. \tag{15}$$

This inequality will be used in Lemma 4 to show that target payoffs remain between 0 and 1.

Suppose that behavior strategies up to period $t - 1$, $\sigma_s^*(a^{s-1}, \bar{a}^{s-1})$ for $s = 1, \ldots, t - 1$, are already specified. This determines the distribution of records of play in the population, $\mu_{t-1}^* \in \Delta(A^{t-1})$. Also suppose that each player is assigned with his target payoff $V_t(a^{t-1})$ as a function of his record of play. Let

$$\mu_{t-1}^H = \sum_{a^{t-1} \in A^{t-1},\, V_t(a^{t-1}) \geq v^*} \mu_{t-1}^*(a^{t-1}), \quad \mu_{t-1}^L = 1 - \mu_{t-1}^H$$

be the masses of players with high and low reputations, respectively.

Now I will specify p_t^H, p_t^L, and $V_{t+1}(a^t)$ for all $a^t \in A^t$. When a player with reputation r is matched with a player with the same reputation level, he must be indifferent between C and D. Therefore, I have

$$(1 - \delta)(p_t^H - (1 - p_t^H)l) + \delta V_{t+1}(a^{t-1}, C) = (1 - \delta)p_t^H(1 + g) + \delta V_{t+1}(a^{t-1}, D) \tag{16}$$

if $V_t(a^{t-1}) \geq v^*$, and

$$(1 - \delta)(p_t^L - (1 - p_t^L)l) + \delta V_{t+1}(a^{t-1}, C) = (1 - \delta)p_t^L(1 + g) + \delta V_{t+1}(a^{t-1}, D) \tag{17}$$

if $V_t(a^{t-1}) < v^*$. The following lemma shows that, as long as these conditions are satisfied together with $0 < p_t^r < 1$, every player has a strict incentive to follow the equilibrium when he faces a player from the opposite level of reputation.

Lemma 3. *Suppose that $g > l$. If (16) is satisfied together with $p_t^H < 1$, then*

$$(1-\delta) + \delta V_{t+1}(a^{t-1}, C) < (1-\delta)(1+g) + \delta V_{t+1}(a^{t-1}, D).$$

Similarly, if (17) is satisfied together with $p_t^L > 0$, then

$$(1-\delta)(-l) + \delta V_{t+1}(a^{t-1}, C) > \delta V_{t+1}(a^{t-1}, D).$$

Proof. This follows from strict submodularity of stage-game payoffs: the more likely a player is to cooperate, the more willing his partner is to defect. See Appendix A.6. □

Also promise-keeping conditions need to be satisfied:

$$\begin{aligned} V_t(a^{t-1}) = \mu_{t-1}^H[(1-\delta)p_t^H(1+g) + \delta V_{t+1}(a^{t-1}, D)] \\ + \mu_{t-1}^L[(1-\delta)(1+g) + \delta V_{t+1}(a^{t-1}, D)] \end{aligned} \tag{18}$$

if $V_t(a^{t-1}) \geq v^*$, and

$$\begin{aligned} V_t(a^{t-1}) = \mu_{t-1}^L[(1-\delta)(p_t^L - (1-p_t^L)l) + \delta V_{t+1}(a^{t-1}, C)] \\ + \mu_{t-1}^H[(1-\delta)(-l) + \delta V_{t+1}(a^{t-1}, C)] \end{aligned} \tag{19}$$

if $V_t(a^{t-1}) < v^*$.

Lemma 4. *Fix any v^* satisfying (15), μ_{t-1}^H, and μ_{t-1}^L. Then there exists a mixing probability $p_t^H \in (0,1)$ such that, for any a^{t-1} with $V_t(a^{t-1}) \in [v^*, 1)$, there exist target payoffs $V_{t+1}(a^{t-1}, C)$, $V_{t+1}(a^{t-1}, D) \in (0,1)$ such that (16) and (18) are satisfied. Similarly, there exists a mixing probability $p_t^L \in (0,1)$ such that, for any a^{t-1} with $V_t(a^{t-1}) \in (0, v^*)$, there exist target payoffs $V_{t+1}(a^{t-1}, C)$, $V_{t+1}(a^{t-1}, D) \in (0,1)$ such that (17) and (19) are satisfied.*

Proof. For any a^{t-1} with $V_t(a^{t-1}) \in [v^*, 1)$, $V_{t+1}(a^{t-1}, C)$ and $V_{t+1}(a^{t-1}, D)$ are determined by (16) and (18):

$$V_{t+1}(a^{t-1}, C) = \frac{V_t(a^{t-1})}{\delta} - \frac{1-\delta}{\delta}[\mu_{t-1}^H p_t^H(1+g) + \mu_{t-1}^L(1+g) - p_t^H g - (1-p_t^H)l]$$

$$V_{t+1}(a^{t-1}, D) = \frac{V_t(a^{t-1})}{\delta} - \frac{1-\delta}{\delta}[\mu_{t-1}^H p_t^H(1+g) + \mu_{t-1}^L(1+g)].$$

Since $V_t(a^{t-1}) \in [v^*, 1)$ and (15) holds, if p_t^H is sufficiently close to 1, I have $0 < V_{t+1}(a^{t-1}, D) < V_{t+1}(a^{t-1}, C) < 1$.

Similarly, for any a^{t-1} with $V_t(a^{t-1}) \in (0, v^*)$, $V_{t+1}(a^{t-1}, C)$ and $V_{t+1}(a^{t-1}, D)$ are determined by (17) and (19):

$$V_{t+1}(a^{t-1}, C) = \frac{V_t(a^{t-1})}{\delta} + \frac{1-\delta}{\delta}[-\mu^L_{t-1}(p^L_t - (1-p^L_t)l) + \mu^H_{t-1}l]$$
$$V_{t+1}(a^{t-1}, D) = \frac{V_t(a^{t-1})}{\delta} + \frac{1-\delta}{\delta}[-\mu^L_{t-1}(p^L_t - (1-p^L_t)l) + \mu^H_{t-1}l - p^L_t g - (1-p^L_t)l].$$

Since $V_t(a^{t-1}) \in (0, v^*)$ and (15) holds, if p^L_t is sufficiently close to 0, I have $0 < V_{t+1}(a^{t-1}, D) < V_{t+1}(a^{t-1}, C) < 1$. □

Proof of Proposition 7. (1) Fix any v^* satisfying (15). For any $x \in (0,1)$, repeatedly applying Lemma 4, I can construct $\{V_t\}$, $\{p^H_t\}$, and $\{p^L_t\}$ that satisfy (16)–(19) with initial condition $V_1(\emptyset) = x$. Now I define strategy σ^* by

$$\sigma^*_t(a^{t-1}, \bar{a}^{t-1})(C) = \begin{cases} p^H_t & \text{if } V_t(a^{t-1}) \geq v^* \text{ and } V_t(\bar{a}^{t-1}) \geq v^*, \\ 0 & \text{if } V_t(a^{t-1}) \geq v^* > V_t(\bar{a}^{t-1}), \\ 1 & \text{if } V_t(a^{t-1}) < v^* \leq V_t(\bar{a}^{t-1}), \\ p^L_t & \text{if } V_t(a^{t-1}) < v^* \text{ and } V_t(\bar{a}^{t-1}) < v^*. \end{cases}$$

Since (18) and (19) are satisfied and target payoffs $\{V_t\}$ are uniformly bounded, $\{V_t\}$ are equal to the value functions under σ^*. By (16), (17), and Lemma 3, σ^* is an equilibrium without history-dependent mixing.

