The Optimal Degree of Cooperation in the Repeated Prisoner's Dilemma with Side Payments

Yuk-Fai Fong, Northwestern University
Jay Surti, International Monetary Fund
The Optimal Degree of Cooperation in the Repeated Prisoners’ Dilemma with Side Payments*

Yuk-fai Fonga,† and Jay Surtib

aManagement and Strategy Department, Northwestern University, 2001 Sheridan Rd., Evanston, IL 60208
bInternational Monetary Fund, 700 19th Street, NW, Washington, DC 20431

Abstract

In the infinitely repeated Prisoners’ Dilemma with side payments, we characterize the Pareto frontier of the set of subgame perfect equilibrium payoffs for all possible combinations of discount factors. Play paths implementing Pareto dominant equilibrium payoffs are uniquely determined in all but the first period. Full cooperation does not necessarily implement these payoffs even when it maximizes total stage game payoffs. Rather, when the difference in players’ discount factors is sufficiently large, Pareto dominant equilibrium payoffs are implemented by partial cooperation supported by repeated payments from the impatient to the patient player. When both players are sufficiently patient, such payoffs, while implemented via full cooperation, are supported by repeated payments from the impatient to the patient player. We characterize conditions under which public randomization has no impact on the Pareto dominant equilibrium payoffs and conditions under which such payoffs are robust to renegotiation.

JEL classification: C72, C73

Keywords: Repeated Prisoners’ Dilemma; Side payments; Differential time preferences; Pareto dominant equilibrium payoffs; Renegotiation-proofness

*An earlier version of this paper was entitled “Bribing to Cooperate”. We are grateful to seminar participants at The Midwest Economic Theory Conference, Spring 2007; Queen’s University; and the University of Toronto; and to Ricardo Alonso, Sandeep Baliga, Eric Chou, Albert Ma, Kathy Spier, and especially three anonymous referees and the anonymous editor-in-charge for helpful comments and suggestions. The views expressed in this paper are the authors’, and should not be attributed to the IMF or IMF policy.

†Corresponding author. Fax: 1 847 467 1777

Email addresses: y-fong@kellogg.northwestern.edu (Y-f. Fong), jsurti@imf.org (J. Surti).
1 Introduction

1.1 Motivation and Overview

Repeated games are a common device for studying the impact of dynamic interaction on the feasibility and limits of cooperation among agents. The use of side-payments to elicit cooperative play has rarely been studied to-date within this framework. In real-life settings however, agents often transfer payments to achieve a mutually agreed upon distribution of the surplus from cooperative interaction, while simultaneously providing incentives to the recipient to behave cooperatively. Examples include monetary transfers from rich to developing countries in the context of environment protection treaties or the opening of domestic markets to international competition, bonuses used to reward employee performance, and briberies in corruption, among others.

Dynamic and repeated games usually analyze situations where players have a common discount factor. This assumption is not universally appropriate as the following examples indicate. An unstable domestic political regime is likely to induce myopia among its leaders, owing to the fear of loss of power, and this has a material impact on how they negotiate in international forums with leaders of stabler political regimes. Employees may be less committed to a firm than are employers, and consequently may discount at a faster rate, future payoffs from the employment relationship. More generally, when the credit market is imperfect, and agents have asymmetric access to it, assuming a distribution of heterogeneous time preferences is more realistic.

We examine a side-payments mechanism players use to transfer payoffs in the repeated prisoners’ dilemma (PD). For every pair of discount factors in the unit square, we characterize play paths implementing payoffs on the Pareto frontier of the set of subgame perfect equilibrium (SPE) outcomes of this augmented, transferable-utility PD.¹

¹Considerable information is available on the Pareto frontier of the repeated PD’s SPE payoffs with common discount factors. Sorin (1986) identifies a bound on the discount factor below which only perpetual non-cooperation is sustainable in SPE, and Mailath, Obara, and Sekiguchi (2003) characterize the maxi-
Focusing on Pareto dominant SPE play paths enables us to uniquely predict the equilibrium outcome(s) in all but the first period. The intuitive conjecture that full cooperation; i.e., perpetual play of the cooperate-cooperate action profile, supported by repeated payments from the patient player (player $P$) to the impatient player (player $I$), necessarily implements such payoffs is not universally valid, even when such behavior maximizes total stage game payoffs. Rather, when the difference in players’ discount factors is sufficiently large, all Pareto dominant SPE payoffs are implemented by partial cooperation; i.e., perpetual play of the cooperate-defect action profile, supported by repeated payments from player $I$ to player $P$. When both players are sufficiently patient, such payoffs, while implemented via full cooperation, are again supported by repeated payments from player $I$ to player $P$. Within each of these transfer protocols, Pareto dominant SPE play paths are differentiated solely by the first period side payment, which can travel in either direction, and whose role is to deliver the bonus payment that implements the selected SPE payoff vector.

The search for Pareto dominant SPE outcomes is facilitated by the key observation that in any given period, the payoffs that player $I$ receives in the future must not exceed that amount which provides him just enough incentive to take the action prescribed by the SPE strategy in that period. Owing to the difference in time preferences, it is always more efficient to let player $I$ receive any additional payoff over and above this amount upfront, instead of in the future. Consequently, when both players are sufficiently patient, Pareto dominant SPE payoffs are achieved via full cooperation, and supported from the second period onward, by repeated side payments from player $I$ to player $P$. A player who is sufficiently patient is willing to take the cooperative action within a full cooperation regime even if (s)he has to make a side payment to the other player. As it is always more
efficient to let player $I$ consume up-front than in the future, efficiency demands that he transfer any future payoff beyond his incentive constraint (IC) to player $P$ (and possibly, be compensated by a one time side payment from player $P$ in the first period).

When player $I$’s patience falls below a critical threshold so that he is unwilling to cooperate in the absence of side payments, yet player $P$ is still willing to cooperate and pay player $I$ to do so as well, Pareto dominant SPE outcomes are still achieved via full cooperation. However, in order to induce player $I$ to cooperate, player $P$ now has to “bribe” player $I$ into cooperation by giving him a side payment each period. This case covers the most intuitive Pareto dominant SPE outcomes.

As player $I$ becomes yet more impatient, however, the per period side-payment required from player $P$ to give him the incentive to cooperate becomes increasingly large. As a result, the intertemporal allocation of instantaneous payoffs that supports full cooperation becomes increasingly inefficient because the payoffs to player $I$ become ever more backloaded. Beyond a point, it is no longer Pareto dominant in equilibrium for both players to cooperate. If only player $P$ cooperates, then from the second period onward, player $I$ can be asked to respond by transferring any payoff received by him in the stage game over and above his individually rational (IR) payoff to player $P$ (and possibly, be compensated by a one time side payment from player $P$ in the first period). Although partial cooperation leads to a lower total stage game payoff than full cooperation, the more favorable intertemporal allocation of instantaneous payoffs makes partial cooperation Pareto dominant.

1.2 Related Literature

**Side-payments mechanisms** Among the papers studying the impact of side-payments on the structure and properties of the set of repeated game equilibria, Baliga and Evans (2000) most closely resembles ours. They analyze a class of common discounting two-
player repeated games with a side-payments mechanism attached to the stage game. They show that when the players’ common discount factor approaches unity, renegotiation-proof SPE necessarily maximize the total stage game payoff in each period. In contrast, we show that maximization of the total stage-game payoff is not renegotiation-proof when players’ time preferences are sufficiently heterogeneous, and is not implementable in SPE when the impatient player is too myopic. Indeed, even for the case where both players are arbitrarily patient (but one more than the other), the disparity in time preferences implies that our Pareto dominant SPE are (Pareto) inefficient despite the fact that total stage game payoffs are maximized in every stage game.

Our paper is also related to the literature on relational incentive contracts. This literature is concerned with relationships where the underlying incentive problem is one-sided and players have identical time preferences. In contrast, in our paper, both players may lack incentives to contribute to the production of surplus in the underlying game. Moreover, the role of the principal (payer of the side payment) may be taken up by either player endogenously in our setting since steady-state transfers may travel either way between the patient and impatient players, depending on their relative patience.

**Differential time preferences** The literature on repeated games with different time preferences is still relatively small. In an important contribution, Lehrer and Pauzner (1999) have studied how players in a repeated game exploit the difference in their time preferences by the intertemporal trade of instantaneous payoffs to enhance efficiency.4

---

3 See Levin (2003), for a general model, and an exhaustive bibliography.

4 Lehrer and Yariv (1999) also analyze two-player repeated games with different discount factors and one-sided incomplete information, where the underlying stage games are zero-sum. As in Lehrer and Pauzner (1999), they restrict attention to situations where both players are infinitely patient. The possibility of patient players’ exploiting their far-sightedness to build a reputation for tough behavior and guarantee a minimum dynamic payoff stream within a setting of heterogeneous discount factors has been analyzed in Fudenberg and Levine (1989), Schmidt (1991), and Celentani et al. (1996), and the references therein. However, the techniques and results presented in these papers are largely unrelated to ours.
Their paper provides the key insight that, by letting the impatient player consume more in the near future and the patient player consume more in the farther future, the set of feasible payoff vectors becomes larger than the convex hull of IR stage game payoffs identified by the folk theorem. They demonstrate that, keeping constant the relative patience of the players, as both become arbitrarily patient, they can achieve outcomes in equilibrium that would be infeasible were their time preferences identical.  

