UNIVERSITÀ DEGLI STUDI FIRENZE
DISEI
DIPARTIMENTO DI SCIENZE PER L'ECONOMIA E L'IMPRESA

WORKING PAPERS - ECONOMICS

Cooperation in Temporary Partnerships

GABRIELE CAMERA, ALESSANDRO GIOFFRÉ

WORKING PAPER N. 15/2024

DISEI, Università degli Studi di Firenze
Via delle Pandette 9, 50127 Firenze (Italia) www.disei.unifi.it

Cooperation in Temporary Partnerships*

Gabriele Camera
ESI, Chapman University

Alessandro Gioffré
University of Florence

April 26, 2024

Abstract

The literature on cooperation in infinitely repeated Prisoner's Dilemmas covers the extreme opposites of the matching spectrum: *partners*, a player's opponent never changes, and *strangers*, a player's opponent randomly changes in every period. Here, we extend the analysis to settings where the opponent changes, but not in every period. In these temporary partnerships, players can deter some deviations by directly sanctioning their partner. Hence, relaxing the extreme assumption of one-period matchings can support some cooperation also off equilibrium because a class of strategies emerges that are less extreme than the typical "grim" strategy. We establish conditions supporting full cooperation as a subgame perfect equilibrium under a social norm that complements direct sanctions with a cyclical community sanction. Though this strategy less effectively incentivizes cooperation, it more effectively incentivizes punishment after a deviation, hence, can be preferable to the grim strategy under certain conditions.

Keywords: prisoner's dilemma, random matching, social norms.

JEL codes: E4, E5, C7

1 Introduction

The literature on cooperation in infinitely repeated Prisoner's Dilemmas covers the extreme opposites of the matching spectrum: *partners*, a player's opponent never changes, and *strangers*, a player's opponent randomly changes in every period.

* We thank participants at the 2023 European Workshop on Economic Theory in Naples, the Guest Co-Editor Luba Petersen, and two anonymous Referees for comments that improved the paper. Gabriele Camera, Economic Science Institute, Chapman University, One University dr., Orange, CA 92866; Tel.: 714-628-2806; e-mail: camera@chapman.edu. Alessandro Gioffré, DISEI, University of Florence; Tel.:+39 055 275 9606; e-mail: alessandro.gioffre@unifi.it

Here, we extend the analysis to settings where the opponent changes, but not in every period.

When individuals interact as partners, cooperation can be easily supported because they can monitor each other's past conduct. In this case, deviations can be deterred by directly sanctioning the counterpart with a suitably long spell of defections. Things, however, are not so simple when individuals interact as strangers. Here, private monitoring shrouds past conduct so players cannot build reputations. Moreover, meetings dissolve after just one interaction, so a player who suffered a defection can neither rely on negative direct reciprocity to deter it, nor indirect reciprocity due to the informational opaqueness of strangers' past conduct. Cooperation is still possible as long as players can tacitly coordinate on some form of community punishment to deter deviations. The literature has focused on a "grim" deterrent—all cooperation forever ceases after observing a single deviation from cooperative play.[1]

The literature has shied away from considering scenarios that fall in-between the two extreme matching situations described above. These scenarios are interesting because they are empirically relevant—we often meet unfamiliar individuals and interact with them for a few consecutive periods before the meeting is dissolved. As a concrete example from the corporate world, consider that employees tend to move from team to team at regular intervals, where they temporarily interact with co-workers who are unfamiliar to them, or have opaque reputations. Corporations, for instance, rely on large multi-unit teams to bring products to market. Employees of large consulting firms also regularly interact within temporary working groups, as they move from project to project. These temporary working groups may comprise unfamiliar workers due to turnover, or because they come from different units or firms. Once the task is completed, the team is broken

[1] Regarding partners' settings, see Friedman (1971) for a theoretical view and Dal Bó (2005) or Blonski et al. (2011) for experimental evidence about discrete-time settings. See Bigoni et al. (2015) and Friedman and Oprea (2012) for theory and experimental evidence for near-continuous time settings and Ghidoni and Suetens (2022) for social dilemmas where actions are asynchronous. Regarding the other extreme of strangers' settings, see (Kandori, 1992) for a theoretical view and Camera and Casari (2009), Duffy and Ochs (2009) or Camera et al. (2012), for experimental evidence. For social dilemmas were actions are not synchronous see Camera et al. (2013).

up and everyone gets reassigned to new tasks and new partners, in a recursive scheme. In this situation, outcomes depend on whether employees are willing and able to cooperate with temporary partners whom they do not know and may not trust. How should we go about supporting cooperation in this case? Since individuals interact with the same counterpart for a few periods, they can observe and directly react to the counterpart's actions. Yet, when the partnership dissolves, if individuals cannot carry their reputation with them, then they effectively enter a new partnership as strangers. The natural question is whether a protracted interaction can facilitate cooperation as compared to strict, one-period meetings. Intuitively, one might imagine that longer meetings could allow players to leverage reciprocity to deter at least some defections.

In this paper, we represent the situation described above by considering a matching process that is random—to motivate partners' unfamiliarity and opaque reputations—and is repeated indefinitely, to motivate the expectation of employment within a firm for the foreseeable future, once a partnership is dissolved. Specifically, we study the existence of cooperative equilibrium for matching setups that lie in-between the two extreme opposites of the matching spectrum, partners and strangers. In our setup, players are periodically randomly re-matched into pairs lasting a fixed and known number of periods. However, individuals can only observe the actions of their current counterpart and cannot carry a reputation with them in subsequent meetings. We call these meetings *temporary partnerships*. It is demonstrated that, under some conditions, values of the discount factor exist that are compatible with a strategy supporting efficient play. This strategy combines a direct, reciprocity-based form of punishment, with an indirect, community-based punishment norm.

The analysis reveals that interacting in temporary partnerships has some advantages. The possibility to engage in direct punishment is a clear advantage, as it improves the incentives for cooperation in all but the very last period of interaction. However, the problem remains of how to incentivize cooperation in the last period of interaction, because when the match dissolves reputation cannot be carried over into the next match. This problem is solved by a community punishment scheme. A second advantage of temporary partnerships is that players can use the periodicity of the matching function as an *explicit* coordination device, which allows them to select cyclical community sanctions less extreme as compared to the classic grim punishment. In this manner, we show that, when players are periodically rematched as strangers, some cooperation can be supported off-equilibrium. This strategy more effectively incentivizes the community sanction (off equilibrium) as compared to the grim strategy. However, it also less effectively incentivizes cooperation (in equilibrium) because continuation payoffs following a deviation are larger, given that not all cooperation is destroyed. It follows that the more lenient strategy is preferable to the grim strategy under certain conditions, but not others. In particular, it allows cooperation to be more effectively sustained, as compared to grim, when players are patient, and the sucker's payoffs is low—which is when the incentive to punish off equilibrium is also low. The rest of the paper is organized as follows. In Section 2, we present the model. In Section 3, we develop our analysis, and Section 4 offers some brief conclusions.

2 Model

There are $N = 2n \geq 4$ infinitely-lived players and time is discrete $t = 1, 2, \ldots$ Every $T \geq 1$ periods players are matched in pairs using a uniform random matching mechanism. This means that at the start of the game, we form n temporary partnerships consisting of two players, which last T periods. Subsequently, we repeatedly form n partnerships every T periods. Hence, whenever we partition the population into partnerships, the probability that player $i = 1, \ldots, N$ meets any player $j \neq i$ is $\frac{1}{N-1}$.

In each period t, every pair (i, j) plays a Prisoner's Dilemma, with action set $\{C, D\}$ ("cooperate" and "defect"). Using standard notation, the stage game payoffs are in the table below.

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

Table 1: The game between player i and j in a period.

As usual, assume $g, l > 0$ and $1 > g - l$ (our results hold also if the second inequality is not fulfilled) so we have a social dilemma. In a period, the outcome (C, C) (= full cooperation) is mutually beneficial and corresponds to the efficient outcome. The outcome (D, D) (= full defection) is inefficient and the unique Nash equilibrium in the one-shot game. As discussed above, this game is played with different players at different points in time, following the random matching process. In what follows we refer to this infinite sequence of one shot games as a *repeated game* (with random matching).