(2) This is proved similarly to Proposition 6 (2). Suppose that σ^* is an equilibrium without history-dependent mixing that sustains cooperation on the equilibrium path. If a player deviates at period t, he will be punished with positive probability at period $t' > t$. Let $a^{t'-1}$ be the deviator's record of play before period t'.

Here I consider two incentive compatibility constraints of a non-deviator. If his partner has followed the equilibrium play, then he weakly prefers C to D:

$$\begin{aligned}(1-\delta)u(C,C) &+ \delta U_{t'+1}(\sigma^*, \sigma^* \mid (C,\dots,C,C),(C,\dots,C),\mu^*) \\ &\geq (1-\delta)u(D,C) + \delta U_{t'+1}(\sigma^*, \sigma^* \mid (C,\dots,C,D),(C,\dots,C),\mu^*),\end{aligned} \quad (20)$$

where μ^*_s puts weight 1 on $(C,\dots,C)$ for any s. Also, he weakly prefers D to C if the partner's record of play is $b^{t'-1}$:

$$\begin{aligned}(1-\delta)u(D,\alpha) &+ \delta U_{t'+1}(\sigma^*, \sigma^* \mid (C,\dots,C,D),(C,\dots,C),\mu^*) \\ &\geq (1-\delta)u(C,\alpha) + \delta U_{t'+1}(\sigma^*, \sigma^* \mid (C,\dots,C,C),(C,\dots,C),\mu^*),\end{aligned} \quad (21)$$

where $\alpha = \sigma^*_{t'}(a^{t'-1}, (C, \dots, C))$ is the deviator's mixed action at period t'. Moreover, since σ^* does not use history-dependent mixing, either (20) or (21) holds with strict inequality. Adding up (20) and (21), I have

$$u(C,C) + u(D,\alpha) > u(D,C) + u(C,\alpha),$$

which does not hold for any α since $g \geq l$.

(3) Since $g = l$, a player's incentive of choosing C over D is not affected by his current partner's action. So there is no payoff-relevant reason for him to vary mixing probabilities based on his partner's record of play. Therefore, the repetition of D is the only equilibrium without history-dependent mixing. ☐

5 Summary Statistics

This section investigates the possibility of cooperation when players use summary statistics of their partners' records of play such as the number of cooperation or the play in the T most recent periods for some finite T. Using summary statistics requires less amount of players' computational abilities and saves social cost of storing data.

5.1 Permutation Invariance

I call a strategy σ *permutation-invariant* if $\sigma_t(a^{t-1}, \bar{a}^{t-1}) = \sigma_t(a^{t-1}, \bar{b}^{t-1})$ whenever $\bar{b}^{t-1}$ is a permutation of $\bar{a}^{t-1}$, i.e., a player's action depends on how many times his current partner has cooperated, but not on at which period the partner cooperated.

The expressions in (6) and (7) in Subsection 3.2 show that, if $g = l$, then linear independent and indifferent equilibria are not permutation-invariant: records in earlier periods have smaller effects on the current play than records in later periods. This observation is generalized to nonlinear independent and indifferent equilibria without any restriction on payoff parameters. The reason is as follows. Any permutation-invariant equilibrium treats recent defection and defection in the distant past equally. Also, since each player is indifferent between cooperation and defection, his continuation payoff after defection is lower than the continuation payoff after cooperation at least by a certain positive amount. Therefore, each time he defects, his continuation payoff differs from the continuation payoff he would get after the record of full cooperation at least by the same amount. This procedure, however, does not continue forever since continuation payoffs are uniformly bounded.

Proposition 8. *There is no permutation-invariant independent and indifferent equilibrium.*

Proof. Suppose that σ^* is a permutation-invariant independent and indifferent equilibrium. Let $p_t(k)$ be the probability that a player cooperates at period t when his partner has cooperated k times, and $V_t(k)$ the continuation payoff of a player at period t who has cooperated k times. Similarly to (5), for any $t \geq 1$ and any $k = 0, \dots, t-1$, I obtain

$$\begin{aligned} V_t(k) &= (1-\delta)[p_t(k) - (1-p_t(k))l] + \delta V_{t+1}(k+1) \\ &= (1-\delta)p_t(k)(1+g) + \delta V_{t+1}(k), \end{aligned}$$

which implies

$$V_{t+1}(k+1) - V_{t+1}(k) = \frac{1-\delta}{\delta}[p_t(k)g + (1-p_t(k))l] \geq \frac{1-\delta}{\delta}\min(g,l).$$

Therefore, I have

$$V_{t+1}(t) - V_{t+1}(0) \geq t \times \frac{1-\delta}{\delta}\min(g,l)$$

for any $t \geq 1$, which contradicts $V_{t+1}(0)$, $V_{t+1}(t) \in [-l, 1+g]$ if t is large enough. □

Note that Proposition 8 leaves it open whether cooperation can be sustained by a permutation-invariant equilibrium that does not satisfy either independence of own play or indifference at all histories. For example, the pairwise grim-trigger strategy is permutation-invariant. Thus, by Lemma 2, cooperation can be sustained by a permutation-invariant equilibrium if $g/(1+g) \leq \delta \leq l/(1+l)$. It is, however, difficult to extend this result to any sufficiently large discount factor because permutation invariance is not preserved under Ellison's trick.

5.2 Bounded Records

A strategy σ has *records of length* T if $\sigma_t(a^{t-1}, \bar{a}^{t-1}) = \sigma_t(a^{t-1}, \bar{b}^{t-1})$ whenever $(\bar{a}_{t-T}, \bar{a}_{t-T+1}, \dots, \bar{a}_{t-1}) = (\bar{b}_{t-T}, \bar{b}_{t-T+1}, \dots, \bar{b}_{t-1})$. Similarly to permutation invariance, it follows from (6) and (7) that linear independent and indifferent equilibria do not have bounded records at least when $g = l$ (except for $\delta = g/(1+g)$). Then I can ask the following question: can nonlinear independent and indifferent equilibria have bounded records? This question is answered in the following proposition.

Proposition 9. (1) *If $g \neq l$, then there is no independent and indifferent equilibrium with bounded records.*

(2) *If* $g = l$ *and* $\delta \geq g/(1+g)$*, then, for any* $x \in [0,1]$*, there is an independent and indifferent equilibrium that has records of length* 1 *with symmetric payoff* x*.*

Proof. See Appendix A.7. □

Proposition 9 shows that unboundedly long records are necessary to keep players indifferent at all histories if $g \neq l$, but records of length 1 are sufficient to sustain cooperation if $g = l$.

In his analysis on prisoner's dilemma, which corresponds to $g = l = 1/3$ in my notation, Rosenthal (1979) argued that cooperation can be sustained only in a knife-edge case. Namely, he showed that there exist pure-strategy equilibria with records of length 1 that sustain cooperation if and only if $\delta = 1/4$. Proposition 9 (2) shows that his construction can be extended to the case of $\delta \geq 1/4$ if mixed strategies are allowed. Also, as I showed at the end of Subsection 3.2, for any $\delta \geq (1+g+l)/(2+g+l) = 5/8$, there exists a pure-strategy independent and indifferent equilibrium that sustains cooperation. To sum up, Rosenthal's non-genericity result is resolved if either mixed strategies or unboundedly long records are allowed.

Note that Proposition 9 deals only with independent and indifferent equilibria. The next proposition shows that, if $g < l$, then, for any sufficiently large discount factor, there exists an equilibrium with bounded records without indifference at all histories that sustains cooperation. The case of $g > l$ is an open question.

