Bribing to Cooperate*

Yuk-fai Fong† and Jay Surti‡

February 14, 2007

Abstract

We analyze how agents use side payments to induce cooperation in the infinitely repeated Prisoners' Dilemma. For every pair of discount factors, we characterize the Pareto frontier of the set of sub-game perfect equilibrium payoffs. Play paths implementing Pareto dominant equilibrium payoffs are uniquely determined in all but the first period. It is an intuitive conjecture that whenever full cooperation maximizes the aggregate instantaneous payoff, it necessarily implements Pareto dominant equilibrium outcomes, supported by repeated bribes from the patient to the impatient player. However, this conjecture is valid only when one player is sufficiently impatient, yet the difference in the players' time preferences is moderate. In contrast, when the difference in discount factors is sufficiently large, all Pareto dominant equilibrium payoffs are achieved by partial cooperation supported by repeated bribes from the impatient player to the patient player. Finally, when both players are sufficiently patient, Pareto dominant equilibrium outcomes are achieved through full cooperation, albeit supported by repeated bribes from the impatient to the patient player. We show that public randomization has no impact upon the Pareto frontier, and characterize conditions for renegotiation-proofness.

*We are grateful to participants of a M&S lunch talk at Northwestern University, of seminars at Queen's University and at the University of Toronto, Ricardo Alonso, Sandeep Baliga, and Eric Chou for helpful comments. The views presented in this paper are the authors' and should not necessarily be attributed to the IMF or IMF policy.

†Northwestern University: Management and Strategy Department, Kellogg School of Management, 2001 Sheridan Road, Evanston, IL 60208; email: y-fong@kellogg.northwestern.edu.

‡International Monetary Fund, 700 19th Street, NW, Washington, DC 20005; email: jsurti@imf.org.

1 Introduction

1.1 Motivation and Overview

In a dynamic setting, 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 environmental protection treaties, or in the context of the opening of domestic markets to international competition; bonuses used to reward employees' performance; and briberies in corruption; among others. While repeated games are a common device to study the impact of dynamic interaction on the feasibility of, and limits to cooperation among agents, side-payments have heretofore been studied sparsely within this framework.

Dynamic games usually analyze situations where all players have a common discount factor. Clearly, this assumption is not universally appropriate as the following examples indicate. An unstable domestic political regime is more likely to induce myopia among its leaders, owing to the fear of loss of power through an assassination or a coup, and this has a material impact on how they negotiate in international fora with leaders from other, stabler political regimes. Employees may be less committed to a firm than employers, and this may be reflected in them discounting at a higher rate future payoffs from the employment relationship. More generally, in environments where the credit market is imperfect and agents have asymmetric access to it, a distribution of heterogeneous time preferences is probably more realistic. We examine a specific side-payments mechanism that players use to transfer payoffs wherein the underlying stage game is the prisoners' dilemma (PD). For every pair of discount factors in the unit square, we characterize the intertemporal structure of play implementing payoffs on the Pareto frontier of the SPE payoffs set of this augmented, transferable utility PD.[1] Focusing attention on these Pareto

[1] A considerable amount of information is available on the Pareto frontier of the repeated PD with com-

dominant SPE paths enables us to uniquely predict the equilibrium outcome(s) in all but the first period.

An intuitive conjecture is that whenever full cooperation maximizes the aggregate stage game payoffs, it necessarily implements Pareto dominant SPE outcomes, supported by repeated bribes from the patient to the impatient player. We find this conjecture to be valid only when the impatient player is sufficiently impatient, yet the difference in the players' time preferences is moderate. In contrast, when the difference in discount factors is sufficiently large, all Pareto dominant SPE payoffs are achieved by partial cooperation, supported from the second period onward, by repeated bribes from the impatient to the patient player. When both players are sufficiently patient, Pareto dominant SPE outcomes are achieved through full cooperation, albeit supported by repeated bribes from the impatient to the patient player starting from the second period. Within each of these transfer protocols, Pareto dominant SPE play paths are differentiated solely by the first period bribe, which can travel in either direction, and whose role is to deliver the bonus payment that implements the selected payoff vector on the Pareto frontier of the SPE payoff set.

The search for Pareto dominant SPE outcomes is facilitated by the key observation that in any given period, the side-payments that the impatient player receives in the future must not exceed the amount which provides him just enough incentive to cooperate. Owing to the difference in time preferences, it is always more efficient to let the impatient player receive any additional payoff over and above this amount up-front, instead of in the future. The cost of achieving full cooperation through such a transfer increases rapidly as the impatient player becomes more myopic, compelling the patient player to offer larger bribes

mon discount factors. Sorin (1986), identifies a bound on the discount factor below which only perpetual non-cooperation can be sustained in SPE, and Mailath, Obara, and Sekiguchi (2003), characterize the maximum SPE payoff obtainable for values of the discount factor that lie between Sorin's bound and a higher bound, above which perpetual cooperation is feasible in SPE. van Damme (1989), examines conditions under which Pareto dominant SPE payoffs are also renegotiation proof.

in the future to provide him the incentive to cooperate today. At the same time, however, an increase in the disparity between the players' time preferences makes it more efficient to frontload the impatient player's payoff, thereby making it increasingly inefficient to reward him with a higher future payoff. Consequently, when the impatient player is sufficiently myopic relative to the patient player, Pareto dominant SPE involve partial cooperation in which only the patient player cooperates. In order to provide her the incentive to do so, the impatient player must pay her every period from the second period onward. Such an arrangement fails to maximize the total stage game payoff in any period, but it allows greater frontloading of the impatient player's payoff, because he only needs to be given a future payoff sufficient to satisfy his (weaker) individual rationality (IR) constraint given that he is not required to cooperate.[2] When both players are excessively patient, each of them is willing to cooperate even if (s)he has to give some bribe to the other player. As it is more efficient to let the impatient player consume up-front than in the future, efficiency demands that he transfer any future payoff beyond his incentive constraint (IC), to the patient player. This explains why efficiency within the SPE payoff space requires that the bribe travel from the impatient to the patient player from the second period onward.[3]

The paper is structured as follows: in the following sub-section, we survey the related literature along two distinct dimensions: dynamic models that blend non-cooperative games with some form of side-payments mechanism, and dynamic games with differential time preferences. In section 2, we describe our side-payments mechanism, which we analyze in Section 3, where we characterize the properties of Pareto dominant SPE play paths. Section 4 discusses the robustness of our results to perturbations of the side-payments

[2] Once the impatient player's discount factor falls below a (still lower) critical threshold, the amount of payment the patient player must make in order to maintain the impatient player's IC under full cooperation violates her own IC, thereby making it impossible to sustain full cooperation in equilibrium.

[3] If the equilibrium under consideration involves a first period transfer from the patient to the impatient player, this can be interpreted as future transfers made by I in exchange for an upfront transfer from the patient player.

mechanism and to refinements of the equilibrium concept. Section 5 concludes. Most of the proofs are relegated to an appendix.

1.2 Related Literature

Side-Payments Mechanisms

Among the papers studying the impact of side-payments mechanism on the structure and properties of set of equilibria in repeated games, Baliga and Evans (2000) is closest to ours.[4] 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, constraining SPE through renegotiation-proofness necessarily restricts attention to payoffs that result from maximization of the total stage game payoff in each period. By contrast, we show that maximization of aggregate instantaneous 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 joint payoffs are maximized in every stage game.

Our paper is also related to the literature on relational incentive contracts.[5] There are, however, important differences between such contracts and our mechanism. The relational

[4] Important contributions in the area of explicit contracting mechanisms include Andreoni and Varian (1999), Barrett (2001), and Jackson and Wilkie (2005). Side-payments are not explicitly contractible in our paper. The role of side payments in initiating cooperation in a variety of informational settings has also been examined, notably, by Camerer (1988), Carmichael and MacLeod (1997), and Bolle (2001). Unlike in these models where a single side payment is used to signal a player's type, in our setting of perfect information, only repeated side-payments can induce cooperation.

[5] See Levin (2003), for a general model, and an exhaustive bibliography of contributions to this important area.

contracting literature is concerned with relationships where the incentive problem is one-sided. In our paper, both players may lack incentives to contribute to the production of surplus in the underlying game. Moreover, as our players have different time preferences, and transfers are made to conduct inter-temporal trade and to induce cooperation, steady-state transfers may travel either way between the patient and impatient players, depending on their relative patience. In other words, the role of the principal (briber), may be taken up by either player endogenously.

Differential time preferences

The literature on repeated games with different time preferences is still relatively small. In an important recent contribution, Lehrer and Pauzner (1999) have studied how players in a repeated game exploit the difference in their time preferences through inter-temporal trade of instantaneous payoffs to enhance efficiency.[6] 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 players, as both become arbitrarily patient,[7] they can achieve outcomes in equilibrium which would be infeasible were their time preferences identical.

Despite our focus on the PD, we generalize Lehrer and Pauzner's analysis by allowing

[6] Lehrer and Yariv (1999) also analyze two-player repeated games with different discounting 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 exploitation of heterogeneous discount factors to build a reputation for tough behavior and guarantee a minimum dynamic payoff stream has also been (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.

[7] 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.

for side-payments and considering all combinations of discount factors in the unit square. Side-payments allow us to identify Pareto dominant SPE play paths which are impossible to support in their absence. For example, in our PD game with side-payments, even when the impatient player is very impatient, SPE exist wherein the patient player cooperates and 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, on any Pareto dominant SPE play path with full or partial cooperation, the impatient player's continuation payoff is pushed down from the second period onward, to either his IC or IR constraint, and play is characterizable in terms of a steady-state in both the bribing and action stages. Without side-payments, this is feasible. Instead for the case where both players are very patient, the impatient player's continuation payoff will decay more gradually to her IC constraint as in Lehrer and Pauzner. The impact of side-payments is also felt in terms of our ability to precisely characterize the Pareto frontier of the SPE payoff set: while Lehrer and Pauzner's analysis covers a much wider class of games than ours, absent side-payments, their characterization theorems are limited to the case where the patient and impatient players are both arbitrarily patient.

2 Model

The underlying game is a two-person infinitely repeated prisoner's dilemma except that in each period, they are allowed to each voluntarily make a transfer of payoff to the other player before players choose whether to defect or cooperate. So, the stage game of the repeated game consists of two stages which we term as the *bribing stage* and the *action stage*.

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

$$\begin{array}{c|c|c|} & C & D \\ \hline C & y, y & 0, z \\ \hline D & z, 0 & x, x \\ \hline \end{array} \tag{1}$$

where $z > y > x > 0$, guaranteeing that (1) is the generic, symmetric PD. The row player, denoted by I, is assumed to be less patient than the column player, denoted by P; i.e., we assume $\delta_P \geq \delta_I$. Finally, we assume that the aggregate instantaneous payoff is the highest under full cooperation, i.e., $2y > z$. This is a very natural assumption in many Prisoners' Dilemma like situations. For example, if we think of (C, C) as a collusive outcome in a duopoly and action D as 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.

