On Optimal Social Convention in Voluntary Continuation Prisoner's Dilemma Games

Filip Vesely
Department of Economics, University of Wisconsin Milwaukee

Chun-Lei Yang
RCHSS, Academia Sinica

August 2016

Abstract

When the repeated prisoner's dilemma setup is generalized to allow for a unilateral breakup, maximal efficiency in equilibrium remains an open question. With restrictions of simple symmetry with eternal mutual cooperation, defection, or (matched) alternation on the equilibrium path, we describe the upper limit of discounted lifetime payoff and construct simple social conventions that, for a large set of parameters, achieve it. While all other well-known equilibrium designs in the literature punish defections with a breakup and thus reach the optimum only in degenerate cases, exploited cooperators in ours allow defectors to compensate them by cooperating more in the future.

1.   Introduction

Gains from cooperation benefit society as a whole. Yet, a self-interested individual can hardly escape the temptations to defect if it is a myopically dominant strategy. Various mechanisms that help incentivize cooperative behavior have been studied in the context of the infinitely repeated prisoner's dilemma (PD) game. It is well known that privately imposed trigger strategies, like *tit-for-tat* (Axelrod, 1984), can sustain eternal cooperation within fixed partnerships.[1] In settings where players are randomly re-matched into fresh partnerships a public record-keeping device (Rosenthal, 1979; Okuno-Fujiwara and Postlewaite, 1995; Takahashi, 2010)[2] or punishment contagion (Kreps et al., 1982; Kandori, 1992; Ellison, 1994) can serve as a similar trigger to sustain cooperation. While the simplicity of fixed and one-shot random matching paradigms is theoretically appealing, they do not accurately capture the realities of the most common partnerships in society such as friendships, marriages or employment and business relationships, which mostly exist on a voluntary basis. It is precisely this voluntary feature, where a defecting partner can unilaterally leave for the matching pool, that makes it impossible for her cooperative counterpart to exercise the sort of direct punishments prescribed by trigger strategies frequently encountered in the fixed partnership paradigm. The difficulty of cultivating cooperation in such relationships is further confounded in a large and complex society where collecting relevant information regarding individuals' past histories is just too costly to be a feasible option.

So, how could cooperation come about in partnerships where continuation is only voluntary? Cooperative behavior builds on reciprocity and thus it can only emerge if the costs of defecting continuously outweigh the benefits. In a voluntary partnership the costs include losing a perhaps cooperative counterpart, the resources and time spent on finding another willing partner in the matching pool, and the forgone business opportunities due to the gradual nature of developing mutual trust in a new stable relationship.[3] As a result, the grim prospect of re-entering the matching pool can serve as a disciplinary device that elicits and sustains long-

---

[1] In environments where signals are noisy, more sophisticated strategies supporting belief-free equilibria can be used to sustain cooperation (see, for example, Mailath and Morris, 2002; Piccione, 2002; Ely, Horner, and Olszewski, 2005).

[2] Also see Dixit (2003) for a discussion about private information intermediaries in the context of an asymmetric PD game.

[3] In Shapiro and Stiglitz (1984), workers receiving efficiency wages do not shirk, due to fear of being caught and fired, in which case they would expect to live with the far lower market clearing wage for extended periods. In Carmichael and MacLeod (1997), inefficient gift exchange between partners before they start to trust each other serves a similar role of making the partnership breakup costly. Ghosh and Ray (1996), Kranton (1996), and Watson (2002) have the partners only gradually increase their level of trust, a costly process a player already in a complete mutual trust stage would be loathe to repeat after being abandoned by the previous betrayed partner.

term cooperation. The re-entry to the matching pool, to some extent, is thus conceptually similar to the trigger strategy in the fixed partnership paradigm.

It is worth noting that, even though in voluntary partnerships it is the implicit offer of future cooperation that allows the threat of separation to elicit partners' current cooperation, an offer of "right off the bat" cooperation can never be sustained as the equilibrium strategy of the whole population. Early cooperation of others indeed encourages players to take advantage of their cooperative counterparts through early defection followed by an escape to the matching pool. Therefore, the option of unilateral separation acts as a double-edged sword that not only enforces but also delays the cooperation.

In order to maximize the lifetime expected payoff of a player in the matching pool, we investigate how the delay in full cooperation or the so-called trust-building phase can be kept at the minimum in a binary-action setting. The following are the specific research questions that we ask: What is the upper bound to the lifetime expected payoff a player in the matching pool can earn under all social conventions of simple symmetry with eternal mutual cooperation, defection, or (matched) alternation? Does there exist an equilibrium distribution that yields players the above maximal payoff?

In order to address these questions, we assume that the economy is populated by a continuum of infinitely lived agents who have to decide if they will stay with the same partners to play a two-person prisoner's dilemma game for at least one more period. While theoretically the population equilibrium may involve very complex behavior and outcome patterns, experience tells us that successful social conventions in human society are rather of simple forms. Thus we focus on stationary population equilibrium, also called the social norm or convention, where once partners have come to trust each other and decided that they would want to stay together forever, their behavior has some symmetry in it, consistent for example with Kant's principle of categorical imperative or the moral maxim of "Do unto others what you want done unto you" (Luke 6:31).[4] In fact, if such a maxim is upheld within the same round, then mutual cooperation or mutual defection would be the expected stable outcome. Extending to two consecutive rounds, it allows for an additional symmetric outcome in which one player cooperates only in even and the other only in odd rounds of their partnership. Let us call it the *social convention of simple symmetry* if players have positive odds to end up in stable partnerships with eternal repetition of exactly one of the three symmetric outcome

---

[4] These are also called "moral Golden Rules".

patterns, called eternal cooperation, eternal defection, and eternal alternation. To our knowledge, this is the only kind of social convention all studies in the related literature have ever dealt with.

Social convention with eternal cooperation is intuitively the form which society will prefer to implement, for efficiency's sake. Eeckhout (2006), Fujiwara-Greve and Okuno-Fujiwara (2009) and Rob and Yang (2010) have designed different variations of this sort in recent studies. However, as we discuss in detail in Section 4 of this paper, their equilibrium designs generically cannot achieve the maximal lifetime expected payoff for social conventions with eternal cooperation. In fact, the literature spots no generic characterization of sustainable frontier of optimal equilibrium payoffs, not to mention any folk-theorem type result, once the more realistic feature of voluntary continuation is introduced into the standard repeated fixed-partnership setup. Our study provides a first such generic result and contributes towards eventually closing this glaring theoretical gap.

In this paper, we construct two classes of social conventions in the form of subgame perfect equilibrium distributions, one each with eternal cooperation or alternation, and show that for a large parameter region of positive measure, the maximal lifetime expected payoff for social conventions of simple symmetry can be achieved. More specifically, in our social convention with eternal cooperation (EC), players play mixed actions at the initial state and, depending on the realized outcome, transition either back to the origin in the case of mutual defection, or to the absorbing state of EC in the case of mutual cooperation, or to the transitional states of alternation otherwise. In the transitional states, the previous defector is to cooperate while the previous cooperator mixes. The state of mutual cooperation is an absorbing one. Thus, the partners are expected to go through finite stretches of alternation between cooperation and defection with positive probability, before ending up with eternal cooperation with probability one. In some sense, alternation is interpretable as "within-partnership punishment" of the defector of the previous round. In our social convention with eternal alternation (EA), by contrast, the partners stay together only after their actions in the stranger's phase are different.

In both types of social convention, the payoff at the origin is the same as the "equilibrium punishing" payoff for the defector from the previous round on the alternation path within the partnership. In some sense, this stick also serves as carrot at the same time, to coax the equilibrium deviator to subject himself to his partner's one-period advantage taking as revenge, instead of heading to the pool himself. It is interesting to note that, in one particular case within the optimal classes of social conventions, the partnership will never voluntarily break up, once

past the very first stranger's phase. Also, all off-equilibrium punishment for deviation can be kept within the partnership, without resorting to breakup.

The rest of the paper is structured as follows. Section 2 discusses the game structure and different concepts of strategies and equilibrium. Section 3 discusses our novel social conventions with either eternal cooperation or eternal alternation, and characterizes the conditions for their optimality among all social conventions of simple symmetry. Section 4 discusses in detail well-known designs of equilibrium in the literature and compares their payoff performance with ours. Section 5 provides further discussions on the model. Section 6 concludes the paper.

2. Basic Model of Voluntary Continuation Prisoner's Dilemma

Let δ be the discount rate in an infinite horizon discrete time environment. There is a continuum of players who are either single or matched in bilateral partnerships that were formed in the past. At the beginning of each period, single players are randomly matched into pairs. All players then simultaneously make a binary decision $a, \tilde{a} \in \{C, D\}$ to cooperate or defect in a PD game with payoffs given by

$$
\begin{array}{ccc}
 & C & D \\
C & 1,1 & -L, 1+G \\
D & 1+G, -L & 0,0
\end{array}
$$

where $L > 0$ and $G > 0$.

After viewing the outcome of the PD game in period t, every player simultaneously makes a second binary decision $b_t \in \{0,1\}$ whether to stay $(= 1)$ with the current partner. By letting $\tilde{b}_t$ denote one's partner's stay decision, the pair is dissolved if and only if $b_t \cdot \tilde{b}_t = 0$, i.e. at least one of them decides to leave.[5] However, if the partnership continues, both players

---

[5] One interpretation of the parameter δ is to assume that each player has a constant survival rate δ per period and, once they die, the current partnership dissolves and a new player of the same behavior trait, i.e., using the same strategy, is reborn in the singles' pool. As players have no say about their survival, for equilibrium analysis, it does not make much difference whether both partners die independently and thus their partnership is exogenously broken, or if their death is somehow correlated as if they are hit by a contagious killer flu. Formally, however, in the case of the exogenous partnership breakup, the surviving widow lives on to enjoy an additional payoff, while in the case of simultaneous death these same payoffs are earned by newborn replacements. Thus as long as the probability of partnership continuation is kept the same, both the old and the new players' best responses are also the same. See also Mailath and Samuelson (2006) for similar reasoning.

proceed to face the same decision problems as in the previous round. For each player, let $S_a = \{CC, CD, DC, DD\}$ and $S_b = \{(1,1), (1,0), (0,1), (0,0)\}$ denote the stage outcome spaces.