Proposition 10. *If* $g < l$ *and* $\delta \geq g(1+l)/[(1+g)l]$*, then there exists an equilibrium that has records of length* T *with symmetric payoff* 1*, where* T *is the smallest integer that satisfies* $\delta^T \leq l/(1+l)$*.*

Proof. Consider a strategy σ such that a player cooperates at period t if $t = 1$ or ($t \geq 2$ and he and his partner cooperated at period $t-1$), and defects otherwise. Note that σ has records of length 1. Similarly to Lemma 2, I can show that σ is an equilibrium if and only if $g/(1+g) \leq \delta \leq l/(1+l)$. Using Ellison's trick as in Proposition 6, I divide the game into T mini-games, and let players follow σ in each mini-game. The modified strategy has records of length T, and is an equilibrium for $\delta \geq g(1+l)/[(1+g)l]$. □

5.3 Second-Order Information

Note that Proposition 9 uses only first-order information. If higher-order information is available, then a finite bound on the length of records may not be restrictive even if $g \neq l$.

Here I assume that, at every period $t \geq 2$, every player i receives a report from the criminal history repository about the realized action *profile* in his current partner's previous match, $(a_{m_t(i),t-1}, a_{m_{t-1}(m_t(i)),t-1})$, before the stage game is played. Compared to the original model, each player has more information in that he obtains a part of second-order information, but less information in that he loses access to the record of his current partner's play more than a period ago.

Proposition 11. *Under the above information structure, if $\delta \geq \max(g/(1+g), l/(1+l))$, then, for any $x \in [0,1]$, there exists an equilibrium with symmetric payoff x.*

Proof. See Appendix A.8 □

If player i at period $t \geq 2$ observes his current partner $m_t(i)$'s action at the previous period (period $t-1$) only, then he does not know what kind of incentive problem his current partner was facing at the previous period. The short-term gain for player $m_t(i)$ to choose defection depends on the mixed action chosen by his partner at that time (player $m_{t-1}(m_t(i))$) unless $g = l$. Without knowing this action, player i cannot make his current partner indifferent at the previous period. This is why independent and indifferent equilibria with records of length 1 exist only when $g = l$ (Proposition 9).[16]

What Proposition 11 shows is the converse of this negative result; player i can make his current partner indifferent if he knows the realization of his current partner's previous partner's mixed action, $a_{m_{t-1}(m_t(i)),t-1}$. Even though player i does not observe the mixing probability of player $m_{t-1}(m_t(i))$, knowing its realization gives him enough instruments to control player $m_t(i)$'s expected continuation payoffs from period $t-1$ on. From player $m_{t-1}(m_t(i))$'s point of view, by using his own action, he can convery necessary information about the incentive problem of his partner (player $m_t(i)$) to his partner's next partner (player i).

The proof of Proposition 11 is based on Ely and Välimäki (2002). In the two-player repeated prisoner's dilemma with private monitoring, Ely and Välimäki constructed a belief-free equilibrium in which each player's mixing probability depends only on his own action and the signal about his opponent's action at the previous period. In the community enforcement model, each player follows a Ely and Välimäki-type equilibrium, although a player's own action at the previous period is replaced by his current partner's previous partner's action at the previous period.

[16]Proposition 9 (1) shows a slightly stronger claim, extending the negative result to records of any finite length.

6 Option to Exit

In decentralized markets without binding contracts, people can stay out of particular business relations if they want. This is equivalent to modifying the stage game as follows:

	C	D	Out
C	$1,1$	$-l, 1+g$	w,w
D	$1+g,-l$	$0,0$	w,w
Out	w,w	w,w	w,w

where w is the reservation payoff that a player obtains outside the matching. I assume that $w < 1$, i.e., the reservation payoff is less than the payoff from cooperation. If a player chooses "Out" at some period, then he stays outside the matching during this period, and comes back at the next period. The option to exit was analyzed also by Dixit (2003b) in the context of the twice repeated prisoner's dilemma.

The next proposition shows that a variant of pairwise grim-trigger strategy can sustain cooperation as an equilibrium outcome if the option to exit is available.

Proposition 12. *Suppose that the stage game is the prisoner's dilemma with the option to exit. If $\delta \geq g/(1+g-w)$, then there exists an equilibrium with symmetric payoff* 1.

Proof. Consider the following strategy σ^*: $\sigma_t^*(a^{t-1}, \bar{a}^{t-1}) = C$ if $a_s \neq D$ and $\bar{a}_s \neq D$ for all $s = 1, \dots, t-1$, and $\sigma_t^*(a^{t-1}, \bar{a}^{t-1}) = Out$ otherwise. In other words, a player cooperates if neither he nor his current partner has ever defected, and opts out otherwise.

First, suppose that neither a^{t-1} nor $\bar{a}^{t-1}$ contains D. Then a player with record a^{t-1} weakly prefers cooperation to defection if and only if

$$1 \geq (1-\delta)(1+g) + \delta w,$$

which holds when $\delta \geq g/(1+g-w)$. Also he prefers cooperation to opting out since $w < 1$.

Second, suppose that a^{t-1} does not contain D but $\bar{a}^{t-1}$ does. Then a player with record a^{t-1} expects his current partner to opt out. So he is indifferent between cooperation and opting out, whereas he gets worse off if he defects.

Lastly, suppose that a^{t-1} contains D. Then no matter what action he chooses, his current and future partners will choose "Out." Thus he is indifferent among three actions. □

In the story behind the game, no trade occurs if at least one of the parties opts out. Then, it may be natural that this situation is recorded as "Out" for both parties even if only one side chooses to exit. Proposition 12 holds under this modification as well.

Proposition 12 exhibits a clear contrast between the prisoner's dilemma games with and without the option to exit. Unlike in the original prisoner's dilemma, it is not difficult to sustain cooperation in the modified prisoner's dilemma if the discount factor is sufficiently large. Neither restrictions on parameters g and l nor Ellison's trick is needed. The reason for these differences is that opting out is a costly action for both parties. Therefore, when a player punishes another player with a bad record by opting out of the match, the former player does not take a risk of being suspected as a deviator.

Although the equilibrium in Proposition 12 is simple, it works only when the outside option is bad enough. If the reservation payoff is higher than $\max(0, 1 - g/l)$, then the interval $[\max(g/(1 + g), l/(1 + l)), g/(1 + g - w))$ is nonempty. If δ is within this interval, then there is an independent and indifferent equilibrium that sustains cooperation (Proposition 2) while the assumption in Proposition 12 is not satisfied. In this case, punishment by defection is severe enough to deter deviations whereas punishment by quitting trade is not.

Appendix

A.1 Proof of Proposition 1

Fix a strategy σ^* that satisfies indifference at all histories. Let μ^* be generated by σ^*. For $t \geq 1$ and a^{t-1}, b^{t-1}, $\bar{a}^{t-1} \in A^{t-1}$, let $p = \sigma_t^*(a^{t-1}, \bar{a}^{t-1})(C)$ and $q = \sigma_t^*(b^{t-1}, \bar{a}^{t-1})(C)$. I will show that $p = q$.

