DISEI - Università degli Studi di Firenze

Working Papers - Economics

Beyond Grim: Punishment Norms in the Theory of Cooperation

Gabriele Camera and Alessandro Gioffré

Working Paper N. 12/2018

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

Beyond Grim: Punishment Norms in the Theory of Cooperation*

Gabriele Camera
ESI, Chapman University
University of Bologna

Alessandro Gioffré
University of Florence

January 30, 2019

Abstract

The theory of repeated games asserts that, when past conduct is unobservable, the efficient outcome is attainable for any given payoff structure, if players are sufficiently patient. Here, we establish a complementary result: the efficient outcome is attainable for any degree of patience, if moving off equilibrium generates limited gains. This result builds on a class of punishment norms less extreme than "grim," which, in fact, may be counterproductive if losses are small, as it *prevents* cooperation among patient players. Our analysis reveals that adoption of moderate punishment schemes can support cooperation when players are impatient, and provides a rationale for the empirical observation that grim punishment is uncommon in laboratory studies of cooperation.

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

JEL codes: E4, E5, C7

1 Introduction

Individuals who interact repeatedly, are patient, and can perfectly monitor each other's past conduct can cooperate rather easily. According to theory, all they need is to punish a defection with a suitably long spell of defections (Friedman, 1971; Fudenberg and Maskin, 1986; Abreu et al., 1990). Things, however, are not

* 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

so simple when private monitoring and random meetings shroud past conduct. Here, individuals are "strangers" who cannot build reputations, so theory asserts that cooperation requires two separate ingredients. Roughly speaking: the common adoption of a *grim* form of punishment—forever ceasing all cooperation—and the deterrent of significant losses for being lenient or forgiving. This ensures that refusing to cooperate initiates a punishment process that is individually rational, and inexorably spreads from person to person until everyone plunges into a permanent uncooperative state (Kandori, 1992).

The extreme severity of this community punishment scheme raises two kinds of issues. One is purely theoretical. Grim punishment can be counterproductive and may in fact destroy cooperative equilibrium. Players who are sufficiently patient would *not* want to carry out the punishment threat when the loss caused by meeting a defector is small (Ellison, 1994). This is why—unless a public coordination device is available—the theory imposes restrictions on off-equilibrium payoffs: it is a way to deter cooperators from condoning defections that cause a small loss. Second, grim punishment is empirically implausible. One would be hard pressed to find societies, present or past, which adopt extreme punishment threats to deter anything but the biggest crimes. Moreover, laboratory evidence reveals that grim punishments are infrequent both among strangers and among partners, while strategies that are lenient or forgiving are much more common (Camera et al., 2012; Fudenberg et al., 2012; Dal Bó and Fréchette, 2011). The question is why.

Standard theory has focused on the grim punishment, and little is known about the feasibility—and the possible advantages—of more moderate punishment schemes. In a way, the theory asserts that communities will tacitly coordinate on the most severe punishment form to deter any kind of deviation, even one that generates only small immediate gains to the deviator (a small temptation payoff). This seems unreasonable and, in fact, is counterproductive when the infractions are minor, i.e., when to small temptation payoffs also correspond small losses (sucker's payoff). The analysis in Ellison (1994) shows that moderate punishments can support full cooperation also when losses are small if players are sufficiently patient. Here, instead, we establish existence of cooperative equilibrium for the converse case where players can be arbitrarily *impatient*. We demonstrate that for any value of the discount factor there exist strategies supporting efficient play as long as the temptation payoff is sufficiently small. To do so, we adopt a new methodology that allows us to provide closed-form solutions of the discount factor boundaries and the continuation payoffs for generic punishment strategies.

Importantly, we show that there may be multiple strategies that support full cooperation independent of the size of losses to defectors. This is accomplished by studying a class of punishment schemes that do not destroy all cooperation off equilibrium and, in fact, are designed to *guarantee* partial cooperation if someone derives a small benefit by deviating from efficient play. In this sense, the motivation for the use of moderate punishments is entirely strategic: patient players should not dissipate the gains from future cooperation to sanction minor infractions.

The argument is developed using a repeated Prisoner's Dilemma with random matching, as done in Kandori (1992) and Ellison (1994). We start by showing that, if the sucker's payoff is small, then a grim strategy cannot generally support full cooperation among patient players. Full cooperation might be attainable only if players are moderately—but not excessively—patient. This result builds on Ellison (1994, Proposition 4), and expands upon it by adopting a methodology allowing us to calculate exact expressions for the discount factor boundaries.

We thus proceed by demonstrating how to solve this equilibrium non-existence problem for (excessively) patient players by reducing the severity of community punishment. This second step builds on Ellison (1994, Lemma 2), which recognizes the importance of softening punishment—spreading it out over time—when players are patient. We expand upon it, by uncovering the relationship between temptation payoff and frequency of punishment, and by determining continuation payoffs as functions of parameters. As the temptation payoff becomes small, the equilibrium interval of the discount factor covers the open unit interval.