Decision a_{t+1} may vary with history $h_t = ((a_\tau, \tilde{a}_\tau), (b_\tau, \tilde{b}_\tau))_{\tau=1}^t \in H_t^a$ of the previous t periods with $(a_\tau, \tilde{a}_\tau) \in S_a$ and $(b_\tau, \tilde{b}_\tau) \in S_b$, while the stay decision b_{t+1} varies with history $l_{t+1} = (h_t, (a_{t+1}, \tilde{a}_{t+1})) \in H_{t+1}^b = H_t^a \times S_a$. Let h_0 denote the empty history at the beginning of time and let $H_0^a = \{h_0\}$. Let $H^a = \bigcup_{t=0}^\infty H_t^a$ and $H^b = \bigcup_{t=1}^\infty H_t^b$.

Definition: Let $x(h)$ denote the probability to play C for any $h \in H^a$ and $y(h)$ the probability to stay for any $h \in H^b$. A ***behavior strategy*** $\sigma: H^a \cup H^b \mapsto [0,1]$ is a pair of mappings $\sigma = (x, y)$ with $x: H^a \mapsto [0,1]$ and $y: H^b \mapsto [0,1]$.

Note that a fixed-partnership can be viewed as a special case of the voluntary continuation framework with the restriction of $y(h) = 1$ for all $h \in H^b$ in the above definition.

Definition: Let $h_{t+k} \backslash h_t = ((a_\tau, \tilde{a}_\tau), (b_\tau, \tilde{b}_\tau))_{\tau=t+1}^{t+k} \in H_k^a$ and $l_{t+k} \backslash h_t \in H_k^b$ denote the k-period continuation histories after h_t, respectively. Then, $\sigma = (x, y)$ is ***partnership independent***, if $b_t \tilde{b}_t = 0$, i.e., the partnership breaks up in period t, implies $x(h_{t+k}) = x(h_{t+k} \backslash h_t)$ and $y(l_{t+k}) = y(l_{t+k} \backslash h_t)$ for all h_t, h_{t+k}, l_{t+k} and all t, k.

In other words, the player's strategy in the new partnership is not affected by his experience in the previous one, as if he is going back to the origin with empty history h_0 each time the old partnership is broken up. Assuming partnership independence, the player no longer needs to worry about histories that contain the event of a breakup, i.e., $b_t \cdot \tilde{b}_t = 0$ for some t. In other words, any strategically relevant history of period t must have $b_\tau = \tilde{b}_\tau = 1, \forall \tau \leq t$. For expositional simplicity, we subsequently concentrate on the simpler form of history recording $h_t = (a_\tau, \tilde{a}_\tau)_{\tau=1}^t \in H_t$ and $H = \bigcup_{t=0}^\infty H_t$ unless noted otherwise.

Thus, the strategically relevant part of any *partnership independent **behavior strategy*** can be restated as $\sigma: H \mapsto [0,1]^2$ with $\sigma(h) = (x(h), y(h)), \forall h \in H$, whenever there is no danger of misunderstanding. Note that consideration of discounted payoffs over different future partnerships will not be affected by this simplified notion of behavior strategy, as we will see below. Technically, the *C/D*-decision part of history recording is exactly the same for H^a, H^b and H. As is common in the literature, for equilibrium analysis we are only interested in strategies of this kind. Furthermore, individual deviations using partnership dependent

strategies trivially will not benefit any player in any equilibrium based on partnership independent strategies.[6]

Let Σ denote the set of all behavior strategies. Let H_∞ denote the set of potential histories of infinite length. For any $h \in H_\infty$, let $\pi(a_t, \tilde{a}_t)$ be the player's stage payoff when he plays a_t and his partner $\tilde{a}_t$ in period t, then $v(h) = \sum_{t=1}^\infty \delta^{t-1}\pi(a_t, \tilde{a}_t)$ is the player's *expected lifetime payoff* from this infinite history with discount rate δ.

Now, assume that the behavior composition in the singles' pool with a continuum of players is *stationary*.[7] The size of the pool is normalized to 1. Let $\Delta(X)$ denote the set of probability measures on any set X. Let $p \in \Delta(\Sigma)$ be the stationary strategy distribution in the pool. For any strategy $\sigma \in \Sigma$, let $P^t(a^t|\sigma, p)$ denote the probability that $a^t \in S_a$ realizes in period t. Then, the expected payoff from the strategy σ given the stationary population p is $v(\sigma; p) = \sum_{t=1}^\infty \sum_{a^t \in S_a} \delta^{t-1} P^t(a^t|\sigma, p)\pi(a^t)$.[8] Now, we can define equilibrium concepts with these notations.

Definition: $p \in \Delta(\Sigma)$ is a *Nash equilibrium distribution* if, for all $\sigma \in \operatorname{supp} p$, we have $v(\sigma; p) \geq v(\sigma'; p)$, for all $\sigma' \in \Sigma$.

Now, let $H(h_t) = \{h \in H : \operatorname{proj}_t h = h_t\}$ denote the set of all *continuation histories* (or paths) at $h_t \in H_t$ and $v(h|h_t) := \sum_{\tau=1}^\infty \delta^{\tau-1}\pi(a_{t+\tau}, \tilde{a}_{t+\tau})$ for all $h \in H(h_t)$ be its expected *continuation payoff* calculated at time t. Let $\sigma_{h_t} := \sigma|_{H(h_t)}$, such that $\sigma_{h_t}(h_k) = \sigma(h_{t+k})$ with $h_{t+k} = (h_t, h_k) \in H_{t+k}$ for any k and $h_k \in H_k$. Due to self-similarity of the events potential anywhere in the game, we always have $H(h_t) = H$. Thus, the expected continuation payoff given h_t can be written as $v(\sigma_{h_t}; p) = \sum_{\tau=1}^\infty \sum_{a^\tau \in S_a} \delta^{\tau-1} P^\tau(a^\tau|\sigma_{h_t}, p)\pi(a^\tau)$. Note that the term $P^\tau(\cdot\,|\sigma_{h_t}, p)$ is adapted to account for updated belief about the partner's strategy type

---

[6] Note although the singles' pool in the equilibria we study is never empty, this may happen in situations when there is no breakup on the equilibrium path such as in one version of all-defection SPED. An off-path breakup may rematch the same partners from the last period with positive likelihood, which could challenge the assumption of partnership independence in degenerate case. This can be avoided by assuming continuous inflow of newborn agents or by introducing exogenous random breakup or death as a substitute for the discount rate. Our model setup is chosen for the sake of simple exposition only. See footnote 6 for more discussion.

[7] Stationarity of the pool population in equilibrium is a common and basic assumption in the literature. It is always *self-consistent* in the following sense. Let q_t be the likelihood that a partnership in the t-th period continues into $t+1$, for some given type σ in a stationary population and $t = 1, 2, \dots$. Given the exogenous breakup rate of $1 - \delta$, the share of σ types whose partnership survived t periods can be recursively calculated at $y_t = \delta q_t y_{t-1}$, for $t = 1, 2, \dots$, with $y_0 = 1$ being the initial size. Thus, after the process has gone on for infinitely many periods already, the total share of σ that is expected to go back to the pool is hence $\sum_{t=1}^\infty y_{t-1}(1 - \delta q_t) = 1$, as all terms but one in the sum cancel each other out. Note that this new solution to the stationarity issue is equally applicable to exogenous-type models like Ghosh and Ray (1996).

[8] Let $\varphi(\sigma, p) \in \Delta(H_\infty)$ denote the distribution over all possible histories of infinite length, induced by the player playing σ within the pool of p. Assume accountability of $\operatorname{supp} \varphi(\sigma, p)$, then $v(\sigma; p) = \sum_{h \in H_\infty} v(h)\, \varphi(h|\sigma, p)$.

based on the personal experience h_t playing σ up until t, which is in general different from the stationary singles' pool population p.

Definition: $p \in \Delta(\Sigma)$ is a *subgame perfect equilibrium distribution* (SPED) if, for all $\sigma \in$ supp p, $v\left(\sigma_{h_t}; p\right) \geq v\left(\tilde{\sigma}_{h_t}; p\right)$, for all $\tilde{\sigma} \in \Sigma$ and $h_t \in H$.

In general, behavior rules and patterns of outcome paths in SPED might be very complex. However, if the set of SPED behavior rules is to realistically establish itself as a social convention within a society, they cannot be too complex, with respect to both the rule description and the pattern of implied long-run outcome paths it aims to achieve. Also, it is inconceivable under such a social convention that the partners in a stable relationship never end up with some sort of symmetric outcome paths, for the sake of fairness and reciprocity common in moral maxims in society. Hence, our main focus is on realistic social conventions as defined below. Let $H_{\sigma,p}(h_t)$ denote the set of all *achievable* continuation paths at h_t for someone with strategy $\sigma \in$ supp p.

Definition: A SPED p is called a *social convention of simple symmetry* if there is at least one $\sigma \in$ supp p such that the player using σ expects with positive likelihood to end up with exclusively experiencing the outcomes of either $h_{CC} = (CC, CC, CC, \dots)$, $h_{DD} = (DD, DD, DD, \dots)$ or $h_{CD} = (CD, DC, CD, DC, \dots)$, after some finite periods of transition. In other words, there is a finite period k where h_k realizes with positive likelihood under $(\sigma; p)$, such that $H_{\sigma,p}(h_k) \subset \{h_{CC}, h_{CD}, h_{DD}\} \subset H_\infty$.

In addition, to avoid the complexity of a general behavioral strategy, the literature has limited itself to the so-called Markov strategy as candidates for equilibrium behavior. Let $\Lambda = \{A_k\}_{k=1}^n$ be a partition of H with $A_k \subseteq H$, $\bigcup_{k=1}^n A_k = H$, $A_k \cap A_l = \emptyset$, $k \neq l$, while $A(h) \in \Lambda$ denotes the state that contains history h.

Definition: A behavior strategy $\sigma: H \to [0,1]^2$ is a *Markov strategy* (or *finite automaton*) if there is a finite partition Λ so that $A(h) = A(h')$ implies $\sigma(h) = \sigma(h')$, for all h, h'. Equivalently, it can be written as $\sigma: \Lambda \to [0,1]^2$.

In fact, any Markov strategy induces a *Markov transition rule,* $\gamma: \Lambda \times S \to \Delta(\Lambda)$ where $\gamma(A'|A; a\tilde{a}, (b, \tilde{b}))$ is the likelihood for the new state to be A', given that the state in the last round is A and the stage outcome is $\{a\tilde{a}, (b, \tilde{b})\}$. In other words, the action at the history (subgame) h_t under such a Markov strategy depends only on the state $A(h_t)$ the player is

currently in, but not on any particular path that led to it.[9] Note also that a Markov strategy is not necessarily partnership-independent. In fact, it is easy to construct (Λ, γ) that is sensitive to partnership changes. For example, the Markov strategy "play C in odd number periods and play D in even number periods while always leaving after one period" is not partnership-independent.