Consider a player with record $\bar{a}^{t-1}$ who faces a player with record a^{t-1}. Since (4) is satisfied with equality, I have

$$\begin{aligned}(1-\delta)(p-(1-p)l) + \delta \sum_{\bar{b}^t \in A^t} U_{t+1}(\sigma^*, \sigma^* \mid (\bar{a}^{t-1}, C), \bar{b}^t, \mu^*)\mu_t^*(\bar{b}^t) \\ = (1-\delta)p(1+g) + \delta \sum_{\bar{b}^t \in A^t} U_{t+1}(\sigma^*, \sigma^* \mid (\bar{a}^{t-1}, D), \bar{b}^t, \mu^*)\mu_t^*(\bar{b}^t). \qquad (22)\end{aligned}$$

Similarly, considering a player with record $\bar{a}^{t-1}$ who faces a player with record

b^{t-1}, I have

$$\begin{aligned}(1-\delta)(q-(1-q)l)+\delta\sum_{\bar{b}^t\in A^t}U_{t+1}(\sigma^*,\sigma^*\mid(\bar{a}^{t-1},C),\bar{b}^t,\mu^*)\mu_t^*(\bar{b}^t)\\ =(1-\delta)q(1+g)+\delta\sum_{\bar{b}^t\in A^t}U_{t+1}(\sigma^*,\sigma^*\mid(\bar{a}^{t-1},D),\bar{b}^t,\mu^*)\mu_t^*(\bar{b}^t).\end{aligned}\qquad(23)$$

Deducting (23) from (22), I can cancel all continuation payoff terms. Thus it holds that

$$(p-(1-p)l)-(q-(1-q)l)=p(1+g)-q(1+g).$$

Rearranging this, I have $(p-q)(g-l)=0$, which implies $p=q$ since $g\neq l$.

A.2 Derivation of Equation (8)

Let σ^* be the linear independent and indifferent equilibrium with symmetric payoff 1. Denote $p_t(\bar{a}^{t-1})=\sigma^*(a^{t-1},\bar{a}^{t-1})(C)$ and $P_t=P_t(\sigma_t^*,\mu_{t-1})$. By independence of own play,

$$\begin{aligned}P_t&=\sum_{a^{t-1}}\mu_{t-1}(a^{t-1})p_t(a^{t-1}),\\ \mu_t(a^{t-1},C)&=\mu_{t-1}(a^{t-1})P_t,\\ \mu_t(a^{t-1},D)&=\mu_{t-1}(a^{t-1})(1-P_t)\end{aligned}$$

for $t\geq T+1$.

For $t\geq T+1$, I can write P_{t+1} as follows:

$$\begin{aligned}P_{t+1}&=\sum_{a^t}\mu_t(a^t)p_{t+1}(a^t)\\ &=\sum_{a^{t-1}}\mu_t(a^{t-1},C)p_{t+1}(a^{t-1},C)+\sum_{a^{t-1}}\mu_t(a^{t-1},D)p_{t+1}(a^{t-1},D)\\ &=\sum_{a^{t-1}}\mu_{t-1}(a^{t-1})P_t\left(\frac{p_t(a^{t-1})}{\delta}-\frac{1-\delta}{\delta}[p_t(a^{t-1})-(1-p_t(a^{t-1}))l]\right)\\ &\quad+\sum_{a^{t-1}}\mu_{t-1}(a^{t-1})(1-P_t)\left(\frac{p_t(a^{t-1})}{\delta}-\frac{1-\delta}{\delta}p_t(a^{t-1})(1+g)\right)\\ &=\left(\frac{1-\delta}{\delta}(g-l)P_t+1-\frac{1-\delta}{\delta}g\right)\sum_{a^{t-1}}\mu_{t-1}(a^{t-1})p_t(a^{t-1})+\frac{1-\delta}{\delta}lP_t\sum_{a^{t-1}}\mu_{t-1}(a^{t-1})\\ &=\left(\frac{1-\delta}{\delta}(g-l)P_t+1-\frac{1-\delta}{\delta}g\right)P_t+\frac{1-\delta}{\delta}lP_t\\ &=P_t-\frac{1-\delta}{\delta}(g-l)P_t(1-P_t).\end{aligned}$$

A.3 Long-Run Stability when $g < l$

Let σ^* be the linear independent and indifferent equilibrium with symmetric payoff 1. Since $g < l$, it is easy to see from (8) that $P_t(\sigma_t^*, \mu_{t-1}) \to 1$ as $t \to \infty$ as long as $P_{T+1}(\sigma_{T+1}^*, \mu_T) > 0$. Now I show that $\sigma_{T+1}^*(a^T, \bar{a}^T) > 0$ for any a^T and $\bar{a}^T$ so that $P_{T+1}(\sigma_{T+1}^*, \mu_T) > 0$ no matter how the shock changes μ_T.

Let $p_t(\bar{a}^{t-1}) = \sigma_t^*(a^{t-1}, \bar{a}^{t-1})$. By Lemma 1, I obtain $p_1(\emptyset) = 1 > 0$. Also, if $p_t(a^{t-1}) > 0$, then

$$p_{t+1}(a^{t-1}, C) \geq p_{t+1}(a^{t-1}, D) = \frac{p_t(a^{t-1})}{\delta} - \frac{1-\delta}{\delta} p_t(a^{t-1})(1+g) > 0$$

since $\delta \geq l/(1+l) > g/(1+g)$. Thus, by induction, I obtain $p_{T+1}(a^T) > 0$ for any a^T.

A.4 Proof of Proposition 3

The proof is almost the same as in Proposition 2. The only difference is that the probability of cooperation needs to be between ε_D and $1 - \varepsilon_C$ at any history. Thus I need to modify Lemma 1 in the following way.

Lemma 5. *If $\varepsilon_C(1+l) + \varepsilon_D(1+g) < 1$ and*

$$\delta \geq \frac{1}{1-\varepsilon_C-\varepsilon_D} \max\left(\frac{(1-\varepsilon_C)g + \varepsilon_C l}{1+g}, \frac{(1-\varepsilon_D)l + \varepsilon_D g}{1+l}\right),$$

then, for every $t \geq 1$, every $a^{t-1} \in A^{t-1}$, and every $V_t(a^{t-1}) \in [\varepsilon_D(1+g), 1-\varepsilon_C(1+l)]$, there exist a probability $p_t(y^{t-1}) \in [\varepsilon_D, 1-\varepsilon_C]$ and target payoffs $V_{t+1}(a^{t-1}, C)$, $V_{t+1}(a^{t-1}, D) \in [\varepsilon_D(1+g), 1-\varepsilon_C(1+l)]$ such that (5) is satisfied.

Proof. Let

$$\begin{aligned} p_t(a^{t-1}) &= \frac{1-\varepsilon_C-\varepsilon_D}{1-\varepsilon_C(1+l)-\varepsilon_D(1+g)}[V_t(a^{t-1}) - \varepsilon_D(1+g)] + \varepsilon_D, \\ V_{t+1}(a^{t-1}, C) &= \frac{V_t(a^{t-1})}{\delta} - \frac{1-\delta}{\delta}[p_t(a^{t-1}) - (1-p_t(a^{t-1}))l], \\ V_{t+1}(a^{t-1}, D) &= \frac{V_t(a^{t-1})}{\delta} - \frac{1-\delta}{\delta} p_t(a^{t-1})(1+g). \end{aligned}$$

By the construction, (5) is satisfied. Since $V_t(a^{t-1}) \in [\varepsilon_D(1+g), 1-\varepsilon_C(1+l)]$, I have $p_t(a^{t-1}) \in [\varepsilon_D, 1-\varepsilon_C]$. Also, by the assumption on δ, I have $V_{t+1}(a^{t-1}, C)$, $V_{t+1}(a^{t-1}, D) \in [\varepsilon_D(1+g), 1-\varepsilon_C(1+l)]$. □

The rest of the proof is the same as in Proposition 2.