The rest of the paper is organized as follows. In Section 2, we present the model. In Section 3, we develop the main result, while Section 4 concludes.

2 Model

There are $N = 2n \geq 4$ infinitely-lived players. In every period $t = 0, 1, 2, \ldots$, players are matched in pairs using a uniform random matching mechanism. Therefore, in each period t, the probability that player i meets any player $j \neq i$ is $\frac{1}{N-1}$. Every pair (i, j) plays a Prisoner's Dilemma, with action set $\{C, D\}$ ("cooperate" and "defect"). To facilitate comparisons between our analysis and existing theory, we retain the same payoff matrix and basic notation as in the classical models in Kandori (1992) and Ellison (1994). 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 stage game between player i and j.

As usual, we assume $g, l > 0$. Full cooperation (C, C) is mutually beneficial and corresponds to the efficient outcome.[1] Full defection (D, D) is the unique Nash equilibrium in the one-shot game. 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

$$\sum_{t=0}^{\infty} \delta^t \pi_i(a_{i,t}, a_{j,t}) \tag{1}$$

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 $(a_{i,t}, a_{j,t})$.

[1] The usual assumption in this game is $1 > g - l$. However, the results in this paper hold also if this inequality is not fulfilled.

There is private monitoring: players can only observe outcomes and actions in their pair and cannot observe the history of play of their opponents.

Since we are interested in studying strategies that support the efficient outcome, we start by considering the trigger strategy proposed in Kandori (1992), which is based on the threat of an unforgiving "grim punishment."

Definition 1 (Grim punishment)**.** *In period $t = 0$, the player cooperates. In period $t \geq 1$, the player cooperates if she has not observed a defection, otherwise if a defection was observed in any period $\tau \leq t$, then the player defects in periods $(\tau + h)_{h=1}^{\infty}$.*

This strategy encompasses two modes of behavior: *cooperation*, when the player has experienced a fully cooperative outcome in every match; *punishment*, if the player defected or suffered a defection at any point in the past. The strategy supports full cooperation by threatening a sanction prescribing defection in every period of the continuation game. A player who is in the grim punishment mode will never revert to cooperating, not even occasionally. This explains why punishment is called grim: punishment is antithetical to any form of future cooperation, even some limited form of cooperation. An illustration is in Figure 1.

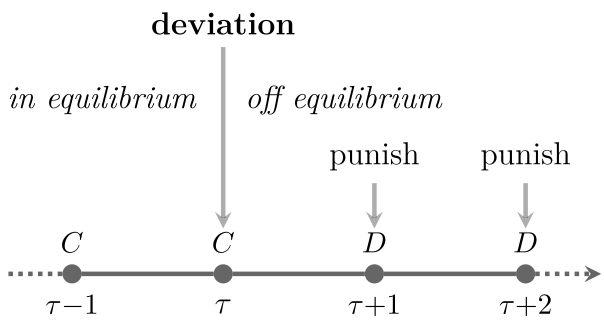


Figure 1: Grim punishment: Defect in every period after a deviation.

We are interested in studying whether or not the efficient outcome can be sustained by punishments that are not so extreme. In particular, punishments that leave the door open to some degree of cooperation in the continuation game. Consequently, we define another strategy that is based on a form of punishment that can be more restrained, if needed.

Definition 2 (Generalized punishment)**.** *Fix $T \geq 1$. In period $t = 0$, the player cooperates. In period $t \geq 1$, the player cooperates if she has not observed a defection; otherwise, if a defection was observed in any period $\tau \leq t$, then the player only defects in periods $(\tau + hT)_{h=1}^{\infty}$ and in no other periods.*

The strategy in Definition 2 is illustrated in Figure 2. As for the case of grim punishment, this strategy also encompasses two modes of behavior: *cooperation*, when the player has experienced a fully cooperative outcome in every match; *punishment*, if the player defected or suffered a defection at any point in the past. However, unlike the grim punishment strategy, full cooperation is supported by threatening a sanction that prescribes defection only in a subsequence of periods of the continuation game. The defections will take place at regular intervals, $T, 2T, 3T, \ldots$, following the first observed defection. In all other periods the player cooperates, even if, the player is in the punishment mode. Notice that the strategy in Definition 1 can be obtained by simply fixing $T = 1$ in Definition 2.