Bribing Stage. At the beginning of each period, both players simultaneously and independently decide on an amount of side payment/ bribe that they wish to offer the opponent.[8] Following Baliga and Evans (2000), we assume that players have large budgets for bribing in each period. In particular, players have enough to bribe any amount that a Pareto dominant *SPE* may prescribe. After the players exchange the side payments (or bribes), they play the stage game (1). Assuming the side payments to be non-negative, every unit of bribe both reduces the giver's utility and increases the recipient's utility by one unit. Denote the bribe given in period t by player i by $b_{i,t} \geq 0$, an action taken by her in period t by $a_{i,t}$, where $t \in \{1, 2, ...\}$ and $i \in \{I, P\}$. Let the net bribe from P to I be $b_t \equiv b_{P,t} - b_{I,t}$.

A period t outcome – at the bribing stage – is defined by $O_t^{\mathcal{B}} := (b_{I,t}, b_{P,t})$, and at the

[8] Since both players discount the future, by assuming that players bribe right before the stage game we provide them with the strongest incentives to bribe. It will soon be clear that the incentive compatibility constraints in the bribing stage becomes trivial because of this assumption. In the conclusion, we provide a brief discussion of how to endogenize the timing of bribes.

action stage, is defined by $O_t^{\mathcal{A}} = (a_{I,t}, a_{P,t})$. A t-history is defined as a play path including all bribes and actions through stage $t-1$; $h_t := \left\{O_s^{\mathcal{B}}, O_s^{\mathcal{A}}\right\}_{s=1}^{t-1}$. In addition, a history through the action stage of period t is $h_t^{\mathcal{B}} := h_t \cup \left\{O_t^{\mathcal{B}}\right\}$. H_t is the set of possible t histories, $H_t^{\mathcal{B}}$ the set of all $h_t^{\mathcal{B}}$, and $H = (\cup H_t) \cup \left(\cup H_t^{\mathcal{B}}\right)$ is the set of all possible histories. We assume perfect monitoring throughout. Let $u_i^{\mathcal{A}}(a_t)$ denote i's stage game payoff under the action profile a_t. The net payoff to player i in period t is $u_i\left(O_t^{\mathcal{B}}, O_t^{\mathcal{A}}\right) := u_i^{\mathcal{A}}(a_t) + (b_{j,t} - b_{i,t})$. For a play path, $\left\{O_t^{\mathcal{B}}, O_t^{\mathcal{A}}\right\}_{t\geq 1} \in H$, the repeated game average payoff to player i at time t is

$$U_{i,t}\left(\left\{O_\tau^{\mathcal{B}}, O_\tau^{\mathcal{A}}\right\}_{\tau\geq t}\right) = (1-\delta_i)\sum_{\tau=t}^{\infty}\delta_i^{\tau-t}u_i\left(O_\tau^{\mathcal{B}}, O_\tau^{\mathcal{A}}\right) \tag{2}$$

As this is a game of complete and perfect information, we use subgame perfect equilibrium (SPE henceforth) as our solution concept.

3 Analysis

The main goal of our analysis is to characterize the set of Pareto dominant SPE payoffs. For clarity of exposition, we limit our attention to pure strategy equilibria; in section 5, we formally argue that this treatment is without loss of generality. Before proceeding to prove the main findings in the model, we would like to 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.[9] In the unique one-shot Nash equilibrium, agents play $b_I = b_P = 0$ in the bribing stage and (D, D) in the action stage, enforcing individually rational payoffs in the supergame, hence representing extremal punishment strategies. Consequently, in order to derive Pareto dominant equilibrium payoffs, we focus attention on these strategies in this section.

[9] See Abreu (1986, 1988).

Note that it is (weakly) Pareto dominated to have both $b_P > 0$ and $b_I > 0$ in a bribing stage for the following reason. Only $b = b_P - b_I$ is relevant for the IRs and ICs in any other stage. Moreover, reducing each player's bribe by the amount $\min\{b_I, b_P\}$ will relax players' ICs to bribe at the current stage because players' deviation payoffs are decreased by $\min\{b_I, b_P\}$. So, in the rest of the paper, we will focus on characterizing the net bribe b and when we state that $b > 0$, it means $b_P = b > 0 = b_I$; when $b < 0$, it means $b_P = 0 < -b = b_I$. Also note that because the extremal punishment strategies are applied and players do not discount between the bribing stage and the action stage in the same period, the designated briber at the bribing stage of every period is willing to bribe as long as a continuation payoff of x (net the bribe) is promised.

Next, we would like to eliminate some action profiles that are clearly Pareto dominated.

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

Proof. See appendix. ■

The basic idea behind part (i) of Lemma 1 is that it is less costly to require the impatient player to cooperate than to require the patient player. Take an arbitrary time period t in which only the impatient player cooperates. Now switch the actions of the two players by asking the patient player to play C (and impatient player to play D) in period t instead. By doing so, the aggregate period t instantaneous payoff remains unchanged. However, by requiring the impatient player to play D in period t, his continuation payoff from period $t+1$ onward can be lowered relative to what had to be guaranteed to him under the action profile $(a_{It}, a_{Pt}) = (C, D)$. This allows the players to conduct intertemporal trade by (a) reducing the period $t + 1$ net bribe from the patient to the impatient player; and (b) compensating the impatient player for this "loss" by increasing the net bribe to him in period 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 net bribes b_t and b_{t+1} that make them both strictly better off.

Part (ii) of the lemma is another straight forward result that if the aggregate instantaneous payoff is higher when both players defect than having one of them cooperate, then partial cooperation will never be observed on any Pareto dominant equilibrium path. In this case, players get a higher aggregate payoff playing (D, D) and, since both players are not required cooperate, the IC constraints are easier to meet. However, we take the view that in many situations, the condition $2x > z$ is *unrealistic*. If we think of action D as pollution, then the aggregate instantaneous payoff should be lower when both players pollute than just one of them does, i.e., $2x < z$. Similarly, this assumption would be unrealistic in the collusive duopoly example of section 2.1, wherein the aggregate industry profits would be higher if just one firm were to cut price, than if both firms do so.

Under the maintained assumption that the play path under consideration is Pareto dominant in the SPE payoff set, we now quantify the "price" player P must pay – as measured by the increase in I's next period continuation payoff – if she would like to achieve full cooperation in a given period.

Lemma 2 *On any Pareto dominant equilibrium play path, for* $t \geq 1$, *(i) if player* I *plays* $a_{It} = D$, *then* $U_{It+1} = x$; *(ii) if* $a_t = (C, C)$, *then* $U_{It+1} = y + \underline{b}$, *where* $\underline{b} \equiv \frac{1-\delta_I}{\delta_I}(z - y) - (y - x)$.

Proof. See appendix. ■

The first part of Lemma 2 establishes that if I is not required to cooperate in period t, the most efficient allocation necessarily involves providing him with exactly his individually rational 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 must be paid to him in period t or earlier instead of $t+1$. However, if I is required to play C together with P, his continuation payoff at time $t+1$ must be increased to at least $y + \underline{b}$. Having to raise

I's continuation payoff at time $t+1$ to above his individually rational level will limit the extent to which players can trade the instantaneous payoffs intertemporally.

Lemma 2 points out the efficiency cost involved when players of different time preferences try to achieve full cooperation. It will soon be clear that because of such tradeoff, partial cooperation may Pareto dominates full cooperation even when full cooperation, which provides a higher aggregate payoff, can be supported as an equilibrium outcome. Such tradeoff is absent in most analyses of relational incentive contracts as these studies are based on models in which players share common discount factors.

Lemma 3 *On every Pareto dominant equilibrium play path, $a_t = a_{t'}$ for all $t, t' \geq 1$.*

Proof. See appendix. ■

Lemma 3 suggests that when side payments are allowed, players do not need to rely on switching among action profiles to split the surplus of the game. In fact, it will be inefficient to do so. Ruling out players' switching among different action profiles over time substantially simplify our analysis. Lemmas 1 and 3 together imply that on a Pareto dominant SPE play path, players play repeatedly one of the following action profiles: (C, C), (D, C), or (D, D), where the play of (D, C) is possible when the reasonable condition $z > 2x$ is satisfied.

3.1 Existence of Cooperative SPNE

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 $\underline{b}$ and $\overline{b}$, where $\underline{b} \leq \overline{b}$, such that there is an SPE in pure strategies, the equilibrium play path of which satisfies:*

$$\begin{aligned} &\text{(i)} \quad \forall t \geq 1;\ a_t \equiv (a_{It}, a_{Pt}) = (C, C)\,; \\ &\text{(ii)} \quad \forall t \geq 2;\ b_{Pt} - b_{It} = b \in [\underline{b}, \overline{b}]; \end{aligned} \qquad (\overrightarrow{CC})$$

if and only if this condition is satisfied:

$$\begin{aligned} & \frac{1}{\delta_I} + \frac{1}{\delta_P} \leq \frac{2}{\delta^*} \\ \Leftrightarrow \quad & \delta_I \geq \frac{\delta_P \delta^*}{2\delta_P - \delta^*}, \end{aligned} \tag{3}$$

where $\delta^* \equiv \frac{z-y}{z-x}$*. The equilibrium is most efficient when* $b = \underline{b}$, $\forall t \geq 2$.

Proof. See appendix. ■

Proposition 1 points out very intuitively how the patient player's excessive patience can compensate the impatient player's excessive myopia. In the rest of the paper, we will abuse notation and use $\overrightarrow{CC}$ to denote a play path which satisfies condition ($\overrightarrow{CC}$). As depicted in Figure 1, $\overrightarrow{CC}$ can be an equilibrium play path if and only if the vector of discount factors lie above the curve $1/\delta_I + 1/\delta_P = 2/\delta^*$. Let r_I and r_P be players' discount rate and $\delta^* = 1/(1 + r^*)$. Then (3) can be rewritten as

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

In other words, it does not matter that one of the players is too myopic, as long as they are *on average* patient enough, full cooperation can still be supported. Specifically, when $\delta_I < \delta^*$, these players can agree to a repeated stream of bribes from the excessively patient to the excessively myopic player, the amount of which is high enough to make defection too costly for the impatient player, but also low enough that the patient player is willing to pay in exchange for the impatient player's cooperation.[10]

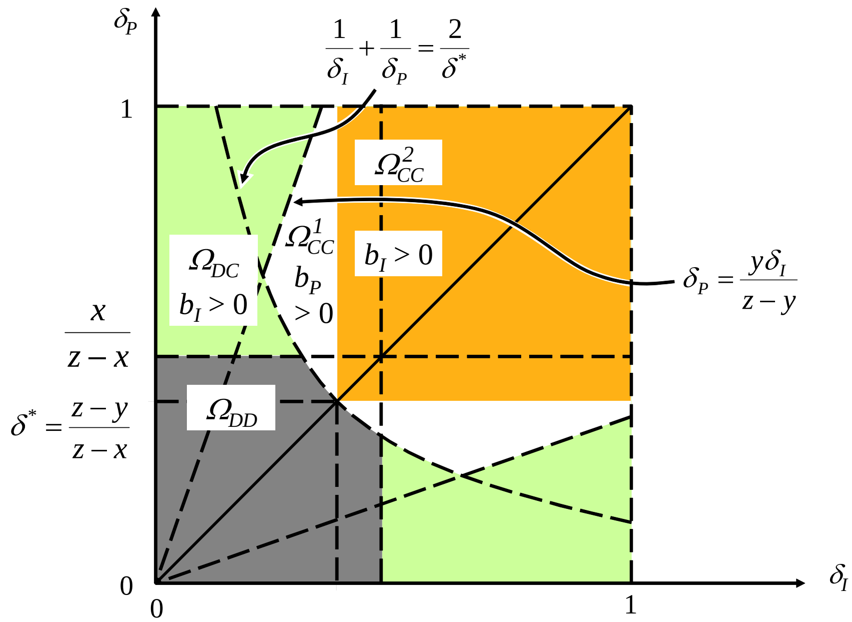

Figure 1: Characterization of Subgame Perfect Equilibria

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

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

$$\begin{array}{ll} \text{(i)} & \forall t \geq 1;\ a_t \equiv (a_{It}, a_{Pt}) = (D, C)\,; \\ \text{(ii)} & \forall t \geq 2;\ b_{Pt} - b_{It} = b \in [x - z, -x/\delta_P]; \end{array} \qquad (\overrightarrow{DC})$$

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

Proof. See appendix. ■

[10] When both players are excessively patient, obviously, (3) will also hold. In that case, it is however most efficient to support full cooperation by having player I repeatedly bribe player P, i.e., $b < 0$. The discussion of this pattern of bribes is postponed to after the statement of the main theorem.