A.5 Proof of Proposition 5

(1) First, I will show that, if there exists an equilibrium σ^* with symmetric payoff x that satisfies independence of own play, then $x \in [\underline{v}, \overline{v}]$. Let

$$\underline{V}_t = \min_{a^{t-1} \in A^{t-1}} U_t(\sigma^*, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu^*),$$
$$\overline{V}_t = \max_{a^{t-1} \in A^{t-1}} U_t(\sigma^*, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu^*),$$

where μ^* is generated by σ^*, and

$$B_t = \{a_t \in A \mid \sigma_t^*(a^{t-1}, \bar{a}^{t-1})(a_t) > 0 \text{ for some } \bar{a}^{t-1} \in A^{t-1}\}.$$

Since σ^* satisfies independence of own play, $\overline{V}_t$ and $\underline{V}_t$ are independent of $\bar{a}^{t-1}$, and B_t is independent of a^{t-1}.

Fix any $\bar{a}^{t-1} \in A^{t-1}$. For any $a^{t-1} \in A^{t-1}$ and any $a_t \in A$, I have

$$\begin{aligned} &U_t(\sigma^*, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu^*) \\ &\quad \geq (1-\delta) u(a_t, \sigma_t^*(\bar{a}^{t-1}, a^{t-1})) + \delta \sum_{\bar{b}^t \in A^t} U_{t+1}(\sigma^*, \sigma^* \mid (a^{t-1}, a_t), \bar{b}^t, \mu^*) \mu_t^*(\bar{b}^t) \\ &\quad \geq (1-\delta) u(a_t, \sigma_t^*(\bar{a}^{t-1}, a^{t-1})) + \delta \underline{V}_{t+1}, \end{aligned}$$

Thus

$$U_t(\sigma^*, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu^*) \geq (1-\delta) \max_{a_t \in A} u(a_t, \sigma_t^*(\bar{a}^{t-1}, a^{t-1})) + \delta \underline{V}_{t+1}$$

for any a^{t-1}, which implies

$$\begin{aligned} \underline{V}_t &= \min_{a^{t-1} \in A^{t-1}} U_t(\sigma^*, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu^*) \\ &\geq (1-\delta) \min_{a^{t-1} \in A^{t-1}} \max_{a_t \in A} u(a_t, \sigma_t^*(\bar{a}^{t-1}, a^{t-1})) + \delta \underline{V}_{t+1} \\ &\geq (1-\delta) \underline{v}(B_t) + \delta \underline{V}_{t+1} \end{aligned}$$

since $\sigma_t^*(\bar{a}^{t-1}, a^{t-1}) \in \Delta(B_t)$. Repeatedly applying this inequality, I obtain

$$x = \underline{V}_1 \geq (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} \underline{v}(B_t). \tag{24}$$

For any $a^{t-1} \in A^{t-1}$ and any $a_t \in B_t$, there exists $\bar{a}^{t-1} \in A^{t-1}$ such that $\sigma_t^*(a^{t-1}, \bar{a}^{t-1})(a_t) > 0$. Thus I have

$$\begin{aligned} &U_t(\sigma^*, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu^*) \\ &\quad = (1-\delta) u(a_t, \sigma_t^*(\bar{a}^{t-1}, a^{t-1})) + \delta \sum_{\bar{b}^t \in A^t} U_{t+1}(\sigma^*, \sigma^* \mid (a^{t-1}, a_t), \bar{b}^t, \mu^*) \mu_t^*(\bar{b}^t) \\ &\quad \leq (1-\delta) u(a_t, \sigma_t^*(\bar{a}^{t-1}, a^{t-1})) + \delta \overline{V}_{t+1}. \end{aligned}$$

Since $\sigma_t^*(\bar{a}^{t-1}, a^{t-1})$ is independent of $\bar{a}^{t-1}$,

$$U_t(\sigma^*, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu^*) \leq (1-\delta) \min_{a_t \in B_t} u(a_t, \sigma_t^*(\bar{a}^{t-1}, a^{t-1})) + \delta \overline{V}_{t+1}$$

for any a^{t-1}, which implies

$$\begin{aligned}\overline{V}_t &= \max_{a^{t-1} \in A^{t-1}} U_t(\sigma^*, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu^*) \\ &\leq (1-\delta) \max_{a^{t-1} \in A^{t-1}} \min_{a_t \in B_t} u(a_t, \sigma_t^*(\bar{a}^{t-1}, a^{t-1})) + \delta \overline{V}_{t+1} \\ &\leq (1-\delta)\overline{v}(B_t) + \delta \overline{V}_{t+1}\end{aligned}$$

since $\sigma_t^*(\bar{a}^{t-1}, a^{t-1}) \in \Delta(B_t)$. Repeatedly applying this inequality, I obtain

$$x = \overline{V}_1 \leq (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} \overline{v}(B_t). \tag{25}$$

Let

$$w(B) = (1-\delta) \sum_{t \geq 1,\, B_t = B} \delta^{t-1}.$$

Then, by (24) and (25),

$$\sum_{B \in \mathcal{P}} w(B) \underline{v}(B) \leq x \leq \sum_{B \in \mathcal{P}} w(B) \overline{v}(B),$$

which implies that $w \in W$. Thus $\underline{v} \leq x \leq \overline{v}$.

Second, pick any $x \in (\underline{v}, \overline{v})$. Then there exist $\varepsilon > 0$, $\delta_0 < 1$, and a periodic sequence $\{B_t\}$ with period K (i.e., $B_s = B_t$ if $s - t$ is divisible by K) such that, for any $\delta > \delta_0$, I have

$$\underline{v}_t \equiv \sum_{B \in \mathcal{P}} w_t(B) \underline{v}(B) \leq x - \varepsilon < x + \varepsilon \leq \sum_{B \in \mathcal{P}} w_t(B) \overline{v}(B) \equiv \overline{v}_t \tag{26}$$

for every $t \geq 1$, where

$$w_t(B) = (1-\delta) \sum_{s \geq t,\, B_s = B} \delta^{s-t}. \tag{27}$$

For each $B \in \mathcal{P}$, pick any

$$\underline{\alpha}(B) \in \arg\min_{\alpha \in \Delta(B)} \max_{a \in A} u(a, \alpha), \qquad \overline{\alpha}(B) \in \arg\max_{\alpha \in \Delta(B)} \min_{a \in B} u(a, \alpha).$$

For any $p \in [0, 1]$, let

$$\alpha(B, p) = p\overline{\alpha}(B) + (1-p)\underline{\alpha}(B)$$

be the convex combination of $\overline{\alpha}(B)$ and $\underline{\alpha}(B)$ with weights p and $1-p$, respectively. Note that $\alpha(B,p) \in \Delta(B)$ for any $p \in [0,1]$ since $\overline{\alpha}(B), \underline{\alpha}(B) \in \Delta(B)$.

I will construct an equilibrium with symmetric payoff x that satisfies independence of own play as follows:

- At period t, players with record a^{t-1} are assigned with target payoff $V_t(a^{t-1})$. Set $V_1(\emptyset) = x$.
- A player who meets another player with record $\bar{a}^{t-1}$ plays mixed action $\overline{\alpha}(B_t)$ with probability $p_t(\bar{a}^{t-1})$ and $\underline{\alpha}(B_t)$ with probability $1-p_t(\bar{a}^{t-1})$, i.e., the former player plays $\alpha(B_t, p_t(\bar{a}^{t-1}))$.
- The target payoff in the next period, $V_{t+1}(a^t)$, is computed recursively from the current target payoff $V_t(a^{t-1})$ and the current action a_t.