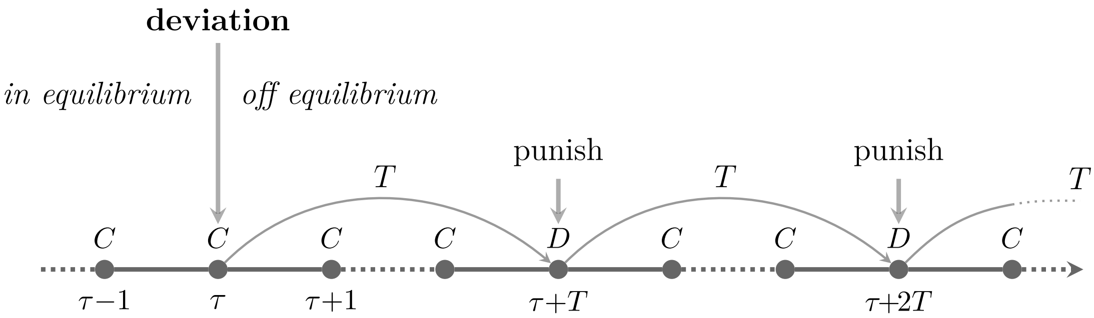


Figure 2: Restrained punishment: Defect every T periods after a deviation.

We say that a strategy is a social norm if it is adopted by every player. The two social norms in Figure 1 and 2 share a fundamental similarity: each norm exploits an implicit threat of community enforcement. The central idea is that this threat should be sufficient to support full cooperation in equilibrium. Independent of the strategy adopted, the observation of a defection moves the economy off equilibrium, as it triggers a switch from the cooperative mode to some punishment mode. Because of random matching and private monitoring, this switch will spread in the economy over time, similarly to a process of "contagion."

In Figure 1 and 2 the long-run consequences of punishment are different. Grim punishment leaves no hope for any kind of future cooperation because the initial defection is never forgiven. The strategy illustrated in Figure 2 supports some cooperation off equilibrium, in the long run since $T > 1$. The initial defection gives rise to punishment at a pre-specified frequency $1/T$, pinning down an infinite sequence of equidistant periods in which punishment must take place. Players will therefore alternate "punishment periods" to "cooperation cycles" in which—so to speak—all defections are periodically forgiven and there is no further contagion.

3 The main result

In the repeated matching game we just described, if players adopt a social norm based on the strategy in Definition 2, then the efficient outcome is a sequential equilibrium for values of δ sufficiently large (Ellison, 1994, Proposition 4). That is to say, for any Prisoner's Dilemma players can always support full cooperation as a sequential equilibrium if they are sufficiently patient.

This section develops a unique result. Arbitrarily patient players can attain full cooperation as a sequential equilibrium, independent of the size of the sucker's payoff; all is needed is a temptation payoff g sufficiently small. As g declines to zero, full cooperation is an equilibrium for any value of players' discount factor; multiple strategies that support efficiency also may exist.

Let players adopt the generalized punishment strategy in Definition 2.

Theorem 1. *For any $\delta \in (0,1)$, there exists $T \geq 1$ such that the generalized punishment strategy supports cooperative equilibrium if g is sufficiently small. Moreover, for some $\delta \in (0,1)$, there exist multiple strategies that support cooperative equilibrium.*

We prove this theorem in three steps. First, we present a new proof to show that a fully cooperative equilibrium always exists for some values of δ when the social norm is based on the strategy in Definition 2, with $T = 1$. We also show that full cooperation may not be an equilibrium for δ sufficiently close to 1. These first two results build on the analysis in Ellison (1994, Proposition 4). To prove them, we demonstrate that, for any l, a social norm based on the strategy in Definition 2, with $T = 1$, supports the efficient outcome in sequential equilibrium on an interval $[\delta_1, \delta_2] \subset (0,1)$ of δ values. However, in this case where no restrictions are imposed on the sucker's payoff l, the equilibrium might not exist unless δ is *bounded away* from one. That is to say, full cooperation might be attainable only if players are moderately—but not excessively—patient. This result is reported in Ellison (1994, Proposition 4); our contribution relative to Ellison (1994) is to formalize this result using a unique methodology, which allows us to calculate exact expressions for the boundaries of the interval of the δ values.

Second, we expand the set of discount factors to include values above δ_2 by working with strategies that reduce the severity of community punishment through an appropriate modulation of its frequency. This step relies on realizing that patient players may be incentivized to punish off equilibrium by appropriately choosing the punishment frequency $1/T$. We demonstrate that a social norm based on the strategy in Definition 2, with $T \geq 1$, supports the efficient outcome in sequential equilibrium on a set $\mathcal{D}$ of δ values, which is generally disconnected and such that $[\delta_1, \delta_2] \subset \mathcal{D} \subset (0,1)$.

Finally, using the two partial results discussed above, we prove our main result. As g becomes small, then the set $\mathcal{D}$ becomes connected and equal to $(0,1)$ as g approaches zero. This step is accomplished by developing a methodology that allows us to determine exact expressions for payoff functions, thus identifying the relationship between the temptation payoff and the frequency of punishment necessary to support cooperation.

We start by characterizing continuation payoffs in and out of equilibrium under the grim strategy in Definition 2, with $T = 1$.

3.1 Cooperation under a grim punishment threat

Suppose that every player adopts the strategy in Definition 2, with $T = 1$. Recall that when $T = 1$, the strategy in Definition 2 is the grim strategy in Definition 1.