Again, in the rest of the paper we use $\overrightarrow{DC}$ to denote a play path which satisfies condition ($\overrightarrow{DC}$). To sustain an *SPE* with just the patient player cooperating, it first requires that partial cooperation is not dominated by both players defecting, which is guaranteed when $z > 2x$. It is also important that the patient player finds it incentive compatible to receive an instantaneous payoff of zero in the action stage. For this, he must be patient enough to weigh the maximum constant stream of bribes I is willing to pay every period, $z - x$, in the future more than the benefit of instantaneous defect, x. This is exactly the condition $\delta_P \geq x/(z-x)$. The set of discount factors which support this equilibrium is also depicted in Figure 1 as the area above the horizontal line $\delta_P = x/(z-x)$. A symmetric result will hold for all values of δ_I and δ_P where $\delta_I > \delta_P$ and $\delta_I \geq x/(z-x)$.

3.2 Characterization of Pareto dominant SPNE

The obvious disadvantage of $\overrightarrow{DC}$ as compared to $\overrightarrow{CC}$ is the smaller amount of joint instantaneous utility generated from the action profile (D, C). However, playing $\overrightarrow{DC}$ allow players to conduct more intertemporal trade of instantaneous payoffs. As we can see from Figure 1, there are regions where both $\overrightarrow{CC}$ and $\overrightarrow{DC}$ are sustainable as SPE play paths. In these regions, players face a trade-off between the higher aggregate instantaneous utility under $\overrightarrow{CC}$ against more intertemporal trade under $\overrightarrow{DC}$.

To derive the Pareto frontier induced by $\overrightarrow{CC}$, consider an equilibrium play path in which player P pays player I the most efficient stationary bribe $\underline{b}$ in every period, including the first. The average payoffs of I and P in this equilibrium are $y + \underline{b}$ and $y - \underline{b}$ and thus the aggregate payoff (without normalization) is $\frac{y+\underline{b}}{1-\delta_I} + \frac{y-\underline{b}}{1-\delta_P}$. Since adjusting the first period bribe does not affect the aggregate payoff, the Pareto frontier induced by $\overrightarrow{CC}$ is characterized by

$$\frac{U_{I1}}{1-\delta_I} + \frac{U_{P1}}{1-\delta_P} = \frac{y+\underline{b}}{1-\delta_I} + \frac{y-\underline{b}}{1-\delta_P}. \tag{4}$$

With a similar approach, we can derive the Pareto frontier induced by $\overrightarrow{DC}$ as

$$\frac{U_{I1}}{1-\delta_I}+\frac{U_{P1}}{1-\delta_P}=\frac{x}{1-\delta_I}+\frac{z-x}{1-\delta_P}. \tag{5}$$

Therefore, $\overrightarrow{CC}$ Pareto dominates $\overrightarrow{DC}$ if

$$\frac{y+b}{1-\delta_I}+\frac{y-b}{1-\delta_P}>\frac{x}{1-\delta_I}+\frac{z-x}{1-\delta_P},$$

or equivalently

$$(z-y)\,\delta_P<y\delta_I. \tag{6}$$

Similarly, $\overrightarrow{DC}$ Pareto dominates $\overrightarrow{CC}$ if $(z-y)\,\delta_P>y\delta_I$.

Intuitively, partial cooperation – supported by the impatient player repeatedly bribing the patient – is more efficient when the patient player's discount factor is relatively large compared to the impatient's, as depicted in Figure 1 by the area above the ray $\delta_P = y\delta_I/(z-y)$. This is because when δ_P is relatively large as compared to δ_I, the benefit from intertemporal trade outweighs the benefit of earning a joint instantaneous payoff of $2y$ instead of just z.

Now, we are ready to fully characterize the Pareto frontier for all possible pairs (δ_I,δ_P) of discount factors. Let $\overrightarrow{DD}$ denote the noncooperative play path:

$$\begin{aligned}&\text{(i)}\quad \forall t\geq 1;\ a_t\equiv(a_{It},a_{Pt})=(D,D)\,;\\&\text{(ii)}\quad \forall t\geq 1;\ b_{Pt}=b_{It}=0,\end{aligned}$$

and ψ denote the Pareto dominant play path(s). Denote different regions of the (δ_I,δ_P)–space by

$$\begin{aligned}\Omega_{CC}&\equiv\cup_{\delta_P\in[\delta^*,1]}\Phi_{CC}(\delta_P)\,,\\\Omega_{DC}&\equiv\cup_{\delta_P\in[\frac{x}{z-x},1]}\Phi_{DC}(\delta_P)\,,\\\Omega_{DD}&\equiv\cup_{\delta_P\in[0,\frac{x}{z-x}]}\Phi_{DD}(\delta_P)\,,\end{aligned}$$

where $\Phi_{CC}(\delta_P)$, $\Phi_{CC}(\delta_P)$, and $\Phi_{CC}(\delta_P)$ are defined as the following ranges of δ_I:

$$\begin{aligned}
\Phi_{CC}(\delta_P) &\equiv \left[\max\left\{\frac{z-y}{y}\delta_P, \frac{\delta_P\delta^*}{2\delta_P-\delta^*}\right\}, \delta_P\right] \times \{\delta_P\}, \\
\Phi_{DC}(\delta_P) &\equiv \left[0, \max\left\{\frac{z-y}{y}\delta_P, \frac{\delta_P\delta^*}{2\delta_P-\delta^*}\right\}\right] \times \{\delta_P\}, \\
\Phi_{DD}(\delta_P) &\equiv \left[0, \min\left\{\frac{\delta_P\delta^*}{2\delta_P-\delta^*}, \delta_P\right\}\right] \times \{\delta_P\}.
\end{aligned}$$

Theorem 1 *The following properties hold in all Pareto dominant SPE. (i) If* $(\delta_I, \delta_P) \in \Omega_{CC}$, *then* $\psi = \overrightarrow{CC}$ *and the Pareto frontier of SPE payoffs is characterized by (4). In this case,* $b_{Pt} - b_{It} = \underline{b}\ \forall t \geq 2$, $\underline{b} < 0$ *if and only if* $\delta_I > \delta^*$, *and*

$$(b_{P1} - b_{I1}) \in \left[-(z-x), \tfrac{\delta_I(1+\delta_P)(y-x)-\delta_P(1-\delta_I)(z-y)}{\delta_I(1-\delta_P)}\right];$$

(ii) If $(\delta_I, \delta_P) \in \Omega_{DC}$, *then* $\psi = \overrightarrow{DC}$ *and the Pareto frontier of SPE payoffs is characterized by (5). In this case,* $b_{It} = z - x > b_{Pt} = 0\ \forall t \geq 2$, *and* $(b_{I1} - b_{P1}) \in [\frac{x-\delta_P(z-x)}{1-\delta_P}, z-x]$; *(iii) Otherwise, if* $(\delta_I, \delta_P) \in \Omega_{DD}$, *then* $\psi = \overrightarrow{DD}$, $U_{I1} = U_{P1} = x$, *and* $b_{It} = b_{Pt} = 0$ $\forall t \geq 1$.

Proof. See appendix. ■

Figure 1 provides a graphical summary of Theorem 1. We have partitioned Ω_{CC} into two subsets: $\Omega^1_{CC} := \{(\delta_I, \delta_P) \in \Omega_{CC} : \delta_I \leq \delta^*\}$, where player P is the perpetual briber; and $(\delta_I, \delta_P) \in \Omega^2_{CC} := \Omega_{CC} \diagdown \Omega^1_{CC}$; i.e., where $\delta_I > \delta^*$ and player I bribes perpetually.[11] In each region above the 45$^{\text{o}}$ line, we have indicated with different colors and shades the regions of the (δ_I, δ_P)–space where $\overrightarrow{CC}$, $\overrightarrow{DC}$, and $\overrightarrow{DD}$ are the most efficient equilibrium play path, and the regions where players I and P play the role of the perpetual briber.

[11] Generally speaking, $\max\left\{\frac{z-y}{y}\delta_P, \frac{\delta_P\delta^*}{2\delta_P-\delta^*}\right\}$ may be greater than or less than δ^* depending on whether $x + y < z$ or $x + y \geq z$. In the latter case, either expression in the max operator is dominated by δ^*. In this event, for all $\delta_I \geq \delta^*$, the Pareto dominant SPE is necessarily $\overrightarrow{CC}$. However, when $x + y < z$, then for $\delta_P > \frac{y}{z-x} \in (\delta^*, 1)$, $\max\left\{\frac{z-y}{y}\delta_P, \frac{\delta_P\delta^*}{2\delta_P-\delta^*}\right\} = \frac{z-y}{y}\delta_P > \delta^*$, and for $\delta_I \in \left(\delta^*, \frac{z-y}{y}\delta_P\right)$, the Pareto dominant SPE is $\overrightarrow{DC}$. Figure 1 assumes parameters such that $x + y > z$.

When only $\overrightarrow{CC}$ is sustainable as an equilibrium play path or when both $\overrightarrow{DC}$ and $\overrightarrow{CC}$ are sustainable but $(z-y)\,\delta_P > y\delta_I$, the most efficient equilibrium path is $\overrightarrow{CC}$. This region is Ω_{CC}. The first period bribe goes either way depending on the allocation of payoffs. When only $\overrightarrow{DC}$ is sustainable or when both $\overrightarrow{DC}$ and $\overrightarrow{CC}$ are sustainable but $(z-y)\,\delta_P < y\delta_I$, the most efficient equilibrium path is $\overrightarrow{DC}$. This region is Ω_{DC}. When (δ_I, δ_P) falls into this region, player I is the perpetual briber but the first period bribe may go either way. In Ω_{DD}, where neither $\overrightarrow{DC}$ nor $\overrightarrow{CC}$ is sustainable, $\overrightarrow{DD}$ is the unique equilibrium play path.[12] We also note that in part (ii) of Lemma 1, we ruled out the possibility that a Pareto dominant equilibrium can admit $a_t = (D, C)$ for some $t > 0$ when $2x > z$. Notice that in this case, $x/(z-x) > 1$, and as we can see in Figure 1, there is no region in which $\overrightarrow{DC}$ is Pareto dominant. We conclude from our analysis that restricting attention to the Pareto dominant *SPE* payoff set uniquely pins down all the outcomes of the game except the first period bribe.