Within the setting of the PD game, allowing for side payments between the players enables us to generalize Lehrer’s and Pauzner’s analysis to all combinations of discount factors in the unit square. Side-payments also allow us to identify Pareto dominant SPE play paths that are impossible to support in their absence. For example, even when the impatient player is very myopic, SPE exist wherein the patient player cooperates. While the impatient player defects, he does reciprocate by transferring to the patient player some of the surplus payoff generated by his defection. In the absence of side-payments, it is impossible to support such a play path for any combination of discount factors. Moreover, adding side-payments vastly enhances the efficiency of intertemporal trade, enabling players to do better than in the PD without side payments. In our paper, sufficiently patient players can conduct some intertemporal trade of instantaneous payoffs while fully cooperating in every period. Without side payments, it is necessary to depart from full cooperation (with a positive probability) to enable intertemporal trade. Also, in the repeated PD with side payments, on any Pareto dominant SPE play path with full or partial cooperation, the impatient player’s continuation payoff is reduced, from the second period onward, to either his IC or IR constraint, and play is characterizable in terms of a steady-state in both side payments and action profiles. Without side payments, this is not Pareto dominant. Instead, in the case where both players are very patient, as in Lehrer

---

5 This is achieved by picking an arbitrary pair of discount factors which are different, then taking a limit which shrinks the duration of each time period between stage games to zero.

6 Section 4.3 discusses how our results extend to application of our side-payments mechanism to repeated iterations of stage games other than the PD.
and Pauzner, the impatient player’s continuation payoff will decay gradually to her IC constraint.

2 Model

The repeated game is a two-person, infinitely repeated PD, except that in each period, players are allowed to each voluntarily make a transfer of payoff to each other before they simultaneously choose whether to defect or cooperate; i.e., the stage game of the repeated game augmented with side payments consists of two sub-stages, the side payment stage and the action stage which immediately follows, with no discounting between the sub-stages within a single period.

**Action Stage** The action stage is the stage game of a repeated PD without transfers:

\[
\begin{array}{ccc}
C & D \\
C & y, y & 0, z \\
D & z, 0 & x, x \\
\end{array}
\]

where \(z > y > x > 0\), guaranteeing that (1) is the generic, symmetric PD. The row player, denoted \(I\), is assumed to be less patient than the column player, denoted \(P\); i.e., we assume \(\delta_I \leq \delta_P\), where \(\delta_i \in (0, 1)\) denotes player \(i\)’s discount factor. We assume that the total stage game payoff is the highest under full cooperation, i.e., \(2y > z\). This is a very natural assumption in many PD-like situations. If we think of \((C, C)\) as a collusive outcome in a duopoly and action \(D\) as initiating a price war, then the duopoly’s aggregate profit is higher under collusion than when one firm cuts price or over-produces. Finally, it will be amply clear to the reader after we prove our main result in Section 3, that this assumption is made to essentially rule out trivial cases.

**Side Payment Stage** At the beginning of each period, both players simultaneously and independently decide on an amount of side payment that they wish to offer the
opponent. Following Baliga and Evans (2000), we assume that players have sufficiently large budgets to pay any amount that a Pareto dominant SPE may prescribe. After the players exchange the side payments, they play game (1). Assuming the side payments to be non-negative, every unit of side payment both reduces the giver’s utility and increases the recipient’s utility by one unit.

Denote the side payment made in period \( t \) by player \( i \) by \( b_{i,t} \geq 0 \), an action taken by her in the period \( t \) play of (1) by \( a_{i,t} \). Let the net side payment from \( P \) to \( I \) be \( b_t = b_{P,t} - b_{I,t} \). A period \( t \) outcome – at the side payment stage – is \( O^B_t := (b_{I,t}, b_{P,t}) \), and at the action stage, is \( O^A_t = (a_{I,t}, a_{P,t}) \). A \( t \)-history is defined as a play path including all side payments and actions through period \( t - 1 \); \( h_t := \{O^B_s, O^A_s\}_{s=1}^{t-1} \). In addition, a history through the action stage of \( t \) is \( h^B_t := h_t \cup \{O^B_t\} \). \( H_t \) is the set of possible \( t \) histories; \( H^B_t \), the set of all \( h^B_t \); and \( H = (\cup H_t) \cup (\cup H^B_t) \), the set of all play paths. We assume perfect monitoring throughout. Let players’ instantaneous payoffs under the action profile \( a_t \) be \( u^A_I(a_t) \) and \( u^A_P(a_t) \), and the sum of these payoffs be denoted \( w(a_t) \). The net payoff to player \( i \) in period \( t \) is \( u_i (O^B_t, O^A_t) := u^A_i(a_t) + (b_{j,t} - b_{i,t}) \). It is clear that side payments in period \( t \) do not affect the total stage game payoff, i.e., \( u_I (O^B_t, O^A_t) + u_P (O^B_t, O^A_t) = u^A_I(a_t) + u^A_P(a_t) = w(a_t) \). For a play path, \( \{O^B_t, O^A_t\}_{t \geq 1} \in H \), the time-average repeated game payoff to player \( i \) at time \( t \) is

\[
U_{i,t}\left(\{O^B_t, O^A_t\}_{t \geq t}\right) = (1 - \delta_i) \sum_{\tau=t}^{\infty} \delta^{\tau-t}_i u_i (O^B_\tau, O^A_\tau)
\]

As perfect monitoring applies, we will use the SPE as our solution concept.

3 Analysis

The main goal of our analysis is to characterize the degree of cooperation and the pattern of side-payments on play paths implementing Pareto dominant SPE payoffs. We limit our
attention to pure strategy equilibria. Before proceeding to prove the main findings in our model, we establish a few building blocks to facilitate the analysis. In the PD with side payments, Nash reversion trigger strategies punish any deviation from the proposed play path by forever returning to the unique, one-shot Nash equilibrium play. In the unique one-shot Nash equilibrium, agents play $b_I = b_P = 0$ at the side-payments stage, and $(D, D)$ at the action stage, enforcing IR payoffs in the supergame. Consequently, in order to derive Pareto dominant SPE payoffs, we focus attention on strategies supported by Nash reversion in this section.

Note that it is (weakly) Pareto dominated to have both $b_P > 0$ and $b_I > 0$ at the side-payment stage for the following reason. Take an arbitrary period $-t$ side-payment stage, only $b_t = b_{P,t} - b_{I,t}$ is relevant for the IRs and ICs for stages other than the current sub-stage. Moreover, reducing $b_{I,t}$ and $b_{P,t}$ by the amount $\min\{b_{I,t}, b_{P,t}\}$ will lower players’ deviation payoffs by $\min\{b_{I,t}, b_{P,t}\}$ and relax their ICs to pay at the current stage. So, in the rest of the paper, we will focus on characterizing side-payments satisfying $\min\{b_{I,t}, b_{P,t}\} = 0$.

Note also that because the extremal punishment strategies are applied, and players do not discount between the side-payment stage and the action stage in the same period, the player designated to make the side-payment in any given period is willing to pay so long as an IR-level of continuation payoff, $x$ (net the bribe), is promised.

Let $W(t)$, which we will refer to as the Total Utility of the supergame starting in period $t$, denote the sum of the players’ present discounted payoffs; i.e.,

$$W(t) = \frac{U_{I,t}}{1 - \delta_I} + \frac{U_{P,t}}{1 - \delta_P}.$$  

(3)

Players can freely split the Total Utility at the side-payment stage of any period $t$, (conditional only on $\min\{U_{I,t}, U_{P,t}\} \geq x$). Moreover, on any SPE play path, the punishment to

---

7 In Section 4, we formally argue that this treatment is without loss of generality for all but a specific subset of discount factors, and characterize the Pareto dominant SPE play paths in correlated strategies for $(\delta_I, \delta_P)$ falling in this subset.

8 With endogenized timing of side-payments, players may discount between the side-payment stage and the action stage. We discuss such an extension in the conclusion.
revert to playing \((D,D)\) forever is sufficient to deter any incentive to deviate. Hence, the strategy profile in any continuation subgame of a Pareto dominant SPE maximizes \(W(t)\) among all possible SPE strategy profiles of the continuation subgame. This will further imply the following:

**Lemma 1** (i) On any Pareto dominant SPE play path, \(W(t) = W(t')\), for all \(t, t' \geq 1\), and (ii) any Pareto dominant SPE payoff vector \((U_{I1}, U_{P1})\) can be implemented by a SPE play path satisfying \(b_{t+1} = b_{t'}\), and \(a_t = a_{t'}\), for all \(t, t' \geq 1\).

