EVEN UP: MAINTAINING RELATIONSHIPS

LARRY SAMUELSON
Yale University

and

ENNIO STACCHETTI
New York University

January 18, 2016

ABSTRACT. When are economic interactions mediated by prices and transfers, and when are they mediated by variations in the continuation payoffs within a relationship? To address this question, we examine a relationship between two players who periodically face opportunities to undertake randomly generated "projects" that impose a cost on one player and a benefit on the other, as well as opportunities to make voluntary transfers to one another. We characterize the set of equilibria, with particular emphasis on the efficient frontier. In the absence of transfers, incentives are often created by adjusting the continuation values of the relationship, sometimes increasing player 1's continuation payoff (to induce player 1 to accept a particularly costly current project) and sometimes increasing player 2's continuation payoff. In contrast, once transfers enter the picture, continuation payoffs maximize the sum of the two players' payoffs, no matter how asymmetric the point on the Pareto frontier under consideration. However, transfers can also be relatively rare, and in particular can occur less often than the adjustments in continuation values that occur without transfers. Players thus rely on transfers to create incentives, but manage to do so while rarely making transfers.

I. Introduction

I.1 Motivation: Maintaining Relationships

In many economic interactions, the incentives to take an action are created by contemporaneous prices, contracts or mechanisms. Consumers purchase goods and pay for them. Firms pay their workers according to contracts specifying wages and commissions. Governments hold procurement auctions and pay the winners. The participants in these encounters need neither know nor care whether they will meet again.

In many other economic encounters, the incentives are created by the shadow of future interactions. A current benefit is countered not by an immediate payment but by offering more lucrative future terms. The compensation for incurring a current cost is not a contemporaneous payment, but more favorable treatment in future interactions. Here, it is important that the parties are engaged in a continuing relationship.

The literature suggests that such relationships are important. Firms in business-to-business relationships typically respond to shocks not by making compensating payments or by appealing to the terms of their contract, but by agreeing that "we'll make it up next time" (e.g., Macaulay (1963) and Paley (1985)). For example, Bernstein (1992) and Richman (2006) describe how diamond merchants rely on relationships rather than contracts or formal agreements to mediate their interactions, and how this in turn has shaped the nature of the diamond industry. The literature on "psychological contracts" (e.g., George (2009) and Rousseau (1995)) argues that interactions within organizations, including notably terms of employment, are typically governed by informal relationships. Ellickson (1991) studies the residents of (primarily rural) Shasta County in California, as well as a variety of other interactions, finding that they rely on the incentives created by their continuing relationships to coordinate their activities and settle disputes, rather than appealing to formal agreements or threats of legal action. More generally, the study of social capital catalysed by Putnam (1995, 2002) stresses the importance of relationships. For example, DiPasquale and Glaeser (1998) present evidence that homeowners invest more than renters in externality-creating amenities, because the lower mobility of the former makes them more likely to enjoy the continued reciprocal benefits of their relationships, a consideration that an appeal to the Coase theorem and the corresponding payments would render irrelevant. Looking backward, Greif (e.g., Greif (1997, 2005) and Greif, Milgrom and Weingast (1994)) highlight the role of repeated interactions in making possible the trade upon which modern economies are built. Cosmides and Tooby (2002a, 2002b) suggest that monitoring relationships was at one point so crucial to our evolutionary success that our brains have developed specialized resources for sustaining relationships.

It is a familiar finding that repeated interactions allow the creation of incentives that cannot be supported in isolated interactions. The well-developed literature on repeated games and reputations is built on this observation (see Mailath and Samuelson (2006, 2013) and Samuelson (2006) for introductions). This literature typically isolates the role of continuation values in creating incentives by working with a stage game that precludes transfers. We are interested in examining the interplay between transfers and continuation values—when are economic interactions mediated by prices and transfers, and when are they mediated by variations in the continuation payoffs within a relationship?

To address these questions, we study repeated interactions in which transfers are possible, but cannot be compelled by recourse to contracts or legal enforcement. Our emphasis on *voluntary* transfers is motivated by the belief that in many interesting circumstances, it is impossible or impractical to compel transfers (or other actions) via contractual or legal means. We discuss this belief at greater length in Section V.2. Here, we note that medieval trade often proceeded in the presence of courts that were too weak for merchants

to reliably count on contract enforcement. Enforcement was undoubtedly all the more problematic in our own evolutionary past. In modern settings, a variety of transactions costs may make formal enforcement impracticable. Richman (2006, p. 18) reports one of the subjects in his study commenting, "The truth is that if someone owes you money, there's no real way to get it from him if he doesn't want to pay you."