To define payoffs in the repeated game, we assume that players have linear preferences and discount future payoffs with a common discount factor $\delta \in (0, 1)$. The payoff to player i in the indefinitely repeated game is therefore

$$\sum_{t=1}^{\infty} \delta^{t-1} \pi_i(a_{i,t}, a_{j,t})$$

where we use $\pi_i(a_{i,t}, a_{j,t})$ to denote the payoff to player i in period t when the action profile is $(a_{i,t}, a_{j,t})$, with $a_{i,t}, a_{j,t} \in \{C, D\}$. It follows that the efficient outcome corresponds to the one in which every player cooperates in every period. Inefficiency can take different gradations, depending on how often defection occurs. We will say that the outcome is (fully) inefficient when every player defects in every period.

There is private monitoring: players can only observe outcomes and actions in their pair and cannot observe the history of play of other pairs. Therefore, the model with $T = 1$ is identical to the model in Kandori (1992) where players interact as strangers in one-period meetings, under private monitoring. We extend that model to include multi-period meetings when $T > 1$. We call these *temporary partnerships*. Each partnership has T *stages*, i.e., T periods of play. The random matching partition generates an infinite number of temporary partnerships $h = 1, 2, \ldots$ each of which lasts T stages, as follows:

$$\text{Periods in the game:} \quad \overbrace{1, \ldots, T}^{\text{partnership 1}}, \overbrace{T+1, \ldots, 2T}^{\text{partnership 2}}, \ldots, \overbrace{(h-1)T+1, \ldots, hT}^{\text{partnership } h>2}, \ldots$$

We study whether the efficient outcome can be sustained as a subgame perfect Nash Equilibrium in this setting. The advantage of a temporary partnership is that *within* it there is perfect monitoring: the opponent's history of actions in previous stages of the partnership is known. This is key as it allows individuals to *directly* sanction a partner who does not cooperate by playing D in all remaining stages of that partnership, if there are any. This might be sufficient to deter defection in early stages of the partnership, but certainly not in the last one because the partnership dissolves. Hence, a sanction in this infinitely repeated matching game cannot rely only on direct punishment. We thus consider a trigger strategy that combines a direct sanction with a *community* sanction that targets *every* future partner—not only the current partner. The community sanction starts as soon as the match dissolves, and cascades through the economy over time.

In other words, direct punishment allows players to immediately sanction a temporary partner who does not cooperate in some stage $\tau < T$. However, since new partnerships are formed, community punishment is needed to remove the incentive to defect because it ensures that such a deviation will be sanctioned by all future partners. The question is thus: what kind of community sanction should we consider? A possibility is a "grim" sanction whereby D is played in every future period, as discussed in Kandori (1992). However, a further advantage of temporary partnerships is that one can limit the community sanction to a subset of future periods. In this manner, a deviation does not destroy all cooperation in *future* partnerships. Below we describe such a strategy using an automaton.

Definition 1 (Direct plus community sanction)**.** *A player can be in one of two states, cooperative or punishment. The player starts in the cooperative state, se-*

lecting C in every stage of every partnership. If D is observed, then the player permanently leaves the cooperative state and enters the punishment state in the next period. Let τ denote a stage where the player experienced asymmetric play. In the punishment state, the player selects (i) D in all stages $t > \tau$ of any partnership where (D, D) is not the outcome in τ (direct sanction), (ii) D in all stages τ of all partnerships (community sanction), and (iii) C in all other circumstances.

We say that the strategy in Definition 1 is a social norm if all players adopt it. This norm encompasses two modes of behavior: *cooperation* and *punishment*. Cooperation involves choosing C in a period. Punishment has two components, direct and indirect. *Direct* punishment applies only to specific opponents, while *community* punishment applies to all opponents. Using direct punishment in some stage $t > \tau$ amounts to selecting D in all stages $\tau+1, \ldots, T$ of *just that* partnership. Using community punishment in some stage τ amounts to selecting D in that stage of *any* partnership; this transmits information about the defection initially observed by the player, to everyone else in the group.

The player starts the game in a cooperative state, and remains in that state unless he experiences asymmetric play, (C, D) or (D, C), in some stage τ. In that case, the player permanently switches to the punishment state in the following period. At that point he will directly punish his opponent (if the partnership has not yet dissolved) and will follow community punishment in every future stage τ. Notice that, once the player is in the punishment state, he will also use direct punishment on any future opponent who *does not* apply community punishment in stage τ (by choosing D in all stages $t > \tau$ of that partnership). This way of sanctioning lack of community punishment allows players to re-coordinate off-equilibrium on punishing only in stage τ. In all stages other than τ, players will select C.

This punishment strategy limits the extent of community punishment. Once the player is off-equilibrium, if the opponent is also seen following the community punishment norm, then both players enjoy (C, C) in all stages of that partnership but the stage(s) where community punishment should be used. Hence, there is *some* cooperation off-equilibrium. This is infrequent if T is small, but substantial if T is large.

We conjecture that it is the threat of switching to the punishment state that supports full cooperation as a subgame perfect Nash equilibrium. Note that the strategy prescribes a switch to a limited punishment scheme, following observation of a defection, as compared to the classic "grim" strategy discussed in Kandori (1992). Because a player in the punishment state defects only in a subset of periods, not all cooperation is destroyed once someone moves off equilibrium. The strategy collapses to the grim strategy for $T = 1$. Table 2 shows the possible play within a generic partnership, in- and off-equilibrium.

	Stages in a Partnership					
	1	2	3	...	$T-1$	T
Equilibrium Play	(C, C)	(C, C)	(C, C)	...	(C, C)	(C, C)
Direct Punish.	(C, C)	(C, D)	(D, D)	...	(D, D)	(D, D)
Community Punish. (a)	(C, C)	(D, C)	(D, D)	...	(D, D)	(D, D)
Community Punish. (b)	(C, C)	(D, D)	(C, C)	...	(C, C)	(C, C)

Table 2: Outcomes in a partnership if a deviation occurred in stage 2.

Notes: Actions (a_i, a_{-i}) refer to player i and his opponent in a period. *Equilibrium play*: player i and his opponent are in the cooperative state and choose C. *Direct punishment*: player i is in the cooperative state and his opponent deviates from equilibrium play in stage 2, so there is asymmetric play. *Community punishment (a)*: player i is in the punishment state and uses community punishment in stage 2, but the opponent is in the cooperative state so there is additional direct punishment. *Community punishment (b)*: both players are in the punishment state, use community punishment in stage 2, hence re-coordinate on cooperation in all subsequent stages.

In the Table there are two players, i and his opponent $-i$. Their actions in a period are (a_i, a_{-i}). The row *Direct Punishment* shows what happens when player i is in the cooperative state and observes a defection in some stage of a partnership (stage 2, in the example). Following the strategy codified by the automaton in Definition 1, player i switches to a punishment state. This implies that he chooses D in all subsequent stages of that partnership, $3, \ldots, T$, and will choose D in stage 2 of all *future* partnerships (as seen in the rows below). Player i will also choose D off the equilibrium path, in all stages of future partnerships where play is asymmetric in stage 2, i.e., when he meets someone who is not in the punishment state or someone who is in the punishment state but (for whatever reason) does choose D in stage 2. In all other instances, the player will cooperate off the equilibrium path with someone who is in the punishment state since the outcome stage 2 allows them to re-coordinate on cooperation in all remaining stages.

Notice that this *does not* imply that players in the punishment state do not react to further deviations from the strategy. If in any future partnership player i were to observe asymmetric play in some stage $\tau' \neq 2$, then he would directly punish the partner and also start a new community punishment in all stages τ' of all future partnerships.

3 Analysis

Here we show that, in the repeated matching game we just described, if players adopt a social norm based on the strategy in Definition 1, then the efficient outcome is a subgame perfect Nash equilibrium for certain values of δ. Formally:

Theorem 1. *For any temporary partnership of duration $T > 1$, there are values $\delta \in (0, 1)$ such that a social norm based on the strategy in Definition 1 supports full cooperation as a subgame perfect Nash equilibrium.*

The proof of this theorem develops in three steps, using the one-shot deviation principle. In the model a period defines a subgame. First, we show that on the equilibrium path taking action D in some period and then reverting to following the strategy in Definition 1 is suboptimal for $\delta \in [\delta_1, 1] \subset (0, 1)$. Second, we show that off the equilibrium path the player has no incentive to choose C in a stage where community or direct punishment should be implemented, if $\delta \in [0, \delta_2] \subset (0, 1)$. Finally, we show that $[\delta_1, \delta_2] \neq \emptyset$.

