

Long-term relationships as safeguards

Rafael Rob · Huanxing Yang

Received: 7 May 2007 / Accepted: 21 October 2008 / Published online: 12 November 2008
© Springer-Verlag 2008

Abstract We analyze a repeated prisoners' dilemma game played in a community setting with heterogeneous types. The setting is such that individuals choose whether to continue interacting with their present partner, or separate and seek a new partner. Players' types are not directly observed, but may be imperfectly inferred from observed behavior. We focus on a class of equilibria that satisfy zero tolerance and fresh start. We find that the punishment for defecting and the reward for cooperating are driven by the formation and the dissolution of long-term, high-paying relationships: an individual who defects, aborts a long-term relationship that he is in, or that he might have entered into, is thrown into short-term interactions with individuals who are likely to defect and, consequently, receives low payoffs. On the flip side, an individual who cooperates, enters into or prolongs a long-term interaction with a partner who cooperates and, consequently, receives high payoffs.

Keywords Community games · Information flows · Heterogeneity of types · Long-term relationships

JEL Classification C73 · C78 · D82

We thank the co-editor, two anonymous referees, and the seminar participants at Indiana, Maryland, Penn, Pittsburgh, and the Decentralization Conference (Purdue, 2003) for helpful comments and suggestions. Rafael Rob acknowledges NSF support under grant number 01-36922.

R. Rob
Department of Economics, University of Pennsylvania, Philadelphia, PA 19104, USA
e-mail: rrob@ssc.upenn.edu

H. Yang (✉)
Department of Economics, The Ohio State University, Columbus, OH 43210, USA
e-mail: yang.1041@osu.edu

1 Introduction

In this paper, we study a prisoners' dilemma game in which individuals interact with varying opponents over time. The environment is such that individuals have partial control over who to interact with. Namely, individuals choose—based on their past experience—whether to prolong the interaction with their present partner for another period, or seek a new partner. In addition, the population of individuals is heterogeneous in that some individuals are strategic players, i.e., they choose actions based on incentives, while others are behavioral types, i.e., are programmed to take a fixed action. Our aim is to study the structure of incentives for a certain class of equilibria in this environment and, in particular, to determine the impact of the “demographics” of the population, i.e., the effect that the type distribution has on equilibrium behavior.

The equilibria we identify have the feature that behavior and, hence, payoffs depend on whether a player is in one of two states. One state is that a player is in an ongoing relationship, i.e., she has interacted with her current partner at least once. The other state is that a player is in a new relationship. Since players are able to condition their behavior on state, being in an ongoing relationship might deliver higher payoffs than being in a new relationship because players cooperate in the former, but defect in the latter. This possibility dictates the structure of equilibrium incentives. In particular, a player in an ongoing relationship, who is “scheduled” to cooperate, has an incentive to do so because, otherwise, the relationship he is in will be terminated and he will be forced to interact with players who are likely to defect and, thereby, inflict low payoffs on him. In addition, a player who is in a new relationship might, nonetheless, cooperate because if he is matched to another player who cooperates, he will enter into a long-term, high-paying relationship. Therefore, cooperation in this setting is a form of investment aimed at creating or maintaining a high-paying status. Our goal is to explore this logic and, more specifically, to pin down the conditions under which this force is sufficient to guarantee that the equilibrium in which strategic players always cooperate—called the good equilibrium—exists and, analogously, to pin down the conditions under which other (pure and mixed strategy) equilibria exist.

One of our main findings is that the fraction of bad-type players (behavioral type that chronically defects) has to be in an intermediate range to sustain the good equilibrium. The reason for this is that if a strategic player causes a long-term relationship to terminate by defecting and if the fraction of bad-type players is sufficiently small, the defector is likely to meet another strategic type fairly quickly and, hence, bounce back to another long-term, high-paying relationship, so he is not made to pay for the infraction. On the other hand, if the fraction of bad-type players is sufficiently large, the time it takes to form a long-term relationship is excessively long, and the cost incurred in the process is excessively high, so strategic players do not try to form such relationships, i.e., they simply defect. A similar principle applies to the bad equilibrium (strategic players always defect), which is shown *not* to exist, if the fraction of good-type players (behavioral type that always cooperates) is in some intermediate range. These results indicate that the type distribution has indeed an impact on the equilibrium behavior. Another result we report is that the problem that the good equilibrium does not exist because the fraction of bad-type

players is too low can be rectified if we allow mixed strategies. Namely, additional bad-type players can be “created” endogenously if some of the strategic players defect.

Although this paper is intended as a theoretical exploration, the forces we identify here are of some relevance in the real world, where long-term relationships are abundant and the length of relationships usually affects the terms of trade. In employment relationships, longer-tenured workers are paid higher wages; in lending markets, borrowers with longer relationships with a bank pay lower interest rates and are less likely to pledge collateral; similar pattern holds for buyer–supplier relationships. A more extensive discussion of this type of institutions in the real world that operate in various contexts may be found in papers by [Johnson et al. \(2002\)](#); [Kali \(1999\)](#); [Kranton \(1996\)](#); [Taylor \(2000\)](#); [Dal Bo \(2007\)](#), and [Yang \(2007\)](#).

Brief literature review This paper relates to several strands of the literature. The first strand is repeated games in a community setting, pioneering papers in this literature being [Kandori \(1992\)](#) and [Ellison \(1994\)](#). Our point of departure from that literature is that we incorporate the decision whether to keep interacting with the same partner, and the heterogeneity of types. More relevant to our setting are the papers by [Datta \(1993\)](#) and [Ghosh and Ray \(1996\)](#), who develop and analyze the “building trust” mechanism. We depart from that literature in that we analyze a wider class of equilibria, fully characterize them, and elucidate on the role of the heterogeneity of types in sustaining good equilibria or eliminating bad equilibria (by contrast, heterogeneity in Ghosh and Ray is used to select an equilibrium, using the criterion that bilateral deviation is not profitable).

A third strand of literature is the literature on reputation, which in a context similar to ours is found in papers by [Watson \(1999, 2002\)](#). These papers study a fixed relationship between two players and, as such, do not consider the possibility of endogenously forming long-term relationships and the impact of the demographics on equilibria. Another relevant paper is [Sobel \(2006\)](#). He focuses, however, on the role of labor market aspects, and does not consider the heterogeneity of types. Another paper that focuses on labor market issues, relating to racial discrimination, is [Eeckhout \(2006\)](#). He does not study, however, the disciplinary role of the heterogeneity of types. Other related papers include [Tirole \(1996\)](#) and [Dixit \(2003\)](#), who study the role of information intermediaries that make information available to players. They, again, do not study the disciplinary role of endogenizing relationships. A recent paper by [Okuno and Fujiwara \(2006\)](#) studied a similar formulation to ours, but from an evolutionary perspective and without heterogeneity of types. [Frank \(1988\)](#) studied a population which consists of one group of agents who always cooperates and another group always defects. When players are paired to play a prisoner’s dilemma game, they have the option to opt out. The difference is that, in his model different group members give off different signals that are informative, while in our model such signals do not exist and players infer their opponents’ types only by the observed actions.¹

¹ A survey of this literature with types of players being observed (or partially observed) can be found in [Sobel \(2005\)](#).

More broadly, our paper relates to three major themes in economic theory. One theme is that one may sustain cooperation in long-run via promises and threats (see [Fudenberg and Maskin 1986](#)). Here we offer a different enforcement mechanism, namely, where no player is specifically called upon to inflict a punishment. Rather, punishment is inherent in the fact that it takes time to rebuild a relationship, during which time costs are incurred. This idea brings us to the second theme, which is the efficiency wage literature (see [Shapiro 1984](#)). In that literature a shirking worker is punished by being unemployed, which is similar to the idea that a player who defects is forced to interact with players who chronically defect. That literature, however, is couched in a competitive (as opposed to a game theoretic) setting, and does not consider the heterogeneity of types. The third theme is the search and matching literature à la [Diamond \(1982\)](#) and [Mortensen \(1982\)](#). In that literature, like here, it takes time to be matched with an acceptable type. On the other hand, that literature does not study the strategic interactions between agents once they are matched.

The rest of the paper is organized as follows. Section 2 introduces our framework. In Sect. 3 we determine when the good equilibrium, in which strategic players always cooperate, exists and how it depends on parameters. In Sect. 4 we do the same thing with respect to the bad equilibrium in which strategic players defect. Section 5 studies mixed strategy equilibria. In Sect. 6 we classify all equilibria, and relate them to parameter values. Section 7 relates social welfare to the heterogeneity of types and Sect. 8 concludes. Some proofs are found in the Appendix, while others are found in a working paper version ([Rob and Yang 2005](#)).