3. Optimal Social Convention of Simple Symmetry

What is the maximal lifetime expected payoff achievable in a SPED in the voluntary continuation PD game? We will provide a partial solution to this open question here.

As a starting observation, note that given any equilibrium distribution p and equilibrium strategy $\sigma \in \operatorname{supp} p$, the expected lifetime payoff in the strangers' pool, $v_0 = v(\sigma; p)$, cannot be higher than the payoff the player expects at any history in the future, i.e., $v_0 \leq \inf_{h \in H} v_h(\sigma; p)$. Otherwise, the player would have a strict incentive to deviate at some history $h' \in H$ by voluntarily going back to the pool instead of sticking to the same partner as prescribed by σ. This limits the extent to which partners can punish each other within the partnership to sustain a high level of cooperation the way the standard trigger-strategy could in the fixed-partnership setup. Note that unconditional defection with or without breakup is always sustainable as a SPED, denoted by σ_{ED}, that yields the initial lifetime payoff $v_{\mathrm{ED}} = 0$ on the equilibrium path of h_{DD}.

Still, the ideal social convention which people intuitively associate with PD is to sustain mutual cooperation forever, after some periods of feeling each other out. Indeed, a feature common to almost all known desirable equilibria in the literature is that agents eventually reach the state of *eternal cooperation* h_{CC}. Once there, a player's expected lifetime payoff becomes $v_{CC} = 1 + \delta + \delta^2 \ldots = (1 - \delta)^{-1}$. For h_{CC} to be sustainable, $v_{CC} \geq v_0$ is a necessary condition to ensure no breakup. On the other hand, deviating to defection yields $1 + G$ in the current period, plus no less than the expected lifetime payoff from re-joining the strangers' pool next period, δv_0. Thus, the condition of no deviation further requires $v_{CC} \geq 1 + G + \delta v_0$. It implies that in the presence of *eternal cooperation* the initial lifetime payoff v_0 must be bounded from above by

$$v_{\mathrm{EC}} = (1 - \delta)^{-1} - G/\delta \tag{1}$$

---

[9] As a matter of fact, any specific description of a Markov strategy starts with a complete transition rule, as it is oftentimes infeasible to start with an explicit partition of H.

Lemma 1: The initial lifetime payoff of any Nash equilibrium distribution with *eternal cooperation* on the path cannot exceed v_{EC}.

Another conceivable simple and symmetric outcome path as social convention is when the partners coordinate in alternately playing C or D, i.e., h_{CD} for one person and the matching mirror history of $h_{DC} = (DC, h_{CD})$ for the partner. The continuation payoffs associated can be easily calculated as $[\delta(1 + G) - L]/(1 - \delta^2)$ and $(1 + G - \delta L)/(1 - \delta^2)$, respectively. We call this state *eternal alternation*. Assuming there is no incentive to break up on this path, it is trivial that the initial lifetime payoff v_0 must be bounded from above by the smaller of these two numbers, namely,

$$v_{\text{EA}} = [\delta(1 + G) - L]/(1 - \delta^2) \tag{2}$$

Lemma 2: The initial lifetime payoff of any Nash equilibrium distribution with *eternal alternation* on the path cannot exceed v_{EA}.

In the following we characterize two classes of Markov strategies that are sustainable as SPEDs, one with eternal cooperation and the other with eternal alternation on the equilibrium path, and achieve the highest possible lifetime expected payoffs as given by Lemmas 1 and 2, respectively.

Let us first discuss when the payoff upper bound v_{EC} can be achieved in a SPED. Consider a 6-state set $\Lambda = \{A_0, A_{DD}, A_{DC}, A_{DC'}, A_{CD}, A_{CC}\}$ with the following actions and transition rules. A_0 is the initial state. Players always return to this state after any partnership breakup or after two defecting partners at A_0 agree to stay together. The outcome of the PD game determines the transition to subsequent states. There, each player first makes the decision on whether to stay, and then on whether to cooperate in the case where the partnership stays intact. If the partner cooperates, defecting players in state A_0/A_{DD} transition to state $A_{DC'}$. In any other state, defectors stay in that state if both defect. Otherwise, players transition to state A_m where $m = DD, DC, CD,$ or CC is the current stage-PD outcome of any state. Players play C with the probabilities $x_0 = 1 - G/[\delta(1 + G)]$ and $x_{CD} = 1 - G/[\delta(1 + L + G - y_{DC}G)]$ in the states A_0 and A_{CD}, respectively. Players choose to stay with probabilities $y_1 = L/G$ and $y_{DC} \in [y_1, 1]$ in the states $A_{DC'}$ and A_{DC}, respectively. In all other situations, players never leave and always cooperate. Fig. 1 is a flow diagram illustration of this strategy and its transition rules. Arrows denote events that can happen on the (equilibrium) play path. Dashed lines represent transitions due to partnership breakups, while solid ones denote transitions of players that agreed to stay and played the PD game with their current counterparts. Off-
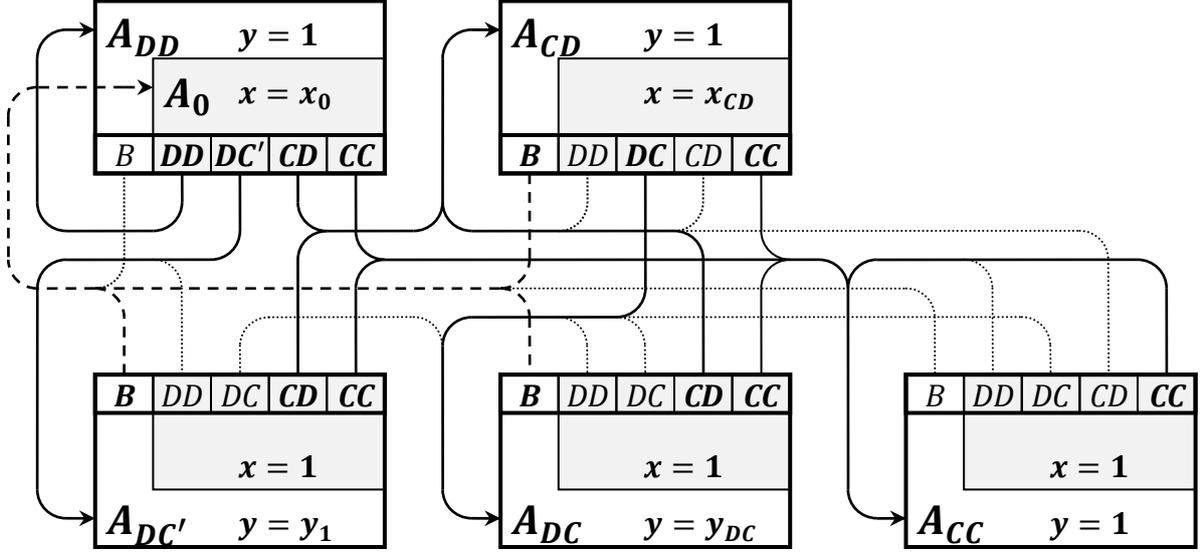
Fig. 1. Eternal Cooperation Markov Strategy σ_{EC}: Each rectangle displays a state of the game with prescribed equilibrium actions (y, x). Boldface letters denote outcomes on the equilibrium path. Dashed arrows signal that breakup and return to the initial state A_0 occurs with a positive likelihood. Solid lines with arrows show state transitions that follow the outcomes of the PD game. Dotted ones signal off-equilibrium transitions.

equilibrium state transitions follow the dotted lines. Note that if a player is in state A_{CD}, then his partner must be in state $A_{DC'}$ or A_{DC}, and vice versa. Once mutual cooperation emerges, both players enter state A_{CC} and cooperation thus lasts forever. We denote this particular class of Markov strategies with *eternal cooperation* as σ_{EC}.

Proposition 1: (1) If $\delta < G/(1+G)$, i.e. $v_{\text{EC}} < 0$, then no SPED with eternal cooperation on the path exists. (2) If $\delta \geq G/(1+L)$, and $\delta \geq G/(1+G)$, then $v_{\text{EC}} = \max\{v_{\text{EC}}, v_{\text{EA}}, v_{\text{ED}} = 0\}$. (3) If $G \geq L$ and $\delta \geq G/(1+L)$, then there exists a Markov strategy with *eternal cooperation* σ_{EC} that is a SPED and yields v_{EC}.

Proof of Proposition 1: To prove (1) recall that any player can always choose to defect and thus earn a lifetime payoff of at least 0. Therefore, in any equilibrium each player must earn at least that much. (2) is straightforward by comparing v_{EC} with v_{EA}.

Now, we prove (3). First, let v_0 denote the expected payoff at the initial state A_0 and v_m with $m \in \{CC, CD, DC, DC', DD\}$ denote the expected payoff at the respective states conditional on the two players staying together to play the PD game. Along the equilibrium path, we have them expressed as follows:

$$v_{CC} = 1 + \delta v_{CC} \tag{3}$$

$$v_{CD} = x_{CD}(1 + \delta v_{CC}) + (1 - x_{CD})[1 + G + y_{DC}\delta v_{DC} + (1 - y_{DC})\delta v_0] \tag{4}$$

$$v_{DC} = v_{DC'} = x_{CD}(1 + \delta v_{CC}) + (1 - x_{CD})[\quad -L + y_{DC}\delta v_{CD} + (1 - y_{DC})\delta v_0] \tag{5}$$

$$v_0 = v_{DD} = x_0^2(1 + \delta v_{CC}) + x_0(1 - x_0)[\quad -L + y_1\delta v_{CD} + (1 - y_1)\delta v_0] +$$
$$+ (1 - x_0)^2\delta v_{DD} + (1 - x_0)x_0[1 + G + y_1\delta v_{DC'} + (1 - y_1)\delta v_0] \tag{6}$$

Note that equation (3) yields $v_{CC} = (1 - \delta)^{-1}$. The equilibrium probabilities to cooperate and to stay, x_0, x_{CD}, y_1 and y_{DC}, as specified in the description of σ_{EC} are chosen so that equations (4), (5) and (6) imply that $v_{CD} = v_{CC}$ and $v_0 = v_{DD} = v_{DC} = v_{DC'} = v_{CC} - G/\delta = v_{EC}$. It is straightforward to check that for any $y_{DC} \in [y_1, 1] = [L/G, 1]$ all four parameters fall within the interval [0,1], which in turn makes σ_{EC} well defined. To prove that the strategy specified in this way constitutes a SPED, by the logic of dynamic programming, we need to check that no player at any state has an incentive to do a one-time deviation from the assigned action.