Figure 2 illustrates how the repeated bribes on the steady state typically change in the impatient player's discount factor.

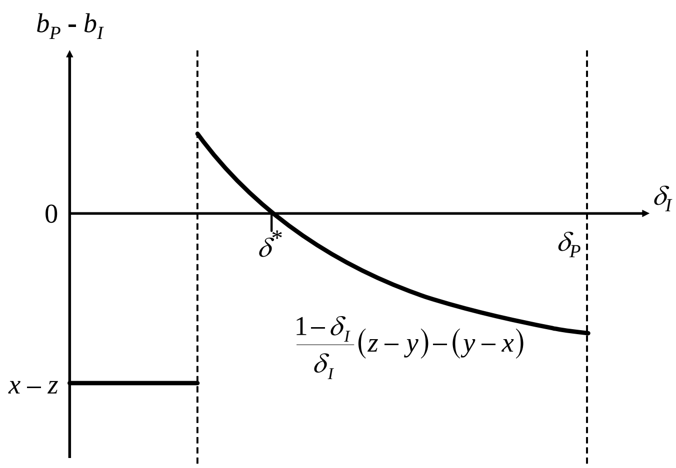

Figure 2: Pattern of repeated bribes, for a fixed $\delta_P \geq x/(z\text{-}x)$

[12] 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 $\overrightarrow{CC}$.

The diagram is drawn based on some $\delta_P \geq x/(z-x)$. When δ_I is relatively small, the most efficient equilibrium outcome will be $\overrightarrow{DC}$, and I pays P a repeated bribe of $z-x$. As δ_I grows beyond a certain level, the most efficient equilibrium outcome becomes $\overrightarrow{CC}$. When δ_I remains below δ^*, it requires P bribing I to give I the incentive to cooperate. Once δ_I goes beyond δ^*, for Pareto dominance, I transfers all the future "excessive" payoffs to P through a bribe which increases in δ_I.

Although unconstrained Pareto efficiency requires both players to cooperate with side-payments used to efficiently split the surplus in each period, the cost of satisfying I's incentive constraint as he gets more myopic makes it mutually beneficial beyond a point, to receive the surplus from partial cooperation rather than full cooperation. This stands in stark contrast with 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 is "ripped off" every period.

4 Discussion

Redundancy of 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 equilibrium set and thus push outward the Pareto frontier of the equilibrium set. In this section, we demonstrate that, for the purpose of identifying the set of Pareto dominant SPE payoffs, extending the strategy space by allowing for public randomization is irrelevant. This suggests that the use of side payments is already a very powerful mechanism.

To accommodate correlated strategies, we assume in each period there is a public randomization device which generates four possible action profiles, $A \equiv \{D, C\} \times \{D, C\}$, whose probabilities, denoted by $\lambda_{i,t}$, $i \in A$ and $t \geq 1$, are chosen based on the history of play. More specifically, the mixes over action profiles are

	C	D
C	$\lambda_{CC,t}$	$\lambda_{CD,t}$
D	$\lambda_{DC,t}$	$\lambda_{DD,t}$

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 bribe amounts. Except the first period net bribe b_1, net bribes in all periods may be contingent on the history of play, including the previous period action profile. Let the net bribes, 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 subgame perfect (correlated) equilibrium, neither player has any incentive to deviate from the recommended action for all histories. Pareto dominant SPE equilibria 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 redundancy of correlated strategies is formally stated as follows.

Proposition 4 *The Pareto dominant equilibrium 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).*

Proof. See appendix. ■

Here we provide the intuition for the proof through a brief sketch of the main ideas. Based on a logic similar to that of Lemma 3, any correlated strategy profile that implements a Pareto dominant equilibrium payoff must necessarily involve a time invariant randomization over the action profiles in the action stage on the equilibrium path. Only the bribe amount may be contingent on the action profile picked by the randomization device in the action stage of the previous period and other parts of the history. For Pareto dominance, when the impatient player is called upon to play C, not knowing whether P

plays C or D, the expected net bribe must be at a level such that his IC constraint binds. Similarly, the expected net bribe must be at a level such that his IR constraint binds when he is called upon to play D.

In turn, this enables us to prove that there exists a payoff equivalent correlated SPE in which the moderator actually "tells" the players which action profile has been selected for play by the public randomization device. The incentive constraints in this correlated SPE are tailored to each action profile separately such that I's incentives are binding, just as in the case of pure strategy profiles. After it is established that the moderator can mandate I's continuation payoff (for e.g., to be $y + \underline{b}$ in case of play of (C, C), and so on and so forth), 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. The magnitude of the impact is linear in the probability weights shifted, and choosing one action profile over the other represents moving the full probability weight from the latter action profile to the former.

Renegotiation-proofness

The full characterization of the set of Pareto dominant SPE equilibrium 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.[13] For every given vector of parameters x, y, and z, we characterize conditions on players' discount factors (δ_I, δ_P) under which the Pareto dominant play paths derived in the previous section are renegotiation proof. We also characterize conditions on x, y, and z under which

[13] As we evaluate renegotiation-proofness against all possible SPE strategy profiles, we are examining whether our Pareto dominant SPE are *strongly* renegotiation-proof in the terminology of Farrell and Maskin (1989).

whenever either $\overrightarrow{CC}$ or $\overrightarrow{DC}$ is Pareto dominant, the play path is also renegotiation proof. Our analysis shows that $\overrightarrow{CC}$ and $\overrightarrow{DC}$ survive renegotiation-proofness for a wide range of parameter values.

We begin with the characterization of conditions under which $\overrightarrow{DC}$ survives renegotiation-proofness. Recall that $\overrightarrow{DC}$ is Pareto dominant when $(\delta_I, \delta_P) \in \Omega_{DC}$, and Ω_{DC} is nonempty if and only if $2x < z$.

Proposition 5 *Let $2x < z$ and $(\delta_I, \delta_P) \in \Omega_{DC}$. (i) If $y \leq z - x + \frac{x^2}{z}$ or $z \leq 3x$, then for all $(\delta_I, \delta_P) \in \Omega_{DC}$, each Pareto dominant SPE payoff implemented by $\overrightarrow{DC}$ survives renegotiation-proofness. (ii) If $y > z - x + \frac{x^2}{z}$ and $z > 3x$, then for $(\delta_I, \delta_P) \in \Omega_{DC}$ satisfying*

$$\frac{\delta_P}{\delta_I} \geq \frac{y-x}{z-y} \text{ or } \frac{\delta_P}{\delta_I} \leq \frac{z-y}{x+y-z}, \tag{7}$$

each Pareto dominant SPE payoff implemented by $\overrightarrow{DC}$ survives renegotiation-proofness. It follows from (i) and (ii) that for all (δ_I, δ_P) at which both $\overrightarrow{CC}$ and $\overrightarrow{DC}$ are SPE play paths and $\overrightarrow{DC}$ Pareto dominates $\overrightarrow{CC}$, the payoffs implemented by $\overrightarrow{DC}$ survive strong renegotiation-proofness.

Proof. See appendix. ■

For the case of discount factors being considered a necessary condition for an *SPE* strategy profile to be renegotiation proof (see for example, 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 $\overrightarrow{DC}$. If not, then the players' can renegotiate away to a Pareto improving *SPE*. The renegotiation-proofness of $\overrightarrow{DC}$ hinges on the existence of renegotiation proof punishments at the bribing stage that pushes the designated briber's payoff to x when he/she refuses to bribe. Once such punishments exist, it is straightforward to construct renegotiation-proof continuation paths that punish deviations at the action stage by any player. For example, if P were to deviate in period t by choosing action D, she will be "punished" by reverting to an *SPE* implemented by $\overrightarrow{DC}$, wherein the period $t+1$ bribes ensure that $U_{P,t+1} = x$. As the resulting deviation payoff of x is strictly less than the *SPE* continuation payoff $\delta_P (z - x)$, the *SPE* is deviation proof at the action stage.

The reader must be wondering what punishments deter deviations at the bribing stage. For example, in the case of the punishment path described above, if P were required to bribe I in period $t+1$, resulting in $U_{P,t+1} = x$, why would she do so? Once more, for any punishment that deters this second deviation to be part of a renegotiation proof *SPE*, play from period $t+2$ must revert once more to $\overrightarrow{DC}$. Punishments to deter deviations at the bribing stage must therefore be designed to deviate from $\overrightarrow{DC}$ only in that period alone. The difficulty with such constructions is most clearly seen by looking at the case where renegotiation-proof deterrence requires imposing an action profile in the period of deviation, that is "unpleasant" for both the deviator and the punisher. This happens when this punishment entails playing (D, D) for one period. The deviator can propose instead to switch to playing (C, C) which is strictly preferred by both players to (D, D). In order for the punisher not to be tempted by this one-shot gain from renegotiation, y must not be too large. This explains the upper bound on y in one of the conditions characterizing part (i) of the proposition: $y \leq z - x + \frac{x^2}{z}$. In fact (7) in part (ii) can also be rewritten into a upper bound on y:

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

where $\rho \equiv \delta_I / \delta_P$. The other sufficient condition for $\overrightarrow{DC}$ to be renegotiation proof for all $(\delta_I, \delta_P) \in \Omega_{DC}$ is $z \leq 3x$. From Proposition 2, we know that for $\overrightarrow{DC}$ to be Pareto dominant, it requires that $z > 2x$. In fact, the larger is z, the larger is the set Ω_{DC}. Basically, when $z \leq 3x$, the set Ω_{DC} is not too large and it is thus possible for $\overrightarrow{DC}$ to be renegotiation proof for all $(\delta_I, \delta_P) \in \Omega_{DC}$. The actual construction of such punishments being delicate - and parameter dependent - is relegated to the appendix. The last part of Proposition 5 highlights that partial cooperation achieved through $\overrightarrow{DC}$ is rather robust to renegotiation. Focusing on situations in which both partial and full cooperation are equilibrium outcomes, whenever partial cooperation Pareto dominates full cooperation, it is also renegotiation proof.

Next, we characterize conditions under which $\overrightarrow{CC}$ can be supported as a renegotiation proof SPE play path. The reader may recall from section 3, that depending upon whether $(\delta_I, \delta_P) \in \Omega^1_{CC}$ or $(\delta_I, \delta_P) \in \Omega^2_{CC}$, player I or player P is the perpetual briber from period 2 onward. Our characterization will cover both cases.

Proposition 6 *Let $(\delta_I, \delta_P) \in \Omega_{CC}$. (i) If $y \leq z - x + \frac{x^2}{z}$ or $y \leq \min\left\{\frac{x+z}{2}, z - \frac{x}{2}\right\}$, then for all $(\delta_I, \delta_P) \in \Omega_{CC}$, each Pareto dominant SPE payoff implemented by $\overrightarrow{CC}$ survives renegotiation-proofness. (ii) Suppose $y > z - x + \frac{x^2}{z}$ and $y > \min\left\{\frac{x+z}{2}, z - \frac{x}{2}\right\}$.*

(ii.a) If achieving a Pareto dominant SPE payoff implemented by $\overrightarrow{CC}$ requires one player to bribe in the first period and the other to bribe in all remaining periods, then the SPE payoff is renegotiation proof if and only if

$$\frac{\delta_P}{\delta_I} \geq \frac{y-x}{z-y} \text{ or } \frac{\delta_P}{\delta_I} \leq \frac{z-y}{x+y-z} \tag{8}$$

and

$$\frac{\delta_P}{\delta_I} \leq \frac{z-y}{y-x} \text{ or } \frac{\delta_P}{\delta_I} \geq \frac{x+y-z}{z-y} \tag{9}$$

where if $2x \leq z$, then (8) implies (9); if $2x \geq z$, then (9) implies (8).