We start by showing how to calculate continuation payoffs on the equilibrium path, and off the equilibrium path.

3.1 Continuation payoffs

Suppose that every player adopts the strategy in Definition 1. On the equilibrium path there is full cooperation and the payoff to any player is

$$v_0 := \frac{1-\delta^T}{1-\delta} \times \frac{1}{1-\delta^T} = \frac{1}{1-\delta}.$$

In equilibrium, the observation of asymmetric play moves the economy off the equilibrium path, triggering a progressive switch from a cooperative to a punishment state in the group. This switch occurs gradually because of the random matching process. This implies that off-equilibrium, all players will eventually alternate "cooperation cycles" to "punishment cycles," which is what we will show in what follows.

Off the equilibrium path, there is someone (possibly everybody) who is in the punishment state. To calculate payoffs off equilibrium, since we have random matching and private monitoring, we must characterize the contagious process of punishment. Suppose for a moment that the population is composed of a generic number $M \geq 4$ of players. Partition the population into what we call *defectors*, who are in the punishment state, and *cooperators*, who are in the cooperation state. According to the strategy in Definition 1 cooperators become defectors at random points in time, via a contagious process which is fully described by the $M \times M$ upper-triangular Markov matrix $\mathcal{Q}_M$, where

$$\mathcal{Q}_M := \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & Q_{22} & 0 & Q_{24} & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & Q_{34} & 0 & Q_{36} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & Q_{M-2,M-2} & 0 & Q_{M-2,M} \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix}. \qquad (1)$$

To describe the elements of $\mathcal{Q}_M$, suppose we are at the end of a period where temporary partnerships are dissolved. Suppose there are $k = 1, \dots, M$ defectors. In the next period, all players are randomly rematched into new temporary partnerships. The generic element $Q_{kk'} = Q_{kk'}(M)$ is the probability to transition from k to $k' \geq k$ defectors next period, when new temporary partnerships are formed.

We can show that

$$Q_{kk'}(M) := \frac{(k'-k)!\binom{k}{k'-k}\binom{M-k}{k'-k}(2k-k'-1)!!(M-k'-1)!!}{(M-1)!!}, \tag{2}$$

where the number of additional defectors created by rematching players into new temporary partnerships is

$$k'-k \in \begin{cases} \{0,2,4,\ldots,\min(k,M-k)\} & \text{if } k = \text{even} \\ \{1,3,5,\ldots,\min(k,M-k)\} & \text{if } k = \text{odd}. \end{cases}$$

To understand $\mathcal{Q}$ and (2), start by noting that since the punishment state is absorbing we have an upper triangular matrix $\mathcal{Q}$ where $Q_{MM} = 1$. Moreover $Q_{kk'} = 0$ when k' is odd since each temporary partnership involving a defector and a cooperator doubles the number of defectors so we cannot ever transition to an odd number of defectors k'. This explains, for instance, why $Q_{23} = 0$. More generally, $Q_{kk'}(M)$ accounts for situations in which there is at least one defector and at most M, i.e., $M \geq k' \geq k \geq 1$. Using the double factorial notation, the number of possible ways to form pairs in an even population of M individuals is $(M-1)!! = (M-1)\cdot(M-3)\cdots 3\cdot 1$. Given the definition $0!! = 1$ and $(-1)!! = 1$, this same notation allows to display the number of temporary partnerships involving two defectors and two cooperators given k current defectors and k' future defectors; these numbers are, respectively, $(2k-k'-1)!!$ and $(M-k'-1)!!$. The expression $(k'-k)!\binom{k}{k'-k}\binom{M-k}{k'-k}$ accounts for all possible temporary partnerships composed by one defector and one cooperator that create $k'-k$ new defectors. The interested reader can find a detailed derivation of $\mathcal{Q}$ in Camera and Gioffré (2014).

Now consider $M = N$. The expected payoff to a generic defector depends on how many defectors are in the economy and the stage(s) in which defections are expected to occur. Hence, suppose off-equilibrium we have k defectors (=players in the punishment state) and $N-k$ cooperators (=players in the cooperative state). Suppose also that the initial deviator moved off-equilibrium in some stage $\tau = 1, \ldots, T$ of some partnership.

When everyone follows the strategy in Definition 1, community punishment must take place in stage τ of every partnership. The expected payoff to someone who is in the punishment state at the start of some partnership can be recursively defined by

$$v_k = \overbrace{\frac{1-\delta^{\tau-1}}{1-\delta}}^{(C,C) \text{ in } t<\tau} + \overbrace{\sigma_k(1+g)\delta^{\tau-1}}^{(D,C) \text{ in } \tau,\ (D,D) \text{ in } t>\tau} + \overbrace{(1-\sigma_k)\delta^{\tau}\frac{1-\delta^{T-\tau}}{1-\delta}}^{(D,D) \text{ in } \tau,\ (C,C) \text{ in } t>\tau}$$
$$+ \delta^T \sum_{k'=k}^{N} Q_{kk'}(N) v_{k'}, \qquad k = 1, \dots, N.$$

Due to uniform random matching, a defector meets a cooperator with probability $\sigma_k := \frac{N-k}{N-1}$. Since community punishment occurs in stage τ of every partnership, the player earns 1 in all preceding stages even when both players are in the punishment state. Here, earnings sum up to $\frac{1-\delta^{\tau-1}}{1-\delta}$. If the opponent is in the cooperative state, then the player earns an additional $\sigma_k(1+g)\delta^{\tau-1}$ because the opponent will cooperate in stage τ and both will defect in all subsequent stages. Instead, if the opponent is also in the punishment state, with probability $1-\sigma_k$, then the player earns an additional $\delta^{\tau}\frac{1-\delta^{T-\tau}}{1-\delta}$ because both players will defect only in stage τ and will cooperate in all subsequent stages.

Letting $v := (v_1, \dots, v_N)^{\mathsf{T}}$, $\sigma := (\sigma_1, \dots, \sigma_N)^{\mathsf{T}}$, and $\mathbf{1} := (1, \dots, 1)^{\mathsf{T}}$, the expression above can be written in vectorial form as

$$v = \left(\frac{1-\delta^T}{1-\delta} - \delta^{\tau-1}\right)\mathbf{1} + \delta^{\tau-1}\left(1+g-\delta\frac{1-\delta^{T-\tau}}{1-\delta}\right)\sigma + \delta^T \mathcal{Q}_N v,$$

which gives us

$$v = (\mathcal{I} - \delta^T\mathcal{Q}_N)^{-1}\left[\left(\frac{1-\delta^T}{1-\delta} - \delta^{\tau-1}\right)\mathbf{1} + \delta^{\tau-1}\left(1+g-\delta\frac{1-\delta^{T-\tau}}{1-\delta}\right)\sigma\right].$$

Given $k = 1, \dots, N$, the generic element v_k of vector v is therefore

$$v_k = e_k^{\mathsf{T}}(\mathcal{I} - \delta^T\mathcal{Q}_N)^{-1}\left[\left(\frac{1-\delta^T}{1-\delta} - \delta^{\tau-1}\right)\mathbf{1} + \delta^{\tau-1}\left(1+g-\delta\frac{1-\delta^{T-\tau}}{1-\delta}\right)\sigma\right], \quad (3)$$

where e_k is the $N-$dimensional column vector with 1 in the k^{th} position and 0 everywhere else.

Now consider v_k in (3). The first component is the payoff $\frac{1-\delta^T}{1-\delta} - \delta^{\tau-1}$ generated in all future partnerships (including the current one) if both players are perfectly coordinated on D in stage τ and C in all other stages. Hence, we have:

$$e_k^\intercal(\mathcal{I}-\delta^T\mathcal{Q}_N)^{-1}\left(\frac{1-\delta^T}{1-\delta}-\delta^{\tau-1}\right)\mathbf{1}=\left(\frac{1-\delta^T}{1-\delta}-\delta^{\tau-1}\right)\times\frac{1}{1-\delta^T}.$$

The second component is the payoff $\delta^{\tau-1}(1+g)$ generated in stage τ of all future partnerships (including the current one) where the opponent is a cooperator (with probabilities listed in σ), and the adjustment $-\delta^\tau\frac{1-\delta^{T-\tau}}{1-\delta}$ because both players defect from stage $\tau+1$ onward, earning 0 in all these stages.