**Proof.** See appendix. ■

Lemma 1 suggests that when side-payments are allowed, players do not need to rely on switching among action profiles to optimally split the surplus of the game. In fact, as we will show below, it is inefficient to do so. Ruling out having to deal with players' switching among different action profiles over time substantially simplifies our analysis. We next eliminate some action profiles that are clearly Pareto dominated.

**Lemma 2** (i) On any Pareto dominant SPE play path, \(a_t \neq (C, D)\) for all \(t \geq 1\). (ii) If \(2x > z\), then on any Pareto dominant SPE play path, \(a_t \neq (D, C)\) for all \(t \geq 1\).

**Proof.** See appendix. ■

The basic idea behind part (i) of Lemma 2 is that it is allocatively more efficient to require the patient player to cooperate than to require the impatient player to do so. Take an arbitrary time period \(t\) in which only the impatient player cooperates. Now ask the patient player to play \(C\) and impatient player to play \(D\) in period \(t\) instead. By doing so, \(w(a_t)\) remains unchanged, but by requiring the impatient player to play \(D\) instead of \(C\), his continuation payoff from period \(t + 1\) onward can be lowered relative to what had to be guaranteed to him under \((a_{I_t}, a_{P_t}) = (C, D)\). This is done by reducing \(b_{t+1}\). The impatient player's loss is compensated by increasing \(b_t\). Since it is efficiency enhancing to let the impatient player consume more up-front and the patient player consume more in
the future, the players can always find adjustments in $b_t$ and $b_{t+1}$ that make them both strictly better off.

Part (ii) of the lemma is equally straightforward in that if $w(a_t)$ is higher when both players defect than when one of them cooperates, then partial cooperation will never be observed on any Pareto dominant SPE path. In this case, a higher $w(a_t)$ is generated by playing $(D, D)$, and since neither player is required to cooperate, their IC constraints are easier to meet.\(^9\)

Lemmata 1 and 2 together imply that in order to characterize Pareto dominant SPE outcomes, it is sufficient to look at Pareto undominated SPE generated by perpetual play of one of the following action profiles: $(C, C)$, $(D, C)$, or $(D, D)$, where the play of $(D, C)$ is possible under the reasonable condition $2x < z$.

Next, we show that

**Lemma 3** On any Pareto dominant SPE play path, for $t \geq 1$, (i) if $a_t = D$, then $U_{t+1} = x$; (ii) if $a_t = (C, C)$, then $U_{t+1} = y + b$, where

$$b \equiv \frac{1 - \delta}{\delta} (z - y) - (y - x). \quad (4)$$

**Proof.** See appendix. \(\blacksquare\)

Part (i) of Lemma 3 establishes that if $I$ is not required to cooperate in period $t$, then the most efficient allocation necessarily involves providing him with exactly his IR level of continuation payoff, $x$, starting from period $t + 1$. Any additional amount rewarded to him over and above $x$ is inefficient as this “extra” utility ought to be paid to him in period $t$ instead, or even earlier, rather than in $t + 1$. According to part (ii) of the lemma, if $I$ is required to play $C$ together with $P$, his continuation payoff starting from period $t + 1$ must

---

\(^9\)We take the view however, that in many situations, the condition $2x > z$ is unrealistic. If we think of action $D$ as pollution, then the total stage game payoff should be lower when both players pollute than when just one of them does; i.e., $2x < z$. Similarly, this assumption would be unrealistic in the collusive duopoly example of Section 2, wherein the aggregate industry profit would be higher if just one firm were to cut price than if both firms were to do so.
be increased at least to a level \( y + b > x \) in order to satisfy his IC. The result demonstrates how raising \( I \)'s continuation payoff at time \( t+1 \) to above his IR or IC level (as appropriate) limits the extent to which players can conduct welfare improving intertemporal trade of instantaneous payoffs. The comparison of parts (i) and (ii) also illustrates the efficiency cost associated with trying to achieve full – as opposed to partial – cooperation when players have different time preferences. Notice that \( \lim_{\delta_I \to 0} y + b = \infty \), illustrating that the cost of increasing cooperation to increase \( w (a_t) \) increases rapidly as the impatient player becomes more myopic. Consequently, and as we will soon demonstrate, when the difference in players’ discount factors becomes sufficiently large, partial cooperation Pareto dominates full cooperation even when the latter, which provides a higher \( w (a_t) \), can be supported as an SPE outcome.

### 3.1 Existence of Cooperative SPE

Now, we are ready to characterize conditions under which some cooperation can be achieved. We first identify the necessary and sufficient conditions for full cooperation to be supported in an SPE.

**Proposition 1** There exist \( b \) and \( \bar{b} \), where \( b \leq \bar{b} \), such that there is an SPE in pure strategies, the equilibrium play path of which satisfies:

\[
\begin{align*}
(i) & \quad \forall t \geq 1; \quad a_t \equiv (a_{Ht}, a_{Pt}) = (C, C) ; \\
(ii) & \quad \forall t \geq 2; \quad b_{Pt} - b_{Ht} = b \in [b, \bar{b}]; \\
\end{align*}
\]

if and only if the following condition is satisfied:

\[
\frac{1}{\delta_I} + \frac{1}{\delta_P} \leq \frac{2}{\delta^*}, \tag{5}
\]

where \( \delta^* \equiv \frac{\bar{z} - \bar{y}}{\bar{z} - x} \). The equilibrium is most efficient when \( b_{Pt} - b_{Ht} = \bar{b}, \forall t \geq 2 \).

**Proof.** See appendix. \( \blacksquare \)
In the rest of the paper, we will abuse notation and use $\overrightarrow{CC}$ to denote a play path that satisfies condition $(\overrightarrow{CC})$. Let $r_I$ and $r_P$ be the players’ discount rate and $\delta^* = 1/(1 + r^*)$. Then (5) can be rewritten as

$$\frac{r_I + r_P}{2} \leq r^*.$$ 

Therefore, Proposition 1 describes very intuitively how the patient player’s excessive patience can compensate for the impatient player’s excessive myopia. It does not matter that one of the players is too myopic; as long as the players are patient enough on average, full cooperation can still be supported as an SPE outcome. Specifically, when $\delta_I < \delta^* < \delta_P$ and (5) is satisfied, these players can support full cooperation by agreeing to a repeated stream of side-payments from player $P$ to player $I$, the amount of which is high enough to make defection too costly for $I$, but also low enough that $P$ is willing to pay in exchange for her opponent’s cooperation.\(^\text{10}\)

Next, we identify necessary and sufficient conditions for partial cooperation to be sustainable. As it turns out, regardless of $I$’s discount factor, as long as $P$ is patient enough, we can at least induce her to cooperate.

**Proposition 2** There exists an SPE in pure strategies the equilibrium play path of which satisfies:

(i) $\forall t \geq 1; a_t \equiv (a_{It}, a_{Pt}) = (D, C)$; 

(ii) $\forall t \geq 2; b_{Pt} - b_{It} = b \in [x - z, -x/\delta_P]$;

if and only if $z > 2x$ and $\delta_P \geq x/(z - x)$. The equilibrium is most efficient when $b_{Pt} - b_{It} = x - z < 0, \forall t \geq 2$.

**Proof.** See appendix. \(\blacksquare\)

Again, in the rest of the paper we use $\overrightarrow{DC}$ to denote a play path that satisfies condition

\(^\text{10}\) When $\delta^* < \delta_I$, (5) will obviously continue to hold. In that case, it is however most efficient to support full cooperation by having player $I$ repeatedly making side-payments to player $P$, i.e., $b < 0$. The discussion of this pattern of side-payments is postponed to after the statement of Theorem 1.
(\overrightarrow{DC}). For the patient player to find it incentive compatible to accept an instantaneous payoff of zero in the action stage, she must prefer receiving the maximum payment I is willing to make every period in the future, \(z - x\), over the benefit of a one-time defection in the current period, \(x\). This is exactly the condition \(\delta_P \geq x/(z - x)\). In order for this condition to be satisfied for \(\delta_P < 1\), the parameters must in addition also satisfy the constraint \(z > 2x\).

### 3.2 Characterization of Pareto dominant SPE

We are now ready to characterize the set of Pareto dominant SPE for all \((\delta_I, \delta_P)\); i.e., over the universal set of discount factors:

\[
\Omega^U \equiv \cup_{\delta_P \in (0, 1)} \Phi(\delta_P),
\]

where \(\Phi(\delta_P) := (0, \delta_P] \times \{\delta_P\}\).

Let \(\overrightarrow{DD}\) denote the noncooperative play path implemented by the strategy profile:

\[
(i) \forall t \geq 1; a_t = (D, D); \quad (\overrightarrow{DD})
\]

\[
(ii) b_{Pt} = b_{It} = 0.
\]

It follows from Lemmata 1 and 2 that for the purpose of characterizing Pareto dominant SPE play paths, it is sufficient to compare the Total Utilities generated by the paths \(\overrightarrow{CC}\), \(\overrightarrow{DC}\), and \(\overrightarrow{DD}\). It is easy to see that whenever \(\overrightarrow{CC}\) or \(\overrightarrow{DC}\) can be supported as a SPE play path, \(\overrightarrow{DD}\) is Pareto dominated, and when neither can be supported as such, \(\overrightarrow{DD}\) is Pareto dominant. Therefore, it remains only to compare the highest Total Utilities generated by \(\overrightarrow{CC}\) and \(\overrightarrow{DC}\) when both of them can be supported as SPE play paths.