First, note that as $v_0 = \min\{v_0, v_{DD}, v_{DC}, v_{DC'}, v_{CD}, v_{CC}\}$, no player strictly prefers to exit her partnership. Furthermore, as $v_0 = v_{DC} = v_{DC'}$, players in states A_{DC} and $A_{DC'}$ are indifferent about whether to leave their counterparts. This guarantees the optimality of all prescribed stay/leave choices. Next, consider players' incentives to cooperate. Playing D yields $1 + G + \delta v_{DC} = (1 - \delta)^{-1} = v_{CC} = v_{CD}$ in both states A_{CC} and A_{CD}, $x_0(1 + G + \delta v_{DC'}) + (1 - x_{CD})(\delta v_{DD}) = v_0 = v_{DD}$ in both states A_0 and A_{DD}, and $x_{CD}(1 + G + \delta v_{DC}) + (1 - x_{CD})(\delta v_{DC}) = v_{DC} = v_{DC'}$ in both states A_{DC} and $A_{DC'}$. Agents are thus always indifferent as to whether to defect or to follow the prescribed C/D choices. To summarize, given x_0, x_{CD}, y_1, and y_{DC} as described and the Markov strategy using these parameters, nobody has a strict incentive to deviate to different actions at any state. This proves that this Markov strategy is a SPED. Moreover, $v_0 = v_{EC}$. □

The novel device to achieve the optimal payoff with eternal cooperation, in actual fact, is to build in a transitional loop of alternations between C and D. This transition allows for differentiated treatment of two players who have been choosing different actions in the previous round. In this way, we manage to find the proper amount of mixing between C and D, as well as between stay and leave at different states, to achieve $v_0 = \min_m v_m$ such that $1 + G + \delta v_0 = v_{CC}$. If the partnership lasts forever, then with probability one it will end up in an *eternal cooperation* continuation path, i.e., A_{CC} is the (only) absorbing state.

As is apparent from the proof, there is a lot of flexibility within the class of σ_{EC}, both on and off the equilibrium path, to sustain v_{EC}. On the equilibrium path, the design works with any $y_{DC} \in [y_1, 1]$, as explicitly stated in Corollary 1 below. An increase in the leave likelihood of the defector of the previous round increases the chance that he is forgiven, which also shortens the expected length of the transitional loop of CD/DC alternation, but increases the probability that the partnership dissolves before the mutual cooperation is established. Note that, with $y_{DC} = 1$, the partnership never breaks up once on the CD/DC path.

Corollary 1: Given that $y_1 = L/G$ and $x_0 = 1 - G/[\delta(1 + G)]$ are fixed, σ_{EC} is a SPED that yields v_{EC} for any $y_{DC} \in [y_1, 1]$ and $x_{CD} = 1 - G/[\delta(1 + L + G - y_{DC}G)]$.

Off the equilibrium path, continuations may vary greatly without destroying the SPED property. The specific form of σ_{EC} as illustrated in Fig. 1 is chosen to put them back into the same state within the partnership as much as possible, as if to make them correct their own mistakes in the very same situation in case partners went astray from the equilibrium path. Note that the partnership is intact following the (equilibrium) outcome of DD in state A_0, even though the players are supposed to go back to state A_0, where they will behave in the same way as if they were strangers.

Fig. 2 illustrates the content of Proposition 1 in the parameter space $\{G, \delta\}$ with $L = 1/3$ fixed. Below the curve $v_{EC} = 0$, the combination of a small discount factor δ and a large payoff from unilateral defection $1 + G$ prevents mutual cooperation from being sustainable in any subgame. Right from the line $v_{EC} = v_{EA}$, alternation dominates cooperation payoff-wise due to the relatively high G that increases v_{EA} but reduces v_{EC}. The $G = L$ line further divides the remaining area into the *submodular* region $\bar{C}$, the region with the darkest shade, and the *supermodular* region $\bar{C}^*$, the region with the medium shade. Our σ_{EC} matches the maximal possible lifetime payoff v_{EC} for *all* parameters inside region $\bar{C}$.

Note that when $G < L$, like in region $\bar{C}^*$, the PD game becomes *supermodular* in terms of Takahashi (2010), or *defensive* in terms of Dixit (2003), in that the more likely a player is to cooperate, the smaller the incentive his partner has to defect.[10]

---

[10] In a random matching PD context with first-order information, Takahashi (2010) shows that it is a lot harder to sustain the all-cooperation population equilibrium if $L \leq G$, i.e. the payoff is *submodular*, with *strict* equilibrium. In a model with two populations, Dixit (2003) finds that under this aggressive condition, *honesty* in one population may induce *dishonesty* in the other. Similarly, in our model the $L \leq G$ condition makes it more tempting to exploit one's counterpart as the cooperation level increases.
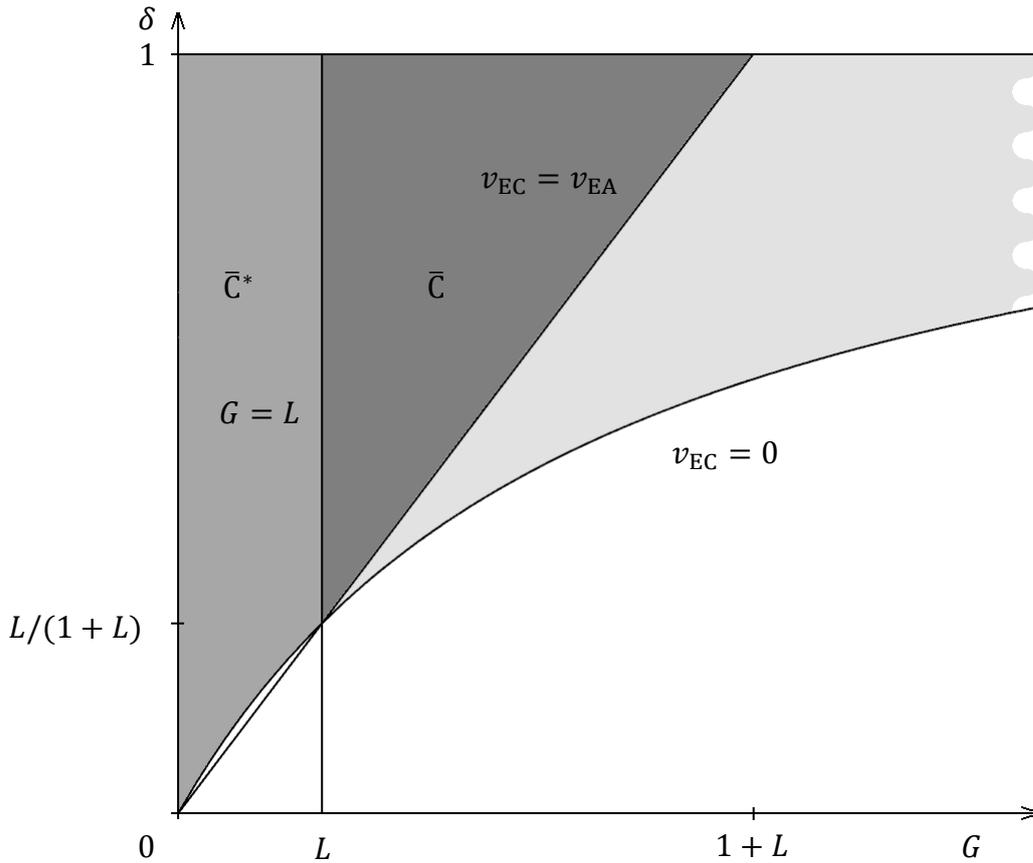
Fig. 2. Parameter Regions $\bar{C}$ and C^* with $L = 1/3$.

Our strategy σ_{EC} as specifically constructed in the proof of Proposition 1 is not well defined in region $\bar{C}^*$, as the parameter $y_1 = L/G$ cannot be smaller than 0. This is because a relatively small G diminishes the benefits of the CD/DC alternations below the level that can sufficiently compensate the cooperating stranger facing defection so that her expected payoff matches the upper bound v_{EC}. Without this optimality requirement, however, this basic structure can generate SPEDs in the region $\bar{C}^*$ as well. For example, by setting $y_1 = 0$ we easily obtain a class of SPED called *cautious* in Section 4. It is, however, an open question whether there are other social conventions of simple symmetry that achieve the payoff upper bound v_{EC}. In the next section we will show that the delayed-cooperation SPED by Fujiwara-Greve and Okuno-Fujiwara (2009) can indeed do so for a degenerate set of parameters. Here, however, their class of strategy can be used in Proposition 2 below to show that the highest-payoff social convention with simple symmetry in region $\bar{C}^*$ must always be a SPED with eternal cooperation.

Proposition 2: If $G \leq L$, and $\delta > G/(1+G)$, then there exists a SPED with eternal cooperation that has a higher initial lifetime payoff than any Nash equilibrium with eternal alternation, i.e., with $v_0 \geq v_{EA}$.