(ii.b) If achieving a Pareto dominant SPE payoff implemented by $\overrightarrow{CC}$ requires player I to bribe in all periods, which is possible only if $(\delta_I, \delta_P) \in \Omega^2_{CC}$, then the SPE payoff is renegotiation proof if and only if (8) holds.

(ii.c) If achieving a Pareto dominant SPE payoff implemented by $\overrightarrow{CC}$ requires player P to bribe in all periods, which is possible only if $(\delta_I, \delta_P) \in \Omega^1_{CC}$, then the SPE payoff is renegotiation proof if and only if (9) holds.

Proof. See appendix. ■

There are many similarities in the conditions for renegotiation-proofness of $\overrightarrow{DC}$ and $\overrightarrow{CC}$ and the intuitions behind these conditions. As in the case of renegotiation-proofness of $\overrightarrow{DC}$, dealing with deviations at the action stage is relatively straightforward; i.e., restart $\overrightarrow{CC}$ in the following period but choose the payoff vector in a way that gives the deviator her minimum payoff x. The key is whether renegotiation-proof punishments exist that deter deviations by players at the bribing stage. Also similar to the case of $\overrightarrow{DC}$, the difficulty lies with the possibility of the deviator proposing to play the action profile (C, C) whenever the deterring punishment calls for an action profile which leads to a low instantaneous payoff for one or both players in that period. It is essential for renegotiation-proofness, therefore, that the individual payoff from full cooperation y be bounded by some value. Part (i) of proposition 6 points out that when $y \leq z - x + \frac{x^2}{z}$ or $y \leq \min\left\{\frac{x+z}{2}, z - \frac{x}{2}\right\}$, renegotiation-proofness of $\overrightarrow{CC}$ holds for all admissible discount factor pairs (δ_I, δ_P). Quite strikingly, $y \leq z - x + \frac{x^2}{z}$ is also a sufficient condition for $\overrightarrow{DC}$ to be renegotiation proof for all discount factors for which $\overrightarrow{DC}$ is Pareto dominant.

Whenever y exceeds $\max\left\{z - x + \frac{x^2}{z}, \min\left\{\frac{x+z}{2}, z - \frac{x}{2}\right\}\right\}$, there are discount factor pairs (δ_I, δ_P) for which renegotiation-proofness fails. The precise conditions on δ_P/δ_I under which renegotiation proof punishments ensure player I's and player P's incentives to bribe are respectively (8) and (9), which can be rewritten into upward bounds on y, namely, $y \leq \max\left\{\frac{x+\rho(z)}{1+\rho}, \frac{z+\rho(z-x)}{1+\rho}\right\}$ and $y \leq \max\left\{\frac{z+\rho(x)}{1+\rho}, \frac{(z-x)+\rho(z)}{1+\rho}\right\}$, where $\rho \equiv \delta_P/\delta_I$. The idea of part (ii) of the proposition is as follows. When it requires both players to bribe, one in the first period and the other in all the subsequent periods, to achieve an allocation of payoffs, both (8) and (9) must be satisfied for renegotiation-proofness. If achieving an allocation of payoffs requires only one of the players to bribe in all periods, then only one of (8) and (9) has to be satisfied.

5 Conclusion

In this paper, we have studied a simple yet very natural mechanism, repeated side-payments, that players may use to resolve conflicts of interest in a long term relationship. When players have different time preferences, side-payments can be used to both, provide incentives to cooperate, and conduct intertemporal trade of instantaneous payoffs. Depending on how different players' time preferences are and whether the less patient player is patient enough, the side payments that help players attain Pareto dominant SPE outcomes, may go in either direction. The steady-state bribes are non-monotone and change nontrivially in the players' discount factors. The use of side payments offers a powerful mechanism that renders publicly observable randomization devices redundant, and in many situations is also 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 bribe amounts affect the Pareto frontier of the set of SPE outcomes. Yet another of our assumptions related to the exogeneity of the timing of the bribe in every period. One way of endogenizing this variable is to assume that players are free to bribe at any point of time between two consecutive stage games and they choose the timing that generates the highest aggregate payoff. Depending on the direction of bribes they agree upon, agents may bribe right before or right after the stage game.[14] Another obvious extension of this study is to model how agents bargain over gain from cooperation. Such analysis will allow us to pin down the first period bribe which is indeterminate in this paper.

[14] The authors' separate analysis on this issue is downloadable through the following URL: http://www.kellogg.northwestern.edu/faculty/fong/htm/timing.pdf

Appendix

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

$$\begin{aligned} & (1-\delta_I)(0) + \delta_I U_{It+1} \geq (1-\delta_I)x + \delta_I x \\ \Leftrightarrow \quad & U_{It+1} \geq \frac{x}{\delta_I}. \end{aligned}$$

Consider an alternative strategy profile in which along the equilibrium path $a'_t = (D, C)$ but $a'_\tau = a_\tau$ for all $\tau \neq t$. Adjust b_{t+1} downward such that $U'_{It+1} = x$, i.e., $\Delta U_{It+1} = U'_{It+1} - U_{It+1} \leq -\frac{x(1-\delta_I)}{\delta_I} < 0$. Since the aggregate payoff in period $t+1$, $\frac{U_{It+1}}{1-\delta_I} + \frac{U_{Pt+1}}{1-\delta_P}$, stays unchanged,

$$\Delta U_{Pt+1} = -\frac{1-\delta_P}{1-\delta_I}\Delta U_{It+1} \geq \frac{1-\delta_P}{\delta_I}x.$$

As $U_{Pt+1} \geq x$, we have

$$\begin{aligned} & U'_{Pt+1} - x \geq \Delta U_{Pt+1} \geq \tfrac{1-\delta_P}{\delta_I}x > \tfrac{1-\delta_P}{\delta_P}x \\ \Leftrightarrow \quad & U'_{Pt+1} \geq \frac{x}{\delta_P}. \end{aligned}$$

Therefore, it is incentive compatible for P to play C at t. Now adjust b_t such that

$$\begin{aligned} & \Delta U_{It} = (1-\delta_I)\Delta b_t + \delta_I \Delta U_{It+1} = 0 \\ \Leftrightarrow \quad & \Delta b_t = -\frac{\delta_I \Delta U_{It+1}}{1-\delta_I}. \end{aligned}$$

Recalling that , $\Delta U_{It+1} < 0$, we have

$$\begin{aligned} \Delta U_{Pt} &= -(1-\delta_P)\Delta b_t + \delta_P \Delta U_{Pt+1} \\ &= (1-\delta_P)\frac{\delta_I \Delta U_{It+1}}{1-\delta_I} - \delta_P \frac{1-\delta_P}{1-\delta_I}\Delta U_{It+1} \\ &= \frac{1-\delta_P}{1-\delta_I}(\delta_I - \delta_P)\Delta U_{It+1} > 0. \end{aligned}$$

In other words, 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. IC of player I and IR of player P imply that $U_{It} \geq x$ and $U_{Pt} \geq x/\delta_P$. Modify this strategy profile by replacing the action profile by (D, D) and lowering the net bribe from P to I by x at time t, but leave the rest of the equilibrium play path identical to the original profile. The new strategy profile (i) yields an increase in U_{Pt} by $(2x - z)$ and keeps U_{It} unchanged; and (ii) relaxes player P's incentive constraint at t by not require her to cooperate. ■

Proof of Lemma 2. In (i) $U_{It+1} < x$ violates individual rationality and in (ii) if $U_{It+1} < y + \underline{b}$, then I is unwilling to play C at time t. If $U_{It+1} > x$ in (i), then there exists $\Delta b_{t+1} < 0$ such that by revising the net transfer from P to I downward by Δb_{t+1}, $U_{It+1} \geq x$ still holds. Adjust b_t upward by $\Delta b_t > 0$ such that U_{It} remain unchanged:

$$\begin{aligned} & \Delta U_{It} = (1 - \delta_I)\, \Delta b_t + \delta_I \Delta b_{t+1} = 0 \\ \Leftrightarrow \qquad & \Delta b_t = -\frac{\delta_I \Delta b_{t+1}}{1 - \delta_I}. \end{aligned}$$

The net change to P's continuation payoff at time t will be:

$$\begin{aligned} \Delta U_{Pt} &= -(1 - \delta_P)\, \Delta b_t + \delta_P \Delta b_{t+1} \\ &= \frac{\delta_I \Delta b_{t+1}}{1 - \delta_I} (1 - \delta_P) + \delta_P \Delta b_{t+1} \\ &= (1 - \delta_P)\, \Delta b_{t+1} \left[\frac{\delta_I}{1 - \delta_I} - \frac{\delta_P}{1 - \delta_P} \right] > 0, \end{aligned}$$

recalling that $\Delta b_{t+1} < 0$. So Pareto improvement can be created. Therefore, it must be that $U_{It+1} = x$.

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., $U_{It+1} \geq y + \underline{b}$. If $U_{It+1} > y + \underline{b}$ in (ii), then there also exists $\Delta b_{t+1} < 0$ such that $U_{It+1} \geq y + \underline{b}$ still holds, and thus Pareto improvement can also be similarly created. Therefore, it must be that $U_{It+1} = y + \underline{b}$. ■

Proof of Lemma 3. Suppose player $i \in \{I, P\}$ is called upon to bribe in the first period. Since in the punishment phase player i's continuation payoff will be x, he/she is willing to bribe any amount so long as his/her equilibrium payoff, net the bribe, is no less than x. Since adjustment of the first period bribe allows players to move from any point to any other on the Pareto frontier of the set of SPE outcomes, every two distinctive Pareto dominant SPE play paths should be identical except the first period bribes are different. Now compare two arbitrary continuation subgames on a Pareto dominant SPE play path starting at the bribing stages of periods t and t', where $1 \leq t < t'$. Although the allocations of continuation payoffs in t and t' may be different, both allocations ought to be on the Pareto frontier of the SPE outcomes set of the continuation subgames. Therefore, the play paths following these continuation subgames are identical except that b_t and $b_{t'}$ may differ. To satisfy this requirement, the action profiles in periods t and t' are identical, i.e., $a_t = a_{t'}$. Since the choice of t and t' is arbitrary, the proof is complete. ■

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

$$(1-\delta_I)\, z + \delta_I\,(x) \quad \leq \quad (1-\delta_I)\, y + \delta_I\,(y+b)\,, \tag{10}$$

$$(1-\delta_P)\, z + \delta_P\,(x) \quad \leq \quad (1-\delta_P)\, y + \delta_P\,(y-b)\,. \tag{11}$$

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

$$\begin{aligned} \underline{b} &= x - y + \frac{1-\delta_I}{\delta_I}\,(z-y)\,, \\ \overline{b} &= y - x - \frac{1-\delta_P}{\delta_P}\,(z-y)\,. \end{aligned}$$

For the existence of some b satisfying both (10) and (11), it requires $\underline{b} \leq \overline{b}$. One can verify with algebra that this requirement is equivalent to (3).