On the equilibrium path there is full cooperation and the payoff to any player is

$$v_0 = \frac{1}{1-\delta}. \tag{2}$$

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

$$\mathcal{Q}_M := \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & Q_{22} & 0 & Q_{24} & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & Q_{34} & 0 & Q_{36} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & Q_{M-2,M-2} & 0 & Q_{M-2,M} \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix}. \tag{3}$$

To describe the elements of $\mathcal{Q}_M$, suppose that at the start of a period there are $k = 1, \dots, M$ defectors. The generic element $Q_{kk'} = Q_{kk'}(M)$ is the probability to transition from k to $k' \geq k$ defectors across two periods. We can show that

$$Q_{kk'}(M) := \frac{(k'-k)!\binom{k}{k'-k}\binom{M-k}{k'-k}(2k-k'-1)!!(M-k'-1)!!}{(M-1)!!}, \tag{4}$$

where the number of additional defectors created in a period is

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

The details of this derivation are in Camera and Gioffré (2014).

Now consider $M = N$. Let v_k denote the expected payoff to a generic defector, when there are k defectors and $N-k$ cooperators. Due to uniform random matching, a defector meets a cooperator with probability $\sigma_k := \dfrac{N-k}{N-1}$. Thus, the expected payoff to the generic defector can be recursively defined by

$$v_k = (1+g)\sigma_k + \delta \sum_{k'=k}^{N} Q_{kk'}(N) v_{k'}.$$

Again, following Camera and Gioffré (2014), v_k can be rewritten as

$$v_k = \frac{\phi_k(\delta)}{1-\delta}(1+g), \tag{5}$$

with

$$\phi_k(\delta) := (1-\delta) e_k^{\mathsf{T}} (\mathcal{I} - \delta \mathcal{Q}_N)^{-1} \sigma,$$

where $\sigma := (\sigma_1, \ldots, \sigma_N)^{\mathsf{T}}$ and e_k is the $N-$dimensional column vector with 1 in the k^{th} position and 0 everywhere else. The right-hand side of (5) is the expected gain that the defector would receive from meeting cooperators. To discuss $\phi_k(\delta)$ note that preferences are linear in payoffs. Hence, we can interpret $1-\delta$ as the probability that the game ends after each period. Under this interpretation, $\phi_k(\delta)$ can be considered the expected number of future meetings with cooperators divided by the expected duration of the continuation game. For convenience we call it the "contact rate."

The following result holds:

Lemma 1. *For all $k = 1, \ldots, N-1$, $\phi_k(\delta)$ is a decreasing function of $\delta \in (0,1)$.*

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

$$\phi_k'(\delta) = -e_k^{\mathsf{T}} (\mathcal{I} - \delta \mathcal{Q}_N)^{-1} [\mathcal{I} - (1-\delta) \mathcal{Q}_N (\mathcal{I} - \delta \mathcal{Q}_N)^{-1}] \sigma,$$

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

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

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

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

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

□

As we show later, it is convenient to rewrite v_k as the sum of two expected payoffs depending on whether the defector earns the temptation payoff $1+g$ in a meeting with a cooperator (with probability σ_k) or the defection payoff 0 in a meeting with another defector (with probability $1-\sigma_k$), i.e.,

$$v_k = \sigma_k\left[1+g+\delta\sum_{k'=k-1}^{N-2} Q_{k-1,k'}(N-2)v_{k'+2}\right] + (1-\sigma_k)\delta\sum_{k'=k-2}^{N-2} Q_{k-2,k'}(N-2)v_{k'+2}, \tag{6}$$

$Q_{kk'}(N-2)$ 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 her opponent. There are two cases to consider.

First, the opponent is a cooperator. Here, next period we expect $k'+2$ defectors. The number 2 includes the defector and her 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 $\sum_{k'=k-2}^{N-2} Q_{k-2,k'}(N-2)v_{k'+2}$.

Therefore, expression (6) simply splits the expected payoff $\sum_{k'=k}^{N} Q_{kk'}(N)v_{k'}$ into two parts:

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

This observation allows us to prove a result that mirrors Ellison (1994, Proposition 4).

Proposition 1. *There exists* $0 < \delta_1 < \delta_2 \leq 1$ *such that for* $\delta \in [\delta_1, \delta_2] \cap (0,1)$ *the strategy in Definition 2, with* $T = 1$*, supports full cooperation as a sequential equilibrium.*

We prove Proposition 1 by establishing three separate results. First, we show that there exists a value $\delta_1 \in (0,1)$ such that equilibrium deviations are suboptimal for $\delta \geq \delta_1$. Then, we show that there exists a value $\delta_2 \in (0,1]$ such that off-equilibrium punishment is incentive compatible for $\delta \leq \delta_2$. Third, we prove that $\delta_1 < \delta_2$ for all payoff matrices of the Prisoner's Dilemma.