2 Model formulation

The environment We consider a community of individuals (or players or agents), modeled as a continuum of measure 1. Time is discrete and the horizon is infinite. Each individual is infinitely lived.

At the beginning of each period, the community is divided into partnerships (or relationships). Then, the following sequence of events occurs. First, each pair of partners plays a prisoners' dilemma game, and each partner chooses either C , which stands for "cooperate," or D , which stands for "defect." The payoff matrix of this game is specified momentarily. Second, after playing this game, each partnership persists with probability ρ , and breaks up exogenously with probability $1 - \rho$. Third, if a partnership persists, then two partners simultaneously decide whether to stay or leave the current relationship. The current partnership continues into the next period if and only if both partners choose to stay.² Finally, all individuals in dissolved partnerships (due to either exogenous or endogenous separation) enter into the unmatched pool, and they are randomly matched at the beginning of the next period. Since there is a countable number of time periods and a continuum of players, no player is ever matched with one of his ex-partners.

² No direct payoffs are associated with the decisions of staying or leaving; their only role is to endogenize the decision whether to interact with the same individual in the next period.

Table 1 Payoff matrix of a type *O* player

	<i>C</i>	<i>D</i>
<i>C</i>	<i>a</i>	$-l$
<i>D</i>	<i>b</i>	0

There are three types of players in the population. There is a measure α of opportunistic (*O*) type players, a measure β of bad (*B*) type players, and a measure γ ($= 1 - \alpha - \beta$) of good (*G*) type players. A type *G* player always chooses *C* in the prisoners' dilemma game, and a type *B* player always chooses *D*. A type *O* player chooses either *C* or *D*, depending on which gives her a higher payoff (which depends on the equilibrium play). The payoff matrix of a type *O* player, considered as a row player, is shown in Table 1. The payoff matrix of a type *G* player is the *C* row of Table 1, and the payoff of a type *B* player is the *D* row.

We assume $0 < a < b, 0 < l$, and $b - l < 2a$. The first two restrictions say that this game, when played by two type *O* players, is a prisoners' dilemma game. The third restriction says that the action profile (*C, C*) maximizes the sum of players' payoffs when the game is played between two type *O* players. The objective of all players is to maximize the discounted sum of payoffs. The discount factor is common to all players and is denoted by δ , where $\delta \in (0, 1)$.

We assume that monitoring is perfect inside each partnership: a player observes his partner's actions—beginning with the date at which this partnership is commenced. However, when a player is matched to a new partner he knows nothing about the partner's past history of actions with other partners. That is, there are no information flows across matches. Also, a player's type is private information. However, players make statistical inferences about types (of other players), based on the actions they observe. In particular, a player observed to choose *C* is known not to be a type *B* player, and a player observed to choose *D* is known not to be a type *G* player. Finally, we assume that the configuration of types, (α, β, γ) , is common knowledge.

Steady-state equilibria In this paper, we focus on a particular class of equilibria, delineated by three properties. The first property is “fresh start”: a player's behavior in a new relationship is independent of his past history. The second property is “zero tolerance”: when a player encounters *D*, he immediately separates from his partner. The third property is “quick familiarity”: a player's behavior within a relationship depends only on whether the partnership has just started, or whether it is an ongoing relationship.

The property of fresh start is in contrast to Kandori's (1992) contagious equilibrium, in which players play grim trigger strategies against the entire population. Note that in Kandori's model players cannot form long-term relationships, as at the beginning of each period the entire population is randomly matched again. In that world, playing grim trigger strategies against the entire population is necessary to induce cooperation, since it is the only way to punish the defectors. In our model players have the potential to form long-term relationships, and this makes grim trigger strategy against the entire population not necessary to induce cooperation. Since players can partially seek whom to interact with, defectors will be punished by interacting with defectors. Actually,

the property of fresh start makes our equilibrium stable compared to the contagious equilibrium (see the discussion in Sect. 3).³

The property of zero tolerance provides the strongest punishment to defectors. Given that partners' actions are perfectly observed, under zero tolerance maximum cooperation can be induced and cooperative long-term relationships will not endogenously break up. However, if partners' actions can only be observed with noise (imperfect monitoring), then zero tolerance might cause cooperative long-term relationships to be broken up prematurely. We conjecture that long-term relationships should allow for some degree of tolerance under imperfect monitoring. The property of quick familiarity is mainly for the elegance of results. More generally, we can allow the first $T + 1$ periods being the "getting familiar phase," with opportunists playing (D, D) and no endogenous separation in the first T periods. The qualitative results of the paper remain the same under this alternative setting.

Another reason for us to focus on this class of equilibria is that the model analysis under these three properties is quite tractable.⁴ Given this, our aim is to provide a complete characterization of equilibria that satisfy these three properties. To simplify the analysis, we further assume that the distribution of types in each phase (defined more precisely below) has settled to a steady state at $t = 0$. For brevity, we call this class of equilibria steady-state equilibria.

Objective of analysis Having delineated the game and the class of equilibria we focus on, we proceed to analyze them. Specifically, for any configuration of parameter values [i.e., some $(a, b, l, \delta, \rho, \alpha, \beta, \gamma)$ -tuple] we determine whether an equilibrium exists, what type of behavior it manifests, and whether it is unique. To this end, we note that some aspects of agents' behavior are already "hard-wired" into our setting. In particular, G and B type players are hard-wired to play C and D , respectively. In addition, we already specified that all player types separate from their current partners if they encounter D (and this behavior is optimal because it gives them a chance to interact with players who play C , which generates higher payoffs).⁵ Given this, the only aspect of behavior that remains to be endogenously determined is the behavior of type O players.

To ease exposition, we use the following terminology: if two partners are about to interact for the first time, we say that they are in the *stranger phase*, denoted S , whereas if they have previously interacted, we say they are in the *friendly phase*, denoted F .⁶ Also, we call a mapping from phases to actions a *behavior pattern*.

³ Another reason we adopt fresh start is that it seems more realistic. When someone is cheated by his partner, it is unlikely that he will revenge (cheat) all his future partners; instead, what he usually does is to break up with his current unfaithful partner, and start a new relationship (with new partners) with good faith, i.e., having a fresh start.

⁴ Note that the equilibria we derive are such that if all agents adopt strategies that satisfy these properties, the remaining agent's *unconstrained* best response is to adopt a strategy that satisfies them as well.

⁵ Type G players always choose C , yet they separate if they encounter D , because they get a higher payoff if they play against an opponent that chooses C as opposed to an opponent that chooses D . In separating, therefore, type G players are seeking future partners that bestow higher payoffs on them, which is driven by discounted payoff maximization.

⁶ The terminology is borrowed from Ghosh and Ray (1996).

3 The good equilibrium

In this section, we analyze a pure strategy equilibrium, referred to as the good equilibrium, in which the behavior pattern of type O players is to play C in both phase S and phase F . That is, type O players behave exactly like type G players.

Steady state This behavior pattern, along with the zero tolerance property, induce a steady-state over the measure of agents in phase S , and its composition. To determine this steady state, note that all type B players are always in phase S . In addition, due to exogenous separation a certain measure of type G and type O players, henceforth called non-bad-type players, are also in phase S . Let $x \in [0, 1 - \beta]$ be the measure of non-bad-type players in phase S . Then, the overall measure of agents in phase S is $x + \beta$, and the overall measure of agents in phase F is $1 - x - \beta$. In the steady state of the good equilibrium x satisfies

$$(1 - \rho)(1 - x - \beta) = x\rho \frac{x}{x + \beta}. \tag{1}$$

To understand (1), note that its left-hand side (LHS) is the measure of agents flowing from phase F into phase S in each period. This “inflow” is simply the probability of exogenous dissolutions, $1 - \rho$, times the measure of agents in phase F , $1 - x - \beta$. The right-hand side (RHS) of (1) is the measure of agents flowing from phase S to phase F in each period. This “outflow” is the product of x , which is the measure of non-bad agents in phase S , the probability that one of these agents is matched with another non-bad agent, which is $\frac{x}{x + \beta}$, and the probability, ρ , that such a match is not exogenously dissolved after the first interaction. In a steady state the inflow equals the outflow. It can be verified that there is a unique $x \in [0, 1 - \beta]$ that solves (1).