To derive the highest Total Utility generated by \(\overrightarrow{CC}\), consider an SPE play path in which player \(P\) makes the most efficient (time-invariant) side-payment, \(b\), to player \(I\) in every period, including the first. The average payoffs of \(I\) and \(P\) in this equilibrium are
y + \hat{b} and y - \hat{b}, and thus the Total Utility is \( \frac{y+\hat{b}}{1-\delta_I} + \frac{y-\hat{b}}{1-\delta_P} \). Since adjusting the first period side-payment does not affect the Total Utility,

\[
\frac{U_{It}}{1-\delta_I} + \frac{U_{Pt}}{1-\delta_P} = \frac{y + \hat{b}}{1-\delta_I} + \frac{y - \hat{b}}{1-\delta_P}, \forall t \geq 1. \tag{6}
\]

With a similar approach, we can derive the Total Utility generated by \( \overrightarrow{DC} \) as

\[
\frac{U_{It}}{1-\delta_I} + \frac{U_{Pt}}{1-\delta_P} = \frac{x}{1-\delta_I} + \frac{z-x}{1-\delta_P}, \forall t \geq 1. \tag{7}
\]

Therefore, when both \( \overrightarrow{CC} \) and \( \overrightarrow{DC} \) are sustainable, \( \overrightarrow{CC} \) Pareto dominates \( \overrightarrow{DC} \) if and only if

\[
\frac{y+\hat{b}}{1-\delta_I} + \frac{y-\hat{b}}{1-\delta_P} \geq \frac{x}{1-\delta_I} + \frac{z-x}{1-\delta_P}, \tag{8}
\]

or equivalently

\[
\delta_I \geq \frac{z-y}{y} \delta_P. \tag{9}
\]

If, on the other hand \( \delta_I \leq \frac{z-y}{y} \delta_P \), then \( \overrightarrow{DC} \) Pareto dominates \( \overrightarrow{CC} \).

Noting that (5) can be rewritten as \( \delta_I \geq \delta_P \delta^*/(2\delta_P - \delta^*) \), we denote the different regions of the \((\delta_I, \delta_P)\)-space by

\[
\Omega_{CC} \equiv \left\{ (\delta_I, \delta_P) \in \Omega^U : \delta_I \geq \max \left\{ \frac{z-y}{y} \delta_P, \frac{\delta_P \delta^*}{2\delta_P - \delta^*} \right\} \right\},
\]

\[
\Omega_{DC} \equiv \left\{ (\delta_I, \delta_P) \in \Omega^U : \delta_I \leq \max \left\{ \frac{z-y}{y} \delta_P, \frac{\delta_P \delta^*}{2\delta_P - \delta^*} \right\} \text{ and } \delta_P \geq \frac{x}{z-x} \right\},
\]

\[
\Omega_{DD} \equiv \Omega^U \setminus (\Omega_{CC} \cup \Omega_{DC}).
\]

The following result formally characterizes the Pareto frontier for all pairs of discount factors in \( \Omega^U \).

**Theorem 1** The following properties hold in all Pareto dominant SPE. (i) If \( (\delta_I, \delta_P) \in \Omega_{CC} \), then \( \overrightarrow{CC} \), supported by \( b_{Pt} - b_{It} = \hat{b} \) for all \( t \geq 2 \), is Pareto dominant and the Pareto frontier of the set of SPE payoffs is characterized by (6). Moreover, \( \hat{b} < 0 \) if and only if \( \delta_I > \delta^* \). (ii) If \( (\delta_I, \delta_P) \in \Omega_{DC} \), then \( \overrightarrow{DC} \), supported by \( b_{It} = z-x > b_{Pt} = 0 \)
∀t ≥ 2, is Pareto dominant and the Pareto frontier of SPE payoffs is characterized by (7).

(iii) Otherwise, if \((\delta_I, \delta_P) \in \Omega_{DD}\), then \(\overrightarrow{DD}\) is Pareto dominant, \(U_{I1} = U_{P1} = x\), and \(b_{It} = b_{Pt} = 0\) ∀t ≥ 1.

**Proof.** See appendix. ■

Figure 1: Characterization of Pareto Dominant Subgame Perfect Equilibria

Figure 1 provides a graphical summary of Theorem 1. In each region above the 45° line, where \(\delta_I \leq \delta_P\), we have indicated the regions of the \((\delta_I, \delta_P)\)-space where \(\overrightarrow{CC}\), \(\overrightarrow{DC}\), and \(\overrightarrow{DD}\) are Pareto dominant paths of play. We also detail whether player I or P is making the side-payment when \((\delta_I, \delta_P)\) fall in each of these regions. When \(\overrightarrow{CC}\) is sustainable as a SPE play path but \(\overrightarrow{DD}\) is not, or when both \(\overrightarrow{DC}\) and \(\overrightarrow{CC}\) are sustainable

---

11Figure 1 assumes that \(x + y > z\). Generally speaking, \(\max\left\{\frac{z-y}{y} \delta_P, \frac{y+z-x}{x+y} \delta_I\right\}\) may be greater than or less than \(\delta^*\) depending on whether \(x + y < z\) or \(x + y \geq z\). In the latter case, either expression in the max operator is dominated by \(\delta^*\). In this event, for all \(\delta_I \geq \delta^*\), the Pareto dominant SPE is necessarily \(\overrightarrow{CC}\). However, when \(x + y < z\), then for \(\delta_P > \frac{z-y}{y} \delta_I \in (\delta^*, 1)\), \(\max\left\{\frac{z-y}{y} \delta_P, \frac{y+z-x}{x+y} \delta_I\right\} = \frac{z-y}{y} \delta_P > \delta^*\), and for \(\delta_I \in (\delta^*, \frac{z-y}{y} \delta_P)\), the Pareto dominant SPE is \(\overrightarrow{DC}\).
but \( \delta_I \geq (z - y) \delta_P / y \), the most efficient SPE path is \( \overrightarrow{CC} \). This region is \( \Omega_{CC} \). When \( \overrightarrow{DC} \) is sustainable and \( \overrightarrow{CC} \) is not, or when both \( \overrightarrow{DC} \) and \( \overrightarrow{CC} \) are sustainable but \( \delta_I \leq (z - y) \delta_P / y \), the most efficient equilibrium path is \( \overrightarrow{DC} \). This region is \( \Omega_{DC} \). In \( \Omega_{DD} \), where neither \( \overrightarrow{DC} \) nor \( \overrightarrow{CC} \) is sustainable, \( \overrightarrow{DD} \) is the unique equilibrium play path and thus also Pareto dominant. Now, it is clear that the analysis is nontrivial only under the assumption \( 2y > z \) we have made throughout; when this assumption was violated, since \( \delta_P \geq \delta_I \), \( (z - y) \delta_P < y \delta_I \) would always hold and \( \overrightarrow{DC} \) would always be more efficient than \( \overrightarrow{CC} \).

As we can see from Figure 1, there are regions where both \( \overrightarrow{DC} \) and \( \overrightarrow{CC} \) are sustainable as SPE paths. When comparing the payoffs under \( \overrightarrow{DC} \) and \( \overrightarrow{CC} \), the obvious disadvantage of \( \overrightarrow{DC} \) is the smaller amount of total stage game payoff generated in each period relative to playing in accordance with \( \overrightarrow{CC} \). However, playing \( \overrightarrow{DC} \) allows players to conduct more intertemporal trade of instantaneous payoffs by reducing player \( I \)’s time-average utility to his IR level, \( x \), from period 2 onward. If player \( I \) is to be compensated, this can be done through a one-time side-payment from player \( P \) in the first period. When the players discount factors are sufficiently different, \( \delta_P \geq \delta_I \), \( (z - y) \delta_P < y \delta_I \) would always hold and \( \overrightarrow{DC} \) would always be more efficient than \( \overrightarrow{CC} \).

Summing up, although unconstrained Pareto efficiency requires both players to cooperate using side-payments to efficiently split the surplus in every period, the cost (in the form of loss in the extent of intertemporal trade) of satisfying \( I \)’s incentive constraint as he gets more myopic makes it mutually beneficial beyond a point, to receive the surplus from \( \overrightarrow{DC} \) and remain at \( \overrightarrow{CC} \). This region is labeled \( \Omega^{2}_{DC} \) in Figure 1. Properties of SPE play paths for \( (\delta_I, \delta_P) \in \Omega^{2}_{DC} \) will be discussed further in Section 4.
partial cooperation rather than full cooperation. This stands in stark contrast to models without side-payments. Absent side-payments, it is impossible to support in equilibrium the play path implied by \(\overrightarrow{DC}\), wherein one of the players “gets ripped off” every period.