3.1.1 Equilibrium deviations

A deviation in equilibrium is never optimal if $v_1 \leq v_0$. Using the definitions of v_0 and v_k, for $k = 1$, we have

$$v_1 - v_0 \leq 0 \qquad \Leftrightarrow \qquad \phi_1(\delta) \leq \frac{1}{1+g}, \tag{7}$$

where the left-hand side is the contact rate if the player is the first to defect and moves the economy off equilibrium. The right-hand side is the ratio between the payoff from full cooperation and the temptation payoff; it measures the relative gain from defecting against a cooperator. Notice that ϕ_1 maps $[0,1)$ into $(0,1]$ and it is a strictly monotone, decreasing function of δ (Camera and Gioffré, 2014). Then, it is invertible, i.e., $\delta = \phi_1^{-1}(x)$ for $x \in (0,1]$. Now, since $\frac{1}{1+g} \in (0,1)$, there exists a value $\delta_1 \in (0,1)$ such that

$$\delta_1 := \phi_1^{-1}\Big(\frac{1}{1+g}\Big). \tag{8}$$

Therefore, monotonicity of ϕ_1 ensures that for all $\delta \in [\delta_1, 1)$ expression (7) is satisfied and deviating in equilibrium is suboptimal.

3.1.2 Off-equilibrium deviations

Suppose that there are k defectors in the economy, and let player i be one of them. If player i deviates by choosing C instead of D, then she loses l in a match with another defector and earns 1 in a match with a cooperator. The expected payoff to deviator i from a one-time deviation is therefore

$$\tilde{v}_k = \sigma_k \overbrace{\left[1 + \delta \sum_{k'=k-1}^{N-2} Q_{k-1,k'}(N-2) v_{k'+1}\right]}^{i \text{ meets a cooperator}} + (1-\sigma_k) \overbrace{\left[-l + \delta \sum_{k'=k-2}^{N-2} Q_{k-2,k'}(N-2) v_{k'+2}\right]}^{i \text{ meets a defector}}. \tag{9}$$

Comparing expression (6) to (9) reveals that deviating by choosing C instead of D affects the expected payoff of player i in two ways. First, it reduces the expected current earnings by $\sigma_k g + (1 - \sigma_k)l$. Second, if meeting a cooperator, it increases the continuation payoff to $v_{k'+1}$ instead of $v_{k'+2}$; it is an increase as v_k falls in k.

Deviating by choosing C instead of D off equilibrium is suboptimal if $\tilde{v}_k \leq v_k$, with $k \geq 2$. Using (6) and (9), this inequality can be rewritten as

$$\sigma_k \delta \sum_{k'=k-1}^{N-2} Q_{k-1,k'}(N-2)(v_{k'+1} - v_{k'+2}) \leq \sigma_k g + (1 - \sigma_k)l. \tag{10}$$

The left-hand side represents the expected gain from slowing down the contagious defection process, while the right-hand side represents the expected loss. Using (5), expression (10) can be rearranged as

$$\frac{\delta}{1-\delta} \sum_{k'=k-1}^{N-2} Q_{k-1,k'}(N-2)[\phi_{k'+1}(\delta) - \phi_{k'+2}(\delta)] \leq \frac{g}{1+g} + \frac{1-\sigma_k}{\sigma_k} \frac{l}{(1+g)}. \tag{11}$$

Since $\phi_k(\delta) - \phi_{k+1}(\delta)$ is decreasing in k (Camera and Gioffré, 2014, Theorem 2) and $\frac{1-\sigma_k}{\sigma_k}$ is increasing in k, the most stringent case for (11) is $k = 2$.

Given that the first row of transition matrix $\mathcal{Q}_M$ has 1 as the second element and 0 otherwise, for any M, then we need

$$\frac{\delta}{1-\delta}[\phi_3(\delta) - \phi_4(\delta)] \leq \frac{g}{1+g} + \frac{1-\sigma_2}{\sigma_2} \frac{l}{(1+g)}. \tag{12}$$

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

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

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

Proof. Suppose that player i moves off equilibrium, i.e., she is the only defector in the economy. Next period there will be two defectors in the economy with certainty. Accordingly, player i's expected payoff must satisfy

$$v_1 = 1 + g + \delta v_2.$$

Using (5) we obtain

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

which, rearranging, gives us the result. $\square$

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

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

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

Therefore, to ensure that (12) holds, it is sufficient to show that

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

The expression can be rearranged as

$$\frac{1 - l/(N-2)}{1+g} \leq \phi_1(\delta) \tag{13}$$

where we note $\dfrac{1-\sigma_2}{\sigma_2} = \dfrac{1}{N-2}$. There are two cases. If $l \geq N-2$ then (13) holds for all $\delta \in [0,1)$ since $\phi_1(\delta) \geq 0$. This is in line with the analysis in Kandori (1992). Instead, if $0 < l < N-2$ then the left-hand side of (13) lies in the unit interval. Hence, by continuity of ϕ_1 there exists a value $\delta_2 \in (0,1)$ such that (13) holds with equality. Since ϕ_1 is a decreasing function of δ we can conclude that (13) is satisfied for all $\delta \in [0, \delta_2]$. It follows that deviating from the proposed punishment is suboptimal off equilibrium for all $\delta \in (0, \delta_2] \cap (0, 1)$, where

$$\delta_2 := \begin{cases} \phi_1^{-1}\left(\frac{1-l/(N-2)}{1+g}\right) & \text{if} \quad 0 < l < N-2 \\ 1 & \text{if} \quad l \geq N-2. \end{cases} \tag{14}$$