In order to provide an incentive for a player to mix actions over B_t at period t, I will find $p_t(a^{t-1})$ and $V_{t+1}(a^{t-1}, \cdot)$ that satisfy

$$V_t(a^{t-1}) = (1-\delta)u(a_t, \alpha(B_t, p_t(\bar{a}^{t-1}))) + \delta V_{t+1}(a^{t-1}, a_t) \tag{28}$$

for any $a_t \in B_t$ and

$$V_t(a^{t-1}) \geq (1-\delta)u(a_t, \alpha(B_t, p_t(\bar{a}^{t-1}))) + \delta V_{t+1}(a^{t-1}, a_t) \tag{29}$$

for any $a_t \in A \setminus B_t$.

Lemma 6. *Fix $\varepsilon > 0$, $\delta_0 < 1$, and a periodic sequence $\{B_t\}$ with period K such that (26) and (27) are satisfied for any $\delta > \delta_0$. Then, for each $k = 1, \dots, K$, there exists $\delta_k < 1$ such that, for every $\delta > \delta_k$, every $t \geq 1$ such that $t-k$ is divisible by K, every $a^{t-1} \in A^{t-1}$, and every $V_t(a^{t-1}) \in [\underline{v}_t, \overline{v}_t]$, there exist a probability $p_t(a^{t-1}) \in [0,1]$ and target payoffs $V_{t+1}(a^{t-1}, a_t) \in [\underline{v}_{t+1}, \overline{v}_{t+1}]$ for $a_t \in A$ such that (28) is satisfied for any $a_t \in B_t$ and (29) is satisfied for any $a_t \in A \setminus B_t$.*

Proof. Let

$$p_t(a^{t-1}) = \frac{V_t(a^{t-1}) - \underline{v}_t}{\overline{v}_t - \underline{v}_t},$$

$$V_{t+1}(a^{t-1}, a_t) = \begin{cases} \frac{V_t(a^{t-1})}{\delta} - \frac{1-\delta}{\delta} u(a_t, \alpha(B_t, p_t(a^{t-1}))) & \text{if } a_t \in B_t, \\ \underline{v}_{t+1} & \text{if } a_t \in A \setminus B_t. \end{cases}$$

By the construction of $V_{t+1}(a^{t-1}, a_t)$, (28) is satisfied for $a_t \in B_t$. To check (29) for $a_t \in A \setminus B_t$, first consider the case of $V_t(a^{t-1}) = \overline{v}_t$. Then I have to show that

$$\overline{v}_t \geq (1-\delta)u(a_t, \overline{\alpha}(B_t)) + \delta \underline{v}_{t+1} \tag{30}$$

for $a_t \in A \setminus B_t$. Since $\overline{v}_t - \underline{v}_{t+1} \geq 2\varepsilon$, there exists $\delta'_t < 1$ such that (30) is satisfied for any $\delta > \delta'_t$. Next consider the case of $V_t(a^{t-1}) = \underline{v}_t$. Then I have to show that

$$\underline{v}_t \geq (1-\delta)u(a_t, \underline{\alpha}(B_t)) + \delta \underline{v}_{t+1} \tag{31}$$

for $a_t \in A \setminus B_t$. By the definition of $\underline{\alpha}(B_t)$, I have $u(a_t, \underline{\alpha}(B_t)) \leq \underline{v}(B_t)$. Also, by (26) and (27), I have $\underline{v}_t = (1-\delta)\underline{v}(B_t) + \delta \underline{v}_{t+1}$. Thus (31) is satisfied. For general $V_t(a^{t-1})$, I can show (29) by adding (30) and (31) with weights $p_t(a^{t-1})$ and $1 - p_t(a^{t-1})$, respectively.

Clearly, I have $p_t(a^{t-1}) \in [0, 1]$. Now I will show that $V_{t+1}(a^{t-1}, a_t) \in [\underline{v}_{t+1}, \overline{v}_{t+1}]$. Since the case of $a_t \in A \setminus B_t$ is obvious, I will focus on the case of $a_t \in B_t$. First, consider the case of $V_t(a^{t-1}) = \overline{v}_t$. In this case, I have to show that

$$\underline{v}_{t+1} \leq \frac{\overline{v}_t}{\delta} - \frac{1-\delta}{\delta} u(a_t, \overline{\alpha}(B_t)) \leq \overline{v}_{t+1}. \tag{32}$$

Since $\overline{v}_t - \underline{v}_{t+1} \geq 2\varepsilon$, there exists $\delta''_t < 1$ such that the first inequality in (32) is satisfied for any $\delta > \delta''_t$. By the definition of $\overline{\alpha}(B_t)$, I have $u(a_t, \overline{\alpha}(B_t)) \geq \overline{v}(B_t)$. Also, by (26) and (27), I have $\overline{v}_t = (1-\delta)\overline{v}(B_t) + \delta \overline{v}_{t+1}$. Thus the second inequality in (32) is satisfied. Next, consider the case of $V_t(a^{t-1}) = \underline{v}_t$. In this case, I have to show that

$$\underline{v}_{t+1} \leq \frac{\underline{v}_t}{\delta} - \frac{1-\delta}{\delta} u(a_t, \underline{\alpha}(B_t)) \leq \overline{v}_{t+1}. \tag{33}$$

By the definition of $\underline{\alpha}(B_t)$, I have $u(a_t, \underline{\alpha}(B_t)) \leq \underline{v}(B_t)$. Also, by (26) and (27), I have $\underline{v}_t = (1-\delta)\underline{v}(B_t) + \delta \underline{v}_{t+1}$. Thus the first inequality in (33) is satisfied. Since $\overline{v}_{t+1} - \underline{v}_t \geq 2\varepsilon$, there exists $\delta'''_t < 1$ such that the second inequality in (33) is satisfied for any $\delta > \delta'''_t$. For general $V_t(a^{t-1})$, I can show $V_{t+1}(a^{t-1}, a_t) \in [\underline{v}_{t+1}, \overline{v}_{t+1}]$ by adding (32) and (33) with weights $p_t(a^{t-1})$ and $1 - p_t(a^{t-1})$, respectively.

Therefore, desired probability $p_t(a^{t-1})$ and target payoffs $V_{t+1}(a^{t-1}, \cdot)$ exist if $\delta > \delta_t \equiv \max(\delta_0, \delta'_t, \delta''_t, \delta'''_t)$. Finally, since $\{B_t\}$ has period K, I can find δ_t that depends only on the remainder when t is divided by K. $\square$

Applying Lemma 6 recursively with initial condition $V_1(\emptyset) = x$, I can construct an equilibrium with symmetric payoff x that satisfies independence of own play for any $\delta > \bar{\delta} \equiv \max_{1 \le k \le K} \delta_k$.