We now discuss how the pattern of side-payments varies with the discount factors. To facilitate this analysis, we divide \(\Omega_{CC}\) into two regions:

\[
\Omega^1_{CC} \equiv \{(\delta_I, \delta_P) \in \Omega_{CC} : \delta_I \leq \delta^* \}, \\
\Omega^2_{CC} \equiv \{(\delta_I, \delta_P) \in \Omega_{CC} : \delta_I > \delta^* \}.
\]

These regions are also labeled in Figure 1. Since the side-payment is used to induce cooperation as well as to conduct intertemporal trade, and Pareto dominant SPE outcomes may be achieved through either partial or full cooperation depending on the relative patience of the players, the side-payment can change in a nonmonotone and discontinuous manner as players’ discount factors change. Figure 2 illustrates how the repeated side-payments on the steady state typically change with the impatient player’s discount factor. The diagram is drawn based on some fixed \(\delta_P \geq x/(z-x)\):

\[
b = x - y + \frac{1-\delta_I}{\delta_I}(z-y)
\]

**Figure 2**: Pattern of repeated side payments, for some fixed \(\delta_P \geq x/(z-x)\)

When \(\delta_I\) is relatively small, the Pareto dominant SPE outcomes are achieved by \(\overrightarrow{DC}\). Player I makes a side-payment of \((z-x)\) every period starting from the second period to
provide player $P$ with the incentive to cooperate and to push down player $I$’s average continuation payoff to $x$ to conduct the maximum amount of intertemporal trade. As $\delta_I$ grows beyond a certain level and becomes relatively close to $\delta_P$, as indicated by $(\delta_I, \delta_P) \in \Omega_{CC}$, the importance of intertemporal trade diminishes and $\overrightarrow{DC}$ becomes Pareto dominated by $\overrightarrow{CC}$. When $\overrightarrow{CC}$ is Pareto dominant, the steady-state side-payment may go either way. When $\delta_I$ is below $\delta^*$; i.e., when $(\delta_I, \delta_P) \in \Omega_{CC}^1$, $P$ makes a side-payment to $I$ to give $I$ the incentive to cooperate; i.e., $b > 0$. Once $\delta_I$ increases beyond $\delta^*$; i.e., when $(\delta_I, \delta_P) \in \Omega_{CC}^2$, both players are sufficiently patient that each is willing to cooperate even if (s)he has to make a side-payment to the other player. To achieve Pareto dominant SPE outcomes, $I$ transfers all the future “excess” payoff to $P$ by a side-payment; i.e., $b < 0$. As $\delta_I$ increases further, the excess payoff becomes larger and it takes a larger side-payment $|b|$ from $I$ to transfer them to player $P$.

We end this section by noting that for a wide range of parameter values, side-payments can be used to raise players’ average payoffs to above $(y, y)$. For example, when $\delta_P \geq \delta_I > \delta^*$, since the less patient player is also excessively patient, even if he is asked to make a side-payment to the patient player every period beginning from the second period, he still has the incentive to cooperate at the action stage. To compensate him for the future side-payments, the patient player can make a one-time payment to him in the first period. Due to the difference in time preference, such intertemporal trade of instantaneous payoffs can raise both players’ average payoffs above $y$.

4 Discussion

Correlated Strategies

In the analysis so far, we have restricted attention to pure strategy equilibria. This begs the question of whether allowing agents to play correlated strategies will expand the
SPE payoff set and thus push outward its Pareto frontier. Define

\[
\Omega_{1DC}^1 = \left\{ (\delta_I, \delta_P) \in \Omega_{DC} : \delta_I \leq \frac{z-y}{y} \delta_P \right\},
\]

\[
\Omega_{2DC}^2 = \left\{ (\delta_I, \delta_P) \in \Omega_{DC} : \delta_I > \frac{z-y}{y} \delta_P \right\}.
\]

In this subsection we demonstrate that, for the purpose of identifying the set of Pareto dominant SPE payoffs, extending the strategy space by allowing for public randomization only enhances Pareto efficiency for \((\delta_I, \delta_P) \in \Omega_{2DC}^2\). Recall that \(\Omega_{DC}^2\) is a specific subset of \(\Omega_{DC}\), where \(\vec{CC}\) is not sustainable as an SPE path in pure strategies, but had it been sustainable, it would have Pareto dominated \(\vec{DC}\).

To accommodate correlated strategies, we assume that in each period, a public randomization device generates one of four possible action profiles, \(A \equiv \{D, C\} \times \{D, C\}\), whose probabilities are denoted by \(\lambda_{it}, i \in A\) and \(t \geq 1\). More specifically, the mixes over the action profiles are

\[
\begin{array}{ccc}
\text{C} & \text{D} \\
\hline
\text{C} & \lambda_{CC,t} & \lambda_{CD,t} \\
\text{D} & \lambda_{DC,t} & \lambda_{DD,t} \\
\end{array}
\]

where \(\lambda_{CC,t}, \lambda_{DC,t}, \lambda_{CD,t}, \lambda_{DD,t} \geq 0\) and \(\lambda_{CC,t} + \lambda_{DC,t} + \lambda_{CD,t} + \lambda_{DD,t} = 1\), for \(t > 0\). Because players are risk neutral, we assume wlog there is no randomization over amounts of side-payments. Except for the first period net side-payment \(b_1\), equilibrium net side-payments in all periods may be contingent on the history of play. Let the net side-payments, contingent on the previous period action profile, be denoted by \(\{b_{CC,t}, b_{DC,t}, b_{CD,t}, b_{DD,t}\}\) for \(t \geq 2\). In a SPE in correlated strategies, neither player has any incentive to deviate from the recommended action for all histories. Pareto dominant SPE continue to be supported by Nash reversion trigger strategies that punish any deviation from the proposed play path by forever returning to the unique, one-shot Nash
equilibrium play. The characterization of the set of Pareto dominant SPE in correlated strategies is formally stated as:

**Proposition 3** For \((\delta_I, \delta_P) \in \Omega^U \setminus \Omega^2_{DC}\), the Pareto dominant SPE outcomes obtained in Theorem 1 are also Pareto dominant in the set of SPE outcomes when we allow for correlated strategy profiles (under assumption of observable correlated strategies). For \((\delta_I, \delta_P) \in \Omega^2_{DC}\), in every Pareto dominant SPE outcome, the public randomization device mixes between \((C, C)\) and \((D, C)\) in each period, assigning probability \(\lambda_{CC}\) to \((C, C)\) and probability \(1 - \lambda_{CC}\) to \((D, C)\), where

\[
\lambda_{CC} = \delta_I \frac{(z - x) \delta_P - x}{(z - y) \delta_P + (z - x - y) \delta_I - (z - x) \delta_I \delta_P}.
\]

For all \(t \geq 1\), if \((C, C)\) is played in period \(t\), then \(b_{t+1} = b_{CC}(\lambda_{CC})\), and if \((D, C)\) is played in period \(t\), then \(b_{t+1} = b_{DC}(\lambda_{CC})\), where

\[
b_{CC}(\lambda_{CC}) = b - (1 - \lambda_{CC})(z - y),
\]

\[
b_{DC}(\lambda_{CC}) = -[(\lambda_{CC}(y - x) + (1 - \lambda)(z - x)].
\]

**Proof.** The proof of this proposition and those of Propositions 4-5 are relegated to a web appendix available for download at

http://www.kellogg.northwestern.edu/faculty/fong/htm/Discussion_proofs.pdf

An important implication of Proposition 3 is that the tradeoff between maximizing total stage game payoff and maximization of intertemporal trade continues to retain its critical role in determining the Pareto dominant SPE outcomes even when correlated strategies are permitted. The only difference is that when \((\delta_I, \delta_P) \in \Omega^2_{DC}\) and players use correlated strategies, another form of partial cooperation, characterized by mixing between \((D, C)\) and \((C, C)\), arises as a Pareto dominant SPE profile.
We provide some intuition for Proposition 3 through a brief sketch of the main ideas. Based on a logic similar to that of Lemma 1, any correlated strategy profile that implements Pareto dominant SPE payoffs must necessarily involve a time invariant randomization over the action profiles in the action stage on the equilibrium path. Only the side payment may be contingent on the action profile picked by the randomization device in the action stage of the previous period.

To achieve Pareto dominant SPE outcomes, when the impatient player is called upon to play $C$, not knowing whether $P$ plays $C$ or $D$, the expected side payment must be at a level such that his IC constraint binds. Similarly, when he is called upon to play $D$, the expected side payment must be at a level such that his IR constraint binds. It turns out that conditional on barely satisfying player $I$’s IC and IR as described, the Total Utility remains linear in the probabilities over the action profiles. Shuffling probabilities between action profiles in a correlated strategy environment has an impact on the Pareto frontier of the SPE payoffs essentially identical to that of choosing among these action profiles for the optimal one, when the continuation play path follows a pure strategy. Moving all the probabilities on to the action profile $(C, C)$ leads to the highest Total Utility when $\delta_I \geq \delta_P (z - y) / y$, and moving all the probability weight on to the action profile $(D, C)$ leads to the highest Total Utility when $\delta_I \leq \delta_P (z - y) / y$. Since $\overrightarrow{CC}$ is indeed a SPE path of play for $(\delta_I, \delta_P) \in \Omega_{CC}$, where $\delta_I \geq \delta_P (z - y) / y$, it must be Pareto dominant. Since $\overrightarrow{DD}$ is indeed a SPE play path for $(\delta_I, \delta_P) \in \Omega_{DD}$, where $\delta_I \leq \delta_P (z - y) / y$, it is Pareto dominant. When $(\delta_I, \delta_P) \in \Omega_{DC}$, there does not exist a SPE equilibrium which places a positive probability on any action profile other than $(D, D)$. So $\overrightarrow{DD}$ continues to be Pareto dominant.