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 $\delta_2 = 1$. For $0 < l < N-2$, we use the definition of δ_2 and δ_1 to derive the following inequality:

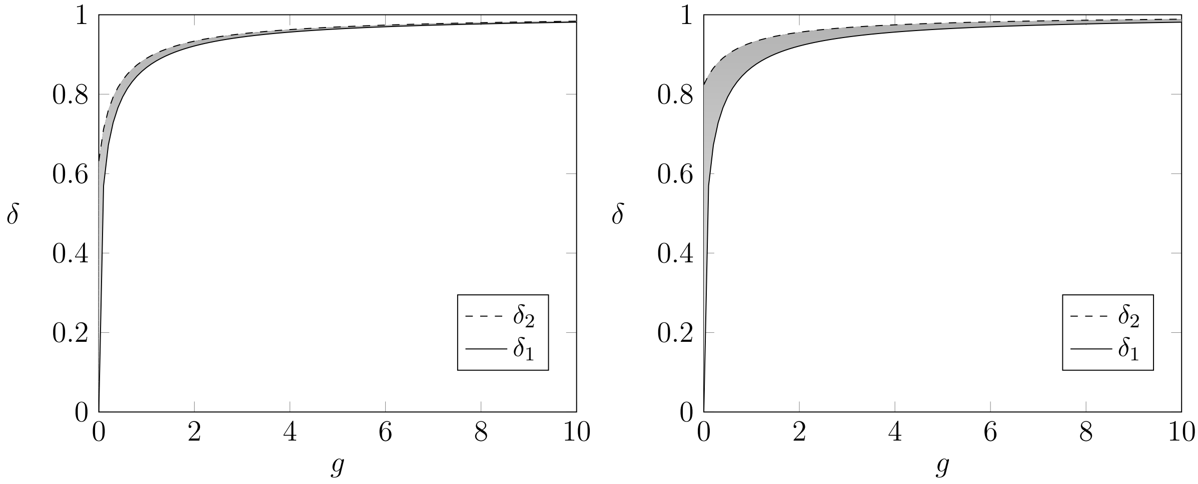
$$\phi_1(\delta_2) = \frac{1 - l/(N-2)}{1+g} < \frac{1}{1+g} = \phi_1(\delta_1).$$

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

The previous analysis also allows us to demonstrate a result that is discussed in Ellison (1994), albeit not formally proved.

Corollary 1. *Full cooperation can be supported independent of l as long as players are not too patient. Otherwise, full cooperation cannot be supported as an equilibrium for l sufficiently small.*

The proof of this corollary is an immediate consequence of expression (12). If δ is close to 1 the left-hand side of this inequality is bounded away from 0, while its right-hand side goes to 0 as g and l become arbitrarily small. Full cooperation is part of a sequential equilibrium for $\delta \in [\delta_1, \delta_2] \cap (0, 1)$. No other condition on the game parameters is needed to ensure existence of cooperative equilibrium. Figure 3 provides an illustration.



(a) $l = 5$

(b) $l = 15$

Figure 3: Discount factor lower bound δ_1 and upper bound δ_2 for $N = 40$

3.2 Cooperation under a generalized punishment threat

Under a grim punishment threat, cooperation is supported for $\delta \in [\delta_1, \delta_2] \cap (0, 1)$. The problem is that δ_2 is generally bounded away from 1 unless we make additional assumptions on the size of the sucker's payoff l. Here we demonstrate that we can partly solve this problem by reducing the severity of community punishment. Doing so allows us expand the set of discount factors to include values above δ_2, in a manner that reflects the intuition in Ellison (1994, Lemma 2).

Let players adopt a social norm based on the strategy in Definition 2, with $T \geq 1$.

Proposition 2. *There exists a superset $\mathcal{D} \supset [\delta_1, \delta_2]$, such that the strategy in Definition 2 supports full cooperation for all $\delta \in \mathcal{D} \cap (0, 1)$ and some $T \geq 1$.*

Proof. If players adopt a social norm based on the strategy in Definition 2, with $T \geq 1$, then this means that they will respond to a defection observed in period τ by punishing only in periods $\tau + T, \tau + 2T, \ldots$ (Figure 2). We can therefore study the problem faced by a player who is considering moving off equilibrium in period τ by splitting the continuation game into two pairwise disjoint supergames S_τ and its complement S_τ^C. The supergame S_τ^C includes all periods greater than τ where the strategy does not call for punishment, i.e., all periods outside of the punishment sequence

$$\boldsymbol{\tau} := (\tau + jT)_{j=1}^{\infty}.$$

The supergame S_τ includes only the periods specified by the punishment sequence $\boldsymbol{\tau}$ where players who are in the punishment mode never cooperate. This means that if we restrict attention to the supergame S_τ, players behave as if using grim punishment.