As (1) shows, x depends on β and ρ , but, since the ensuing analysis focuses mostly on the role of β , we consider x as a function of β only, writing it as $x = X(\beta)$. Given $X(\beta)$ and β we define the variable $y = Y(\beta) \equiv \beta/X(\beta)$, which reflects the composition of type B players versus non-bad-type players in phase S . Given the behavior pattern we focus on, y also reflects the composition of *behavior* in phase S , i.e., the ratio of the measure of agents choosing D to the measure of those choosing C . We next state a simple but useful property of $Y(\beta)$.

Lemma 1 $Y(\beta)$ is increasing in β , ranging from zero to infinity, as β varies from 0 to 1.

Proof See the Appendix. □

Value functions Given the behavior pattern prescribed by the good equilibrium and given the steady state corresponding to it, we define the value functions for type O players. Let V_F and V_S be the discounted payoffs in phases F and S , respectively. Let V_F^d be the discounted payoff when in phase F , deviating to D , and returning to prescribed behavior (i.e., C) thereafter, a one-shot deviation. And let V_S^d be the discounted payoff to a one-shot deviation when in phase S . The equations defining these values are:

$$V_F = a + \delta[\rho V_F + (1 - \rho)V_S] \tag{2}$$

$$V_S = \frac{x}{x + \beta} \{a + \delta[\rho V_F + (1 - \rho)V_S]\} + \frac{\beta}{x + \beta}(-l + \delta V_S) \tag{3}$$

$$V_F^d = b + \delta V_S \tag{4}$$

$$V_S^d = \frac{x}{x + \beta}(b + \delta V_S) + \frac{\beta}{x + \beta}(0 + \delta V_S). \tag{5}$$

To understand how these equations are formed, consider the RHS of (2). The payoff V_F is the sum of two terms: the period payoff a (since all agents in phase F play C), and the continuation payoff: with probability ρ the partnership continues and a type O player gets δV_F ; with probability $1 - \rho$ the partnership dissolves and a type O player gets δV_S . The remaining three equations are based on a similar logic.

Incentive constraints Now we determine the conditions under which type O players have the incentive to play C in each period. For that, the following two incentive constraints must be satisfied:

$$\text{No deviation in phase } F : 0 \leq V_F - V_F^d; \tag{6}$$

$$\text{No deviation in phase } S : 0 \leq V_S - V_S^d. \tag{7}$$

Analysis of these incentive constraints gives the first result.

Lemma 2 (i) *Inequality (6) is redundant if (7) is satisfied.* (ii) *The good equilibrium exists if, and only if,*

$$b - a \leq \frac{\beta}{x + \beta} \delta \rho b - \frac{\beta}{x} (1 - \delta \rho) l. \tag{8}$$

Proof (i) From (4) and (5), we have

$$V_S^d = \frac{x}{x + \beta} V_F^d + \frac{\beta}{x + \beta} \delta V_S.$$

Subtracting this last equation from (3), we get

$$0 \leq V_S - V_S^d \Leftrightarrow 0 \leq \frac{x}{x + \beta} (V_F - V_F^d) - \frac{\beta}{x + \beta} l.$$

Since $0 < l$, the above result shows that (7) implies (6).

(ii) Subtracting (5) from (3), we get

$$V_S - V_S^d = \frac{x}{x + \beta} (-b + V_F - \delta V_S) - \frac{\beta}{x + \beta} l.$$

From (2) we have

$$V_F - \delta V_S = a + \delta \rho (V_F - V_S).$$

Substituting the last equation into the one just before it, we get

$$0 \leq V_S - V_S^d \Leftrightarrow (b - a) + \frac{\beta l}{x} \leq \delta\rho(V_F - V_S). \tag{9}$$

Solving for $V_F - V_S$ from (2) and (3) and substituting the result into (9), we obtain (8). □

Lemma 2 indicates that it is “safer” to play C in phase F than in phase S . Indeed, in phase F a type O player is sure to encounter C from her partner, resulting in a payoff of a , while in phase S she may encounter D , resulting in a payoff of $-l$. Therefore, if it pays to play C in phase S , it certainly pays to play C in phase F . The inequality (9) ensures that type O players have the incentive to play C in phase S . The LHS of (9) is the current period loss of playing C , while the RHS is the long-run reward for playing C , which is proportional to $V_F - V_S$. Therefore, for the good equilibrium to exist the difference between V_F and V_S has to be big enough.

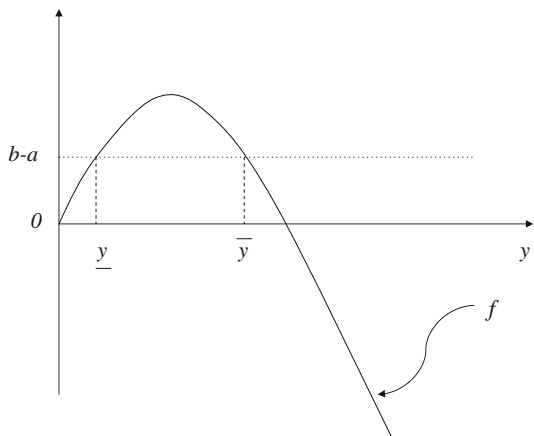
Existence of the good equilibrium Inspection of condition (8) reveals that it depends on all parameter values. We wish to isolate the role that the heterogeneity of types plays, i.e., the role that (α, β, γ) plays regarding to the existence of the good equilibrium. To this end, we use the definition $y \equiv \frac{\beta}{x}$ to rewrite (8) as

$$b - a \leq \frac{y}{1 + y} \delta\rho b - y(1 - \delta\rho)l \equiv f(y). \tag{10}$$

We define the RHS of (10) as $f(y)$, since it will be used frequently in the analysis. Figure 1 shows one possibility for what the graph of f looks like.

Inspecting (10) we see that its LHS, $b - a$, is positive and independent of y . On the other hand, its RHS is strictly concave in y , goes to 0 as y goes to 0, and goes to $-\infty$ as y goes to ∞ (see Fig. 1). Also, f is uniquely maximized at

Fig. 1 The graph of function f



$$y^* = \sqrt{\frac{\delta\rho b}{(1 - \delta\rho)l}} - 1.$$

Consequently, for the good equilibrium to exist, two conditions must hold: $0 < y^*$, and $b - a \leq f(y^*)$. The first condition is necessary because, if $y^* \leq 0$, then f is strictly decreasing and $f(y) \leq 0$ for all $0 \leq y$, so obviously there is no $0 \leq y$ for which $0 < b - a \leq f(y)$. The second condition is necessary because, if the inequality were reversed, $f(y^*) < b - a$, there would again no $y \geq 0$ for which $b - a \leq f(y)$. After some manipulations, we eliminate the endogenous variable y , and write the two conditions in terms of model primitives only:

$$(1 - \delta\rho)l \leq \delta\rho b \quad \text{and} \quad 4\delta\rho b(1 - \delta\rho)l \leq [a + (1 - \delta\rho)(l - b)]^2. \quad (11)$$

This analysis shows that (11) is a necessary condition for the existence of the good equilibrium. Condition (11) is also sufficient. Indeed, if (11) is satisfied, then, as shown in Fig. 1, there is an interval of y 's, call it $[y, \bar{y}]$, where (10) holds and the good equilibrium exists. The cutoffs y and \bar{y} are the small and the large roots of the equation $f(y) = b - a$, which are independent of (α, β, γ) (because f is). Since, as per Lemma 1, y is strictly increasing in β , $y \in [y, \bar{y}]$ is equivalent to $\beta \in [\underline{\beta}, \bar{\beta}]$, where $\underline{\beta}$ is defined by $y = Y(\underline{\beta})$, and $\bar{\beta}$ is defined by $\bar{y} = Y(\bar{\beta})$. Moreover, $[\underline{\beta}, \bar{\beta}]$ does not include 0 or 1. This is because when $\beta = 0$, $y = 0$, and $f(0) = 0$. And, when $\beta = 1$, $y = \infty$, and $f(\infty) = -\infty$. Either way, (10) does not hold. Finally, observe that criterion (10) is independent of γ . Summarizing our analysis, we have the following result.