The only region of $(\delta_I, \delta_P)$ in which correlated strategies actually help create Pareto improvement is $\Omega^2_{DC}$. When $(\delta_I, \delta_P) \in \Omega^2_{DC}$, since $\delta_P \leq y\delta_I / (z - y)$, $\overrightarrow{CC}$ leads to a higher Total Utility than does $\overrightarrow{DC}$. However, since $1/\delta_I + 1/\delta_P \geq 2/\delta^*$, $\overrightarrow{CC}$ is not sustainable.


as an SPE outcome in pure strategies. Nevertheless, by mixing over \((C, C)\) and \((D, C)\) in each period, we can pool \(P\)'s ICs for \(\overrightarrow{CC}\) and \(\overrightarrow{DC}\), using the slack in the IC for \(\overrightarrow{DC}\) to help satisfy the IC for \(\overrightarrow{CC}\). Such mixing over \((C, C)\) and \((D, C)\) is also a form of a partial cooperation which represents the trade off between maximizing total stage game payoff and maximizing intertemporal trade.

**Renegotiation-proofness**

The characterization of the set of Pareto dominant SPE outcomes conducted in the previous section is based on the assumption that cooperation, partial or full, is supported by trigger strategies. As a result, the equilibria giving rise to the play paths \(\overrightarrow{CC}\) and \(\overrightarrow{DC}\) are not renegotiation-proof. In this subsection, we explore how to modify the punishment paths to make these play paths renegotiation proof, while continuing to restrict attention to pure strategy SPEs.\(^{13}\) We identify a wide range of parameter values under which the Pareto dominant play paths derived in the previous section are robust to renegotiation. More precisely, for every given vector of parameters \(x, y, \) and \(z,\) we characterize conditions on players’ discount factors \((\delta_I, \delta_P)\) under which \(\overrightarrow{CC}, \overrightarrow{DC},\) and the entire set of Pareto dominant SPE payoffs implemented by these play paths are renegotiation-proof.

We begin with the characterization of conditions under which \(\overrightarrow{CC}\) and the entire set of Pareto dominant SPE payoffs implemented by \(\overrightarrow{CC}\) survive renegotiation-proofness.

**Proposition 4** Let \((\delta_I, \delta_P) \in \Omega_{CC}\). (i) If \(y \leq \max \left\{ z - x + \frac{x^2}{z}, \min \left\{ \frac{z + x}{2}, z - \frac{x}{2} \right\} \right\}, \) then for all \((\delta_I, \delta_P) \in \Omega_{CC}, \overrightarrow{CC}\) and the entire set of Pareto dominant SPE payoffs implemented by \(\overrightarrow{CC}\) survive renegotiation-proofness. (ii) Suppose \(y > \max \left\{ z - x + \frac{x^2}{z}, \min \left\{ \frac{z + x}{2}, z - \frac{x}{2} \right\} \right\}. \) Then \(\overrightarrow{CC}\) and the entire set of Pareto dominant SPE payoffs implemented by \(\overrightarrow{CC}\) survive

\(^{13}\)Since correlated strategies may create Pareto improvement only when \((\delta_I, \delta_P) \in \Omega_{DC}^2,\) our analysis of renegotiation-proofness for \((\delta_I, \delta_P) \in \Omega_U \setminus \Omega_{DC}^2\) generalizes to the case where players are allowed to use correlated strategies.
renegotiation-proofness for all \((\delta_I, \delta_P) \in \hat{\Omega}_{CC}\) where
\[
\hat{\Omega}_{CC} \equiv \left\{ (\delta_I, \delta_P) \in \Omega_{CC} : \frac{\delta_P}{\delta_I} \notin \left( \frac{z-y}{x+y-z}, \frac{y-x}{z-y} \right) \cup \left( \frac{z-y}{y-x}, \frac{x+y-z}{z-y} \right) \right\}.
\]

Next, we show that \(\overrightarrow{DC}\) and the entire set of Pareto dominant SPE payo
ffs implemented by \(\overrightarrow{DC}\) survives renegotiation-proofness for all \((\delta_I, \delta_P) \in \Omega_{DC}\).

**Proposition 5** For all \((\delta_I, \delta_P) \in \Omega_{DC}, \overrightarrow{DC}\) and the entire set of Pareto dominant SPE payo
ffs implemented by \(\overrightarrow{DC}\) survive renegotiation-proofness.

A necessary condition for an SPE strategy profile to be renegotiation-proof (Farrell
and Maskin (1989)) is that in punishing a deviation in period \(t \geq 1\), the play path - period
\(t+1\) onward - must return to the Pareto dominant SPE play path. If not, then the players
can renegotiate away to a Pareto improving SPE play path.

The renegotiation-proofness of the entire set of Pareto dominant SPE implemented by
either \(\overrightarrow{CC}\) or \(\overrightarrow{DC}\) hinges on the existence of renegotiation-proof punishments at the side-
payment stage that pushes the player’s payoff to the IR level, \(x\), when (s)he is designated
to make a side-payment but refuses to do so. Once such punishments exist, renegotiation-
proof continuation paths that punish deviations at the action stage by any player can be
constructed as follows. If player \(i \in \{I, P\}\) deviates in the action stage of period \(t\) from the
equilibrium/punishment path action profile, she will be punished by reverting to a Pareto
dominant SPE play path, wherein the period \(t+1\) side-payment ensure that \(U_{it+1} = x\).

Since this is what Nash reversion would have brought about, the SPE is deviation proof at
the action stage as long as renegotiation-proof punishments that are severe enough exist
at the side-payment stage.

The actual construction of punishments that push a deviator at the side payment stage
to his/her IR level being delicate - and sometimes parameter dependent - is relegated to
the web appendix. But here we sketch some of the basic ideas behind it. The key to
enforcing the IR payoff \((x)\) for one of the players at the side payment stage in period \(t\) is to require the other player to play \(D\) as soon as the player designated to receive the IR payoff refuses to make the side payment that implements the IR payoff. As long as the punishment path following the refusal to pay begins with the other player playing \(D\), the deviator’s period \(t\) instantaneous payoff will be no larger than \(x\) and there always exists a side payment amount in period \(t+1\) such that the deviator will receive a period \(t\) average payoff no larger than \(x\).

One difficulty with such constructions is that the deviator in period \(t\) can propose instead to switch to playing \((C, C)\) which may be mutually beneficial to both players, as compared to having at least the non-deviating player playing \(D\). In order for the punisher not to be tempted by this one-shot gain coming from playing \((C, C)\), \(y\) must not be too large. This explains the upper bound on \(y\) in Proposition 4(i) for the renegotiation proofness of the entire set of Pareto dominant SPE payoffs implemented by \(\overrightarrow{CC}\). In fact, this also explains Proposition 4(ii), as the restriction of \((\delta_I, \delta_P)\) to \(\hat{\Omega}_{CC}\) in Proposition 4(ii) can also be rewritten as a upper bound on \(y\):

\[
y \leq \min \left\{ \max \left\{ \frac{x + \rho z}{1 + \rho}, \frac{z + \rho (z - x)}{1 + \rho} \right\}, \max \left\{ \frac{z + \rho x}{1 + \rho}, \frac{z - x + \rho z}{1 + \rho} \right\} \right\},
\]

where \(\rho \equiv \delta_P/\delta_I\).

When \(\overrightarrow{DC}\) is Pareto dominant, it is easier to discipline a player’s incentive to make a side payment, rendering the entire set of Pareto dominant SPE payoffs implemented by \(\overrightarrow{DC}\) renegotiation-proof for all \((\delta_I, \delta_P) \in \Omega_{DC}\). This is partly due to the fact that playing \((C, C)\) is either not as efficient as playing \((D, C)\), (when \((\delta_I, \delta_P) \in \Omega_{DC}^1\)), or simply not sustainable, (when \((\delta_I, \delta_P) \in \Omega_{DC}^2\)), so that it becomes harder for the punished player following his/her refusal to make the side payment to renegotiate with the punishing player by proposing to start a punishment path with \((C, C)\).

**Generalization**
We believe the main insights presented in this paper are applicable to a larger class of repeated games augmented with side payments. In this subsection we discuss the generalizability of the following features of Pareto dominant SPE play paths on the repeated PD with side payments: the tradeoff between maximizing the total stage game payoff and maximizing intertemporal trade; the dispensability of public randomization; and the stationarity of Pareto dominant SPE play paths. Since the formal results on the robustness of the equilibria to renegotiation proofness are technical by nature, we abstain from speculating their generalization to other games.