There is again a well developed literature on such "relational contracts" or "relational incentive contracts" (e.g., Baker, Gibbons and Murphy (2002), Levin (2003), and Pearce and Stacchetti (1998)). The standard result in this literature is that contractual payments are routinely supplemented by voluntary payments. In practice, however, transfers often appear to be used only under exceptional circumstances. For example, Ellickson (1991) describes the residents in his study as managing the incidents in which they impose costs and benefits on one another without recourse to formal agreements or threats of legal action and also typically without making transfers. Instead, their behavior follows an *even up* policy (Ellickson, 1991, pp. 225–229) built around the following principles:

- The participants in a relationship maintain a sense of the balance in their relationship. As long as this balance does not stray too far from even, the participants make no transfers and choose their actions so as to maximize joint welfare. Most notably, transfers are a relatively rare occurrence.
- If the balance strays too far from even, then one participant makes a transfer to the other that restores the balance, returning the relationship to the first point.
- Failures to perform the duties of a relationship prompt responses from the participants in the relationship, notably in the form of declining costly actions that would be of value to one's partner.
- The responses to failures tend to be temporary, lasting until balance can be restored and the relationship brought back into balance.

Fiske (1992) refers to "equality matching, " in which people continually confer costs and benefits on one another without making transfers, as one of four basic forms of social interaction. The etiquette industry goes to great lengths to manage relationships without transfers. Our goal is thus to examine a relationship in which voluntary transfers are always feasible, but are in equilibrium only sparingly used.

I.2 Results: A Preview

We examine the simplest possible interaction in which the creation of incentives gives rise to a tradeoff between voluntary transfers and continuation values. Two players periodically face opportunities to accept randomly generated "projects" that impose a cost on one player and a benefit on the other.

We set the stage in Section III by examining a repeated interaction *without* transfers. The stage game in this interaction has a unique equilibrium in which no projects are undertaken. The repeated interaction has a stationary-outcome equilibrium that yields higher payoffs—the players accept all projects that are not too costly, switching in the event

of a deviation to an "autarky" equilibrium in which no projects are accepted. However, the players can achieve even higher payoffs in an equilibrium in which a player on the negative end of a particularly costly project is rewarded with a higher continuation value. In this case, continuation payoffs are routinely varied along the equilibrium path to create incentives. The forces at work here are familiar from the theory of repeated games.

Section IV turns to our case of primary interest, that in which the stage game of Section III is augmented by the opportunity to make voluntary transfers. We present an example showing that the Pareto frontier of payoffs can lie strictly outside the frontier without transfers. We find that no matter what weighted sum of payoffs for the two players is being maximized at a particular point on the frontier, the continuation value in every subsequent period maximizes the equally-weighted sum of payoffs. Hence, along the equilibrium path continuation values are never shifted to create incentives to undertake projects. Nonetheless, there are equilibria in which transfers are rare (in a sense we develop below), and come into play only when particularly expensive projects arise. Hence, despite the fact that continuation values are not shifted to create incentives, transfers may be rarely made.

This equilibrium captures the basic features of the even up strategy examined by Ellickson (1991). First, under ordinary circumstances, in which the relationship is not too unbalanced, no transfers are made and the continuation payoffs that lie along the equilibrium path do not respond to current actions. Second, the relationship is thrown off balance when a particularly expensive project arises. The response here is an immediate transfer that puts the relationship back on track. Third, the failure to execute an equilibrium action, whether to undertake a project or make a transfer, causes the relationship to enter a punishment phase in which projects are declined that would otherwise have been undertaken. Finally, though we do not pay attention to this in the analysis, this punishment phase need not be permanent, so that the relationship can get back on track.

These two models together provide insight into the tradeoff between transfers and continuation values. In the absence of transfers, continuation values are often adjusted, in the process giving rise to a tradeoff between current decisions and continuation values. Some current projects that would increase the target weighted sum of payoffs are rejected, even though one could create the incentives to accept them, because these incentives require too costly an adjustment in continuation payoffs. In contrast, once transfers enter the picture, continuation payoffs maximize the (unweighted) sum of the two players' payoffs, no matter how asymmetric the point on the Pareto frontier under consideration. Moreover, the tradeoff between accepting current projects and continuation values disappears. Current projects are judged only in terms of their contribution to the current objective. Transfers are used to create incentives, but may be rare.