For sufficiency, it is easy to verify that when (3) is satisfied, then the action profit (C, C) can be supported in equilibrium by a trigger strategy and the expectation that a net bribe $b \in [\underline{b}, \overline{b}]$ 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

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

for $b \in [\underline{b}, \overline{b}]$, the IRs at the bribing stage are automatically satisfied under (3) 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 bribe amount. ■

Proof of Proposition 2. The necessity of $z > 2x$ is due to part (ii) of Lemma 1. For it to be incentive compatible for P to play C, it requires that

$$\begin{aligned} (1 - \delta_P) x + \delta_P x &\leq (1 - \delta_P)(0) + \delta_P b_{It} \\ \Leftrightarrow \qquad b_{It} &\geq \frac{x}{\delta_P}. \end{aligned}$$

For it to be incentive compatible for I to give the bribe, it requires that $z - b_{It} \geq x$, i.e., $b_{It} \leq z - x$. Therefore, the existence of some b satisfying both P's IC to play C and I's IC to give the bribe, it requires $z - x \geq x/\delta_P$. 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. 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 bribe amount. ■

Proof of Theorem 1. First, we can summarize the discussion preceeding (6) 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 $(z - y)\delta_P > y\delta_I$, and $\overrightarrow{CC}$ Pareto dominants $\overrightarrow{DC}$ if $(z - y)\delta_P < y\delta_I$.*

It follows from Lemma 1 and 4 that the only possible Pareto dominant play paths are $\overrightarrow{CC}$, $\overrightarrow{DC}$, and $\overrightarrow{DD}$. It is straightforward that whenever $\overrightarrow{CC}$ or $\overrightarrow{DC}$ can be supported, $\overrightarrow{DD}$ is Pareto dominated. Combining this observation with Propositions 1 - 2 and Lemma A1, the followings are immediate: $\psi = \overrightarrow{CC}$ if $(\delta_I, \delta_P) \in \Omega_{CC}$ because $\overrightarrow{CC}$ Pareto dominates $\overrightarrow{DC}$; $\psi = \overrightarrow{DC}$ if $(\delta_I, \delta_P) \in \Omega_{DC}$ because either $\overrightarrow{DC}$ Pareto dominates $\overrightarrow{CC}$ or $\overrightarrow{CC}$ is not feasible; and $\psi = \overrightarrow{DD}$ if $(\delta_I, \delta_P) \in \Omega_{DD}$ because neither $\overrightarrow{CC}$ nor $\overrightarrow{DC}$ is feasible. The bounds on $b_{I1} - b_{P1}$ when $\psi = \overrightarrow{CC}$ are derived from the following IR constraints in period 1:

$$\begin{aligned} U_{I1} &= (1-\delta_I)\left(y + (b_{P1} - b_{I1})\right) + \delta_I \left(y + \underline{b}\right) \geq x, \\ U_{P1} &= (1-\delta_P)\left(y - (b_{P1} - b_{I1})\right) + \delta_P \left(y - \underline{b}\right) \geq x. \end{aligned}$$

The first inequality pins down $(b_{P1} - b_{I1}) \geq -(z-x)$ and the second inequality pins down $(b_{P1} - b_{I1}) \leq \frac{\delta_I(1+\delta_P)(y-x) - \delta_P(1-\delta_I)(z-y)}{\delta_I(1-\delta_P)}$. Similarly, the bounds on $b_{I1} - b_{P1}$ when $\psi = \overrightarrow{DC}$ are derived from the following IR constraints in period 1:

$$\begin{aligned} U_{I1} &= (1-\delta_I)\left(z - (b_{I1} - b_{P1})\right) + \delta_I x \geq x, \\ U_{P1} &= (1-\delta_P)\left(b_{I1} - b_{P1}\right) + \delta_P \left(z - x\right) \geq x. \end{aligned}$$

This completes the proof of the theorem. ■

Proof of Proposition 4. Adapting the proof of Lemma 3, we observe that in the game with transferable utilities, Pareto dominance of the whole game requires Pareto dominance in very period on the equilibrium path. It also implies that the probability mix over the four action profiles in each period is stationary, although different histories (particularly, different realizations of the randomization in the action stage) may lead to different bribe amounts in the following period.

Now consider for some point in the (δ_I, δ_P)–plane, a correlated strategy profile σ that implements a payoff vector that is Pareto dominant in the *SPE* payoff set. The equilibrium path of this strategy is represented canonically by a periodic sequence of possible bribes $b_1, \{b_{CC,t}, b_{DC,t}, b_{CD,t}, b_{DD,t}\}$, $t > 1$, and corresponding stationary action set probability matrices

$$
\begin{array}{c|c|c|}
 & C & D \\
\hline
C & \lambda_{CC} & \lambda_{CD} \\
\hline
D & \lambda_{DC} & \lambda_{DD} \\
\hline
\end{array}
\tag{12}
$$

for all $t \geq 1$.

Adapting Lemma 2, it follows that in (12), maintaining I's incentive constraints requires that she receives expected payoff of:

$$
\begin{aligned}
\frac{\lambda_{CC}}{\lambda_{CC}+\lambda_{CD}}\left(y+\underline{b}\right)+\frac{\lambda_{CD}}{\lambda_{CC}+\lambda_{CD}}\frac{x}{\delta_I}; &\quad \text{if called upon to play action } C, \\
x; &\quad \text{if called upon to play action } D,
\end{aligned}
\tag{13}
$$

and this applies to *all* periods $t > 0$. Let the incentive constraints in the action stage of period t be enforced by some net bribes from P to I from period 2 onward, leading to the following expected bribes:

$$
\begin{aligned}
\frac{\lambda_{CC}}{\lambda_{CC}+\lambda_{CD}}b_{CC,t+1}+\frac{\lambda_{CD}}{\lambda_{CC}+\lambda_{CD}}b_{CD,t+1}; &\quad \text{if called upon to play action } C, \\
\frac{\lambda_{CC}}{\lambda_{CC}+\lambda_{CD}}b_{DC,t+1}+\frac{\lambda_{CD}}{\lambda_{CC}+\lambda_{CD}}b_{DD,t+1}; &\quad \text{if called upon to play action } D.
\end{aligned}
$$

If, however, I were to somehow "know" which action profile has been picked by the correlation device when she is called upon to play C (alt: D), then her incentive constraints would require that she receive in the following period continuation payoffs:

$$
\begin{aligned}
y+\underline{b}; &\quad \text{if } (C,C), \\
\frac{x}{\delta_I}; &\quad \text{if } (C,D), \\
x; &\quad \text{if } (D,C) \text{ or } (D,D),
\end{aligned}
\tag{14}
$$

for all $t \geq 1$. Suppose the incentive constraints are enforced by some net bribes from P to I denoted by $(b^P_{CC,t+1}, b^P_{CD,t+1}, b^P_{DC,t+1}, b^P_{DD,t+1})$. Combining (13) and (14), we obtain that players have the same expected payoffs and the expected bribes are the same in that period. It follows immediately that there exists a payoff equivalent correlated SPE strategy profile that replaces the side payments $(b_{CC,t+1}, b_{CD,t+1}, b_{DC,t+1}, b_{DD,t+1})$ with $(b^P_{CC,t+1}, b^P_{CD,t+1}, b^P_{DC,t+1}, b^P_{DD,t+1})$, leaving all else identical to σ. Starting from $t = 2$, we replace sequentially the bribes in every period in the same manner and the end product is a payoff equivalent correlated SPE with the side payments $(b^P_{CC,t}, b^P_{CD,t}, b^P_{DC,t}, b^P_{DD,t})$, for all $t \geq 2$. Stationarity of the probability mix in (12) and stationarity of the possible continuation payoffs of player I in (14) imply the stationarity of the set of side payments. So the time subscripts for these bribes can be dropped. Let the transformed strategy profile be denoted by $\sigma^{(P)}$.

Henceforward, we focus attention on the properties of this payoff equivalent SPE strategy profile, $\sigma^{(P)}$. In $\sigma^{(P)}$, continuation payoffs of the players are tailored to satisfy the incentive constraints generated by the action profile that is actually selected by the randomization device, (and in this sense, it is "as if" the moderator of the game tell each player the realization of the randomization in each period). Hence, moving probability weights from an action profile having positive weight to another one also having positive weight in the first period generates a new SPE. Subsequently, we can move probability weights the same way in all the following periods and the strategy profile remains an SPE strategy profile. In other words, when we move around the probability weights among action profiles with positive weights in the stationary probability mix in (12), the resulting strategy profiles still constitute an SPE. By direct computation, the Pareto frontier of the set of SPE payoffs is then characterized by the following equation:

$$\frac{U_{I1}}{1-\delta_I} + \frac{U_{P1}}{1-\delta_P} = \frac{\lambda_{CC}\,(y+\underline{b}) + \lambda_{CD}\,(x/\delta_I) + (\lambda_{DC}+\lambda_{DD})\,x}{1-\delta_I}$$
$$+\frac{\lambda_{CC}\,(2y) + (\lambda_{CD}+\lambda_{DC})\,z + \lambda_{DD}\,(2x) - [\lambda_{CC}\,(y+\underline{b}) + \lambda_{CD}\,(x/\delta_I) + (\lambda_{DC}+\lambda_{DD})\,x]}{1-\delta_P},$$

the right hand side of which is linear in the probabilities, λ_{CC}, λ_{CD}, λ_{DC}, and λ_{DD}. This linearity property and Pareto dominance of $\sigma^{(P)}$ in the set of SPE strategy profiles imply that shifting weights among the positive probabilities should neither shift the frontier inward nor outward. Therefore, there is no loss in generality in shifting all the probability weights to one of the action profiles. This finishes the proof of Proposition 4. ■

Proof of Proposition 5. We start by noting that given a unilateral deviation in period t from any fixed, arbitrary play path implemented by $\overrightarrow{DC}$, renegotiation-proofness of the punishments corresponding to this deviation must entail reverting to $\overrightarrow{DC}$ by period $t+1$ (with an appropriately scaled period $(t+1)$ bribe vector in order to ensure deterrence against the deviation). For if not, then by the Pareto dominance of play paths in $\overrightarrow{DC}$ (Theorem 1), the proposed punishment is not renegotiation-proof, a contradiction. With this fact in our pocket, we now proceed with the proof:

Step 1 (Deviations at the action stage in period $t \geq 1$). First assume that if player P chooses an action other than C, players can simply restart an equilibrium play path implemented by $\overrightarrow{DC}$ from period $(t+1)$ corresponding to $U_{Pt+1} = x$. So proving the result is reduced to constructing renegotiation-proof punishments that deter deviations from the equilibrium path of play at the bribing stage, thereby verifying that the assumption we make here is implied by renegotiation-proofness at the bribing stage.

Step 2 (Deviations by player I at the bribing stage in period $t \geq 1$ when $b_{It} > 0 = b_{Pt}$).