The most important insight of our paper is that when agents have different time preferences, Pareto dominant SPE outcomes exhibit tradeoff between maximizing the total stage game payoff and maximizing intertemporal trade. It is clear that if the action profile(s) maximizing total stage game payoffs in the underlying (stage) game is also a static Nash equilibrium, then such trade-off is absent. However, this tradeoff still arises in a wide range of games. We conjecture that whenever the action profile(s) maximizing total stage game payoffs in the underlying (stage) game is not a static Nash equilibrium, and a player with the incentive to deviate in the one-shot game becomes sufficiently impatient, achieving Pareto dominant SPE outcomes requires that players settle on a suboptimal total stage game payoff. The following (non-zero-sum) variant of the matching pennies game provides an example of this phenomenon:

\[
\begin{array}{c|cc}
 & H & T \\
\hline
H & (8, -5) & (-5, 5) \\
T & (-5, 10) & (5, -5)
\end{array}
\]

where I is the row player and P the column player. In this example, the total stage game payoff is maximized if players play \((T, H)\) every period. However, given that \(P\) plays \(H\), requiring \(I\) to play \(T\) entails giving him a large future payoff in order to prevent his temptation to deviate and earn an instantaneous gain of 13 units at the action stage. If \(P\)
is very patient, i.e., \( \delta_P \sim 1 \), but \( I \) is sufficiently myopic, this will result in a very inefficient intertemporal allocation of payoffs which may not even be sustainable. Playing \((H,H)\) results in 2 units lower of total stage game payoff as compared to playing \((T,H)\). However, doing so also removes player \( I \)'s incentive to deviate in the action stage, meaning that from the second period onward, his continuation payoff can be pushed down to his IR level. In this context, when the impatient player is required to receive low future payoffs, it also means that he is likely to be the one making side payments to the patient player on the steady state equilibrium play path.\(^{14}\) Quite intuitively, the very same tradeoff between total stage game payoff and intertemporal trade also exists when the underlying stage game is an asymmetric Prisoners’ Dilemma.

A second important implication of our analysis was that when players can use side payments, they often do not have to use correlated strategies to achieve Pareto dominant SPE outcomes. In fact, for some classes of games, side payments render correlated strategies redundant for the entire class. For instance, this is true for all games in which the action profile maximizing the total stage game payoff is also a Nash equilibrium of the underlying game. Examples of such games include battle of sexes and the chicken game. The reason for the redundancy of public randomization lies in the absence of a trade-off between maximization of total stage game payoffs and maximizing intertemporal trade.

For illustration consider the following variant of the battle of the sexes.

<table>
<thead>
<tr>
<th></th>
<th>( L )</th>
<th>( R )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>(2,3)</td>
<td>(0,0)</td>
</tr>
<tr>
<td>( B )</td>
<td>(0,0)</td>
<td>(3,1)</td>
</tr>
</tbody>
</table>

In this case, not only does the action profile \((T, L)\) maximize total stage game payoffs, it is also a (Pareto dominant) Nash equilibrium of the stage game. Many \( 2 \times 2 \) coordination

\(^{14}\)The exception to the rule arises when the distribution of payoffs resulting from the chosen mix over action profiles is sufficiently unfavorable to the impatient player such that he has no excess payoff to give the patient one.
games share this property, whereas the PD and matching pennies violate this property. It is easily shown that all Pareto dominant SPE of the infinitely repeated version of this coordination game with side payments involve perpetual play of \((T, L)\) on the equilibrium path. Moreover, owing to the difference in incentives in the underlying stage games, Pareto dominant SPE of repeated coordination games augmented with our side payments mechanism have the property that public randomization is redundant from the perspective of improving on the Pareto frontier of the (mixed) strategy SPE payoff set.

Finally, recall that the stationarity of side payments and action profiles (or correlated mixes over action profiles where applicable), along Pareto dominant SPE play paths, were derived (Lemma 1) without reference to the underlying game being played. Hence, this property extends to Pareto dominant SPE of repeated, two player games generally.

5 Conclusion

When players have different time preferences, side-payments can be used in a long-term relationship to both provide incentives to cooperate and conduct intertemporal trade of instantaneous payoffs. Depending on how different players’ time preferences are and how myopic the less patient player is, Pareto dominant SPE outcomes may be achieved by full or partial cooperation. The steady-state side payment that help attain these outcomes is non-monotone and its magnitude and direction change nontrivially in players’ discount factors. The use of side payments renders publicly observable randomization devices redundant except for a specific subset of discount factors, and in many situations the SPE outcomes are robust to renegotiation.

We make the important assumption that agents have sufficient endowments to finance the bribes required to achieve the most efficient equilibrium outcomes. By focusing on this special case, our study provides a useful benchmark (for comparison with cases where
agents’ budgets are more limited). It will be interesting for future research to study how realistic constraints on players’ abilities to make side-payments affect the Pareto frontier of the set of SPE outcomes. Yet another of our assumptions is related to the exogeneity of the timing of the side-payments in every period. One way of endogenizing this variable is to assume that players are free to make the side-payment at any point in time between two consecutive action stages and that they choose the timing that generates the highest Total Utility. Depending on the amount and direction of side-payments they agree upon, agents may pay right before or right after the action stage.\textsuperscript{15} Another obvious extension is to model how agents bargain over the gains from cooperation. Such an analysis will allow us to pin down the first period side-payment, which is indeterminate in this paper.

\textsuperscript{15}The authors’ separate analysis of this issue is available at http://www.kellogg.northwestern.edu/ faculty/fong/htm/timing.pdf.
Appendix

Proof of Lemma 1. (i) Wlog, suppose $W(t) < W(t')$ for some $t \neq t'$ in a Pareto dominant SPE play path $\sigma$. Consider an alternative strategy profile $\tilde{\sigma}$ in which along the equilibrium path $\tilde{a}_\tau = a_\tau$ and $\tilde{b}_\tau = b_\tau$ for all $\tau < t$, but the play path of the continuation subgame which begins at the action stage of period $t$ is replaced by the play path of the continuation subgame which begins at the action stage of period $t'$. Since $W(t') > W(t)$, an appropriate adjustment of the side payment $b_t$ can lead to some $(\tilde{U}_{I1}, \tilde{U}_{P1}) > (U_{I1}, U_{P1})$. It follows from $\tilde{a}_\tau = a_\tau$ and $\tilde{b}_\tau = b_\tau$ for all $\tau < t$ that $(\tilde{U}_{I1}, \tilde{U}_{P1}) > (U_{I1}, U_{P1})$. This contradicts the assumption that $\sigma$ is Pareto dominant.

(ii) Suppose $(a_t, b_{t+1}) \neq (a_{t+1}, b_{t+2})$ for some $t \geq 1$ but $W(t) = W(t+1) = W(t+2)$. By replacing $(a_{t+1}, b_{t+2})$ by $(a_t, b_{t+1})$, the new strategy profile is still sustainable and $W(t+1)$ remains at $W(t)$. Therefore, $W(1)$ can be achieved by a strategy profile in which $(a_t, b_{t+1}) = (a_1, b_2)$ for all $t \geq 1$.

Proof of Lemma 2 (i) By contradiction, suppose $a_t = (C, D)$ for some $t \geq 1$ on a Pareto dominant SPE play path. For $a_t = (C, D)$ to be mutually incentive compatible, $P$’s IR and $I$’s IC to play $C$ must be satisfied at $t$: i.e., respectively, $U_{ Pt+1 } \geq x$ and

$$ (1 - \delta_I) (0) + \delta_I U_{ It+1 } \geq (1 - \delta_I) x + \delta_I x \Rightarrow U_{ It+1 } \geq \frac{x}{\delta_I}. $$

Consider an alternative strategy profile in which along the equilibrium path $a'_t = (D, C)$ but $a'_\tau = a_\tau$ for all $\tau \neq t$ and $b'_\tau = b_\tau$ for all $\tau \neq t, t+1$. Adjust $b_{t+1}$ downward such that $U_{ It+1 }' = x$. Such adjustment leaves the aggregate payoff in period $t+1$ unchanged,
implying that
\[ \frac{U'_{It+1}}{1 - \delta_I} + \frac{U'_{Pt+1}}{1 - \delta_P} = \frac{U_{It+1}}{1 - \delta_I} + \frac{U_{Pt+1}}{1 - \delta_P} \]
\[ \Delta U_{Pt+1} = -\frac{1 - \delta_P}{1 - \delta_I} \Delta U_{It+1} \]
\[ = \frac{1 - \delta_P}{1 - \delta_I} (U_{It+1} - x) > 0. \]

Since \( U'_{It+1} \) and \( U'_{Pt+1} \) are both at least \( x \), it is incentive compatible to make the side payment in period \( t + 1 \). Also raise \( b_t \) by \( \delta_I (x - U_{It+1}) \) so that \( U_{It} \) remains unchanged.