(2) For the negative case, suppose that there exists an equilibrium with symmetric payoff x that satisfies independence of own play. Let

$$\underline{V}_t = \min_{a^{t-1}} U_t(\sigma^*, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu^*),$$
$$\overline{V}_t = \max_{a^{t-1}} U_t(\sigma^*, \sigma^* \mid a^{t-1}, \bar{a}^{t-1}, \mu^*),$$

where μ^* is generated by σ^*, and

$$B_t = \{a_t \in A \mid \sigma_t^*(a^{t-1}, \bar{a}^{t-1})(a_t) > 0 \text{ for some } \bar{a}^{t-1} \in A^{t-1}\}.$$

Similarly to the proof of Part (1), I obtain

$$\sum_{B \in \mathcal{P}} w_t(B)\underline{v}(B) \le \underline{V}_t \le \overline{V}_t \le \sum_{B \in \mathcal{P}} w_t(B)\overline{v}(B),$$

where

$$w_t(B) = (1-\delta) \sum_{s \ge t,\, B_s = B} \delta^{s-t}.$$

Since $\underline{v}(B) \ge \overline{v}(B)$ for every $B \in \mathcal{P}$, I have $\underline{V}_t = \overline{V}_t$. Since the continuation payoff is constant, the mixed action at period $t-1$ is a symmetric Nash equilibrium of the stage game. Thus x is a convex combination of symmetric Nash equilibrium payoffs of the stage game.

If $\delta \ge 1/2$, then it is well known that I can sustain any payoff in $[\min NE, \max NE]$ by altering the best and the worst symmetric Nash equilibria over time.

A.6 Proof of Lemma 3

If (16) is satisfied, then

$$\begin{aligned}
&(1-\delta) + \delta V_{t+1}(a^{t-1}, C) \\
&\quad = (1-\delta)(1-p_t^H)(1+l) + (1-\delta)(p_t^H - (1-p_t^H)l) + \delta V_{t+1}(a^{t-1}, C) \\
&\quad = (1-\delta)(1-p_t^H)(1+l) + (1-\delta)p_t^H(1+g) + \delta V_{t+1}(a^{t-1}, D) \\
&\quad < (1-\delta)(1-p_t^H)(1+g) + (1-\delta)p_t^H(1+g) + \delta V_{t+1}(a^{t-1}, D) \\
&\quad = (1-\delta)(1+g) + \delta V_{t+1}(a^{t-1}, D),
\end{aligned}$$

where the inequality follows from $g > l$ and $p_t^H < 1$.

Similarly, if (17) is satisfied, then

$$
\begin{aligned}
&(1-\delta)(-l)+\delta V_{t+1}(a^{t-1},C) \\
&\quad = -(1-\delta)p_t^L(1+l)+(1-\delta)(p_t^L-(1-p_t^L)l)+\delta V_{t+1}(a^{t-1},C) \\
&\quad = -(1-\delta)p_t^L(1+l)+(1-\delta)p_t^L(1+g)+\delta V_{t+1}(a^{t-1},D) \\
&\quad > -(1-\delta)p_t^L(1+g)+(1-\delta)p_t^L(1+g)+\delta V_{t+1}(a^{t-1},D) \\
&\quad = \delta V_{t+1}(a^{t-1},D),
\end{aligned}
$$

where the inequality follows from $g > l$ and $p_t^L > 0$.

A.7 Proof of Proposition 9

I use the following lemma to prove Part (1).

Lemma 7. *Suppose that $g \neq l$. Fix any $s \geq 1$. For $t \geq s+1$, if $p_t(a^{t-1})$ and $V_{t+1}(a^t)$ satisfy (5) for every a^{t-1}, and $V_{t+1}(a^t)$ is independent of a_s, then $p_t(a^{t-1})$ and $V_t(a^{t-1})$ are also independent of a_s.*

Proof. Since $g \neq l$, it follows from (5) that $p_t(a^{t-1})$ is uniquely determined by $V_{t+1}(a^{t-1},C)$ and $V_{t+1}(a^{t-1},D)$ as follows:

$$
p_t(a^{t-1}) = \frac{\delta}{1-\delta}\frac{V_{t+1}(a^{t-1},C)-V_{t+1}(a^{t-1},D)}{g-l} - \frac{l}{g-l}.
$$

Since $V_{t+1}(a^{t-1},C)$ and $V_{t+1}(a^{t-1},D)$ are independent of a_s, so is $p_t(a^{t-1})$. By (5), $V_t(a^{t-1})$ is also independent of a_s. ☐

In a proof by contradiction, I assume that σ^* is an independent and indifferent equilibrium with records of length T, but not of length $T-1$. If $T=0$, then it implies that σ^* is the repetition of D. This contradicts the indifference property of independent and indifferent equilibria. Thus I can assume $T \geq 1$ without loss of generality.

For any a^{t-1}, since σ^* has records of length T, $V_{t+1}(a^t)$ is independent of a_{t-T}. By Lemma 7, $p_t(a^{t-1})$ is also independent of a_{t-T}, which contradicts the assumption that σ^* does not have records of length $T-1$.

I prove Part (2) by constructing equilibria explicitly. Since $\delta \geq g/(1+g)$, for any $x \in [0,1]$, I can find a real number y such that

$$
0 \leq y \leq x \leq y + \frac{1-\delta}{\delta}g \leq 1.
$$

For any $t \geq 1$ and any a^{t-1}, define $V_t(a^{t-1})$ and $p_t(a^{t-1})$ as follows:

$$V_t(a^{t-1}) = \begin{cases} x & \text{if } t = 1, \\ y + \frac{1-\delta}{\delta} g & \text{if } t \geq 2 \text{ and } a_{t-1} = C, \\ y & \text{if } t \geq 2 \text{ and } a_{t-1} = D, \end{cases}$$

$$p_t(a^{t-1}) = \begin{cases} \frac{x - \delta y}{(1-\delta)(1+g)} & \text{if } t = 1, \\ \frac{y}{1+g} + \frac{1}{\delta}\frac{g}{1+g} & \text{if } t \geq 2 \text{ and } a_{t-1} = C, \\ \frac{y}{1+g} & \text{if } t \geq 2 \text{ and } a_{t-1} = D, \end{cases}$$

It is easy to check that (5) is satisfied, $0 \leq V_t(a^{t-1}) \leq 1$, and $0 \leq p_t(a^{t-1}) \leq 1$ for all a^{t-1}. Thus the strategy profile induced by this $\{p_t\}$ is an independent and indifferent equilibrium that achieves $V_1(\emptyset) = x$.

A.8 Proof of Proposition 11

For any $x \in [0, 1]$, I will construct an equilibrium with symmetric payoff x as follows:

- At period 1, players are assigned with target payoff x. At period $t \geq 2$, players who chose a and whose previous partner chose $\bar{a}$ at period $t-1$ are assigned with target payoff $V(a, \bar{a})$. I will construct an equilibrium under which all players obtain their target payoffs as their continuation payoffs.
- At period 1, each player chooses action C with probability p_1 and action D with probability $1 - p_1$. At period $t \geq 2$, if a player meets another player who chose a and whose previous partner chose $\bar{a}$ at period $t-1$, then he chooses action C with probability $p(\bar{a}, \bar{\bar{a}})$ and action D with probability $1 - p(\bar{a}, \bar{\bar{a}})$. Note that $p(\bar{a}, \bar{\bar{a}})$ is stationary and independent of his and his previous partner's actions at period $t-1$.

To accomplish the construction, I specify p_1 and $\{(V(a, \bar{a}), p(a, \bar{a}))\}_{a, \bar{a} \in A}$ so that each player becomes indifferent between C and D at all histories.