Letting

$$\phi_k(\delta^T):=(1-\delta^T)e_k^\intercal(\mathcal{I}-\delta^T\mathcal{Q}_N)^{-1}\sigma,\qquad \text{for } k=1,\ldots,N$$

this stream of payoffs can be written as

$$e_k^\intercal(\mathcal{I}-\delta^T\mathcal{Q}_N)^{-1}\delta^{\tau-1}\left(1+g-\delta\frac{1-\delta^{T-\tau}}{1-\delta}\right)\sigma=\frac{\delta^{\tau-1}}{1-\delta^T}\phi_k(\delta^T)\left(1+g-\delta\frac{1-\delta^{T-\tau}}{1-\delta}\right).$$

Given the above, v_k can be written as

$$v_k=\left(\frac{1-\delta^T}{1-\delta}-\delta^{\tau-1}\right)\times\frac{1}{1-\delta^T}+\frac{\delta^{\tau-1}}{1-\delta^T}\phi_k(\delta^T)\left(1+g-\delta\frac{1-\delta^{T-\tau}}{1-\delta}\right).\tag{4}$$

The following result holds:

Lemma 1. *Let $z:=\delta^T\in(0,1)$. For all $k=1,\ldots,N-1$, $\phi_k(z)$ decreases in z.*

Proof. The first derivative of $\phi_k(z)$, for all $k=1,\ldots,N-1$ is

$$\frac{d\phi_k(z)}{dz}\equiv\phi_k'(z)=-e_k^\intercal(\mathcal{I}-z\mathcal{Q}_N)^{-1}[\mathcal{I}-(1-z)\mathcal{Q}_N(\mathcal{I}-z\mathcal{Q}_N)^{-1}]\sigma,$$

where we have defined $\mathcal{A}(z):=\mathcal{I}-z\mathcal{Q}_N$, and we used $\frac{d\mathcal{A}^{-1}(z)}{dz}=-\mathcal{A}^{-1}(z)\frac{d\mathcal{A}(z)}{dz}\mathcal{A}^{-1}(z)$. To prove that $\phi_k'(z)<0$ if $k<N$, notice that $(1-z)(\mathcal{I}-z\mathcal{Q}_N)^{-1}\sigma\leq\sigma$ (with strict inequality for all $k\leq N-1$) since $(1-z)(\mathcal{I}-z\mathcal{Q}_N)^{-1}\mathbf{1}=\mathbf{1}$ and $\sigma_k\in\sigma$ is decreasing in k.[2] Therefore, $\mathcal{Q}_N(1-z)(\mathcal{I}-z\mathcal{Q}_N)^{-1}\sigma\leq\mathcal{Q}_N\sigma\leq\mathcal{I}\sigma$. But then we also have

$$(\mathcal{I}-z\mathcal{Q}_N)^{-1}(1-z)\mathcal{Q}_N(\mathcal{I}-z\mathcal{Q}_N)^{-1}\sigma\leq(\mathcal{I}-z\mathcal{Q}_N)^{-1}\sigma,$$

[2] Each element of matrix $(\mathcal{I}-z\mathcal{Q}_N)^{-1}$ is non-negative and its rows sum to $(1-z)^{-1}$, hence $(1-z)(\mathcal{I}-z\mathcal{Q}_N)^{-1}\mathbf{1}=\mathbf{1}$, where $\mathbf{1}=(1,1,\ldots,1)^\intercal$.

which holds with strict inequality if $k \leq N-1$, i.e.,

$$e_k^\mathsf{T}(\mathcal{I} - z\mathcal{Q}_N)^{-1}(1-z)\mathcal{Q}_N(\mathcal{I} - z\mathcal{Q}_N)^{-1}\sigma < e_k^\mathsf{T}(\mathcal{I} - z\mathcal{Q}_N)^{-1}\sigma, \quad k = 1, \ldots, N-1.$$

□

It is convenient to rewrite v_k as the sum of two expected payoffs depending on whether the defector meets a cooperator (with probability σ_k) or a defector (with probability $1 - \sigma_k$), i.e.,

$$\begin{aligned} v_k = & \frac{1-\delta^{\tau-1}}{1-\delta} + \sigma_k \left[(1+g)\delta^{\tau-1} + \delta^T \sum_{k'=k-1}^{N-2} Q_{k-1,k'}(N-2) v_{k'+2} \right] \\ & + (1-\sigma_k) \left[\delta^\tau \frac{1-\delta^{T-\tau}}{1-\delta} + \delta^T \sum_{k'=k-2}^{N-2} Q_{k-2,k'}(N-2) v_{k'+2} \right]. \end{aligned} \tag{5}$$

$Q_{kk'}(N-2)$ is the element of the transition matrix $\mathcal{Q}_M$ when $M = N-2$, because we are considering the $n-1$ matches other than the one between the defector and his opponent. There are two cases to consider.

First, the opponent is a cooperator. Here, we expect $k'+2$ defectors when players are rematched into new temporary partnerships. The number 2 includes the defector and his opponent. The number k' depends on the remaining $n-1$ random matches between $k-1$ defectors and $N-k-1$ cooperators. In this case the expected continuation payoff is $\sum_{k'=k-1}^{N-2} Q_{k-1,k'}(N-2) v_{k'+2}$.

Second, the opponent is a defector. Here, in the remaining $n-1$ matches there are $k-2$ defectors and $N-k$ cooperators. In this case, the continuation payoff is $\sum_{k'=k-2}^{N-2} Q_{k-2,k'}(N-2) v_{k'+2}$.

Expression (5) splits the expected continuation payoff $\sum_{k'=k}^{N} Q_{kk'}(N) v_{k'}$ into two parts:

$$\begin{aligned} \sum_{k'=k}^{N} Q_{kk'}(N) v_{k'} = & \sigma_k \sum_{k'=k-1}^{N-2} Q_{k-1,k'}(N-2) v_{k'+2} \\ & + (1-\sigma_k) \sum_{k'=k-2}^{N-2} Q_{k-2,k'}(N-2) v_{k'+2}. \end{aligned}$$

We proceed by deriving conditions such that deviating from equilibrium play is suboptimal, and deviating from the sanction off-equilibrium is suboptimal.

3.1.1 Equilibrium deviations

Consider one-period deviations on the equilibrium path. We derive a condition ensuring that deviating in any stage τ is suboptimal, in equilibrium. This is so when $v_1 \leq v_0$. Using the definitions of v_0 and v_1 we have

$$\left(\frac{1-\delta^T}{1-\delta} - \delta^{\tau-1}\right) \times \frac{1}{1-\delta^T} + \frac{\delta^{\tau-1}}{1-\delta^T}\phi_k(\delta^T)\left(1+g-\delta\frac{1-\delta^{T-\tau}}{1-\delta}\right) \leq \frac{1-\delta^T}{1-\delta} \times \frac{1}{1-\delta^T},$$

$$\Rightarrow \qquad \phi_1(\delta^T)\left[1+g-\delta\frac{1-\delta^{T-\tau}}{1-\delta}\right] \leq 1.$$

Since $\delta\frac{1-\delta^{T-\tau}}{1-\delta} \geq 0$ ($=0$ if the initial deviation occurs in $\tau = T$) the most restrictive case is $\tau = T$, which gives us the following sufficient condition

$$\phi_1(\delta^T)(1+g) \leq 1.$$

Intuitively, if the player does not deviate in stage T—where he does not suffer the consequences of direct punishment but only faces future community punishment—then he will certainly not deviate in stages before T, because in those stages he is immediately punished by his opponent and, in addition, he also faces future community punishment. As ϕ_1 is invertible we get:

$$\delta_1^T := \phi_1^{-1}\Big(\frac{1}{1+g}\Big).$$

Notice that ϕ_1 maps $[0,1)$ into $(0,1]$ and it is a strictly monotone, decreasing function of δ^T (Lemma 1).[3] It follows that ϕ_1 is invertible, i.e., $\delta^T = \phi_1^{-1}(x)$ for $x \in (0,1]$. Since $\frac{1}{1+g} \in (0,1)$, there exists a value $\delta_1 \in (0,1)$ such that

$$\delta_1^T := \phi_1^{-1}\Big(\frac{1}{1+g}\Big).$$

Monotonicity of ϕ_1 ensures that for all $\delta^T \in [\delta_1^T, 1)$ the inequality $\phi_1(\delta^T) \leq \frac{1}{1+g}$ holds. It follows that if $\delta \in [\delta_1, 1)$, then equilibrium deviations are suboptimal in any stage of any partnership, where $\delta_1 = \left[\phi_1^{-1}\Big(\frac{1}{1+g}\Big)\right]^{1/T}$.