Proposition 1 *Hold all parameter values other than (α, β, γ) constant. Then: (i) The existence of the good equilibrium does not hinge on γ . (ii) If (11) is not satisfied, then there is no β for which the good equilibrium exists. (iii) If (11) is satisfied, then the good equilibrium exists if, and only if, $\beta \in [\underline{\beta}, \bar{\beta}]$, where $0 < \underline{\beta} < \bar{\beta} < 1$, $\underline{\beta}$ and $\bar{\beta}$ being the roots of $f(Y(\beta)) = b - a$.*

The main insight from Proposition 1 is that for the good equilibrium to exist the measure, β , of type B players must not be too small or too large. If β is too small, say $\beta = 0$, the fraction of type B players in phase S is zero, which implies that behavior (under the hypothesized equilibrium strategy) in phase S is the same as behavior in phase F . But, then, there is no punishment for playing D , and no reward for playing C . If a type O player chooses D in phase F , he goes into phase S , where he encounters the same behavior he encountered in phase F , and receives the same payoff, which means he is not punished. For the same reason, there is no reward to play C in phase S . Therefore, if $\beta = 0$, $V_F = V_S$ and the good equilibrium unravels. At the other end of the spectrum, if the measure of type B players is too large, the probability of being matched with a non-bad type in phase S is next to nil, which destroys the incentive to play C , and the good equilibrium unravels again. Only if the proportion of type B players is in some intermediate range, not too small to reduce the effectiveness of

punishment in phase F , and not too large to discourage cooperation in phase S , does the good equilibrium exist.⁷

Another way to think about the structure of incentives in the good equilibrium is as follows. Due to endogenous formation of long-term relationships, phase S is “contaminated” by a disproportionately large measure of the bad-type players because they never leave this phase. But this induces type O players to choose C , because choosing D means going to (or staying at) phase S , interacting with bad-type players with a non-negligible probability, and receiving low payoffs. Without a critical mass of type B players this threat is not strong enough, and the good equilibrium does not exist.⁸

Stability Now we compare the stability of the good equilibrium to that of the contagious equilibrium à la [Kandori \(1992\)](#). (In the contagious equilibrium, a player defects forever following a defection by himself or by one of his partners.)

So far, we have assumed that monitoring is perfect within a relationship. Consider now the possibility of observational errors: a player observes her partner to play D (C) with probability $\varepsilon > 0$, even though the partner actually chose C (D). Then, no matter how small ε is, an observational error eventually occurs, i.e., some player is erroneously observed to play D . Once that happens, a contagious process is set in motion under the contagious equilibrium, whereby more and more players defect, so cooperation in the community breaks down. By contrast, consider the good equilibrium in our setting. This equilibrium continues to exist under the presence of observational errors—for conditions analogous to (11), and as long as ε is small enough (one has to appropriately modify the steady-state condition and the incentive constraints to account for the observational errors). More importantly, cooperation does not break down *globally* in this equilibrium. Intuitively, in the good equilibrium an agent that observes his partner playing D separates from the partner, and both get a fresh start in a new relationship next period. In these new relationships, each player ignores the past and expects (rationally) that playing C bears a chance of being rewarded in the future. Thus, the effect of an observational error is local; it does not trigger the spread of uncooperative behavior, and has no effect on global behavior in the community. This difference between the good equilibrium and the contagious equilibrium comes from the fact that we endogenize separations and restart of relationships, which is exactly what ‘contains’ the impact of observational errors.

Let us mention at this juncture that [Ellison \(1994\)](#) proposed—within the context of the contagious equilibrium—a different way to contain the spread of defection. In Ellison’s framework the contagious equilibrium is made resilient if players have access to a public randomization device. Such device allows the severity of punishments

⁷ If there is a “getting acquainted” phase consisting of T periods of playing (D, D) (without endogenous separation) before the stranger (or testing) phase starts, then one can show that the good equilibrium exists under a smaller set of parameter values compared to the setting that there is no such a phase. The reason is that with T periods of the “getting acquainted” phase, the reward for playing C at the stranger phase ($T + 1$) becomes smaller, because although the average amount of time spent in the friendly phase is the same, this is followed by a longer stretch of time in which one’s opponents are playing D .

⁸ The result that the existence of the good equilibrium does not depend on the measure of type G players is that type G players and type O players behave alike in the good equilibrium.

to be adjusted and coordinated based on the outcome of a device that everyone in the community can perfectly observe. By contrast, such device is not necessary in our framework. Instead, the threat of terminating a relationship and the consequent interaction with bad-type players are sufficient to enforce cooperative behavior.

Comparative statistics Since Inequality (8) determines the existence of the good equilibrium, one can readily use it to derive comparative statics results. It can be shown that an increase in δ increases the RHS of (8), which means that the good equilibrium exists under a wider set of circumstances. This is consistent with the result in the literature of repeated games with imperfect monitoring (Abreu et al. 1990) that the set of equilibrium payoffs expands as δ increases. On the other hand, the effect of the persistence probability, ρ , is not so conventional, and is, in fact, non-monotonic. In one sense, an increase in ρ , “should be” equivalent to an increase in δ because it prolongs the longevity of relationships and, as such, should have a positive effect. What we find, instead (under a mild extra restriction), is that the effect is non-monotonic.

Proposition 2 *Assume $(1 - \delta)(a + l) < b < \frac{a+l}{1-\delta}$,⁹ and a good equilibrium exists for some value of ρ . Then, there exist a $\underline{\rho}$ and a $\bar{\rho} \in (0, 1)$, where $\underline{\rho} < \bar{\rho}$, so that the good equilibrium exists if, and only if, $\rho \in [\underline{\rho}, \bar{\rho}]$.*

Proof See the working paper. □

The intuition is that an increase in ρ has two effects. The first effect is what we mentioned earlier: an increase in ρ prolongs the expected amount of time spent in phase F and, thus, makes it more rewarding to play C in phase S . The second effect is that an increase in ρ reduces the measure of non-bad-type players in phase S . As a result, a type O player is less likely to be matched with a non-bad type in phase S , which makes it less rewarding to play C in that phase. These two effects work in opposite directions. It turns out that when ρ is small the first effect dominates, whereas when ρ is large the second effect dominates. Thus, in a community setting, a small possibility of exogenous turnover ($1 - \rho$) may help, rather than hinder, cooperation. Another way to look at this is that turnover introduces “fluidity” into the system,¹⁰ enabling movements from phase S to phase F and, thereby, generating incentives to play C in phase S .

4 The bad equilibrium

In this and the next section, we expand our results to other steady-state equilibria. To start with, we study a pure strategy equilibrium, that we call the bad equilibrium, in which type O players play D in phase S . Given zero tolerance, type B and type O players, henceforth called non-good-type players, are always in phase S . On top of those there is a certain measure of type G players in phase S because of exogenous dissolutions. Let $x \in [0, \gamma]$ be the measure of type G in phase S . Then, the steady-state condition corresponding to the bad equilibrium is

⁹ This assumption is satisfied if δ is large enough or if $b = a + l$.

¹⁰ When $\rho = 1$ agents in phase S are “stuck” in phase S , so there is no long-term reward for playing C .

$$(1 - \rho)(\gamma - x) = x\rho \frac{x}{x + 1 - \gamma}. \tag{12}$$

Analogous to (1), the solution to (12) determines x as a function of γ , which we continue to call $X(\gamma)$. Likewise, we let the ratio of non-good-type players to good-type players in phase S be $y = Y(\gamma) \equiv \frac{1-\gamma}{X(\gamma)}$, which, as before, is also the ratio of the measure of agents choosing D to the measure of those choosing C in phase S . Similar to the good equilibrium, Y is strictly decreasing in γ , approaches 0 as γ goes to 1, and approaches ∞ as γ goes to 0.

Since the hypothesized behavior pattern of type O players here is such that they play D in phase S , they are never in phase F . Nevertheless, to check whether this strategy is part of an equilibrium, the choice in phase F has to be specified. Obviously, there are two possible specifications: either play D , or play C in phase F . We analyze these two cases in turn.

Case 1 Type O players play D in phase F

We first define value functions. The notation is similar to that of the previous section, except that the hypothesized behavior pattern in the bad equilibrium is different. Making the requisite adjustments, the new value functions are:

$$V_F = b + \delta V_S \tag{13}$$

$$V_S = \frac{x}{x + 1 - \gamma} b + \delta V_S \tag{14}$$

$$V_F^d = a + \delta[\rho V_F + (1 - \rho)V_S] \tag{15}$$

$$V_S^d = \frac{x}{x + 1 - \gamma} \{a + \delta[\rho V_F + (1 - \rho)V_S]\} + \frac{1 - \gamma}{x + 1 - \gamma} (-l + \delta V_S). \tag{16}$$

Given these value functions, the incentive constraints are:

$$\text{No deviation in phase } F : 0 \leq V_F - V_F^d; \tag{17}$$

$$\text{No deviation in phase } S : 0 \leq V_S - V_S^d. \tag{18}$$

Analyzing these constraints, we have the following result (a proof is found in the working paper version).