Proof: Divide region $\bar{C}^*$ into an infinite number of partitions $\bar{C}_N^*$, $N \in \{1,2,\dots\}$ such that $\bar{C}_N^* = \{(\delta, G, L) \text{ s.t. } L > 0, \delta \in (0,1), \text{ and } G \in (0,L) \cap (\gamma(N-1), \gamma(N)]\}$, where $\gamma(j) = \sum_{i=0}^{j} \delta^i - 1$. In each partition $\bar{C}_N^*$ there exists a *delayed cooperation* equilibrium with N periods of defection, as previously described in Fujiwara-Greve and Okuno-Fujiwara (2009). In such an equilibrium, all players defect during the initial N periods and cooperate thereafter; players leave if and only if their counterparts deviate from this pattern. It yields an initial lifetime payoff $\delta^N/(1-\delta) = [\delta + \delta\gamma(N) - \gamma(N-1)]/(1-\delta^2)$. Because $G \in (\gamma(N-1), \gamma(N)]$, thus $\delta^N/(1-\delta) > (\delta + \delta G - G)/(1-\delta^2)$. Finally, as $G \in (0,L)$ the right-hand side is greater than v_{EA}. $\square$

Let us now discuss when the payoff upper bound v_{EA} can be achieved in a SPED. Consider a 5-state set $\Lambda = \{A_0', A_{DD,CC}', A_{CD}', A_{DC}', A_{DC'}'\}$ with the following actions and transition rules. A_0' is the initial state and players also transition to this state in the case of a partnership breakup. If players do not break up, with outcomes CD/DC, cooperators always transition to state A_{CD}', while their defecting counterparts transition to state A_{DC}' ($A_{DC'}'$) from states A_0' and $A_{DD,CC}'$ (otherwise). Players that defected or cooperated mutually stay in their previous round states or, in the case of being in the initial state A_0', they transition to state $A_{DD,CC}'$. Fig. 3 is a flow diagram illustration of these transitions. Players always defect in state A_{CD}' and always cooperate in states A_{DC}' and $A_{DC'}'$. In states A_0' and $A_{DD,CC}'$ players play C with a probability $x_0' = [\delta(1+G) - L]/[(1+\delta)(1+G)]$. In state $A_{DC'}'$ players stay with probability $y_1' = [\delta G(1+G) + L(1+L)](1+\delta)/[\delta(1+G+L)^2]$. This is the only situation where players voluntarily leave the partnerships; in all other cases they always stay. Note that once the two players find themselves in states A_{CD}' and A_{DC}', then the alternation between CD and DC continues forever. We denote this particular class of Markov strategy with *eternal alternation* σ_{EA}.
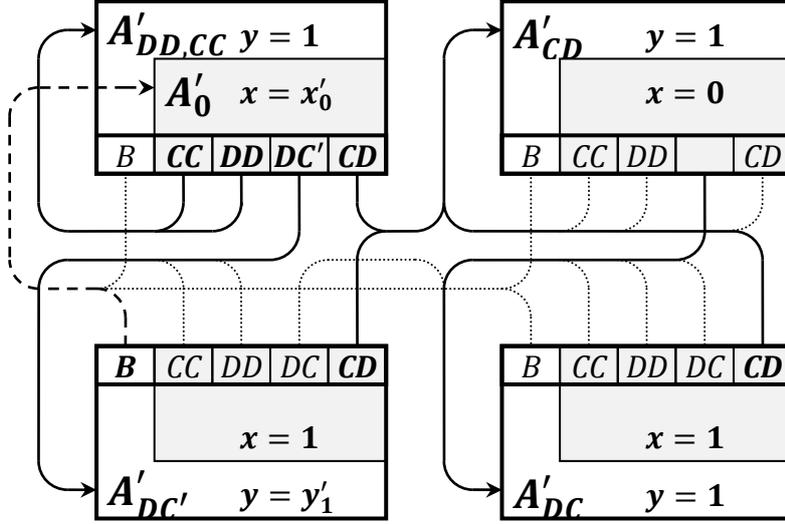
Fig. 3. Eternal Alternation Markov Strategy σ_{EA}

Proposition 3: (1) If $\delta < L/(1 + G)$, i.e., $v_{\text{EA}} < v_{\text{ED}}$, then no SPED with eternal alternation on the path exists. (2) If $\delta \leq G/(1 + L)$ and $\delta \geq L/(1 + G)$, then $v_{EA} = \max\{v_{\text{ED}}, v_{\text{EC}}, v_{\text{EA}}\}$. (3) If $\delta \in [L/(1 + G), \min\{G/(1 + L), (1 + L)/G\}]$, then there exists a Markov strategy with *eternal alternation* σ_{EA} that is a SPED and yields v_{EA}.

Proof of Proposition 3: Proofs to (1) and (2) are straightforward. To prove (3), let v'_m denote the expected payoff along the PD stage games on the equilibrium path:

$$v'_{CD} = 1 + G + \delta v'_{DC} \tag{7}$$

$$v'_{DC} = v'_{DC'} = -L + \delta v'_{CD} \tag{8}$$

$$v'_0 = v'_{DD} = {x'_0}^2(1 + \delta v'_{DD}) + x'_0(1 - x'_0)[\quad -L + y'_1 v'_{CD} + (1 - y'_1)v'_0] + $$
$$+ (1 - x'_0)^2 \delta v'_{DD} + (1 - x'_0)x'_0[1 + G + y'_1 v'_{DC'} + (1 - y'_1)v'_0] \tag{9}$$

Equations (7) and (8) yield the payoffs from eternal alternations: $v'_{DC} = v'_{DC'} = v_{\text{EA}}$ and $v'_{CD} = [1 + G - \delta L]/(1 - \delta^2)$. The equilibrium probabilities x'_0 and y'_1 are chosen so that in state A'_0 cooperating and defecting players earn the same: $x'_0(1 + \delta v'_{DD}) + (1 - x'_0)[-L + y'_1 v'_{CD} + (1 - y'_1)v'_0] = (1 - x'_0)\delta v'_{DD} + x'_0[1 + G + y'_1 v'_{DC'} + (1 - y'_1)v'_0]$; and so that a defector facing a cooperator is indifferent between starting a new partnership in the state A'_0 and continuing the partnership in states $A'_{DD,CC}$, A'_{DC} and $A'_{DC'}$: $v'_0 = v'_{DD,CC} = v'_{DC} = v'_{DC'} = $
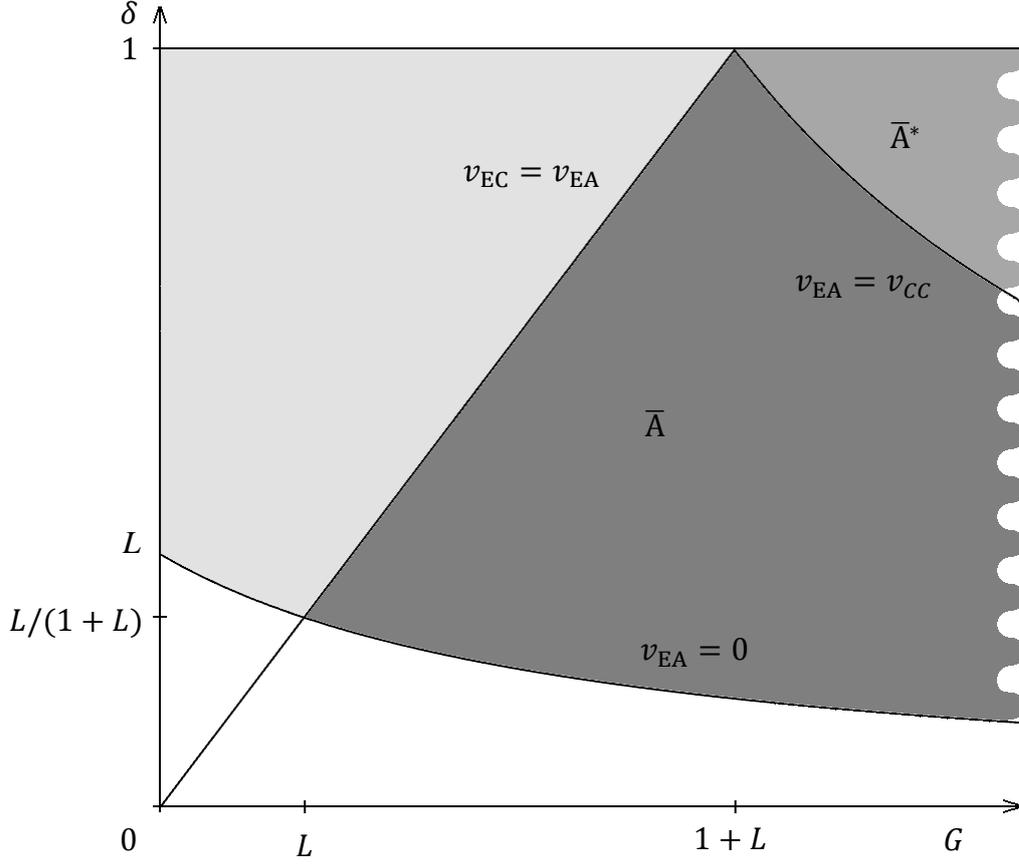
Fig. 4. Parameter Regions $\overline{A}$ and $\overline{A}^*$ with $L = 1/3$.

v_{EA}. Our strategy σ_{EA} exists only if $x_0' \geq 0$ and $y_1' \geq 0$. These conditions are satisfied if $\delta \leq L/(1 + G)$ and $\delta \leq (1 + L)/G$, respectively.

We complete the proof that σ_{EA} is a SPED by showing that no player has an incentive to deviate from the eternal alternation path in the remaining states. In states A_{DC}' and $A_{DC'}'$ a player does not benefit from playing D, as the resulting mutual defection yields her a zero payoff and only delays her lifetime payoff v_{EA} for a period. A similar argument applies to deviation to C in state A_{CD}'. If $\delta \leq G/(1 + L)$, then securing the mutual cooperation payoff of 1 is no better than a period equilibrium payoff of $(1 - \delta)v_{CD}'$. Finally, $v_0' = \min\{v_0', v_{DD,CC}', v_{DC}', v_{DC'}', v_{CD}'\}$ implies that players do not ever benefit from leaving their current partnerships. It follows that σ_{EA} is a SPED with $v_0' = v_{\mathrm{EA}}$. $\square$

Fig. 4 illustrates Proposition 3 with $L = 1/3$ fixed. Line $v_{\mathrm{EA}} = 0$ separates the bottom region where the small discount factor δ discourages a player from cooperating against a defector in the hope of being rewarded later on. To the left of the $v_{\mathrm{EC}} = v_{\mathrm{EA}}$ line, *eternal*

cooperation payoff-wise dominates just as in Fig. 2. The remaining top-right region is further divided by line $v_{EA} = v_{CC}$. Our EA strategy σ_{EA} matches the maximal possible lifetime payoff v_{EA} for all parameters inside the region $\overline{A}$. In region $\overline{A}^*$ in Fig. 4 there does not exist any SPED with *eternal alternation* that yields v_{EA}. This is because the presence of the alternation path requires that the partners must first determine who will cooperate in the odd periods, and who in the even periods. Thus it must be preceded by at least one round of PD in which mixed strategies are used to determine which player will be the defector in the next round. It follows that a player who is willing to cooperate earns at most v_{EA} when matched with a defecting counterpart, and no more than $1 + \delta v_{EA}$ when matched with a cooperating one. If $v_{EA} > v_{CC}$ then $1 + \delta v_{EA} < v_{EA}$ and thus the cooperating player's lifetime payoff can never match the upper bound v_{EA}.

Proposition 4: If $\delta > (1 + L)/G$, then there exists a SPED with eternal alternation that has a higher initial lifetime payoff than any Nash equilibrium with eternal cooperation, i.e., with $v_0 \geq v_{EC}$.