Now notice that $\phi_1^{-1}\left(\frac{1}{1+g}\right)$ is increasing and strictly concave in g.

[3]A formal proof is in Camera and Gioffré (2014) where, however, temporary partnerships are restricted to lasting 1 period ($T = 1$).

3.1.2 Out of equilibrium deviations

Suppose that k players are in the punishment state and let player i be one of them. We wish to establish that punishment is credible off the equilibrium path. That is, it is optimal not to delay the switch from a cooperative to a punishment state, by choosing D as prescribed by the strategy in Definition 1, off the equilibrium path. Given the unimprovability principle, we consider one-shot deviations.

Consider the optimality of direct punishment. It should be clear that it is suboptimal to choose C in any stage following asymmetric play because those stages call for direct punishment. A one-shot deviation to C is suboptimal because the opponent, after seeing asymmetric play, enters the punishment state and selects D following the prescribed direct punishment norm.

Now consider the optimality of community punishment. Suppose we are in stage τ where community punishment must take place, and player i is in the punishment state but chooses C instead of D. That is, the player wants to delay community punishment—in case he meets a cooperator—and reverts to following the punishment strategy from the next period.

The payoff from this deviation from the community punishment norm depends on the state of the opponent. If the opponent is in the punishment state (with probability $1 - \sigma_k$), then by deviating to C player i earns $-l$ in stage τ. In the remaining stages, player i earns 0 since in stage τ there is asymmetric play and direct punishment ensues. By deviating to C, player i cannot coordinate on D in stage τ with his opponent, and so he loses the cooperation payoff in all subsequent stages. Moreover, as the current opponent is in the punishment state, avoiding community punishment does not slow down the contagion process. Hence, there is no advantage from this deviation in terms of the continuation payoff to player i from future partnerships.

Now consider the case where the opponent is in the cooperative state, with probability σ_k, By avoiding community punishment in stage τ, player i earns 1 since the outcome is (C, C). If there are no additional stages, then this is the best-case scenario for player i since the current opponent is in the cooperative state and will remain in that state. Hence, deviating to C instead of applying community punishment in stage $\tau = T$ slows down the contagion process and increases the continuation payoff to player i from future partnerships. By contrast, if there are additional stages in the partnership, in stage $\tau + 1$ player i goes back to following the punishment norm. Consequently, he implements direct punishment because in stage τ the outcome *was not* (D, D). This has three consequences. First, player i earns $1 + g$ in stage $\tau + 1$ (the outcome is (D, C)) and 0 in all future stages (if any) since his opponent observes asymmetric play in $\tau + 1$ and uses direct punishment. Second, deviating from the community punishment norm in stage τ does not slow down the contagion process as the opponent switches to the punishment state following asymmetric play in stage $\tau + 1$. Third, deviating from the community punishment norm in stage τ *expands* community punishment to an additional stage $\tau + 1$, since asymmetric play now occurred in that stage also. This additional community punishment will start an additional contagion process of defection, which lowers the continuation payoff of player i as compared to not deviating in τ. Note that if in equilibrium player i cannot gain from deviating in some stage $\tau + 1$, earning $1 + g$ and triggering community punishment in that stage, then this is also suboptimal off equilibrium.[4] Hence, the best-case scenario for a deviation from the community punishment norm is for $\tau = T$.

Given the above, the expected payoff to player i, who is in the punishment state and deviates to C in the community punishment stage $\tau = T$, is

$$\tilde{v}_k = \left(\frac{1-\delta^{T-1}}{1-\delta}\right) + \sigma_k \overbrace{\left[\delta^{T-1} + \delta^T \sum_{k'=k-1}^{N-2} Q_{k-1,k'}(N-2)v_{k'+1}\right]}^{i \text{ deviates to C in } \tau = T \text{ \& meets a cooperator}}$$
$$+ (1-\sigma_k) \overbrace{\left[-l\delta^{T-1} + \delta^T \sum_{k'=k-2}^{N-2} Q_{k-2,k'}(N-2)v_{k'+2}\right]}^{i \text{ deviates to C in } \tau = T \text{ \& meets a defector}}.$$

[4]Furthermore, if g is sufficiently large, then there is clearly an incentive to community punish in stage τ as compared to delay the g payoff to stage $\tau + 1$.

Using (5) for $\tau = T$ we have

$$v_k = \frac{1-\delta^{T-1}}{1-\delta} + \sigma_k\Big[(1+g)\delta^{T-1} + \delta^T \sum_{k'=k-1}^{N-2} Q_{k-1,k'}(N-2)v_{k'+2}\Big] + (1-\sigma_k)\Big[\delta^T \sum_{k'=k-2}^{N-2} Q_{k-2,k'}(N-2)v_{k'+2}\Big].$$

Comparing expression $\tilde{v}_k$ to v_k reveals that deviating from community punishment in stage $\tau = T$ by choosing C instead of D affects the expected payoff of player i in two ways. First, it reduces the expected earnings in the current partnership if the opponent is in the punishment state and increases them if the opponent is in the cooperative state. Second, if meeting a cooperator, it increases the continuation payoff to $v_{k'+1}$ instead of $v_{k'+2}$ because the cooperator does not switch to the punishment state; it is an increase because v_k falls in k.

Deviating by choosing C in stage $\tau = T$ is thus suboptimal if $\tilde{v}_k \leq v_k$, with $k \geq 2$. Using the relevant expressions, this inequality is rewritten as

$$\sigma_k \delta \sum_{k'=k-1}^{N-2} Q_{k-1,k'}(N-2)(v_{k'+1} - v_{k'+2}) \leq \sigma_k g + (1-\sigma_k)l. \tag{6}$$

The left-hand side represents the expected change in continuation payoffs in future partnerships, from slowing down the contagious punishment process. The right-hand side represents the expected loss or gain in the current partnership.

We have

$$v_k - v_{k+1} = \frac{\delta^{T-1}}{1-\delta^T}(1+g)[\phi_k(\delta^T) - \phi_{k+1}(\delta^T)].$$

Hence, we rearrange (6) as

$$\frac{\delta^T}{1-\delta^T} \sum_{k'=k-1}^{N-2} Q_{k-1,k'}(N-2)[\phi_{k'+1}(\delta^T) - \phi_{k'+2}(\delta^T)] \leq \frac{g}{1+g} + \frac{1-\sigma_k}{(1+g)\sigma_k} l.$$

Since $\phi_k(\delta) - \phi_{k+1}(\delta)$ is largest for $k = 2$ (Camera and Gioffré, 2014, Theorem 2) and $\frac{1-\sigma_k}{\sigma_k}$ is smallest for $k = 2$, the most stringent case is $k = 2$. In that case, we have $Q_{1,k'}(N-2) = 1$ for $k' = 2$ and $Q_{1,k'}(N-2) = 0$ for $k' \neq 2$ since all elements in the first row of $\mathcal{Q}_M$ are 0 except the second, which is 1.

It follows that, for $k = 2$, the inequality above becomes

$$\frac{\delta^T}{1-\delta^T}[\phi_3(\delta^T) - \phi_4(\delta^T)] \leq \frac{g}{1+g} + \frac{1-\sigma_2}{(1+g)\sigma_2} l. \tag{7}$$

To demonstrate that this inequality holds for some δ^T we need an additional piece of information.

Lemma 2. *For any* $\delta^T \in [0, 1)$ *we have*

$$\frac{\delta^T}{1-\delta^T}[\phi_1(\delta^T) - \phi_2(\delta^T)] = 1 - \phi_1(\delta^T).$$

Proof. Suppose that player i moves off equilibrium in stage T of a partnership (he is the only defector). Next period—when new partnerships are formed—there will be two defectors with certainty. Accordingly, player i's expected payoff must satisfy

$$v_1 = \frac{1-\delta^{T-1}}{1-\delta} + \delta^{T-1}(1+g) + \delta^T v_2.$$

Using (4) to derive v_1 and v_2, the equality above can be written as

$$\frac{\phi_1(\delta^T)}{1-\delta^T} = 1 + \frac{\delta^T \phi_2(\delta^T)}{1-\delta^T},$$

which gives us the desired result. □

Now we show that there exists a value $\delta_2 \in (0, 1]$ such that cooperating off equilibrium in stage T is suboptimal for all $\delta \in (0, \delta_2] \cap (0, 1)$.