Lemma 3 (i) *Inequality (18) is redundant if (17) is satisfied.* (ii) *A bad equilibrium in which type O players defect in phase F exists if, and only if,*

$$\frac{1 - \gamma}{x + 1 - \gamma} \delta \rho b \leq b - a. \tag{19}$$

Although Lemma 3 is the analogue of Lemma 2, two differences should be noted. First, the binding incentive constraint here is in phase F , not in phase S . Second, $b - a$ has to be bigger, not smaller, than some threshold value. This is due to the fact that in the bad equilibrium opportunists are supposed to defect, not cooperate.

Case 2 Type O players play C in phase F

We carry out similar analysis as in the previous case. For brevity, we just report the result (a proof is found in the working paper version).

Lemma 4 *A bad equilibrium in which type O players play C in phase F exists if, and only if,*

$$\frac{1-\gamma}{x+1-\gamma}\delta\rho b - \frac{1-\gamma}{x}(1-\delta\rho)l \leq b-a \leq \frac{1-\gamma}{x+1-\gamma}\delta b. \quad (20)$$

Unlike in Lemmas 2 and 3, no deviation in phase F does not imply no deviation in phase S , and no deviation in phase S does not imply no deviation in phase F . That is why two inequalities (rather than one) have to be satisfied in condition (20).

Combining Lemmas 3 and 4, we see that a bad equilibrium exists if, and only if,

$$\frac{1-\gamma}{x+1-\gamma}\delta\rho b - \frac{1-\gamma}{x}(1-\delta\rho)l \leq b-a. \quad (21)$$

Existence of the bad equilibrium As we did with the good equilibrium, we transform condition (21) to a condition that involves only the primitive data. To this end we rewrite the LHS of (21) in terms of y :

$$\frac{y}{1+y}\delta\rho b - y(1-\delta\rho)l \leq b-a. \quad (22)$$

As can be readily seen, (22) is similar to (10), with $1-\gamma$ replacing β and reversing the inequality. Thus, following the analysis leading up to Proposition 1, we derive the following result.

Proposition 3 *Hold all parameter values other than (α, β, γ) constant. Then: (i) The existence of the bad equilibrium does not hinge on β . (ii) If (11) is not satisfied, then the bad equilibrium exists for any γ . (iii) If (11) is satisfied, then the bad equilibrium exists if, and only if, $\gamma \in [0, \underline{\gamma}] \cup [\bar{\gamma}, 1]$, where $\underline{\gamma}$ and $\bar{\gamma}$ are found by solving $f(Y(\gamma)) = b-a$, and are such that $0 < \underline{\gamma} < \bar{\gamma} < 1$.*

Although Proposition 3 is analogous to Proposition 1, one feature of it merits discussion and comparison to the traditional theory of repeated games. Namely, Proposition 3 shows that the bad equilibrium does not exist for some parameter configurations. This contrasts with the theory of repeated games, where an indefinite repetition of a Nash equilibrium (the bad equilibrium in our context) is the simplest equilibrium to construct. This is still true in our context if we consider a community setting with good-type players, but without endogenously formed long-term relationships. Therefore, Proposition 3 shows that with endogenously formed relationships, a new force comes into play: an opportunist may cooperate in phase S in the hope of hooking up with a good type, entering into phase F , and enjoying high future payoffs. Therefore, having good-type players and the possibility of forming long-term relationships may

destroy the bad equilibrium. Proposition 3 pins down the set of circumstances under which this force is sufficiently strong that the bad equilibrium does not exist.

To be more specific about this set of circumstances, Proposition 3 shows that the bad equilibrium does not exist if γ is in some intermediate range. If γ is small, all opportunists playing D in phase S is an equilibrium because the probability of meeting a good type is too small. If γ is big, all opportunists playing D in phase S is again an equilibrium, since the difference between the continuation payoffs in phase F and phase S is too small. Thus, in both cases the bad equilibrium exists. However, if γ is in some intermediate range, opportunists in phase S have a reasonable chance of meeting a good type, and opportunists in phase F enjoy a significantly higher continuation payoff than in phase S , so they cooperate. Thus, the bad equilibrium does not exist when γ is in this range.

A convenient feature of Propositions 1 and 3 that we are going to exploit later is that there is a duality between the existence of the good equilibrium and the non-existence of the bad equilibrium. The incentive for an opportunist to cooperate in phase S (which is what it means for the good equilibrium to exist, or for the bad equilibrium not to exist) depends on the proportion of agents cooperating in that phase. Since this proportion is strictly decreasing in β in the good equilibrium and strictly increasing in γ in the bad equilibrium, there is a duality between β and γ : if the good equilibrium exists for some β , then the bad equilibrium does not exist for $\gamma = 1 - \beta$, and if the bad equilibrium does not exist for some γ , then the good equilibrium exists for $\beta = 1 - \gamma$. One implication of this property is that $\underline{\beta} = 1 - \underline{\gamma}$, and $\underline{\beta} = 1 - \bar{\gamma}$.¹¹

5 The mixed strategy equilibrium

In this section, we study mixed strategy equilibria in which the behavior pattern of type O players is to mix instead of playing pure strategies. Since opportunists may mix in either or both phases, there are several types of mixed behavior patterns to consider. As we show in the working paper version, however, several of these behavior patterns do not give rise to equilibria, or give rise to equilibria that are payoff equivalent to equilibria we already considered. The only mixed behavior pattern that gives rise to a novel equilibrium is the one where type O players mix in phase S and play C in phase F . Consequently, we focus now on this behavior pattern, investigating the circumstances under which it is part of an equilibrium. As a matter of notation, we let $\lambda \in (0, 1)$ be type O players' probability of playing D in phase S .

Steady state and value functions In a mixed strategy equilibrium all three types of players behave differently. This requires the introduction of additional notation. Let x_α be the measure of type O players, and let x_γ be the measure type G players in phase S . The steady state of a mixed strategy equilibrium is characterized by a pair $(x_\alpha, x_\gamma) \in [0, \alpha] \times [0, \gamma]$, which satisfies

¹¹ Another feature of this duality is that the presence of bad-type players gives rise to the good equilibrium, while the presence of good-type players does not. Analogously, the presence of good-type players eliminates the bad equilibrium, while the presence of bad-type players does not.

$$(1 - \rho)(\alpha - x_\alpha) = (1 - \lambda)x_\alpha\rho \frac{(1 - \lambda)x_\alpha + x_\gamma}{x_\alpha + x_\gamma + \beta}; \tag{23}$$

$$(1 - \rho)(\gamma - x_\gamma) = x_\gamma\rho \frac{(1 - \lambda)x_\alpha + x_\gamma}{x_\alpha + x_\gamma + \beta}. \tag{24}$$

Let $z \equiv x_\alpha + x_\gamma$ be the measure of non-bad-type players in phase S , and $x \equiv (1 - \lambda)x_\alpha + x_\gamma$ be the measure of non-bad-type players that play C in phase S . Then, $\beta + z$ is the overall measure of players in phase S , and $\frac{\beta+z-x}{x}$ is the ratio of the measure of agents playing D to the measure of agents playing C in phase S .¹²

The value functions of type O players, defined under this mixed behavior pattern, are:

$$V_F = a + \delta[\rho V_F + (1 - \rho)V_S^C], \tag{25}$$

$$V_S^C = \frac{x}{z + \beta}V_F + \frac{z + \beta - x}{z + \beta}(-l + \delta V_S^C), \tag{26}$$

$$V_F^d = b + \delta V_S^C,$$

$$V_S^D = \frac{x}{z + \beta}(b + \delta V_S^C) + \frac{z + \beta - x}{z + \beta}(0 + \delta V_S^C),$$

where the superscripts (C or D) on V_S refer now to (candidate) equilibrium behavior, rather than to deviation from such behavior, while the superscript (d) on V_F continues to refer to deviation.