Look at the following four possible punishments:

$$a_t = (D, D) \text{ and } U_{It+1} = x \tag{$I1$}$$

$$a_t = (C, D) \text{ and } U_{It+1} = \frac{x}{\delta_I} \tag{$I2$}$$

$$a_t = (C, C)\,; U_{It+1} \geq y + \underline{b} \text{ and } U_{Pt+1} \geq y - \overline{b} \tag{$I3$}$$

and

$$a_t = (D, C) \text{ and } U_{It+1} \geq x. \tag{$I4$}$$

By definition, only ($I1$) and ($I2$) can push I's post-deviation payoff to x. Since I's continuation payoff starting from period two is x on path $\overrightarrow{DC}$, only ($I1$) or ($I2$) may constitute a renegotiation-proof punishment deterring I from deviating at the bribing stage. More specifically, if for certain values of (δ_I, δ_P), either $(I1)$ or $(I2)$ (or both) are Pareto undominated by $(I3)$ and $(I4)$, then there exists a deviation-deterring, renegotiation-proof punishment; if both $(I1)$ and $(I2)$ are Pareto dominated by either $(I3)$ or $(I4)$, then any renegotiation punishment (be it $(I3)$ or $(I4)$) necessarily leads to a payoff for I higher than his equilibrium payoff, x, and he therefore will prefer not to bribe.

Denote by $\mathcal{P}$ a "Pareto correspondence": for any subset, Λ, of *SPE* strategies of the supergame, $\mathcal{P}(\Lambda)$ denotes those strategies in Λ that are Pareto undominated by any other strategy in Λ. Our objective in this step is to identify conditions under which either (i) $(I1) \in \mathcal{P}((I1),(I3),(I4))$ conditioned on $(I1) \in \mathcal{P}((I1),(I2))$; or (ii) $(I2) \in \mathcal{P}((I2),(I3),(I4))$ conditioned on $(I1) \in \mathcal{P}((I1),(I2))$.

Claim 1: $(I1) = \mathcal{P}((I1),(I2))$ if $\frac{\delta_P}{\delta_I} > \frac{z-x}{x}$, and $(I2) = \mathcal{P}((I1),(I2))$ when $\frac{\delta_P}{\delta_I} < \frac{z-x}{x}$.

Proof: Conditioned on $U_{It}((I2)) = U_{It}((I1)) = x$, we have

$$\begin{aligned}\frac{U_{Pt}((I2)) - U_{Pt}((I1))}{1-\delta_P} &= (z-x) - \delta_P \frac{\frac{x}{\delta_I} - x}{1-\delta_I} \\ &= (z-x) - \frac{\delta_P}{\delta_I} x.\end{aligned}$$

Claim 2: Suppose $(I1) = P((I1),(I2))$.[15] Then $(I1) \in P((I1),(I3),(I4))$ if and only if $\frac{\delta_P}{\delta_I} \geq \frac{y-x}{z-y}$.

Proof: Note that $u_{Pt}((I1)) - u_{Pt}((I4)) = x$ and conditioned on $U_{It}((I4)) \geq U_{It}((I1)) = x$, $U_{Pt}((I4)) - U_{Pt}((I1)) \leq x - z < 0$, which proves that $I1 \in \mathcal{P}((I1),(I4))$. Next, conditioned on offering player I the lowest payoff in $(I3)$, i.e., $U_{It}(I3) = (1-\delta_I) y + \delta_I (y + \underline{b})$,

$$\begin{aligned}\frac{U_{Pt}((I3)) - U_{Pt}((I1))}{1-\delta_P} &= (y-x) - \delta_P \frac{y + \underline{b} - x}{1-\delta_I} \\ &= (y-x) - \frac{\delta_P}{\delta_I}(z-y).\end{aligned}$$

[15] This claim holds irrespective of this assumption. It is added for the expositional clarify of the proof of the proposition.

Therefore, $I1 \in \mathcal{P}((I1),(I3))$ if and only if $\frac{\delta_P}{\delta_I} \geq \frac{y-x}{z-y}$. This proves the claim.

Claim 3: Suppose $(I2) = P((I1),(I2))$. Then $(I2) \in P((I2),(I3),(I4))$ if and only if $(x+y-z)\,\delta_P/\delta_I \leq (z-y)$.

Proof:

$$\begin{aligned}
\frac{U_{Pt}((I4)) - U_{Pt}((I2))}{1-\delta_P} &\leq -z - \delta_P \frac{x - x/\delta_I}{1-\delta_I} \\
&= \frac{\delta_P}{\delta_I} x - z
\end{aligned}$$

Therefore, $(I2) \in \mathcal{P}((I2),(I4))$ if and only if $\delta_P/\delta_I \leq z/x$, which is satisfied whenever $(I2) = P((I1),(I2))$, i.e., $\delta_P/\delta_I \leq (z-x)/x$. Next,

$$\begin{aligned}
\frac{U_{Pt}((I3)) - U_{Pt}((I2))}{1-\delta_P} &\leq (y-z) - \delta_P \frac{y + \underline{b} - x/\delta_I}{1-\delta_I} \\
&= \frac{\delta_P}{\delta_I}(x+y-z) - (z-y)
\end{aligned}$$

Therefore, $(I2) \in \mathcal{P}((I2),(I3))$ if and only if $(x+y-z)\,\delta_P/\delta_I \leq (z-y)$. This completes the proof.

Note that $\frac{z-x}{x} \geq \frac{y-x}{z-y}$ if and only if $z - y \geq \frac{z-x}{x}(x+y-z)$; both conditions are identical to

$$y \leq (z-x) + \frac{x^2}{z}.$$

Therefore, if $y \leq (z-x) + \frac{x^2}{z}$, then $\frac{\delta_P}{\delta_I} > \frac{z-x}{x}$ implies $\frac{\delta_P}{\delta_I} > \frac{y-x}{z-y}$ and $\frac{\delta_P}{\delta_I} < \frac{z-x}{x}$ implies $\frac{\delta_P}{\delta_I}(x+y-z) < z-y$. In other words, for all $(\delta_I, \delta_P) \in \Omega_{DC}$, if $(I1) = \mathcal{P}((I1),(I2))$, then $(I1) \in \mathcal{P}((I1),(I3),(I4))$ and if $(I2) = \mathcal{P}((I1),(I2))$, then $(I2) \in \mathcal{P}((I2),(I3),(I4))$.

For the remainder of the proof, focus on the case that $y > (z-x) + \frac{x^2}{z}$. Take the point where $1/\delta_I + 1/\delta_P = 2/\delta^*$ and $\delta_P = x/(z-x)$ intersect in Figure 1. We can verify that at this point $\delta_P/\delta_I = (2x+y-z)/(z-y)$. Therefore, if

$$\begin{aligned}
& \frac{y-x}{z-y} \leq \frac{2x+y-z}{z-y} \\
\Leftrightarrow \quad & z \leq 3x,
\end{aligned}$$

then for all $(\delta_I, \delta_P) \in \Omega_{DC}$, $\frac{\delta_P}{\delta_I} > \frac{y-x}{z-y}$. Besides, $y > (z-x) + \frac{x^2}{z}$ also implies $\frac{y-x}{z-y} > \frac{z-x}{x}$. Therefore, according to claims 1 and 2, $(I1) \in \mathcal{P}((I1), (I2), (I3), (I4))$. Up to here we have proved that under the assumption of part (i) of the proposition, there exists a renegotiation-proof punishment that deters I's incentive to deviate at the bribing stage.

Now suppose instead the assumption of part (ii) holds. Recall that $y > (z-x) + \frac{x^2}{z}$ implies

$$\frac{z-y}{x+y-z} < \frac{z-x}{x} < \frac{y-x}{z-y}.$$

It follows immediately that there exists renegotiation proof punishment to ensure player I's incentive to bribe the amount $(z-x)$ if and only if

$$\frac{\delta_P}{\delta_I} \geq \frac{y-x}{z-y} \text{ or } \frac{\delta_P}{\delta_I} \leq \frac{z-y}{x+y-z}.$$

Since $\frac{y-x}{z-y} > \frac{y}{z-y}$, it immediately follows that all $(\delta_I, \delta_P) \in \Omega_{DC}$ in the region where $\overrightarrow{DC}$ Pareto dominates $\overrightarrow{CC}$, i.e., satisfy $\frac{\delta_P}{\delta_I} \geq \frac{y}{z-y}$ and $1/\delta_I + 1/\delta_P \leq 2/\delta^*$, also satisfy $\frac{\delta_P}{\delta_I} \geq \frac{y-x}{z-y}$. So, for these discount factors, there exists renegotiation proof punishment to ensure player I's incentive to bribe the amount $(z-x)$.

Step 3 (Deviations by player P at the bribing stage in period $t=1$ when $b_{P1} > 0 = b_{I1}$). Suppose player P is required to bribe a positive amount such that $U_{Pt} = x$. To ensure player P's incentive to bribe this amount, we propose the punishment that if player P refuses to bribe the specified amount in $t=1$, players will play (D, C) in the stage game of $t=1$ and $\overrightarrow{DC}$ starting from $t=2$ in which period player P will be required to bribe an amount such that $\tilde{U}_{P2} = x/\delta_P$. Also, in every subsequent continuation subgame, as long as player P has not paid the bribe, players continue to play (D, C) and in the period after P is required to bribe the amount such that $\tilde{U}_{Pt} = x/\delta_P$. In this punishment path, it is incentive compatible for player P to play C. Also since player P's instantaneous payoff is 0 and that of player I's is z in the stage game and the only way of renegotiation is to increase I's instantaneous payoff in t and increase P's continuation payoff in period

$t + 1$, renegotiation is infeasible post the bribing stage. With the same logic, the same punishment proposed will provide an enough incentive for player P to pay a lower amount of bribe.

Step 4 (Verify that assumption in step 1 is satisfied). In fact, using the punishment strategy in step 3, we can make player P bribe at the bribing stage of any period t an amount large enough that her continuation payoff is $U_{Pt} = x$. This verifies the assumption in step 1 that if P plays D in the action stage of any period t, the two players can enforce a renegotiation proof punishment that leads to $U_{Pt+1} = x$.

This completes the proof of Proposition 5. ■

Proof of Proposition 6. As in the case of Proposition 5, we note that after a unilateral deviation in some period t from any fixed, arbitrary play path implemented by $\overrightarrow{CC}$, renegotiation-proofness of the play path entails reverting to $\overrightarrow{CC}$ by period $t + 1$ (with an appropriately scaled period $(t + 1)$ bribe vector in order to ensure deterrence against the deviation).

Step 1 (Deviations at the action stage in period $t \geq 1$). First assume that if player $i \in \{I, P\}$ chooses an action other than C, players can simply restart $\overrightarrow{CC}$ in period $(t + 1)$ but with $U_{it+1} = x$. Proving renegotiation-proofness is, hence, reducible to proving that we can construct renegotiation-proof punishments that deter deviations from $\overrightarrow{CC}$ at the bribing stage, thereby verifying the assumption we make here.

Step 2 (Deviations by player I at the bribing stage in period $t \geq 1$ when $b_{It} > 0 = b_{Pt}$).