Proof: Consider the following SPED: strangers defect with probability $(1 + G)^{-1}$ and cooperate otherwise; partners defect if they cooperated last period and vice versa. Cooperators stay with defectors but leave cooperators. Defectors stay with cooperators with probability $(1 + L)/(\delta G)$ and leave defectors always. This equilibrium strategy yields the initial lifetime payoff $(1 - \delta)^{-1}$ which is greater than v_{EC}. □

Note, that this strategy looks structurally similar to our σ_{EA} as illustrated in Fig. 3. The difference is that once the CD/DC phase starts in our current best EA strategy, then it continues forever. In the strategy of Proposition 4, the current defector (cooperator to-be) always leaves with a constant positive probability and thus the alternation is with positive probability not eternal. In fact, without this modification, σ_{EA} in its originally specified form would not be well-defined. In this equilibrium, the payoff from CD/DC alternation is strictly smaller than v_{EA}. Combining all the above results, we straightforwardly have the following conclusion.

Theorem 1: Aside from the parameter regions $\delta > G/(1 + G)$ for $G < L$ and $\delta > (1 + L)/G$ for $G > 1 + L$ (regions $\overline{C}^*$ and $\overline{A}^*$ in the figures), the payoff upper bound $v_{max} = \max\{0, v_{EC}, v_{EA}\}$ for *social conventions of simple symmetry* can always be realized, with either σ_{EC}, σ_{EA}, or σ_{ED}.
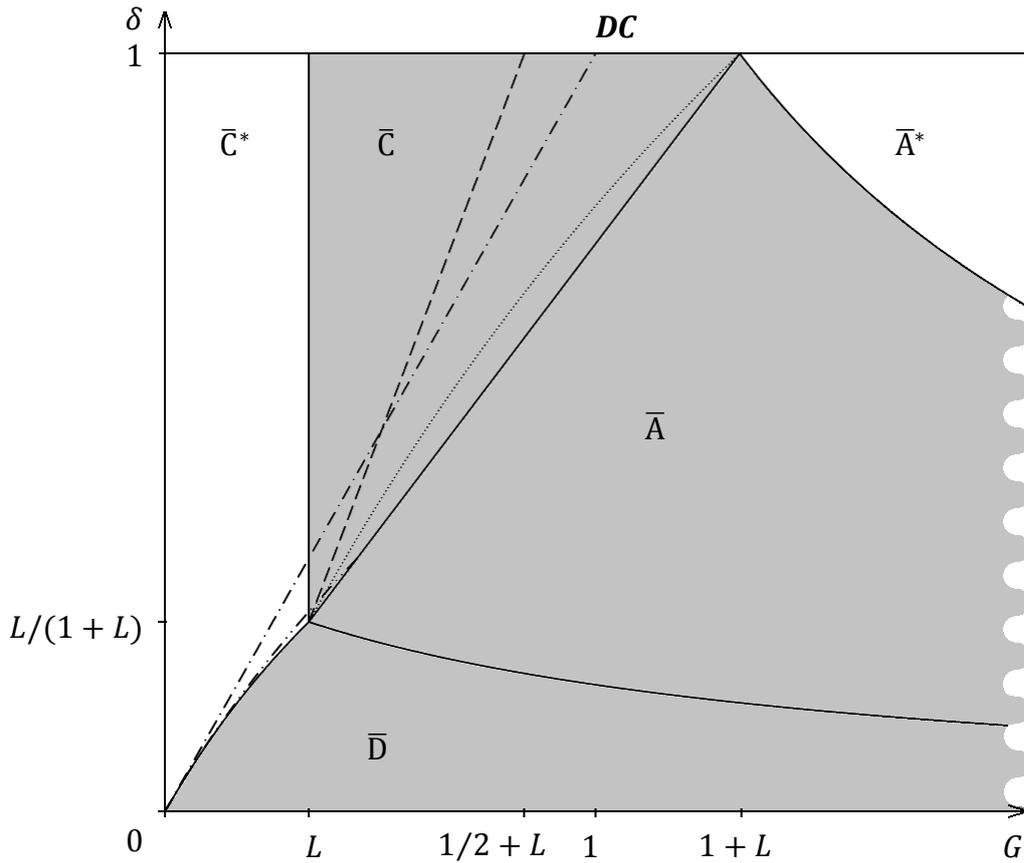
Fig. 5. Parameter Regions $\overline{C}$, $\overline{A}$ and $\overline{D}$ with $L = 1/3$.

Fig. 5 shows all three regions where equilibrium strategies σ_{EC}, σ_{EA} and always-defect match the upper bound v_{max} in grey. The additional dashed and dotted lines illustrate the limited sets of parameters where various previously known equilibria do so as well. These equilibria will be briefly discussed in the next section. Note that while in most meaningful interpretations of the PD game the total surplus in CD must be non-greater than that in CC, i.e., $G \le L + 1$,[11] this condition does not guarantee that the optimal SPED must be from eternal cooperation. In fact, for sufficiently low discount rate, the σ_{EA} strategy beats the EC one for optimality even in this condition, because the needed transition phase for the former is much shorter, due to lower off-equilibrium payoffs, as illustrated in the proofs of Prop. 1 and 3.

4. Further Discussions

Comparison with other SPED in the literature. As discussed in details in the **Appendix**, we

---

[11] Characterization of behavior in the area of $G > 1 + L$ that includes the whole $\overline{A}^*$ area only serves the purpose of theoretical completeness.

can characterize the conditions under which each of the well-known SPED in the literature achieves the upper bound for equilibrium payoff. As illustrated in Fig. A in Appendix, this occurs only in a non-generic subset of the parameter conditions where our new classes of SPED achieve payoff optimality. In this sense, our optimality result is a significant improvement in the literature.

Beyond simple symmetry. We speculate intuitively that the established upper bound $v_{\max}$ is also the overall upper bound beyond social conventions of simple symmetry, which can be shown for many specific long-run paths. For illustration, consider the symmetric path $(CD(n), DC(n))$ where the cycle develops around first playing CD, and then DC, for n consecutive times each. Following the proof logic for Lemma 2, the binding condition for no deviation is given by the willingness to stay (and cooperate) when the player expects to wait out the maximal n periods of defections by the counterpart. This yields $v_{EA(n)} = \frac{(1-\delta^{n+1})(\delta^n+\delta^n G-L)}{(1-\delta)(1-\delta^{2n})}$, which reaches its maximum at $n = 1$. For general paths, we are yet to find a generic formal proof.

In addition, outside the parameter region stated in Theorem 1 where our EC or EA strategies work, it is unclear whether $v_{\max}$ can be realized, beyond a degenerate set of parameters where delayed cooperation does it (see Proposition 5). It is left for future work to explore what would be the "second best" there.

Indifference, belief free equilibrium, and robustness. One unique feature apparent from our proofs is that players are indifferent between actions C and D in all states under σ_{EC}. In addition, aside from the CC and CD states, the players are also indifferent between staying and leaving. In this sense, our σ_{EC} is reminiscent of the so-called belief-free equilibrium in the fixed partnership PD paradigm with public or private randomization (e.g., Ely et al., 2005) or community information (Takahashi, 2010).

In fact, since in the case of deviation our EC/EA strategies require the players to stay and return to the previous state, they can easily handle public random noises to the PD actions: they do not need to know whether the counterpart intentionally deviated from the SPE since there is no difference between on and off equilibrium behavior, given the specified state transition rules. In other words, their actions are independent of any belief that can be derived from past play histories or even whether they are on or off the equilibrium path. In this spirit, we conjecture that our EC/EA strategies still constitute SPE once the noise to actions is introduced in an explicit model setup.

Breakup and neutral stability One seminal contribution by Fujiwara-Greve and Okuno-Fujiwara (2009) is to introduce the notion of *Neutrally Stable Distribution* in the voluntary continuation PD paradigm. Vesely and Yang (2010) generalize their definition to include invasion with mixed strategies and demonstrate that their bimorphic delayed cooperation SPED is no longer neutrally stable. It turns out that a seemingly innocuous breakup such as off-equilibrium punishment in traditional trigger strategies may invite secret-handshake types of invasions that undermine the equilibrium in an evolutionary environment. This insight can be interpreted as evolutionary justification for designs of social conventions that adhere to the maxim "the grass is not always greener on the other side".

Randomization and cheap talk. One intriguing idea to solve the optimal SPED problem may be the following. Assume there is an external, public randomizing device to coordinate actions in the partnership, so that both simultaneously play C with probability x as well as D with $1 - x$ based on realizations of the device. If the device signals C, the state transitions to EC right away and forever ignores the device. If the signal is D, then both play D and restart the same process as if they were strangers. It can be shown that a randomized setup like this may indeed generically sustain v_{EC}, say, with x^*, and nobody has an incentive to deviate if they have no other alternative but to blindly trust this external device. Note, however, for any randomization device $x \in (0, x^*]$ sustained this way in the population, the players in a partnership have strict incentive to agree to switching to any alternative device $x' \in (x, 1]$. The reason is that for x to be sustainable nobody shall have incentive to deviate to defection *after* realization of the C signal, which is fully determined by the expected payoff in the singles' pool calculated under the assumption that all other people adhere to this x device, but independent of the rate for the C-signal in the specific device used within one's partnership. If they had the liberty to opt for another device, they would have a strong incentive to abandon the suggested one. However, if everybody is free-riding the system in this way, it collapses. In other words, no public randomization equilibrium x is robust if private cheap talk is allowed, as it is in real applications. Our EC and EA conventions are robust against this type of cheap talk, in contrast.[12]

Renegotiation and bilateral rationality. The above discussion on cheap talk naturally leads to the issue of renegotiation-proof equilibrium. Note that none of the SPEDs discussed in the

---

[12] Note that allowing for cheap talk in a static game selects the Pareto frontier among all Nash or correlated equilibria. But this is not the relevant issue here, as any correlation device x fails to survive the simple and naïve cheap talk exercise. And this problem is not related to renegotiation-proofness issues discussed subsequently. The application of randomized device as suggested to the VCPD setup suffers a serious consistency problem.

binary-action VCPD models in the literature satisfies the *bilateral rationality* condition, which Ghosh and Ray (1996) imposed on their *social equilibrium* and requires renegotiation-proofness on the equilibrium path. Since *DD* is on the equilibrium path of every SPED that sustains mutual cooperation in the long run, if the partners were to renegotiate and decide to skip the painful trust-building rituals and to switch to the EC path right away, nobody would have an incentive to deviate from this new agreement just as the partners would not individually deviate from the EC path in the equilibrium in the first place. This line of reasoning is, however, undermined if there are a certain number of myopic players in the singles' pool as assumed in Ghosh and Ray (1996). Assuming that such social equilibrium existed in binary-action setups, the efficiency it achieves would still be bound by $v_{\max}$ that we derived here.