Using Lemma 2 and recalling that $\phi_k(\delta^T) - \phi_{k+1}(\delta^T)$ is decreasing in k, we have

$$\frac{\delta^T}{1-\delta^T}[\phi_3(\delta^T) - \phi_4(\delta^T)] < 1 - \phi_1(\delta^T).$$

Therefore, to ensure that (7) holds, it is sufficient to show that

$$1 - \phi_1(\delta^T) \leq \frac{g}{1+g} + \frac{1-\sigma_2}{\sigma_2}\frac{l}{(1+g)}.$$

Notice that $\sigma_2 = \frac{N-2}{N-1}$, hence, by continuity of ϕ_1, the inequality above holds for all $\delta^T \in (0, \delta_2^T] \cap (0,1)$, where

$$\delta_2^T := \begin{cases} \phi_1^{-1}\left(\frac{1-l/(N-2)}{1+g}\right) \in (0,1) & \text{if} \quad 0 < l < N-2 \\ 1 & \text{if} \quad l \geq N-2. \end{cases}$$

It follows that deviating off-equilibrium is always suboptimal if $l \geq N-2$, and otherwise it is suboptimal if $\delta < \delta_2$ where $\delta_2 = \left[\phi_1^{-1}\left(\frac{1-l/(N-2)}{1+g}\right)\right]^{1/T}$.

To conclude we demonstrate the following:

Lemma 3. *We have $\delta_1 < \delta_2$ for all $l > 0$ and $\delta_2 \to \delta_1$ as $l \to 0$.*

Proof. For $l \geq N-2$ the proof is obvious since $\delta_2^T = 1$. For $0 < l < N-2$, we use the definition of δ_2^T and δ_1^T to derive the following inequality:

$$\phi_1(\delta_2^T) = \frac{1-l/(N-2)}{1+g} < \frac{1}{1+g} = \phi_1(\delta_1^T).$$

Since $\phi_1(\delta^T)$ is decreasing in $\delta^T \in (0,1)$ then $\phi_1(\delta_2^T) < \phi_1(\delta_1^T)$ implies $\delta_2 > \delta_1$, and we immediately have $\delta_2 \to \delta_1$ as $l \to 0$. □

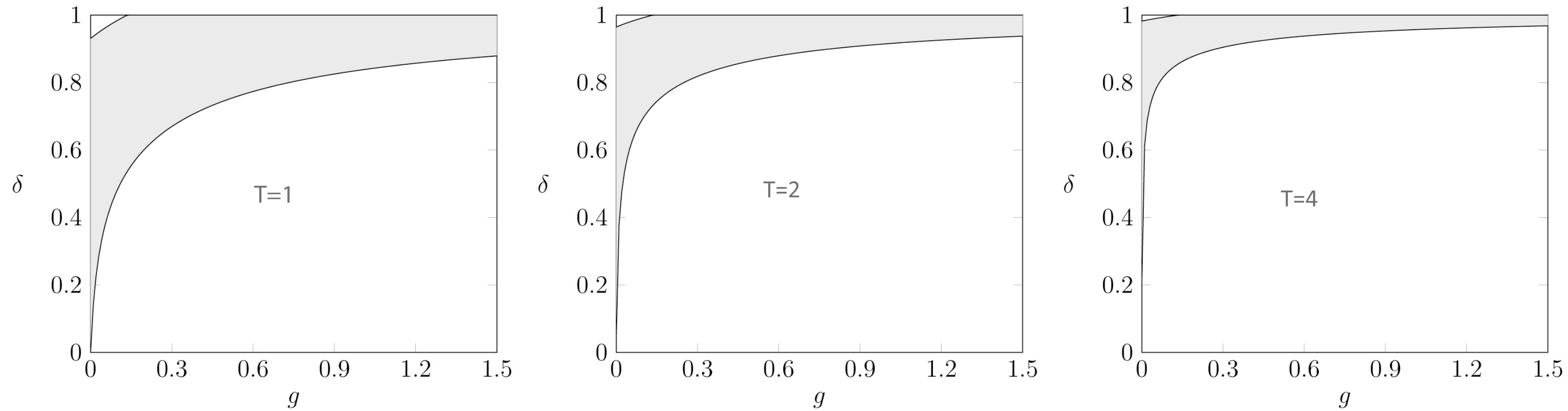
It should also be clear that it is suboptimal to choose D in any stage where the social norm calls for cooperation. Here, a one-shot deviation to D is suboptimal because it is suboptimal on the equilibrium path; the player does not have an incentive to trigger community punishment in that stage.

Having established existence of equilibrium with the strategy proposed in Definition 1, we now discuss how the set of parameters supporting equilibrium compares as we vary the length of the partnership.

3.2 How a partnership affects existence of equilibrium.

Here we show how—given a social norm based on the strategy in Definition 1—the subset of parameters supporting cooperative equilibrium varies as the length of the partnership T varies. In this manner we can see how decreasing the frequency of switching counterparts $1/T$ affects the incentives to cooperate in equilibrium, and to punish off equilibrium. We will consider three different scenarios: $T = 1, 2, 4$ with the aid of Figure 1.

Figure 1: Parameters supporting full cooperation as an equilibrium.



Notes: In all three panels $l = 6, N = 20$. Moving left to right we have $T = 1, 2, 4$ (no temporary partnerships, short and longer temporary partnerships). The lower boundary of the shaded area corresponds to δ_1; the upper boundary corresponds to the δ value satisfying the optimality condition (7) with equality (this is an exact value for incentive compatibility, while δ_2 is a sufficient condition).

The $T = 1$ scenario corresponds to the standard extreme case where meetings last only one period. Hence, we do not have temporary partnerships. Here, the strategy in Definition 1 corresponds to the standard grim strategy, as proposed in Kandori (1992), because direct sanctions are impossible and only the community sanction is viable. In the $T = 2$ scenario, players are in very short partnerships, lasting only two periods. In $T = 4$ the duration of partnership doubles.

The shaded areas in Figure 1 show where the strategy in Definition 1 supports full cooperation as g varies from 0 to 1.5. Two main findings emerge from this numerical analysis. First, as we increase the number of consecutive meetings with a given player—the length of a partnership T—the lowerbound threshold δ_1 increases. Intuitively, as the frequency of switching counterparts decreases the frequency of the community sanction also decreases. Consequently, the continuation payoff associated with a deviation in the last period of a partnership increases with the length of the partnership. The incentive for cooperation thus requires a more patient player. Second, the upper bound threshold δ_2 also increases as the frequency of switching counterparts falls. Intuitively, as T increases the loss from stopping cooperation in the community punishment stage of all partnerships is also smaller. This makes community punishment cheaper and, hence, more easily implementable.

As we move from the extreme case where counterparts change in every period of play ($T = 1$), to less extreme cases where they change less frequently ($T = 2, 4$), two things happen. First, incentivizing cooperation in equilibrium using the threat of community punishment becomes harder. Consequently, the threshold discount factor δ_1 increases—choosing C in equilibrium is now optimal for a smaller set of parameters δ (players must be more patient). Second, incentivizing community punishment off equilibrium becomes easier so that the threshold discount factor δ_2 also increases—choosing D off equilibrium in the community punishment stage is now optimal for a larger set of parameters δ. In the figure, for $T = 4$ this is true for larger δ values, as compared to $T = 1, 2$.

The insight is that creating temporary partnerships instead of having constant rematching is not helpful to support cooperation if players are impatient. The reason is that the last period in a partnership is the weak link, in which the incentives to cooperate and to punish are affected by the time discount factor. Therefore, we have a tradeoff: frequent rematching helps supporting cooperation if players are impatient but removes the incentives to follow a community sanction if players are patient. As a result, our analysis reveals that setting up temporary partnerships among strangers—instead of relying on constant rematching—is useful when players are patient. The natural question is whether this result depends on our use of a sanction that combines direct punishment a temporary partner seen to deviate from the prescribed strategy, with a limited form of community punishment of every future partner. Would the adoption of a grim strategy help support cooperation in temporary partnerships?

3.3 Grim strategy analysis

Here we show that the grim strategy typically can support cooperation in temporary partnerships only if g is sufficiently small. To see this, we start by defining the grim strategy.