Section V considers extensions designed to probe the robustness of our findings. Section V.1 by considers the case in which the marginal utility of receiving a transfer is diminishing in the size of the transfer, so that utility is not perfectly transferable. Perhaps paradoxically, the result of making transfers less effective in moving utility from one player to the other is that, in equilibrium, transfers occur in every period. Section V.2 considers

an example in which the players are initially uncertain about the distribution of possible projects. The interesting observation here is that equilibrium has a "starting small" characteristic. Equilibrium play is initially characterized by the acceptance of only relatively inexpensive projects. Encouraging news about the distribution of such projects can lead the relationship to grow and embrace more expensive projects. Discouraging news may or may not reduce the scale of projects undertaken. Sections V.3 and V.4 turn to imperfect monitoring. As is familiar from theories of repeated games, imperfect monitoring blurs the line between equilibrium play and the supporting punishments. The basic forces and construction explored in Section III reappear, though with more complication.

II Related Literature

We have already noted that our work is connected to the (large) literatures on repeated games and relational contracts. We concentrate here on the papers in this literature whose structure is most similar to ours.

II.1 Favors

Bramoullé and Goyal (2013) examine a model in which people are divided into two groups. In each period, one individual is chosen to be the "principal" and another individual is chosen to be the "expert." A unit surplus is created if the principal hires the expert to pursue a joint venture, and a smaller surplus is created if the principal hires any other individual for the joint venture. The efficient outcome calls for the principal to always hire the expert. In contrast, a group practices "favoritism" if principals from that group always hire agents from that group, in the process hiring the expert if (but only if) the latter happens to be drawn from the principal's group. The basic result is that under certain conditions a group can gain from practicing favoritism (while the other group loses). Both groups lose if both practice favoritism.

Leo (2014) examines a "volunteer's dilemma" in which two players face a task in each period that gives a payoff of 1 to each player if completed and a payoff of 0 if not completed. In each period each player independently draws a privately-observed cost of completing the task. Any function $h : \{0, 1, 2, \dots\} \to \{1, 2\}$ gives rise to an equilibrium in which player $h(t)$ completes the task in period t, but such equilibria are inefficient because they make no use of the information about which player has the lower cost. Leo (2014) examines the extent to which the two players can reduce the average cost of performing the task by using messages to allocate responsibility for performing the task.

Bowen, Kreps and Skrzypacz (2013) examine a model in which the incentive to undertake randomly generated and costly favors are provided by community enforcement. The costs and benefits of the favor are observed by the person bearing the cost and the recipient of the benefits, but not by the remainder of the community. Interest centers of the effects of allowing the recipient to either excuse (ex ante) the person incurring the cost from undertaking the favor, or to forgive (ex post) the person for not doing so.

Unlike Bramoullé and Goyal (2013), our setting involves no conception of a group and emphasizes the extent to which an interaction can bring mutual gains rather than achieving gains at the expense of the other group. The draws of the cost levels in Leo (2014) play a role quite similar to the randomly drawn projects of our paper. Unlike Bowen, Kreps and Skrzypacz (2013), we have no surrounding community. The most important difference is that there are no transfers in Bramoullé and Goyal (2013), Leo (2014) and Bowen, Kreps and Skrzypacz (2013), and no sense of the tradeoff between transfers and continuation payoffs in creating incentives that lies at the center of our analysis.

II.2 Private Information

Mobius (2001) and Hauser and Hopenhayn (2005) examine somewhat different versions of a continuous-time model with two players. A pair of independent Poisson processes, one for each player, generate opportunities for the player in question to generate a favor. Favors are either fixed in size or divisible, in which case the player receiving the opportunity chooses the scale of the favor, with a favor of scale x produce a cost cx to the person granting the favor and benefit bx (with $b > c$) to the recipient.

If the arrival of favor opportunities were observed, then the efficient equilibrium in this game would be straightforward. Every favor would be granted at the maximum feasible scale, with deviations deterred by (for example) reversion to an autarkic outcome in which no subsequent favors are granted. Interest centers on the case in which the arrival of favor opportunities is privately observed. Mobius (2001) constructs a simple mechanism that relies on the running calculation of a state variable k^t, defined as the number of times player 1 has granted a favor in the interval $[0, t)$ minus the number of times player 2 has granted a favor. There is a constant K such that player 1 grants a favor at time t (if given the opportunity) if and only if $k^t < K$, and player 2 does so if and only $k^t > -K$.

The equilibrium examined by Mobius (2001) allows favors to often be granted, but does not in general lie on the Pareto frontier of equilibrium payoffs. Hauser and Hopenhayn (2005) characterize this frontier. They first show that the frontier is self generating. The equilibria they construct are conceptually similar to that of Mobius, but the counterpart of the state variable k^t is now a continuation payoff drawn from the frontier of such payoffs. Each favor granted by player 2 pushes the continuation payoff upward and to the left on this frontier, increasing 2's continuation payoff by just enough to create the incentive to grant the favor, while each favor granted by player 1 does the reverse. Favors are granted at the maximum feasible scale unless continuation payoffs cannot be moved far enough (because an axis intervenes) to create the requisite incentive. Moreover, in the absence of favors, continuation payoffs drift back toward the center.