There are several reasons why we prefer not to impose this condition on our solution. First, our setting, like others in the binary-action framework, has an evolutionary underpinning, where players can be considered automata happily following simple rules – the social convention. Second, allowing people to be sophisticated and to renegotiate based on the common knowledge of mutual sophistication opens up a whole world of indeterminacy that is hard to deal with. If one opens the negotiation with "let bygones be bygones, and the grass is not always greener on the other side", then the same line can be used again by the person who cheats on the partner in the first period of their new agreement. Furthermore, by the same logic based on which they deviated from the initial universal social convention, the sucker would again buy in. In the end, there would be no consistently convincing argument to justify their initial deviation from the norm in the first place. In some sense, the very same maxim of "the grass is not always greener on the other side" that is behind our EC solution is only cheap talk in this context. Sometimes, you just have to cut your losses.

This being said, our optimal EC equilibrium is constructed so that the sucker cannot be strictly swayed by such cheap talk on or off the equilibrium path, as her prescribed equilibrium strategy already yields the same payoff as on the EC path. The logical indeterminacy occurs only in the case of *DD*.

5.   Concluding Remarks

A maximally sustainable level of efficiency is one of the primary issues, if not *the* very issue, to be solved in theoretical models in economics. In the repeated Fixed-Partnership PD paradigm, a special case of the voluntary continuation one, full efficiency of mutual

cooperation forever can be achieved in SPE. One concern associated with Folk theorem types of results is to find the lower bound of the discount rate, so that all-C can still be sustained. Abreu (1988), for example, looks into the *optimal punishment strategy* that can sustain the lowest discount rate for this purpose. As a more natural and realistic extension of the FPPD, however, the voluntary continuation paradigm has the distinctive feature that full efficiency is never sustainable and that there is a glaring theoretical gap on the systematic characterization of a maximally achievable level of efficiency. The main culprit can be found in the fact that the feature of unilateral breakup is a double-edged sword. It can be used by honest partners as a means of punishment of the norm deviator, as well as by the deviator as a path to evade the wrath of such betrayed partners.

We offer a first attempt to partially close this gap in the present paper, under the binary-action setting. By focusing on SPEDs that sustain the intuitively appealing long-run paths, we first establish an upper bound for any sustainable lifetime payoff. We then show that this upper bound can be achieved for a large parameter set, and particularly in all but an extreme region of submodular conditions, with surprisingly simple behavior rules such as our Markov strategies σ_{EC} and σ_{EA}. In the spirit of Abreu (1988), we implicitly employ some kind of optimal punishment strategy. The novel feature in our design of optimal social convention is to keep the partnership as long as possible by giving the exploiter of the previous round a chance to repent and to make good with unconditional cooperation in the next round. It is interesting to note that once the novel behavioral patterns of our EC and EA equilibrium are pinned down with the help of the associated flow diagrams, the actual proofs are technically rather simple. As a byproduct of the strategy design, these optimal equilibria are also "belief free" and robustly applicable in a noisy condition.

Appendix: Comparison with well-known SPED in the Literature

In the main text of the paper, we characterized two classes of Markov SPED that achieve the highest possible lifetime payoff among all SPED that either have eternal mutual cooperation CC or eternal alternation of CD/DC on the path. Fig. A illustrates the lifetime payoffs on an example where $L = 1/3$ and $\delta = 3/4$. The upper bounds v_{EC} and v_{EA} are the two straight dashed lines that intersect at $G = \delta(1 + L)$. The lifetime payoff for our EC strategy is represented by a solid line v_{EC} for $G \in [L, \delta(1 + L)]$, and that for our EA strategy is v_{EA} for $G \in [\delta(1 + L), (1 + L)/\delta]$.

In the Appendix, we wish to compare these with SPEDs that the recent literature has discussed. It turns out that all relevant equilibrium strategies ever discussed in the literature can be captured with the following class of Markov strategies,[13]

$$\left\{ x_0; (x_m : y_m)_{m \in S_a \times \mathcal{T}} \right\} \in [0,1] \times ([0,1] \times [0,1])^{S_a \times \mathcal{T}}$$

with $\mathcal{T}$ denoting some partition of the time space. Table A1 summarizes all those strategies for comparison illustrated in Fig. A. There are four classes of known equilibria aside from ours that sustain cooperation to some level.

The first well-known class is called the *delayed cooperation* equilibrium, which has been analyzed by Fujiwara-Greve and Okuno-Fujiwara (2009) in great detail. In monomorphic form, this requires players in any partnership to defect $N > 0$ periods first, before simultaneously switching to *eternal cooperation* thereafter. For $N = 1$, this is also Eeckhout's (2006) equilibrium $(0, (0,0,1))$. In its polymorphic form, there are multiple types of players in the population. Type i players always defect until period N_i, and switch to the cooperation thereafter. If two players of different types i and j, where $i < j$, are matched in a partnership, the round N_i outcome of CD is followed by players' decisions to break up the partnership.

---

[13] Note that finer partition of the history space enables direct comparison of all existing SPED strategies in the literature. However, the necessary number of states for each SPE strategy is generally smaller than indicated in this formulation. For example, behavior variation in the delayed cooperation equilibrium by Fujiwara-Greve and Okuno-Fujiwara (2009) can be captured with 4 instead of 25 states as indicated here. In fact, had we started with this more encompassing form of optimal EC strategy, the task of checking equilibrium conditions for 17 states, instead of 6 as in Fig. 1, might have proven to be too overwhelming for us to achieve Prop. 1. In addition, our preferred version of σ_{EC} where partners return to previous equilibrium states once they find themselves off the path, is infeasible in this form of presentation.

Table A1: **SPED strategies in the literature**. B = leave, and play D if one mistakenly stayed. In the case of a single number for an outcome state, it refers to $p(C)$ with $p(\text{Stay}) = 1$ being omitted.

$p(C):p(\text{Stay})$	h_0	DD	CD	DC	CC	t	SPE payoff
Optimal EC	x_0 $-$	B	x_{CD}	$1:y_1$ $1:y_{DC}$	1	$=1$ >1	v_{EC}
Optimal EA	x'_0 $-$	B	0	$1:y'_1$ 1	B	$=1$ >1	v_{EA}
Delayed cooperation	0 $-$ $-$	0 1 B	B	B	B B 1	$<n$ $=n$ $>n$	v_D^n
Cautious	x_0^G, x_0^B	B	B	B	1	$-$	v_C^G, v_C^B
Forgiving	x_0^F $-$	B	1 B	1 B	1	$=1$ >1	v_F
Anti-symmetry	x_0^A $-$	B	1 B	1 B	0 1	$=1$ >1	v_A

The lifetime payoff of the monomorphic *delayed cooperation* equilibrium with N periods of delay equals $v_D^N = \delta^N/(1-\delta)$. Thus, as the two horizontal dotted lines v_D^1 and v_D^2 in Fig. A illustrate, v_D^N can only match the upper bound v_{EC} if $G = \delta + \delta^2 + \ldots \delta^N$. For example, v_D^2 is the best monomorphic equilibrium at $G = \delta + \delta^2$. Parameters where $v_D^1 = v_{\max}$ and $v_D^2 = v_{\max}$ are represented in Fig. A by dot-dashed and double-dot-dashed lines that start at $(G, \delta) = (0,0)$. Being a *pure* strategy, the monomorphic *delayed cooperation* strategy can match the v_{EC} even inside the supermodular region $\overline{C}^*$.

Fujiwara-Greve and Okuno-Fujiwara (2009) showed that, if a bimorphic equilibrium with a shorter delay exists, then it yields a higher payoff than the best monomorphic one. The dotted line $v_D^{1,2}$ in Fig. A refers to the bimorphic form of delayed cooperation equilibria with types 1 and 2 and illustrates this phenomenon as its curve slightly cuts into the $G > \delta$ area, where a one-period delay is no longer a SPED. Note that, in doing so, $v_D^{1,2}$ never reaches the level of v_{EC}, and hence cannot show up in Fig. A. Indeed, it can be shown below that no polymorphic delayed cooperation equilibrium can match v_{EC}.

Fig. A. Comparison of Lifetime Payoffs for Equilibria with $L = 1/3$ and $\delta = 3/4$.

Proof of Suboptimality of polymorphic delayed-coopeartion SPED Let N be the number of initial periods of defections of the type that starts cooperation the soonest. The initial lifetime payoff of that type is thus $v_I = \delta^N[x(1-\delta)^{-1} + (1-x)(-L + \delta v_I)]$, where $x \in (0,1)$ is the fraction of the population of this type over the total population. The lifetime payoff of the other types that defect for at least $N + 1$ periods equals $v_{II} = \delta^N[x(1 + G + \delta v_{II}) + (1-x)\delta\tilde{v}]$, where $\tilde{v}$ is the lifetime utility after the $N + 1$ periods of mutual defections. Substituting v_{EC} for v_I and v_{II} implies that $v_{EC} = \delta^N[x(1-\delta)^{-1} + (1-x)(\delta/(1-\delta) - L - G))] = \delta^N[x(1-\delta)^{-1} + (1-x)\delta\tilde{v}]$. Thus, as in polymorphic equilibrium $x < 1$, the assumption that v_I and v_{II} match the optimal v_{EC} implies that $\tilde{v} = (1-\delta)^{-1} - (G + L)/\delta$. Because $\tilde{v}$ is smaller than v_{EC}, then after $N + 1$ periods of mutual defections no player prefers to stay on the assumed equilibrium path. □

The second class called *cautious* equilibria corresponds to Eeckhout's (2006) equilibrium $(\sigma, (1,0,0))$ as well as that of Rob and Yang (2010). Here, players mix in the first period and continue with cooperation only after CC, i.e., cautiously punishing both unilateral

and mutual defections by breaking up the partnership. Fig. A confirms the result by Rob and Yang (2010) that for $G < \delta$ there is always a *good* and a *bad* type of *cautious* equilibrium with payoffs shown as v_C^G and v_C^B. For $G > 0$ they never reach v_{EC} and thus they do not appear in Fig. A.