Incentive constraints This mixed behavior pattern is an equilibrium if, and only if, the following constraints are satisfied

$$\text{No deviation in phase } F: 0 \leq V_F - V_F^d \Leftrightarrow \frac{b - a}{\delta\rho} \leq V_F - V_S; \tag{27}$$

$$\text{Indifference in phase } S: V_S^D = V_S^C \Leftrightarrow V_F - V_S = \frac{b - a}{\delta\rho} + \frac{(z + \beta - x)l}{\delta\rho x}. \tag{28}$$

Since the RHS of (28) exceeds the LHS of (27), it suffices to require (28), which we rewrite (after solving for V_F and V_S) as

$$\frac{xa - (1 - \delta\rho)(z + \beta - x)l}{(z + \beta)(1 - \delta\rho) + \delta\rho x} = \frac{xb}{z + \beta}. \tag{29}$$

As before, letting $y \equiv \frac{\beta+z-x}{x}$, Eq. (29) is rewritten as

$$b - a = \frac{y}{1 + y}\delta\rho b - y(1 - \delta\rho)l \equiv f(y). \tag{30}$$

¹² $\beta + z$ is the analogue of $\beta + x$ in the good equilibrium and $1 - \gamma + x$ in the bad equilibrium; $\frac{\beta+z-x}{x}$ is the analogue of $\frac{\beta}{x}$ in the good equilibrium and $\frac{1-\gamma}{x}$ in the bad equilibrium.

Existence of mixed strategy equilibria We note that (30) is the same as (10), except that an equality is in place of the inequality. This narrows down the set of y 's that can be associated with a mixed strategy equilibrium to at most two values, \underline{y} and \bar{y} , which are the small and the large roots of (30). From the discussion in Sect. 3 we know that if (11) is not satisfied, there are no roots to Eq. (30) and, hence, no mixed strategy equilibria. Therefore, to proceed, we assume that (11) is satisfied.

Analogous to previous notation, the dependence of y on λ is denoted as $y = Y(\lambda)$. Observe now that when $\lambda = 0$, $y = \frac{\beta}{x_G}$, where x_G satisfies the steady-state condition of the good equilibrium, (1), and that when $\lambda = 1$, $y = \frac{1-\gamma}{x_B}$, where x_B satisfies the steady-state condition of the bad equilibrium, (12). Furthermore, straightforward calculations show that for any (α, β, γ) , $\frac{\beta}{x_G} < \frac{1-\gamma}{x_B}$, and that $Y(\lambda)$ is strictly increasing in λ .¹³ Therefore, as λ varies over $[0, 1]$, the value of y varies over $[\frac{\beta}{x_G}, \frac{1-\gamma}{x_B}]$. Combining this with the fact that the y associated with any mixed strategy equilibrium is either \underline{y} and \bar{y} , we conclude that a mixed strategy equilibrium exists if, and only if, at least one of \underline{y} or \bar{y} is in $(\frac{\beta}{x_G}, \frac{1-\gamma}{x_B})$, and that a mixed strategy equilibrium is unique if exactly one of \underline{y} or \bar{y} is in $(\frac{\beta}{x_G}, \frac{1-\gamma}{x_B})$.

To be more precise about the set of circumstances under which a mixed strategy equilibrium exists, consider the condition $\frac{\beta}{x_G} < y < \frac{1-\gamma}{x_B}$. The LHS of this condition is equivalent to $\beta < \underline{\beta}$ and the RHS is equivalent to $\gamma < \bar{\gamma}$; this follows from the monotonicity of $\frac{\beta}{x_G}$ in β , and $\frac{1-\gamma}{x_B}$ in γ , and from the definitions of $\underline{\beta}$ and $\bar{\gamma}$. If this condition is satisfied, i.e., if $(\beta, \gamma) \in [0, \underline{\beta}) \times [0, \bar{\gamma})$, a $\lambda \in (0, 1)$ can be found which gives rise to a mixed strategy equilibrium in which the ratio of the measures of agents choosing D to agents choosing C is \underline{y} . Likewise, the condition $\frac{\beta}{x_G} < \bar{y} < \frac{1-\gamma}{x_B}$ is equivalent to $\beta < \bar{\beta}$ and $\gamma < \underline{\gamma}$, and when this condition is satisfied, a $\lambda \in (0, 1)$ can be found which gives rise to a mixed strategy in which the ratio of the measures of agents choosing D to agents choosing C is \bar{y} . This gives us a complete characterization of when mixed strategy equilibria exist as a function of underlying parameters. We summarize this analysis as follows.

Proposition 4 *Hold all parameter values other than (α, β, γ) constant. Then, if (11) is violated, there are no mixed strategy equilibria. If (11) holds, then: (i) A mixed strategy equilibrium exists if, and only if, $(\beta, \gamma) \in [0, \underline{\beta}) \times [0, \bar{\gamma}) \cup [0, \bar{\beta}) \times [0, \underline{\gamma})$. (ii) A mixed strategy equilibrium is unique if, and only if, $(\beta, \gamma) \in [0, \underline{\beta}) \times [0, \bar{\gamma})$ or $(\beta, \gamma) \in [0, \bar{\beta}) \times [0, \underline{\gamma})$, but not both. (iii) In any mixed strategy equilibrium the ratio of the measure of agents playing D to the measure of agents playing C in phase S is either \underline{y} or \bar{y} , where \underline{y} and \bar{y} are the small and the large roots of $f(y) = b - a$.*

Now let us comment on the procedure to construct mixed strategy equilibria in general, and how it relates to the pure strategy equilibria we studied in Sects. 3 and 4. To be concrete we make these comments for parameter configurations in the domain $(\beta, \gamma) \in [0, \underline{\beta}) \times (\underline{\gamma}, \bar{\gamma})$. We know—from Propositions 1 and 3—that no pure strategy equilibria exists for such parameter values.

¹³ This is parallel to the property that y is increasing in β for the good equilibrium, and in $1 - \gamma$ for the bad equilibrium.

1. If all opportunists play C , $y < \underline{y}$ (because $\beta < \underline{\beta}$), which implies that an opportunist is better off playing D . On the other hand, if all opportunists play D , $\underline{y} < y < \bar{y}$ (because $\underline{\gamma} < \gamma < \bar{\gamma}$), which implies that an opportunist is better off playing C . As usual, the existence of such “cycle” suggests that a mixed strategy equilibrium may be found by letting *some* opportunists play C and others play D in phase S .
2. One way to think about the mixed strategy equilibrium is that it endogenizes the measure of bad-type players. Indeed, there is a measure β of bad-type players to begin with, but the measure of agents that play D is actually $\underline{\beta}$, where $\beta < \underline{\beta} = \beta + z - x$. This, in effect, means that the measure of bad-type players is endogenously increased via uncooperative behavior of opportunists. Alternatively, one may think of the mixed strategy equilibrium as endogenously increasing the measure of good-type players from γ to $\bar{\gamma}$.
3. Once the measures of behavioral types are endogenously increased in this way, we can think of the mixed strategy equilibrium as replicating the good (bad) equilibrium in a fictional community with $\underline{\beta}$ bad ($\bar{\gamma}$ good) type players. Either way, the measure of agents in phase S is $\beta + z$ and the ratio of the measure of agents playing D to the measure of agents playing C in phase S is $\frac{\beta+z-x}{x} = \frac{\underline{\beta}}{x(\underline{\beta})}$. These two variables, $\beta + z$ and $\frac{\beta+z-x}{x}$, are independent of the particular value that (β, γ) assumes, as long as $(\beta, \gamma) \in [0, \underline{\beta}] \times (\underline{\gamma}, \bar{\gamma})$. Therefore, if we define aggregate behavior as this pair of variables, we see that aggregate behavior in the community, at this mixed strategy equilibrium, is the same for all $(\beta, \gamma) \in [0, \underline{\beta}] \times (\underline{\gamma}, \bar{\gamma})$.

Likewise, mixed strategy equilibria over other regions in the parameter space are equivalent to pure strategy equilibria (good or bad) in fictional communities with $\underline{\beta}$ or $\bar{\beta}$ bad-type players, or $\underline{\gamma}$ or $\bar{\gamma}$ good-type players. As stated earlier, what mixed strategies do is to (endogenously) increase the measure of bad-type players to $\underline{\beta}$ or $\bar{\beta}$ and the measure of good-type players to $\underline{\gamma}$ or $\bar{\gamma}$. This trick works whenever there are sufficiently many opportunists to increase the measure of behavioral types to the requisite critical values. Obviously, this trick does not work to decrease the measures of bad or good-type players.

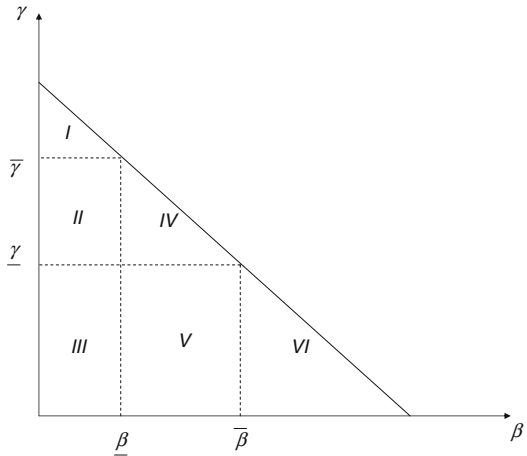
6 Summary

Here we summarize the existence of steady-state equilibria for any parameter configuration (α, β, γ) . Since $\alpha + \beta + \gamma = 1$, it is convenient to represent the various (α, β, γ) -triples in the simplex $\beta + \gamma \leq 1$, which is shown in Fig. 2.