We focus once more on the four punishments $(I1) - (I4)$ defined in the proof of Proposition 5, except that here one period after this punishment started, players return to $\overrightarrow{CC}$ instead of $\overrightarrow{DC}$. As in that case, the only two punishments that can successfully deter deviations are $(I1)$ and $(I2)$. As the essentials of the arguments leading to this conclusion are identical to Proposition 5, we leave the details to the reader here. The reader should note that claims 1 - 3 of that result continue to apply here as well, providing the basic building blocks for the construction of renegotiation-proof punishments that deter I's deviations at the bribing stage. Just as in that case, $y \leq z-x+\frac{x^2}{z}$ implies that $\frac{z-x}{x} \in \left[\frac{y-x}{z-y}, \frac{z-y}{x+y-z}\right]$, and hence, from claims 1 - 3, also implies the existence of a renegotiation-proof punishment deterring I from deviating at the bribing stage in any $t \geq 1$ for all $(\delta_I, \delta_P) \in \Omega_{CC}$. Note that we focus on $2x < z$ in Proposition 5. Based on $z < 2x$, i.e., $\frac{z-x}{x} < 1$, $y \leq z - x + \frac{x^2}{z}$ implies $\frac{y-x}{z-y} \leq \frac{z-x}{x} < 1 \leq \frac{\delta_P}{\delta_I}$ and thus $(I1) \in \mathcal{P}\left((I1), (I3), (I4)\right)$ for all $(\delta_I, \delta_P) \in \Omega_{CC}$. For the case where $y > z - x + \frac{x^2}{z}$, or wherein $\frac{y-x}{z-y} > \frac{z-x}{x} > \frac{z-y}{x+y-z}$, we note that $\frac{\delta_P}{\delta_I}$ is minimized when $\delta_I = \delta_P$; and hence, it follows that $y \leq \frac{z+x}{2}$ is also sufficient to guarantee that for all $(\delta_I, \delta_P) \in \Omega_{CC}$, $\frac{\delta_P}{\delta_I} > \frac{y-x}{z-y}$, and hence, renegotiation proof punishments exist that deter I from refusing to bribe. Another condition for existence of renegotiation proof punishments is $\frac{z-y}{x+y-z} \geq \frac{y}{z-y}$, which is however contradicted by $\frac{z-y}{x+y-z} < \frac{y-x}{z-y} < \frac{y}{z-y}$. Summing up, there exist renegotiation proof punishments that deter I from refusing to bribe for all $(\delta_I, \delta_P) \in \Omega_{CC}$ if:

$$y \leq z - x + \frac{x^2}{z} \text{ or } y \leq \frac{z+x}{2}. \tag{15}$$

When this condition fails, such renegotiation proof punishments still exist for $(\delta_I, \delta_P) \in \Omega_{CC}$ satisfying (8).

Step 3 (Deviations by player P at the bribing stage in period $t \geq 1$ when $b_{Pt} > 0 = b_{It}$). We start by constructing four possible punishments for a deviation by player P akin to $(I1) - (I4)$:

$$a_t = (D, D) \text{ and } U_{Pt+1} = x, \tag{P1}$$

$$a_t = (D, C) \text{ and } U_{Pt+1} = \frac{x}{\delta_I}, \tag{P2}$$

$$a_t = (C, C)\,; U_{It+1} \geq y + \underline{b} \text{ and } U_{Pt+1} \geq y - \overline{b}, \tag{P3}$$

and

$$a_t = (C, D) \text{ and } U_{Pt+1} \geq x. \tag{P4}$$

It is directly computed that $(P1) \in \mathcal{P}(\{(P1),(P4)\})$. We can also argue that neither $(P3)$ nor $(P4)$ can deter P's deviations at the bribing stage. This follows from the fact that if this were the case, then following a deviation in period t, the players would agree to $(P3)$ or $(P4)$ and upon reaching period $t+1$, P would refuse to bribe again and propose the same punishment path. Hence, proving the existence of renegotiation-proof punishments is equivalent to providing conditions under which $\{(P1),(P2)\} \subset \mathcal{P}(\{(P1),(P2),(P3),(P4)\})$. We now proceed in a manner exactly akin to the proof of Step 2 in Proposition 5, but leave the proofs to the reader.

Claim 4: $(P1) = P(\{(P1),(P2)\})$ iff $\frac{\delta_P}{\delta_I} < \frac{x}{z-x}$, and $(P2) = P(\{(P1),(P2)\})$ iff $\frac{\delta_P}{\delta_I} > \frac{x}{z-x}$.

Claim 5: Suppose $(P1) = P(\{(P1),(P2)\})$. Then $(P1) \in P(\{(P1),(P3),(P4)\})$ if and only if $\frac{\delta_P}{\delta_I} \leq \frac{z-y}{y-x}$.

Claim 6: Suppose $(P2) = P(\{(P1),(P2)\})$. Then $(P2) \in P(\{(P2),(P3),(P4)\})$ if and only if $\frac{\delta_P}{\delta_I} \geq \frac{x+y-z}{z-y}$.

Based on an argument similar to those used in step 2 and the proof of Proposition 5, it can be proved that there exists a renegotiation-proof punishments that deter P from deviating at the bribing stage of $t \geq 1$ over the entire set Ω_{CC} of admissible discount factors if and only if $\frac{x+y-z}{z-y} \leq \frac{z-y}{y-x}$ (i.e., $y \leq z - x + \frac{x^2}{z}$) or $\frac{x+y-z}{z-y} \leq 1$ (i.e., $y \leq z - \frac{x}{2}$). Combining this with (15) as showed in step 2, we have established that under the assumption in part (i) of the proposition, for all $(\delta_I, \delta_P) \in \Omega_{CC}$, there exist renegotiation proof punishments which deter both players from refusing to bribe. More generally, according to claims 4-6, such renegotiation proof punishments exist so long as $(\delta_I, \delta_P) \in \Omega_{CC}$ satisfies (9).

When both (8) and (9) hold, as in (ii.a) of the proposition, any equilibrium payoff which requires both players to bribe is renegotiation proof at the bribing stage. Notice that $y < \frac{z+x}{2}$ implies $\frac{\delta_P}{\delta_I} \geq \frac{y-x}{z-y}$ because $\frac{y-x}{z-y} < \frac{(z+x)/2-x}{z-(z+x)/2} = 1 \leq \frac{\delta_P}{\delta_I}$. Besides, $2x \leq z$ if and only if $\frac{y-x}{z-y} \geq \frac{x+y-z}{z-y}$. In other words, if $2x \leq z$, then $\frac{\delta_P}{\delta_I} \geq \frac{y-x}{z-y}$ implies $\frac{\delta_P}{\delta_I} \geq \frac{x+y-z}{z-y}$.

Besides, $\frac{\delta_P}{\delta_I} \leq \frac{z-y}{x+y-z}$ implies $\frac{x+y-z}{z-y} < 1 < \frac{\delta_P}{\delta_I}$. This proves that (8) implies (9) if $2x \leq z$. Similar arguments show that if $2x \geq z$, then (9) implies (8). When $(\delta_I, \delta_P) \in \Omega^2_{CC}$, player I will be bribing for all $t \geq 2$. If player P is also not required to bribe in the first period, then only (8) is required to supply renegotiation proof punishments at the bribing stage. When $(\delta_I, \delta_P) \in \Omega^1_{CC}$, player P will be bribing for all $t \geq 2$. If player I is also not required to bribe in the first period, then only (9) is required to supply renegotiation proof punishments at the bribing stage. This completes the proof of step 3.

Step 4 (Verify that assumption in step 1 is satisfied). In fact, using the punishment strategy in step 3, we can make either player bribe at the bribing stage of any period t an amount large enough that his/her continuation payoff is $U_{it} = x$, $i \in \{I, P\}$. This verifies the assumption in step 1 that if any player plays D in the action stage of any period t when he/she is required to play C, the two players can enforce a renegotiation proof punishment that leads to $U_{it+1} = x$, $i \in \{I, P\}$. ■

References

[1] Abreu, D., (1986), “Extremal Equilibria of Oligopolistic Supergames”, *Journal of Economic Theory*, 39, pp. 191-225.

[2] Abreu, D., (1988), “On the Theory of Infinitely Repeated Games With Discounting”, *Econometrica*, 56, pp. 383-96.

[3] Abreu, D., D. Pearce, and E. Stachetti, (1990), "Toward a Theory of Discounted, Repeated Games With Imperfect Monitoring", *Econometrica*, 58, pp. 1041-63.

[4] Andreoni, J. and H. Varian, (1999), “Preplay Contracting in the Prisoners’ Dilemma”, *Proceedings of the National Academy of Sciences, USA*, 96, pp. 10933-8.

[5] Baliga, S. and R. Evans, (2000), “Renegotiation in Repeated Games With Side Payments”, *Games and Economic Behavior*, 33, pp. 159-76.

[6] Barrett, S., (2001), “International Cooperation for Sale”, *European Economic Review*, 45, pp. 1835-50.

[7] Bolle, F., (2001), “Why to Buy Your Darling Flowers: On Cooperation and Exploitation”, *Theory and Decision*, 50, pp. 1-28.

[8] Camerer, C., (1988), “Gifts as Economic Signals and Social Symbols”, *American Journal of Sociology*, 94, pp. S180-S214.

[9] Carmichael, L. and B. MacLeod, (1997), “Gift Giving and the Evolution of Cooperation”, *International Economic Review*, 38, pp. 485-509.

[10] Celentani, M., D. Fudenberg, D. Levine, and W. Pesendorfer, (1996), “Maintaining a Reputation Against a Long-Lived Opponent”, *Econometrica*, 64, pp. 691-704.

[11] Farrell, J. and E. Maskin, (1989), “Renegotiation in Repeated Games”, *Games and Economic Behavior*, 1, pp. 327-60.

[12] Fudenberg, D. and D. Levine, (1989), "Reputation and Equilibrium Selection in Games With a Patient Player", *Econometrica*, 57, pp. 759-78.

[13] Jackson, M. and S. Wilkie, (2005), "Endogenous Games and Mechanisms: Side Payments Among Players", *Review of Economic Studies*, 72, pp. 543-66.

[14] Levin, J., (2003), "Relational Incentive Contracts", *American Economic Review*, 93, pp. 835-57.

[15] Lehrer, E. and A. Pauzner, (1999), "Repeated Games With Differential Time Preferences", *Econometrica*, 67, pp. 393-412.

[16] Lehrer, E. and L. Yariv, (1999), "Repeated Games With Incomplete Information on One Side: The Case of Different Discounting Factors", *Mathematics of Operations Research*, 24, pp. 204-18.

[17] Mailath, G., I. Obara, and T. Sekiguchi, (2002), "The Maximum Efficient Equilibrium Payoff in the Infinitely Repeated Prisoners' Dilemma", *Games and Economic Behavior*, 40, pp. 99-122.

[18] Schmidt, K., (1993), "Reputation and Equilibrium Characterization in Games With Conflicting Interests", *Econometrica*, 61, pp. 325-51.

[19] Sorin, S., (1986), "On Repeated Games With Complete Information", *Mathematics of Operations Research*, 11, pp. 147-60.

[20] Stahl II, D., (1991), "The Graph of the Prisoners' Dilemma Supergame Payoffs as a Function of the Discount Factor", *Games and Economic Behavior*, 3, pp. 368-84.

[21] Van Damme, E., (1989), "Renegotiation-Proof Equilibria in Repeated Prisoners' Dilemma", *Journal of Economic Theory*, 47, pp. 206-17.