Similarly to the proof of Proposition 2, I have the following promise-keeping and indifference conditions:

$$\begin{aligned} x &= p_1[(1-\delta) + \delta V(C, C)] + (1 - p_1)[-(1-\delta)l + \delta V(C, D)] \\ &= p_1[(1-\delta)(1+g) + \delta V(D, C)] + (1 - p_1)\delta V(D, D) \end{aligned}$$

and

$$\begin{aligned} V(a, \bar{a}) &= p(a, \bar{a})[(1-\delta) + \delta V(C, C)] + (1 - p(a, \bar{a}))[-(1-\delta)l + \delta V(C, D)] \\ &= p(a, \bar{a})[(1-\delta)(1+g) + \delta V(D, C)] + (1 - p(a, \bar{a}))\delta V(D, D) \end{aligned}$$

for any a, $\bar{a} \in A$. These conditions are satisfied if

$$\begin{aligned} p_1 &= x, \\ p(C,C) &= V(C,C) = 1, \\ p(C,D) &= V(C,D) = \frac{1-\delta}{\delta} l, \\ p(D,C) &= V(D,C) = 1 - \frac{1-\delta}{\delta} g, \\ p(D,D) &= V(D,D) = 0. \end{aligned}$$

Since $x \in [0,1]$ and $\delta \geq \max(g/(1+g), l/(1+l))$, I have p_1, $p(a,\bar{a}) \in [0,1]$ for any a, $\bar{a} \in A$.

References

[1] Alós-Ferrer, C., 1999. Dynamical Systems with a Continuum of Randomly Matched Agents. *Journal of Economic Theory*, 86 (2), 245–267.

[2] Bhaskar, V., 1998. Informational Constraints and the Overlapping Generations Model: Folk and Anti-Folk Theorems. *Review of Economic Studies*, 65, 135–149.

[3] Bhaskar, V., Mailath, G. J., Morris, S., 2006. Purification in the Infinitely-Repeated Prisoners' Dilemma. mimeo.

[4] Bhaskar, V., van Damme, E., 2002. Moral Hazard and Private Monitoring. *Journal of Economic Theory*, 102, 16–39.

[5] Bolton, G. E., Katok, E., Ockenfels, A., 2005. Cooperation among Strangers with Limited Information about Reputation. *Journal of Public Economics* 89, 1457–1468.

[6] Dal Bó, P., 2004. Social Norms, Cooperation and Inequality. Forthcoming in *Economic Theory.*

[7] Dellarocas, C., 2003. The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms. *Management Science*, 49 (10), 1407–1424.

[8] Dixit, A., 2003a. On Modes of Economic Governance. *Econometrica*, 71 (2), 449–481.

[9] Dixit, A., 2003b. Trade Expansion and Contract Enforcement. *Journal of Political Economy*, 111 (6), 1293–1317.

[10] Duffie, D., Sun, Y., 2004. The Exact Law of Large Numbers for Independent Random Matching. mimeo.

[11] Duffie, D., Sun, Y., 2006. Existence of Independent Random Matching. Forthcoming in *the Annals of Applied Probability*.

[12] Ellison, G., 1994. Cooperation in the Prisoner's Dilemma with Anonymous Random Matching. *Review of Economic Studies*, 61 (3), 567-588.

[13] Ely, J. C., Hörner, J., Olszewski, W., 2004. Dispensability of the randomization device in "Belief-free Equilibria in Repeated Games". http://www.kellogg.northwestern.edu/faculty/horner/ehoappendix.pdf.

[14] Ely, J. C., Hörner, J., Olszewski, W., 2005. Belief-Free Equilibria in Repeated Games. *Econometrica*, 73 (2), 377–415.

[15] Ely, J. C., Välimäki J., 2002. A Robust Folk Theorem for the Prisoner's Dilemma. *Journal of Economic Theory*, 102 (1), 84–105.

[16] Engelmann, D., Fischbacher, U., 2004. Indirect Reciprocity and Strategic Reputation Building in an Experimental Helping Game. mimeo.

[17] Feldman, M., Gilles, C., 1985. An Expository Note on Individual Risk without Aggregate Uncertainty. *Journal of Economic Theory*, 35, 26–32.

[18] Greif, A., 1993. Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition. *American Economic Review*, 83 (3), 525–548.

[19] 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 (5), 912–950.

[20] Harrington, J. E., 1995. Cooperation in a One-Shot Prisoners' Dilemma. *Games and Economic Behavior*, 8, 364–377.

[21] Harsanyi, J., 1973. Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points. *International Journal of Game Theory*, 2, 1–23.

[22] Hörner, J., Olszewski, W., 2006. The Folk Theorem for Games with Private Almost-Perfect Monitoring. Forthcoming in *Econometrica*.

[23] Judd, K. L., 1985. The Law of Large Numbers with a Continuum of IID Random Variables. *Journal of Economic Theory*, 35, 19–25.

[24] Kandori, M., 1992. Social Norms and Community Enforcement. *Review of Economic Studies*, 59 (1), 63–80.

[25] Kandori, M., Obara, I., 2006. Efficiency in Repeated Games Revisited: The Role of Private Strategies. *Econometrica*, 74 (2), 499–519.

[26] Klein, D. B., 1992. Promise Keeping in the Great Society: A Model of Credit Information Sharing. *Economics and Politics*, 4, 117–136.

[27] Kreps, D. M., Wilson, R., 1982. Sequential Equilibria. *Econometrica*, 50 (4), 863–894.

[28] McLennan, A., Sonnenschein, H., 1991. Sequential Bargaining as a Noncooperative Foundation for Walrasian Equilibrium. *Econometrica*, 59 (5), 1395–1424.

[29] Milgrom, P. R., North, D. C., Weingast, B. R., 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.

[30] Milinski, M., Semmann, D., Bakker, T. C. M., Krambeck, H.-J., 2001. Cooperation through Indirect Reciprocity: Image Scoring or Standing Strategy? *Proceedings of the Royal Society* B, 268, 2495–2501.

[31] Nowak, M. A., Sigmund, K., 1998a. Evolution of Indirect Reciprocity by Image Scoring. *Nature*, 393, 573–578.

[32] Nowak, M. A., Sigmund, K., 1998b. The Dynamics of Indirect Reciprocity. *Journal of Theoretical Biology*, 194, 561–574.

[33] Nowak, M. A., Sigmund, K., 2005. Evolution of Indirect Reciprocity. *Nature*, 437, 1291–1298.

[34] Okuno-Fujiwara, M., Postlewaite, A., 1995. Social Norms and Random Matching Games. *Games and Economic Behavior*, 9, 79–109.

[35] Panchanathan, K., Boyd, R., 2003. A Tale of Two Defectors: The Importance of Standing for the Evolution of Indirect Reciprocity. *Journal of Theoretical Biology*, 224, 115–126.

[36] Piccione, M., 2002. The Repeated Prisoner's Dilemma with Imperfect Private Monitoring. *Journal of Economic Theory*, 102 (1), 70–83.

[37] Rosenthal, R. W., 1979. Sequences of Games with Varying Opponents. *Econometrica*, 47 (6), 1353–1366.

[38] Rosenthal, R. W., Landau, H. J., 1979. A Game-Theoretic Analysis of Bargaining with Reputations. *Journal of Mathematical Psychology*, 20, 233–255.

[39] Sun, Y., 2006. The Exact Law of Large Numbers via Fubini Extension and Characterization of Insurable Risks. *Journal of Economic Theory*, 126, 31–69.

[40] Tirole, J., 1996. Theory of Collective Reputations (with Applications to the Persistence of Corruption and to Firm Quality). *Review of Economic Studies*, 63 (1), 1–22.

[41] Yamamoto, Y., 2006. A Characterization of Belief-Free Payoffs in Repeated Games with Almost-Perfect Monitoring. mimeo.