Next, notice that
\[ U_{it} = (1 - \delta_i) u_{It} + \delta_i U_{it+1}, \; i \in \{I, P\}. \]
This implies that
\[ W_t = (u_{It} + u_{Pt}) + \delta_I \frac{U_{It+1}}{1 - \delta_I} + \delta_P \frac{U_{Pt+1}}{1 - \delta_P}. \]
Also notice that \( u'_{It} + u'_{Pt} = u_{It} + u_{Pt} = z \). These changes lead to
\[ \Delta W_t = \Delta (u_{It} + u_{Pt}) + \delta_I \frac{\Delta U_{It+1}}{1 - \delta_I} + \delta_P \frac{\Delta U_{Pt+1}}{1 - \delta_P} \]
\[ = \frac{\delta_P - \delta_I}{1 - \delta_I} (U_{It+1} - x) > 0. \]
Since \( U_{It} \) remains unchanged, \( \Delta U_{Pt} = (1 - \delta_P) \Delta W > 0 \). Once again, it is incentive compatible to make side payment in period \( t \) and Pareto improvement can be created.

(ii) Suppose \( 2x > z \) and in equilibrium a strategy profile in which \( a_t = (D, C) \), for some \( t \), is played. Modify this strategy profile by replacing the action profile in the period-\( t \) action stage by \( a'_t = (D, D) \) and raising the net side payment from \( P \) to \( I \) by \( z - x \) in the same period, but leaving the rest of the equilibrium play path identical to the original profile. The new strategy profile yields an increase in \( U_{Pt} \) by \( (2x - z) \) \([U'_{Pt} - U_{Pt} = (x - 0) - (z - x)]\) and keeps \( U_{It} \) unchanged. Therefore, the incentive to make side payment in period \( t \) remains intact and, by not requiring player \( P \) to cooperate, it also relaxes her incentive constraint in the action stage of the same period. The fact that Pareto improvement can be created renders the original strategy profile Pareto dominated. \]
**Proof of Lemma 3** (i) If $U_{It+1} < x$, then individual rationality is violated. If $U_{It+1} > x$, then there exists $\Delta b_{t+1} = x - U_{It+1} < 0$ and $\Delta b_t = -\delta_I \Delta b_{t+1} > 0$ such that by revising the net transfer from $P$ to $I$ in period $t+1$ downward by $\Delta b_{t+1}$ and raising the net transfer from $P$ to $I$ in period $t$ by $\Delta b_t$, $U'_{It+1} = x$ and $U'_{It} = U_{It}$. Adapting (10) in the proof of Lemma 2,

$$
\Delta W_t = \frac{\Delta U_{It}}{1 - \delta_I} + \Delta u_{Pt} + \delta_P \frac{\Delta U_{Pt+1}}{1 - \delta_P} = 0 - \Delta b_t - \delta_P \frac{(1 - \delta_P) \Delta b_{t+1}}{1 - \delta_P} = (\delta_P - \delta_I) (U_{It+1} - x) > 0.
$$

So Pareto improvement can be created. Therefore, Pareto dominance requires that $U_{It+1} = x$.

(ii) It is incentive compatible for $I$ to cooperate if and only if $(1 - \delta_I) y + \delta_I U_{It+1} \geq (1 - \delta_I) z + \delta_I (x)$, i.e., $U_{It+1} \geq y + b$. If $U_{It+1} < y + b$, then $I$ is unwilling to play $C$ at time $t$. If $U_{It+1} > y + b$, then there also exists $\Delta b_{t+1} < 0$ such that $U_{It+1} \geq y + b$ still holds, and thus Pareto improvement can also be created just as in the case that $a_{It} = D$ and $U_{It+1} > x$. Therefore, it must be that $U_{It+1} = y + b$. ■

**Proof of Proposition 1.** In order to support an equilibrium play path characterized by $(CC)$, it requires that both $I$ and $P$ find it incentive compatible to always play $C$, i.e.,

$$
(1 - \delta_I) z + \delta_I (x) \leq (1 - \delta_I) y + \delta_I (y + b), \quad (1 - \delta_P) z + \delta_P (x) \leq (1 - \delta_P) y + \delta_P (y - b).
$$

Let these inequalities hold, respectively at $b = \underline{b}$ and at $b = \bar{b}$. Then

$$
\underline{b} = x - y + \frac{1 - \delta_I}{\delta_I} (z - y), \quad \bar{b} = y - x - \frac{1 - \delta_P}{\delta_P} (z - y).
$$
For the existence of some $b$ satisfying both (11) and (12), it requires $\bar{b} \leq \overline{b}$. One can verify with algebra that this requirement is equivalent to (5).

For sufficiency, it is easy to verify that when (5) is satisfied, then the action profit $(C, C)$ can be supported in equilibrium by a trigger strategy and the expectation that a net side payment $b \in [\underline{b}, \overline{b}]$ will be paid in every future period. Besides, since

$$y - b \geq y - \overline{b} = x + \frac{1 - \delta_P}{\delta_P} (z - y) > x$$

and

$$y + b \geq y + \overline{b} = x + \frac{1 - \delta_I}{\delta_I} (z - y) > x$$

for $b \in [\underline{b}, \overline{b}]$, the IRs at the bribing stage are automatically satisfied under (5) regardless of the sign of $b$.

Since the maximum amount of intertemporal trade between the first and all future periods is allowed when $b = \underline{b}$ for all $t \geq 2$, the most efficient outcome is achieved by such net side payment.

**Proof of Proposition 2.** For it to be incentive compatible for $P$ to play $C$, it requires that

$$(1 - \delta_P) x + \delta_P x \leq (1 - \delta_P) (0) + \delta_P b_{it}$$

$$\Leftrightarrow b_{it} \geq \frac{x}{\delta_P}.$$ 

For it to be incentive compatible for $I$ to pay, it requires that $z - b_{it} \geq x$, i.e., $b_{it} \leq z - x$. Therefore, there exists $b_{it}$ satisfying both $P$’s IC to play $C$ and $I$’s IC to make the side payment only if $z - x \geq x/\delta_P$. Also, this condition is satisfied for some $\delta_P \in (0, 1)$ only if $z > 2x$. For sufficiency, it is easy to verify that when $z > 2x$ and $\delta_P \geq x/(z - x)$, a play path characterized by $(\overrightarrow{DC})$ can be supported as an equilibrium play path by trigger strategies with some $b_{it} \in [x/\delta_P, z - x]$. Since the maximum amount of intertemporal trade between the first and all future periods is allowed when $b_{it} = z - x$ for all $t \geq 2$, the most efficient outcome is achieved by such amount of side payment.
Proof of Theorem 1. First, we can summarize the discussion up to (9) as

**Lemma A1** Suppose \( \delta_P \geq x/(z - x) \) and \( \delta_I \geq \delta_P \delta^*/(2\delta_P - \delta^*) \) so that \( \overrightarrow{DC} \) and \( \overrightarrow{CC} \) are both SPE play paths. Then \( \overrightarrow{DC} \) Pareto dominates \( \overrightarrow{CC} \) if \( (z - y)\delta_P \geq y\delta_I \), and \( \overrightarrow{CC} \) Pareto dominates \( \overrightarrow{DC} \) if \( (z - y)\delta_P \leq y\delta_I \).

(i) Suppose \( (\delta_I, \delta_P) \in \Omega_{CC} \). In this case, according to Proposition 1, \( \overrightarrow{CC} \) is sustainable as a SPE play path because \( \delta_I \geq \delta_P \delta^*/(2\delta_P - \delta^*) \) and it is most efficient when \( b_{Pt} - b_{It} = b \) \( \forall \ t \geq 2 \). If \( \overrightarrow{DC} \) is also sustainable, then according to Lemma A1, \( \overrightarrow{CC} \) Pareto dominates \( \overrightarrow{DC} \) because \( \delta_I \geq (z - y)\delta_P/y \). It can be directly verified that \( b < 0 \) if and only if \( \delta_I > \delta^* \).

(ii) Suppose \( (\delta_I, \delta_P) \in \Omega_{DC} \). In this case, according to Proposition 2, \( \overrightarrow{DC} \) is sustainable because \( \delta_P \geq x/(z - x) \) and it is most efficient when \( b_{It} = z - x \) and \( b_{Pt} = 0 \) \( \forall \ t \geq 2 \). Also, when \( (\delta_I, \delta_P) \in \Omega_{DC} \), either \( \overrightarrow{CC} \) is not sustainable when \( \delta_I \leq \delta_P \delta^*/(2\delta_P - \delta^*) \), or, according to Lemma A1, \( \overrightarrow{DC} \) Pareto dominates \( \overrightarrow{CC} \) when \( \delta_I \leq (z - y)\delta_P/y \). Therefore, \( \overrightarrow{DC} \) is Pareto dominant.

(iii) Suppose \( (\delta_I, \delta_P) \in \Omega_{DD} \). In this case neither \( \overrightarrow{CC} \) nor \( \overrightarrow{DC} \) is sustainable, so \( \overrightarrow{DD} \) is Pareto dominant. This completes the proof of the theorem. \( \blacksquare \)
References