Proof of Suboptimality of Cautious Equilibria Let $x \in (0,1)$ be the fraction of players that cooperate in the stranger's round. The initial lifetime payoffs of cooperating and defecting players are $v_I = x(1 - \delta)^{-1} + (1 - x)(-L + \delta v_I)$ and $v_{II} = x(1 + G) + \delta v_{II}$ respectively. Assuming that both these payoffs equal v_{EC}, this implies that $x = 1 - G/[\delta(1 + G + L)]$ in the case of the cooperator, but a smaller $x = 1 - G/[\delta(1 + G)]$ in the case of the defector. □

In the third class called *forgiving* equilibria, players mix in the first period and start a mutual cooperation after CC as well as after CD and DC, as if *forgiving* the defection of a single player. An example of this is Eeckhout's (2006) equilibrium $(\sigma, (1,1,0))$. In the fourth class called *anti-symmetric* equilibria, players mix in the first period and start a mutual cooperation only after CD and DC, as if they punish each other for the initial symmetry. An example of this is Eeckhout's (2006) equilibrium $(\sigma, (0,1,0))$. Fig. A shows that the equilibrium payoffs v_F and v_A of both these equilibria can reach the $v_{EC} = v_{max}$. The dashed line in Fig. A shows that the *forgiving* equilibrium matches the upper bound for $\delta = (2G - L)/(1 + L)$; the dotted line shows that the *anti-symmetric* one does so for $\delta = (3G - L)/(2 + L + G)$. The proofs are straightforward and omitted here.

Proposition 5: (1) The monomorphic *delayed cooperation* equilibrium strategy achieves $v_{max}(= v_{EC})$, if and only if $\delta \geq G/(1 + L)$ and $G = \delta + \dots \delta^N$. (2) The *forgiving* equilibrium achieves $v_{max}(= v_{EC})$, if and only if $G \in [L, 1/2 + L)$ and $\delta = (2G - L)/(1 + L)$. (3) The *anti-symmetric* equilibrium achieves $v_{max}(= v_{EC})$, if and only if $G \in [L, 1 + L)$ and $\delta = (3G - L)/(2 + L + G)$. (4) However, parameters satisfying (1), (2) and (3) are degenerate in the sense that they have a measure of zero in the total parameter space.

Proof: (1) Consider the Delayed Cooperation Equilibrium with N periods of defection. The initial lifetime payoffs of all agents is $v_0 = \delta^N/(1 - \delta) = (1 - \delta)^{-1} - 1 - \delta - \dots \delta^{N-1}$. □

(2) Consider the Forgiving Equilibrium. Let $x \in (0,1)$ be the fraction of players that cooperate in the stranger's round. The initial lifetime payoffs of cooperating and defecting players are $v_I = (1 - \delta)^{-1} - (1 - x)(1 + L)$ and $v_{II} = x(G + (1 - \delta)^{-1}) + (1 - x)\delta v_{II}$, respectively. These payoffs equal each other as well as v_{EC} if and only if $x = 1 - G/[\delta(1 + L)]$, $\delta = (2G - L)/(1 + L)$ and $G \in [L, 1/2 + L)$. □

(3) Consider the Anti-Symmetric Equilibrium. Let $x \in (0,1)$ be the fraction of players that cooperate in the stranger's round. The initial lifetime payoffs of cooperating and defecting players are $v_I = x(1 + \delta v_I) + (1 - x)(-L + \delta(1 - \delta)^{-1})$ and $v_{II} = x(G + (1 - \delta)^{-1}) + (1 - x)\delta v_{II}$, respectively. These payoffs equal each other as well as v_{EC} if and only if $x = (G - 2GL)/(3G - L)$, $G \in [L, 1 + L)$ and $\delta = (3G - L)/(2 + L + G)$. $\square$

Compared to these four classes, our Markov strategy with *eternal cooperation* incorporates a lot more fine-tuning to sustain the maximal possible lifetime payoff for social conventions of simple symmetry. In particular, it employs (matched) alternation between cooperation and defection as a form in which the temporary exploiter compensates the temporary sucker to keep the partnership intact, before they inevitably reach the promised land of eternal mutual cooperation. In fact, the EC equilibrium path may endogenously remain for a very long time on the transition path of alternation, getting shorter only when L approaches G. This matches quite well what Lei, Vesely, and Yang (in preparation) observed in a recent experiment, with $(G, L; \delta) = (1, .67; .875)$, that a lot of cooperative partnerships stably build up after CD outcomes.

In our EA strategy agents pursue an arrangement that is superior to the mutual cooperation for given parameters. In general, our optimal EC/EA equilibria induce a lower expected rate of the worst outcome DD during the transition phase, compared to others. Further illustration of behavior characteristics of different equilibria based on numerical examples can be found in Table A2 below.

Table A2. Equilibrium Comparison: Numerical Examples

G	1/3	2/3*	3/4*	751/1000	25/27*	1	16/9
v_{EA}	1.52 [5, 34]	2.10 [5, 36]	2.24 [5, 36]	2.24 [5, 36]	2.54 [5, 37]	<u>2.67</u> [5, 37]	<u>4.00</u> [5, 39]
v_{EC}^1	<u>3.56</u> [81, 8]	<u>3.11</u> [57, 16]	<u>3.00</u> [50, 17]	<u>3.00</u> [50, 17]	<u>2.77</u> [34, 22]	<u>2.67</u> [25, 25]	
$v_{EC}^{y_1}$	<u>3.56</u> [81, 8]	<u>3.11</u> [59, 14]	<u>3.00</u> [54, 15]	<u>3.00</u> [54, 15]	<u>2.77</u> [43, 16]	<u>2.67</u> [39, 17]	1.63 [10,13]
v_C^G	3.28 [71, 11]	2.22 [33, 17]					
v_C^B	0.72 [7, 11]	1.33 [14, 14]					
v_A	3.23 [73, 8]	<u>3.11</u> [67, 8]					
v_D^1	3.00 [75, 0]	3.00 [75, 0]	<u>3.00</u> [75, 0]				
$v_D^{1,2}$	2.27 [51, 6]	2.63 [61, 4]	2.92 [72, 1]	2.97 [74, 0] / 2.94 [73, 1]			
v_F	2.80 [61, 9]	2.78 [57, 9]	2.77 [57, 9]	2.77 [57, 9]	<u>2.77</u> [55, 9]		
v_D^2	2.25 [56, 0]	2.25 [56, 0]	2.25 [56, 0]	2.25 [56, 0]	2.25 [56, 0]	2.25 [56, 0]	

Note $L = 1/3$ and $\delta = 3/4$ are fixed for varying G. v_Y^X refers to discounted lifetime payoffs of different SPEDs as defined above in Section 4. $[p, q]$ refer to discounted average probabilities in %, i.e., expected length $* 100(1 - \delta)$, of mutual (CC) and unilateral (CD) cooperation. Hence, $1 - p - 2q$ and q are the discounted expected probabilities of the DD and DC outcomes on the equilibrium path. v_{EC}^X for Markov strategy σ_{EC} refers to cases of $y_{DC} = 1$ and $y_{DC} = y_1$. In the case of $v_D^{1,2}$ with $G \in (3/4, 4(667 - 15\sqrt{1817})/147)$, there are two different SPEDs. Empty entries refer to no existence of the corresponding type of SPED. The highest numbers are underlined. (*) refers to examples of the degenerate parameters characterized in Proposition 5.

References

Abreu, Dilip, (1988): "On the Theory of Infinitely Repeated Games with Discounting", *Econometrica*, **56**(2), 383-396.

Axelrod, R., (1984): *The Evolution of Cooperation*. Basic Books, New York.

Carmichael L. and B. MacLeod, (1997): "Gift Giving and the Evolution of Cooperation", *International Economic Review*, **38**, 485-509.

Dixit, A., (2003): "On Modes of Economic Governance", *Econometrica*, **71**, 449–481.

Eeckhout, J., (2006): "Minorities and Endogenous Segregation", *Review of Economic Studies*, 73, 31-53.

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

Ely, J. C., J. Horner, and W. Olszewski, (2005): "Belief-Free Equilibria in Repeated Games", *Econometrica*, **73** (2), 377-415.

Fujiwara-Greve, T. and M. Okuno-Fujiwara, (2009): "Voluntarily Separable Repeated Prisoner's Dilemma", *Review of Economic Studies,* **76**, 993-1021.

Ghosh, P. and D. Ray, (1996): "Cooperation in Community Interaction without Information Flows", *Review of Economic Studies*, **63**, 491-519.

Kandori, M., (1992): "Social Norms and Community Enforcement", *Review of Economic Studies*, **59** (1), 63-80.

Kranton, R., (1996): "The Formation of Cooperative Relationships", *Journal of Law, Economics, & Organization* **12**, 214-233.

Kreps, D., P. Milgrom, J. Roberts, and R. Wilson, (1982): "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma", *Journal of Economic Theory*, **27**, 245-252.

Lei, V., F. Vesely, and C.-L. Yang, "Experiment on Voluntary Partnership Prisoner's Dilemma", in preparation.

Mailath, G.L. and S. Morris, (2002): "Repeated Games with Almost-Public Monitoring", *Journal of Economic Theory*, **102**(1), 189-228.

Mailath, G.L. and L. Samuelson, (2006): *Repeated Games and Reputations*, Oxford University Press.

Okuno-Fujiwara, M. and A. Postlewaite, (1995). "Social Norms and Random Matching Games", *Games and Economic Behavior*, **9**, 79-109.

Piccione, M., (2002): "The Repeated Prisoner's Dilemma with Imperfect Private Monitoring", *Journal of Economic Theory*, **102** (1), 70-83.

Rob, R. and H. Yang, (2010): "Long-Term Relationships as Safeguards", *Economic Theory*, **43**, 143-166.

Rosenthal, R. W., (1979): "Sequences of Games with Varying Opponents", *Econometrica*, **47**, 1353-1366.

Shapiro, C. and J. E. Stiglitz, (1984): "Equilibrium Unemployment as a Worker Discipline Device", *American Economic Review*, **74**, 433-444.

Takahashi, S. (2010): "Community Enforcement when Players Observe Partners' Past Play", *Journal of Economic Theory*, **145** (1), 42-62.

Vesely, F. and C.-L. Yang, (2010): "On Optimal and Neutrally Stable Population Equilibrium in Voluntary Partnership Prisoner's Dilemma Games", http://ssrn.com/abstract=1541684.

Watson, J., (2002): "Starting Small and Commitment", *Games and Economic Behavior*, **38**, 176-199.