To elaborate on what Fig. 2 shows, let us first consider the existence of pure strategy equilibria. We know from Propositions 1 and 3 that the good equilibrium exists if and only if $\beta \in [\underline{\beta}, \bar{\beta}]$, and that the bad equilibrium exists if and only if $\gamma \notin (\underline{\gamma}, \bar{\gamma})$. Also, due to duality, $\underline{\gamma} = 1 - \bar{\beta}$ and $\bar{\gamma} = 1 - \underline{\beta}$. Because of this, the simplex $\beta + \gamma \leq 1$ is partitioned into six regions.¹⁴ In regions I, III, and VI, the bad equilibrium exists, while the good equilibrium does not exist. In region IV, the good equilibrium exists, while

¹⁴ In most statements below a region is understood as the interior of a region.

Fig. 2 Classification of equilibrium outcomes



the bad equilibrium does not exist. In region V, both the good and the bad equilibria exist. In region II, neither the good nor the bad equilibrium exists.

Let us turn now to mixed strategy equilibria. Based on Proposition 4 and the discussion following it, we have the following summary. There are two mixed strategy equilibria in region III because we can either increase β to $\underline{\beta}$ or increase γ to $\bar{\gamma}$. On the other hand, there are no mixed strategy equilibria in regions I, IV, or VI because neither β nor γ can be increased to bring them into a region in which a pure strategy equilibrium exists. Finally, a unique mixed strategy equilibrium exists in region II because, although both β or γ may be increased, the two increases lead to equivalent equilibria (corresponding to $\underline{\gamma}$). And, likewise, a unique mixed strategy equilibrium exists in region V (corresponding to $\bar{\gamma}$).

We summarize the existence of pure and mixed strategy equilibria in Table 2.

7 Welfare

In this section, we construct measures of social welfare at certain steady-state equilibria, and show how they relate to the configuration of types. We already know from the analysis in Sect. 6 that some (α, β, γ) configurations give rise to multiple equilibria,

Table 2 Characterization of equilibria

Regions	Pure strategy equilibria	Mixed strategy equilibria
I	Bad equilibrium	None
II	None	One replicating $\underline{\gamma}$
III	Bad equilibrium	Two
IV	Good equilibrium	None
V	Both equilibria	One replicating $\bar{\gamma}$
VI	Bad equilibrium	None

so numerous welfare measures may be calculated. To limit the number of cases to report, we focus on two calculations. In the first calculation we fix the measure of good-type players at zero, $\gamma = 0$, and compute welfare as a function of β at the best equilibrium. Then, in the second calculation, we fix the measure of bad-type players at zero, $\beta = 0$, and compute welfare as a function of γ at the worst equilibrium.¹⁵ Our measure of welfare is the total per-period payoff to the whole community at the equilibrium in question. Since the overall measure of agents is one, this is the same as the average per-period payoff.

Welfare as a function of β Suppose $\gamma = 0$. Then, following the analysis in Sect. 6, we have a tripartite partition. When $\beta < \underline{\beta}$ (region III), three equilibria exist and the best equilibrium is the mixed strategy equilibrium replicating y . When $\underline{\beta} \leq \beta \leq \bar{\beta}$ (region V), two equilibria exist and the best equilibrium is the good equilibrium. When $\bar{\beta} < \beta$ (region VI), the unique steady-state equilibrium is the bad equilibrium.

Altogether, social welfare takes the following form:

$$W(\beta) = \begin{cases} (1 - z - \beta)a + (\beta + z - x)\frac{x}{z+\beta}b + x[\frac{x}{z+\beta}a - \frac{\beta+z-x}{z+\beta}l] & \text{if } \beta < \underline{\beta} \\ (1 - x - \beta)a + \beta\frac{x}{x+\beta}b + x[\frac{x}{x+\beta}a - \frac{\beta}{x+\beta}l] & \text{if } \underline{\beta} \leq \beta \leq \bar{\beta}, \\ 0 & \text{if } \bar{\beta} < \beta \end{cases} \tag{31}$$

where x in the second line comes from the solution to (1), and x and z in the first line come from the solution to (23) and (24).

To elaborate on how (31) is arrived at, consider the middle term, which applies to the range $\underline{\beta} \leq \beta \leq \bar{\beta}$. Then, as stated above, welfare is evaluated at the good equilibrium. Opportunists in this equilibrium get a period payoff of a in phase F , and get either a or $-l$ in phase S , depending on whom they meet. Bad-type players get either b or 0 , depending again on whom they meet. Using the measures of agents at each phase (which come from the solution to the steady-state equation), we take the average over these payoffs, and get the reported expression.

Analyzing Eq. (31) we derive the following result, which is graphically illustrated in the left panel in Fig. 3.

Proposition 5 (i) When $\beta < \underline{\beta}$, $W(\beta)$ is constant; (ii) when $\underline{\beta} \leq \beta \leq \bar{\beta}$, $W(\beta)$ is strictly decreasing and is, hence, maximized at $\underline{\beta}$; (iii) when $\bar{\beta} < \beta$, $W(\beta)$ is zero.

Proof See the Appendix. □

The reason W is zero for $\bar{\beta} < \beta$ is that welfare is evaluated at the bad equilibrium, where all agents play D and collect zero. The reason W decreases for $\underline{\beta} \leq \beta \leq \bar{\beta}$ is that welfare is evaluated at the good equilibrium at which having more bad-type players is not necessary to induce opportunists to play C , but only reduces the average level of cooperation and, hence, the average payoff in the community. Finally, the reason

¹⁵ These two calculations relate to our previous results that the presence of type B players can support the good equilibrium and the presence of type G players can upset the bad equilibrium.

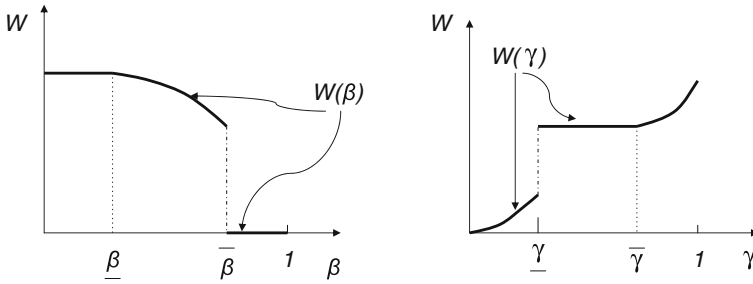


Fig. 3 Welfare measures

welfare is constant for $\beta \leq \bar{\beta}$ is that welfare (for each β in this range) is measured at the mixed strategy equilibrium replicating \underline{y} , leading to the same aggregate behavior and aggregate payoff.

An interesting feature of Fig. 3 is that welfare decreases discontinuously at $\beta = \bar{\beta}$. The reason for this is that an equilibrium sustaining some cooperation can be achieved for $\beta \leq \bar{\beta}$, but not for β slightly above $\bar{\beta}$ (for $\beta > \bar{\beta}$, the only equilibrium is the bad one). Therefore, as β crosses $\bar{\beta}$, an infinitesimal increase in β has a quantum effect on the degree of cooperation in the community and on welfare.

Welfare as a function of γ Let us turn now to the case $\beta = 0$. As γ varies over $[0, 1]$, the worst equilibrium varies as follows: When $\gamma \in [0, \underline{\gamma}]$ or $\gamma \in [\bar{\gamma}, 1]$, the worst equilibrium is the bad equilibrium; and, when $\gamma \in (\underline{\gamma}, \bar{\gamma})$, the unique equilibrium is the mixed strategy equilibrium replicating \underline{y} . Evaluating welfare at these equilibria, we get

$$W(\gamma) = \begin{cases} x(-l) + (\gamma - x)a + (1 - \gamma)\frac{x}{x+1-\gamma}b & \text{if } \gamma \leq \underline{\gamma} \text{ or } \bar{\gamma} \leq \gamma \\ (1 - z)a + (z - x)\frac{x}{z}b + x[\frac{x}{z}a - \frac{z-x}{z}l] & \text{if } \underline{\gamma} < \gamma < \bar{\gamma}. \end{cases} \quad (32)$$

Analyzing this welfare function we derive the following result, which is analogous to Proposition 5, and is proven in the working paper version.