This last observation suggests that we can apply the same technique developed in Section 3 to define payoffs in and out of equilibrium in the supergame S_τ. The only adjustment we have to make is the discount factor, which is now δ^T due to the selector T used to pinned down the sequence through the frequency of punishment $1/T$. We provide this analysis in what follows.

Consider the supergame S_τ. Let $v_0(\boldsymbol{\tau})$ denote the full cooperation payoff and let $v_k(\boldsymbol{\tau})$ denote the payoff to a defector when there are k defectors. Given the discussion above, along the sequence $\boldsymbol{\tau}$ players behave as if they have adopted a social norm of grim punishment (as in Definition 1). If so, the equilibrium payoff function is $v_0(\boldsymbol{\tau}) = \dfrac{1}{1-\delta^T}$. The off-equilibrium payoff earlier defined in (5) implies

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

Using the previous analysis, moving off equilibrium on date τ is suboptimal if $v_1(\boldsymbol{\tau}) \leq v_0(\boldsymbol{\tau})$, which implies

$$\phi_1(\delta^T) \leq \frac{1}{1+g} = \phi_1(\delta_1).$$

The above inequality holds for all $\delta \geq \delta_1^{\frac{1}{T}}$ because ϕ_1 is a decreasing function of δ.

Similarly to what done in the proof of Proposition 1 punishing off equilibrium in the sequence $\boldsymbol{\tau}$ is optimal if $\tilde{v}_k(\boldsymbol{\tau}) \leq v_k(\boldsymbol{\tau})$ for all k, which implies

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

Notice that the above inequality is satisfied for all $\delta \leq \delta_2^{\frac{1}{T}}$.

For any given $T \geq 1$, consider $[\delta_1^{\frac{1}{T}}, \delta_2^{\frac{1}{T}}]$. This interval is non-empty (Lemma 3). Therefore, for any $\delta \in [\delta_1^{\frac{1}{T}}, \delta_2^{\frac{1}{T}}] \cap (0,1)$ there exists a fully cooperative equilibrium supported by the strategy in Definition 2. Letting $\mathcal{D} := \bigcup_{T=1}^{\infty} [\delta_1^{\frac{1}{T}}, \delta_2^{\frac{1}{T}}]$, we can support a fully cooperative equilibrium for all $\delta \in \mathcal{D} \cap (0,1)$ for an appropriately chosen punishment frequency $1/T$. □

Note that if the frequency $1/T$ is sufficiently small the set $\mathcal{D}$ includes discount factors arbitrarily close to 1. Hence, we may support cooperative equilibrium, independent of l, even if agents are very patient.

3.3 Cooperative equilibrium for arbitrarily patient players

An issue here is that the set $\mathcal{D}$ can be generally disconnected, so that it might not include the value taken by δ in the game. We now find a sufficient condition under which $\mathcal{D}$ is connected.

Lemma 4. *If $\delta_1^{\frac{1}{2}} \leq \delta_2$, then the set $\mathcal{D} = [\delta_1, 1]$.*

Proof. We claim that if $\delta_1^{\frac{1}{2}} \leq \delta_2$ then $\delta_1^{\frac{1}{T+1}} \leq \delta_2^{\frac{1}{T}}$ for any $T \geq 2$. In this case $\mathcal{D}$ is connected. To prove the claim suppose $\delta_1^{\frac{1}{T}} \leq \delta_2^{\frac{1}{T-1}}$; we want to show that $\delta_1^{\frac{1}{T+1}} \leq \delta_2^{\frac{1}{T}}$. We have

$$\delta_1^{\frac{1}{T}} \leq \delta_2^{\frac{1}{T-1}} \Rightarrow \delta_1^{\frac{1}{T}} \leq \delta_2^{\frac{T+1}{T^2}} \Rightarrow \delta_1^{\frac{1}{T+1}} \leq \delta_2^{\frac{1}{T}}$$

where the second inequality holds because $\frac{1}{T-1} > \frac{T+1}{T^2}$ for $T \geq 2$ and the third inequality holds because

$$\delta_1^{\frac{1}{T+1}} = (\delta_1^{\frac{1}{T}})^{\frac{T}{T+1}} \leq (\delta_2^{\frac{T+1}{T^2}})^{\frac{T}{T+1}} = \delta_2^{\frac{1}{T}}.$$

Noting that $\lim_{T\to\infty} \delta_2^{\frac{1}{T}} = 1$ concludes the proof. □