Definition 2 (Grim strategy)**.** *A player can be in one of two states, cooperative or punishment. The player starts in the cooperative state, where C is selected in every stage of every partnership. The player leaves the cooperative state and permanently switches to the punishment state if D is observed. In the punishment state, the player selects D in all periods.*

There are advantages and disadvantages from following this strategy as compared to the one in Definition 1. A key advantage is a greater incentive to cooperate in equilibrium because the community sanction is harsher. Under grim, a deviation pushes the entire economy to permanent full defection, imposing a lower payoff as compared to the off-equilibrium payoff sustainable under the strategy in Definition 1. This implies that deviating in equilibrium is suboptimal for a lower discount factor as compared to δ_1. To calculate such a discount factor threshold, note that for a player who is off-equilibrium at the start of a partnership we have the following:

$$v_k = \sigma_k(1+g) + \delta^T \sum_{k'=k}^{N} Q_{kk'}(N) v_{k'}, \qquad k = 2, \dots, N.$$

This is because a player who is in a defection state must punish in every period, under the grim strategy. Hence,

$$v_k = e_k^\intercal (\mathcal{I} - \delta^T \mathcal{Q}_N)^{-1}(1+g)\sigma,$$

which using the same definitions for $\phi_k(\delta^T)$ as before gives us

$$v_k = \frac{\phi_k(\delta^T)}{1-\delta^T}(1+g).$$

As before, deviating in any stage $\tau < T$ is suboptimal in equilibrium because the deviator is immediately punished with D. Hence, we consider a deviation in the very last stage, T. Here deviating is suboptimal if $1 + g + \delta v_2 \leq 1 + \delta v_0$

$$\begin{aligned} & \delta \frac{\phi_2(\delta^T)}{1-\delta^T}(1+g) \leq \frac{\delta}{1-\delta} - g, \\ \Rightarrow \qquad & \phi_2(\delta^T)(1+g) \leq (1-\delta^T)\frac{\delta(1+g)-g}{\delta(1-\delta)}. \end{aligned}$$

This last inequality gives us a value for δ_1' making cooperation incentive compatible in equilibrium for $\delta \geq \delta_1'$.

A key disadvantage of the grim strategy is a lower incentive to follow the community sanction off-equilibrium, because now its cost is larger as compared to the cost of the sanction implied by the strategy in Definition 1. This disadvantage is especially important when in large groups because the lost income from the community sanction is proportionately larger. Moreover, in the initial phases of the contagion process a player who has observed a defection has a great incentive to cooperate in the first stage of a new partnership. The reason is that the population is mostly composed of cooperators at this point, so it makes sense to try to earn income from a cooperative outcome in stage 1 and then revert to the grim sanction, earning the temptation payoff in stage 2. This incentive to avoid the community sanction is especially large when g is small, which is when cooperation has a large value relative to the temptation payoff.

To see this, suppose we are off-equilibrium. According to the grim strategy, punishment takes place in every period of every partnership. Now suppose that the player who is in a punishment state chooses to deviate from this strategy in stage 1 of a partnership, and revert to it in stage 2. This implies that contagion will surely occur in this meeting so the continuation payoff will be unaffected by the deviation. It follows that off-equilibrium, the condition for optimality of punishment in stage 1 of a partnership is:

$$\sigma_k[1+\delta(1+g)] + (1-\sigma_k)(-l) \leq \sigma_k(1+g) \qquad \Rightarrow \qquad \delta \leq \frac{l}{\sigma_k(1+g)} + \frac{g-l}{1+g}$$

Note that a player that is in a punishment state would not deviate from punishment in the last stage of a partnership. Intuitively, if the player did not deviate in earlier stages, then his opponent is certainly in a punishment state. Hence, deviating from the punishment strategy in stage T cannot stop contagion and can only lower the current payoff of the deviator. Instead, if the player deviated in some earlier stage (cooperating instead of defecting) he has now reverted to playing the grim strategy.[5] This is why we must focus on deviations from punishment in the very first stage of a partnership.

Since σ_k decreases in k and $\sigma_2 = \frac{N-2}{N-1}$ then the sufficient and necessary condition is

$$\delta \leq \delta_2' := \frac{(N-1)l}{(N-2)(1+g)} + \frac{g-l}{1+g},$$

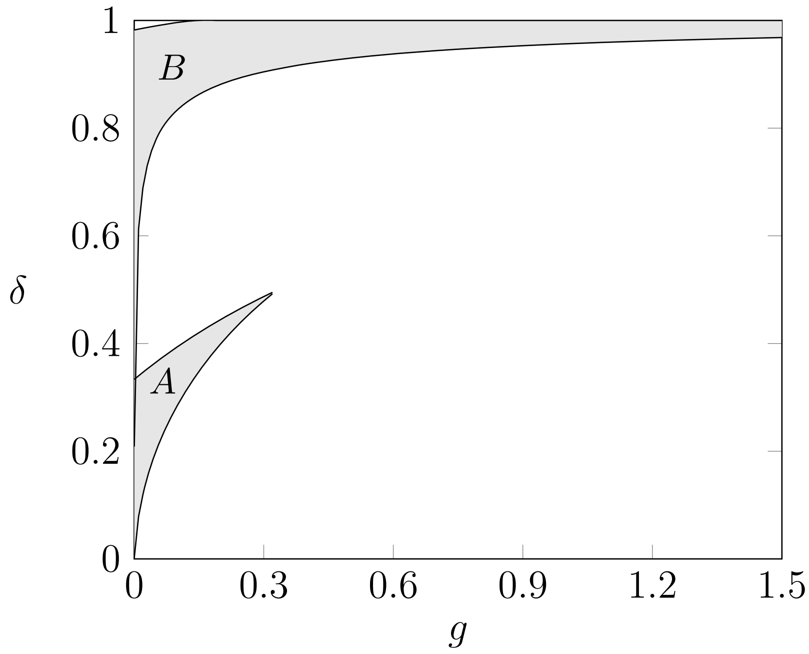
since grim punishment must be optimal for all possible states $k = 2, \ldots, N$. Intuitively, one cannot gain from cooperating in the first stage of a partnership, and then go back to punishment when she is sufficiently impatient.

An example is reported in Figure 2. For g small the grim strategy helps supporting full cooperation if players are impatient (area A), while the strategy in Definition 1 is helpful if players are patient (area B). As g increases and becomes sufficiently large, roughly grows above 0.32, the grim strategy cannot support full

[5]There are multi-period deviations, which may prevent contagion. For example, cooperate in stage 1 of a partnership (instead of defecting as prescribed by the punishment norm) and keep cooperating until the end of the partnership if the counterpart is in a cooperative state in stage 1. We do not consider multi-period deviations for the proposed strategy.

cooperation for any values of δ, while the strategy in Definition 1 still can if players are sufficiently patient.[6]

Figure 2: Parameters supporting full cooperation.



Notes: In the figure $l = 6, N = 20, T = 4$. The shaded areas identify parameters supporting full cooperation; area A considers the grim strategy in Definition 2, while area B considers the strategy in Definition 1.

4 Conclusions

This study has shown that forming temporary partnerships in an infinitely repeated social dilemma has two benefits. An obvious benefit is that it is possible to immediately and directly punish a temporary partner who deviates from equilib-

[6]This analysis also suggests that increasing the community punishment spell above one stage is not necessarily helpful to support cooperation. To see this, suppose community punishment occurs in 2 or more stages. This induces a greater incentive to cooperate in equilibrium because continuation payoffs are smaller off equilibrium. Deviating will thus be unprofitable for a lower discount factor as compared to δ_1. This would push the lower boundary of area B below the one depicted in Fig. 2, bridging the gap between areas A and B. However, multiple community punishment stages also lower the incentive to punish off-equilibrium, because the cost of punishing increases. This is especially important when there are many possible cooperators, which is when punishing off equilibrium would require a lower threshold as compared to δ_2. This would push the upper boundary of area B below the one in Fig. 2, reducing the set of parameters supporting cooperation.

rium play in some early stage of the partnership. Although this is not necessarily sufficient to deter deviations in those stages, this kind of direct punishment is helpful because it reduces the incentives to deviate in early stages as compared to later stages of the partnership. In fact, the worse-case scenario concerns the very last stage of interaction with a counterpart, i.e., right before the meeting ends. In that stage, a deviator cannot be directly punished and so community punishment must be used to incentivize cooperation.