Proposition 6 (i) When $\gamma \in [0, \underline{\gamma}] \cup [\bar{\gamma}, 1]$, $W(\gamma)$ is increasing in γ ; (ii) when $\gamma \in (\underline{\gamma}, \bar{\gamma})$, $W(\gamma)$ is constant in γ .

Intuitively, as γ increases the average cooperation level in the bad equilibrium increases and, thus, social welfare increases. In the mixed strategy equilibrium replicating \underline{y} , aggregate behavior is constant (i.e., independent of γ) and, thus, the social welfare in that equilibrium is constant too.

The relationship between social welfare at the worst equilibrium and γ is plotted in the right panel of Fig. 3. Analogous to the best equilibrium, social welfare has an upward jump at $\underline{\gamma}$. This is because the bad equilibrium no longer exists when γ is infinitesimally bigger than $\underline{\gamma}$.

8 Conclusion

We study how forming long-term relationships endogenously can provide incentives for strategic players to cooperate in community games with heterogeneous players. We focus on a class of equilibria that satisfy fresh start, zero tolerance and quick familiarity. The type distribution of players affects equilibrium behavior. Specifically, the good equilibrium in which strategic players always cooperate exists if and only if the fraction of bad-type players is in an intermediate range. In contrast to Kandori's contagious equilibrium, the good equilibrium in our model is stable with respect to observational errors. On the other hand, the bad equilibrium in which strategic players defect does not exist if the fraction of good-type players is in an intermediate range. We characterize the existence of all pure strategy and mixed strategy equilibria under all possible configurations of the type distribution, and compare the welfare across different equilibria.

In our model the type distribution is exogenously given. One possible extension is to endogenize the type distribution. Specifically, suppose initially all players are of the bad type, and each individual player has the option of becoming an opportunistic type by investing in skills and incurring a positive investment cost. Then the resulting type distribution is commonly observed and the community game begins. In this setting, players' incentive to invest in the initial investment game depends on the payoff difference between an opportunistic type and a bad type in the equilibrium of the ensuing community game. Interestingly, compared to the social optimum, in equilibrium players might overinvest (the number of opportunistic players is more than the socially optimal number) in the investment game. This is due to the fact that to sustain the good equilibrium in the community game, a certain portion of players must behave like the bad type in the stranger phase. If the number of bad-type players is less than this threshold, then opportunistic players must play a mixed strategy in new relationships, which means that the investment in skills is socially "wasted." A formal analysis of this extension can be found in the working paper version.

Appendix

Proof of Lemma 1

Proof The solution to (1) is

$$x = \frac{(1 - \rho)(1 - 2\beta) + \sqrt{(1 - \rho)^2 + 4\beta(1 - \beta)\rho(1 - \rho)}}{2}. \quad (33)$$

Dividing (33) by β we get

$$\frac{1}{y} = \frac{x}{\beta} = \frac{(1 - \rho)(\frac{1}{\beta} - 2) + \sqrt{\frac{(1 - \rho)^2}{\beta^2} + 4(\frac{1}{\beta} - 1)\rho(1 - \rho)}}{2}. \quad (34)$$

Since all terms in (34) decrease in β , $y(\beta)$ increases in β . Moreover, when $\beta \rightarrow 0$, $x/\beta \rightarrow \infty$, and $y \rightarrow 0$. On the other hand, when $\beta \rightarrow 1$, $x \rightarrow 0$, and $y \rightarrow \infty$. \square

Proof of Proposition 5

Proof (i) In steady state, by abusing notation (both z and x are functions of β),

$$(1 - \rho)(1 - \beta - z) = x\rho \frac{x}{z + \beta}$$

$$\Leftrightarrow \frac{1 - \rho}{\rho} \left(\frac{1}{x} - \frac{\beta + z}{x} \right) = \frac{x}{z + \beta} \tag{35}$$

But we know that, when $\beta \leq \underline{\beta}$, in the mixed strategy equilibrium $\frac{\beta+z-x}{x} = \underline{y}$ is independent of β . Therefore, from (35) x is also independent of β . As a result, $\beta + z$ is also independent of β . Since $W(\beta)$ is only a function of x and $z + \beta$ [see Eq. (31)], we reach the conclusion that $W(\beta)$ is constant when $\beta \leq \underline{\beta}$.

(ii) First we show that $1 - x - \beta$, the measure of agents in phase F , is strictly decreasing in β . Suppose not, that is, suppose there exist a β' and a β'' in $[\underline{\beta}, \bar{\beta}]$ so that $\beta' < \beta''$ and yet $1 - x' - \beta' \leq 1 - x'' - \beta''$, where x' (x'') is the steady state x under β' (β''). Then from (1)

$$x' \frac{x'}{x' + \beta'} \leq x'' \frac{x''}{x'' + \beta''},$$

which is equivalent to

$$x' \frac{1}{1 + \beta'/x'} \leq x'' \frac{1}{1 + \beta''/x''}.$$

But $\beta'/x' < \beta''/x''$ since y is increasing in β . Therefore, we must have $x' < x''$, which implies

$$1 - x'' - \beta'' < 1 - x' - \beta',$$

a contradiction.

Lemma 1 shows that $\frac{x}{x+\beta}$, the probability of being matched with a non-bad type in phase S , is decreasing in β . From expression (31) we see that by increasing β , the average payoff in phase S decreases and the weight placed on this payoff increases. Hence the total social welfare in the community must decrease. \square

References

Abreu, D., Pearce, D., Stacchetti, E.: Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica* **58**, 1041–1063 (1990)

Dal Bo, P.: Social norms, cooperation and inequality. *Econ Theory* **30**, 89–105 (2007)

Datta, S.: Building trust. Working paper, London School of Economics, Mimeo (1993)

Diamond, P.: Aggregate demand management in search equilibrium. *J Polit Econ* **90**, 881–894 (1982)

Dixit, A.: On modes of economic governance. *Econometrica* **71**, 449–481 (2003)

Eeckhout, J.: Minorities and endogenous segregation. *Rev Econ Stud* **73**, 31–53 (2006)

Ellison, G.: Cooperation in the prisoners-dilemma with anonymous random matching. *Rev Econ Stud* **61**, 567–588 (1994)

- Frank, R.: *Passions Within Reason*. New York: Norton (1988)
- Fudenberg, D., Maskin, E.: The folk theorem in repeated games with discounting or with incomplete information. *Econometrica* **54**, 533–554 (1986)
- Ghosh, P., Ray, D.: Cooperation in community Interaction without information flows. *Rev Econ Stud* **63**, 491–519 (1996)
- Johnson, S., McMillan, J., Woodruff, C.: Courts and relational contracts. *J Law Econ Organ* **18**, 211–277 (2002)
- Kali, R.: Endogenous business networks. *J Law Econ Organ* **15**, 615–636 (1999)
- Kandori, M.: Social norms and community enforcement. *Rev Econ Stud* **59**, 63–80 (1992)
- Kranton, R.: The formation of cooperative relationships. *J Law Econ Organ* **12**, 214–233 (1996)
- Mortensen, D.: Property rights and efficiency in mating, racing, and related games. *Am Econ Rev* **72**, 968–979 (1982)
- Okuno-Fujiwara, M., Fujiwara-Greve, T.: Voluntarily separable prisoners' dilemma. Working paper, University of Tokyo, Mimeo (2006)
- Rob, R., Yang, H.: Long-term relationships as safeguards. Working paper, Ohio State University, Mimeo (2005)
- Shapiro, C., Stiglitz, J.: Equilibrium unemployment as a worker disciplinary device. *Am Econ Rev* **74**, 433–444 (1984)
- Sobel, J.: Interdependent preferences and reciprocity. *J Econ Lit* **43**, 392–436 (2005)
- Sobel, J.: For better or forever: formal versus informal enforcement. *J Labor Econ* **24**, 271–297 (2006)
- Taylor, C.: The old-boy network and the young-gun effect. *Int Econ Rev* **41**, 871–891 (2000)
- Tirole, J.: A theory of collective reputations. *Rev Econ Stud* **63**, 1–22 (1996)
- Watson, J.: Starting small and renegotiations. *J Econ Theory* **85**, 52–90 (1999)
- Watson, J.: Starting small and commitment. *Games Econ Behav* **38**, 176–199 (2002)
- Yang, H.: Nonstationary relational contracts. Working paper, Ohio State University, Mimeo (2007)