Given the results above we can finally prove Theorem 1,

Proof of Theorem 1. Suppose that for some values of the parameters of the game $[\delta_1, \delta_2] \cap [\delta_1^{\frac{1}{2}}, \delta_2^{\frac{1}{2}}] \neq \emptyset$, i.e., $\delta_1^{\frac{1}{2}} \leq \delta_2$. Then from the proof of Lemma 4 it follows that set $\mathcal{D}$ defined in the previous section is connected and $\mathcal{D} = [\delta_1, 1]$. Therefore, for any $\delta \in \mathcal{D} \cap (0,1)$ cooperative equilibrium can be sustained using the strategy in Definition 2 for some appropriate T. Moreover, since $[\delta_1^{\frac{1}{T}}, \delta_2^{\frac{1}{T}}] \cap [\delta_1^{\frac{1}{T+1}}, \delta_2^{\frac{1}{T+1}}] \neq \emptyset$ for all T in this case, we have that for $\delta \in [\delta_1^{\frac{1}{T}}, \delta_2^{\frac{1}{T}}] \cap [\delta_1^{\frac{1}{T+1}}, \delta_2^{\frac{1}{T+1}}]$ cooperative equilibrium can be sustained using strategies in Definition 2, for either T or $T+1$ (multiple strategies).

To show that $[\delta_1, \delta_2] \cap [\delta_1^{\frac{1}{2}}, \delta_2^{\frac{1}{2}}] \neq \emptyset$ we need $\delta_1^{\frac{1}{2}} \leq \delta_2$, since we already know $\delta_1 < \delta_2 \leq \delta_2^{\frac{1}{2}}$. If $l \geq N-2$ the nonempty condition always holds because $\delta_2 = 1$, see expression (14). If $0 < l < N-2$ then in order to ensure that $\delta_1^{\frac{1}{2}} \leq \delta_2$, we need δ_1 sufficiently small.

Using the definition of δ_1 and δ_2 the expression $\delta_1^{\frac{1}{2}} \leq \delta_2$ gives

$$\left[\phi_1^{-1}\left(\frac{1}{1+g}\right)\right]^{\frac{1}{2}} \leq \phi_1^{-1}\left(\frac{1-l/(N-2)}{1+g}\right),$$

and since $\lim_{x \to 1} \phi_1^{-1}(x) = 0$, this inequality holds for g sufficiently small.

Finally, using the definition of δ_1 and δ_2, we have $\lim_{g \to 0} \delta_1 = 0$ and $\lim_{g \to 0} \delta_2 > 0$, because $\phi_1^{-1}(x)$ is a decreasing continuous function of x that maps $(0,1)$ into $(0,1)$. This implies that as $g \to 0$ then $\mathcal{D} = (0,1)$ and, thus, for any value of the discount factor the cooperative equilibrium can be sustained by the generalized punishment in Definition 2, for some given T. □

4 Conclusions

Consider a group of anonymous individuals who face an indefinite sequence of Prisoner's Dilemma games in random pairs and under private monitoring. The theory of social norms developed in Kandori (1992) suggests that these strangers can attain the efficient, fully cooperative outcome by using a strategy that brings the entire group into a terminal state of full defection, if anyone acts uncooperatively. The incentive to follow such an indiscriminate and extreme sanction hinges on the size of the loss imposed by defectors on cooperators—the so-called "sucker's payoff." Ellison (1994) argues that if this payoff is small, then cooperative equilibrium cannot be sustained if players are too patient, but can be re-established if punishment is made infrequent. Intuitively, a patient individual may prefer to forgive a defection that causes a small temporary loss, and carry on cooperating, instead of participating in bringing about a grim state of full defection.

We have shown that players can attain full cooperation when a deviation corresponds to a minor infraction, i.e., both the temptation to defect as well as the damage it causes are sufficiently small. Here, for any value of the discount factor, the strategy ensures that players have an incentive to punish because punishment is moderate and does not fully destroy future cooperation. Moreover, we have shown that there could be multiple strategies, which support efficient play. Each of these strategies gives rise to a norm of punishment characterized by a specific frequency of defection and, consequently, various cooperation levels off-equilibrium.

References

Abreu, D., D. Pierce and E. Stacchetti. 1990. Toward a Theory of Discounted Repeated Games with Imperfect Monitoring. *Econometrica*, 58, 1041-1063.

Camera, G., Casari, M., and Bigoni, M. 2012. Cooperative strategies in anonymous economies: an experiment. *Games and Economic Behavior* 75, 570-586.

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

Dal Bó, P., and G. Fréchette, 2011. The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence. *American Economic Review*, 101(1), 411-429.

Ellison, G., 1994. Cooperation in the prisoner's dilemma with anonymous random matching. *Review of Economic Studies* 61, 567-88.

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

Fudenberg, D. and Maskin, E. 1986. The Folk Theorem in Repeated Games with Discounting or with Incomplete Information. *Econometrica* 50, 533-554.

Fudenberg, D., G.D. Rand, and A. Dreber. 2012. Slow to anger and fast to forgive: Cooperation in an uncertain world. *American Economic Review*, 102(2), 720-749.

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