Another benefit of temporary partnerships is that players can use the commonly known periodicity T of the matching function as an *explicit* coordination device. This allows players to deter defections by coordinating on an intuitive community sanction: if someone deviates in some stage, then the player who observes the deviation will never cooperate in that stage of *any* future partnership. This community sanction is cyclical because it requires punishment to occur every T periods, hence it has the benefit of supporting partial cooperation even off the equilibrium path. We have shown that players who are periodically arranged into temporary partnerships can support full cooperation as a subgame perfect equilibrium by coordinating on sanctions that combine direct punishment of a current counterpart, with a community sanction of *all* future counterparts. The sanction is milder than the grim sanction typically adopted in these indefinitely repeated matching games. Two partners who have coordinated on this community sanction will cooperate in all stages of their temporary partnership but the community punishment stage.

A main insight is that there are pros and cons to adopting this milder form of community punishment. The disadvantage is a higher discount factor threshold δ_1. In other words, incentivizing cooperation requires players that are more patient as compared to using a grim sanction. The advantage is that a milder sanction more easily motivates punishment as compared to the harsh grim sanction, which, in fact, may not motivate punishment at all if the gain from taking advantage of a cooperator is large (the temptation payoff component g). In that case, adopting a grim punishment is not incentive compatible, unlike the milder community sanction considered in this paper. Note that the incentive-compatibility condition supporting cooperation in equilibrium is the same as the threshold δ reported in Kandori (1992), re-scaled to the power $1/T$. Intuitively, one can interpret our T-period repeated matching game as a composition of T "virtual" 1-period repeated matching games (each corresponding to a partnership stage). Community punishment happens only in the virtual game $\tau = 1, \ldots, T$ where someone deviates from the proposed strategy. It does not spread to any of the other virtual games—hence the similarity with the condition on δ in Kandori (1992). Yet, the condition for incentive-compatibility off-equilibrium is quite different from the one reported in Kandori (1992), which involves the parameter l. This is because punishment in our setting does not spread to all periods, so players must be sufficiently impatient to prefer to punish. This gives us an additional condition on the discount factor δ.

To be sure, such a milder punishment scheme is also possible without temporary partnerships. However, in that case there is no explicit reference point available to players. When matches only last one period, infinite punishment periodicities should also be possible. Yet, coordinating on a cyclical punishment scheme seems problematic in this case because players must be able to do so *tacitly*. It remains to be seen if these theoretical insights hold up to empirical scrutiny, which we hope to assess in a laboratory setting in the future. Indeed, note that in our setup the temporary partnerships allow players to form a temporary reputation so it would be interesting to compare results of this setting with indefinitely repeated PD experiments where players can form a reputation thanks to a public monitoring mechanism (e.g., see Stahl, 2013; Camera and Casari, 2018). In those experiments players carry their reputation with them as they move from random meeting to random meeting. This contrasts with our design where subjects can only form a reputation within a temporary partnership but *cannot* carry their reputation into subsequent temporary partnerships. Since no mechanism exists that allows transmission of information to future partners, we expect lower cooperation as compared to designs with public monitoring of defections.[7]

[7] There is evidence of a positive effect of public monitoring of past defections on cooperation. In

A natural question is how our results would change if we endogenized the duration of a partnership T. For instance, suppose players could terminate a partnership at some cost and be immediately rematched. Would this help cooperation? The answer is likely negative: adding an outside "exit" option (terminating the match early) should weaken the incentive for cooperation. Since no rational player would have an incentive to break up a cooperative relationship, we would observe temporary partnerships *only* off equilibrium. By contrast, in our setup partnerships are temporary also in equilibrium, and off equilibrium they may involve both defectors and cooperators. Without the ability carry reputation from meeting to meeting, introducing a breakup option would simply weaken the incentive to cooperate in equilibrium, and to punish off equilibrium. Someone who deviates in equilibrium could limit the duration of the punishment spell, exiting the meeting to try to match up with another cooperator. Cooperators who suffer a defection also have an incentive to break up a meeting, to avoid a string of (low) defection payoffs and, instead, seek a new, and possibly cooperative, partner. Hence, if the breakup option is generally "in the money" (e.g., no breakup costs), then economic incentives for cooperation should decline as compared to our setup. Data from indefinitely repeated experiments with break-up options seem to support this conclusion (Honhon and Hyndman, 2020; Wilson and Wu, 2017).

In fact, given these previous experiments, it would be interesting to further develop our understanding of cooperation in temporary partnerships by bringing the setup proposed in this paper to the laboratory and combine it with computational experiments. Laboratory experiments would allow us to document possible differences in human behavior in cooperation tasks based on temporary partnerships without break-up options vs. constant one-period random rematching. Complementing these experiments with computational work that leverages machine-learning algorithms could be useful to study how individuals learn to

Stahl (2013), subjects can detect if their current counterpart ever defected with a cooperator in the past and learn to use this reputational mechanism to support cooperation as they gain experience with the task. Yet, public monitoring is not always effective. In Camera and Casari (2018), players observe statistics of past behavior of their current counterpart that are created either by an endogenous or an exogenous monitoring mechanism. Here, cooperation does not improve as compared to a treatment where no monitoring is possible.

coordinate on cooperation in environments where they are constantly randomly rematched as opposed to when they face temporary partnerships. It would be especially interesting to adopt Individual Evolutionary Learning algorithms, as for instance discussed in Arifovic and Ledyard (2012, 2018), to study if and how individuals might be able to support efficient outcomes by learning to coordinate on a punishment scheme consisting of a mix of direct and cyclical community punishment, as suggested in our analysis.

References

Arifovic, Jasmina and John Ledyard. 2012. Individual evolutionary learning, other-regarding preferences, and the voluntary contributions mechanism. *Journal of Public Economics*, 96, 808-823.

Arifovic, Jasmina and John Ledyard. 2018. Learning to alternate. *Experimental Economics* 21, 692-721.

Bigoni, M, M. Casari, A. Skrzypacz, and G. Spagnolo. 2015. Time horizon and cooperation in continuous time. *Econometrica* 83(2), 587-616.

Blonski, Matthias, Peter Ockenfels, and Giancarlo Spagnolo. 2011. Equilibrium Selection in the Repeated Prisoner's Dilemma: Axiomatic Approach and Experimental Evidence. *American Economic Journal: Microeconomics* 3 (3), 164-192.

Camera, G., and M. Casari. 2009. Cooperation among Strangers under the Shadow of the Future. *The American Economic Review* 99(3), 979-1005.

Camera, G., and M. Casari (2018). Monitoring institutions in indefinitely repeated games. *Experimental Economics* 21, 673-691.

Camera, G., Casari, M., and M. Bigoni. 2012. Cooperative Strategies in Anonymous Economies: an Experiment. *Games and Economic Behavior* 75, 570-586.

Camera, G., Casari, M., and M. Bigoni. 2013. Money and Trust among Strangers. *Proceedings of the National Academy of Sciences* 110(37), 14889-14893.

Camera, G., Gioffré, A., 2014. A tractable analysis of contagious equilibria. *Journal of Mathematical Economics* 50, 290-300.

Dal Bó, P.. 2005. Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games. *American Economic Review* 95 (5), 1591-1604.

John Duffy, and Huan Xie. 2016. Group size and cooperation among strangers. *Journal of Economic Behavior & Organization* 126, 55-74.

Duffy, J., and Ochs, J., 2009. Cooperative behavior and the frequency of social interaction. Games an Economic Behavior 66 (2), 785-812.

Friedman, James W. 1971. A Non-Cooperative Equilibrium for Supergames. *Review of Economic Studies*, 38(113), 1-12.

Friedman, Dan, and Ryan Oprea (2012). A Continuous Dilemma. *American Economic Review*, 102 (1), 337-363.

Ghidoni, R. and S. Suetens. The effect of sequentiality on cooperation in repeated games. *American Economic Journal: Microeconomics*, 14(4), 58-77.

Honhon, D. and K. Hyndman (2020). Flexibility and Reputation in Repeated Prisoner's Dilemma Games. *Management Science* 66(11), p. 4998-5014.

Kandori, M., 1992. Social norms and community enforcement. *Review of Economic Studies* 59, 63-80.

Stahl, Dale O. 2013. An experimental test of the efficacy of a simple reputation-mechanism to solve social dilemmas. *Journal of Economic Behavior & Organization* 94, 116-124.

Wilson A. J., and H. Wu (2017) At-will relationships: How an option to walk away affects cooperation and efficiency. *Games and Economic Behavior* 102, p. 487-507.