Continuation payoffs take center stage in both Mobius (2001) and Hauser and Hopenhayn (2005), while the key difficulty in the analysis arises out of the fact that an opportunity to confer a favor is private information. In contrast, the arrival of a favor is public information in our model, while interest centers on the ability to rely on transfers (or not) rather than continuation payoffs in creating incentives. We believe that private information about opportunities to generate gains from trade is often important, but we

also believe that the familiarity created by a continuing relationship may often make it a good approximation to think of such opportunities as being commonly observed.

Abdulkadiroğlu and Bagwell (2013) examine a repeated interaction between two players in which privately-observed opportunities again play a central role, but with a more complicated stage game. In each period, either one player is given an income of 1, the other player is given an income of 1, or neither player is given an income. The receipt of income is privately observed. If a player receives income, then that player may choose to exhibit trust by investing any amount $x \in [0,1]$ with the other player. The value of x is publicly observed. This investment succeeds with probability $q < 1$, in which case it produces an income of kx, or fails, in which case it produces nothing. The outcome of the investment is privately observed by the recipient of the investment. If the investment is successful, then the recipient can transfer some or all of the returns back to the investor, while keeping the remainder. The efficient outcome is for income to always be fully invested (since $qk > 1$). Given the complexity of the stage game, a characterization of the efficient frontier of equilibria in the repeated game is elusive. Abdulkadiroğlu and Bagwell (2013) construct equilibria in which players who receive income invest a portion of that income, the size of which is decreasing in the length of time since the opposing player last invested. Each investment shifts the continuation payoffs along a line in the direction of the player making the investment.

Once again, our model differs in that the arrival of a project (the counterpart of the receipt of income in Abdulkadiroğlu and Bagwell (2013)) is commonly observed. While we concentrate on the case in which the costs and benefits of a project are also commonly observed, we consider extensions in which these are privately observed. One might suspect that this would pose obstacles similar to those encountered by Abdulkadiroğlu and Bagwell (2013), but even here we are able to characterize the efficient frontier. The ability to observe the *existence* of an opportunity (as in our model) thus appears to be more important than the ability to observe the details of these opportunities.

II.3 The Role of Limits

Miller (2009) examines a repeated game in which players in each period first receive (possibly correlated) private information, then communicate, then choose their actions, and then make monetary transfers. The central result of his paper is a folk theorem. His analysis proceeds by first viewing the choice of communication and actions in each period as a mechanism design problem, assuming that the designer can commit to the transfers that characterize the mechanism. Once this problem is solved, one must next confront the fact that in the repeated interaction, incentives must be created to ensure that the transfers emerging from the mechanism design stage are indeed carried out. Miller (2009) shows that these incentives are straightforward, as long as the stage games (and hence the induced mechanisms) are bounded and the players are allowed to become arbitrarily patient.

The analysis of our setting would similarly be straightforward if we assumed that the measure μ generating projects had bounded support, and that players are arbitrarily pa-

tient. In this case, we could readily construct equilibria in which every profitable project is accepted, with incentives crated by the prospect of reversion to a continuation path in which no projects are accepted. There would be no need for transfers, nor for variations in continuation payoffs along the equilibrium path. The subtleties and the interest in our problem arises because we work with the reverse configuration of limits. We do not assume that μ has bounded support, while working with fixed discount factors rather than arbitrary patience.

III. The Model

This section presents a preliminary model, distinguished by the absence of transfers, that illustrates many of our ideas in a simple setting and that allows us to subsequently isolate the role of transfers.

III.1 The Stage Game

There are two players, 1 and 2. The stage game has an extensive form. First, Nature chooses a *project* $x \in \mathbb{R}^2$ according to a probability measure μ. Players 1 and 2 observe Nature's choice and then simultaneously choose $(a_1, a_2) \in \{0, 1\}^2$. Player i's utility function is given by

$$u_i(x_1, x_2, a_1, a_2) = a_1 a_2 x_i.$$

We think of a project as imposing a cost on one player and a benefit on the other. This may represent an opportunity for one player to grant a costly favor to the other player, or for one player to realize a benefit that imposes a cost on the other. Nature's draw determines who bears the cost and who reaps the benefit of the project, as well as the magnitude of these costs and benefits. The project is accepted, and the payoffs (x_1, x_2) realized, if and only if both players choose to accept (i.e., if $a_i = 1$, $i = 1, 2$). Otherwise, the project is rejected and both players receive payoff 0.

Let $X = \text{supp}(\mu)$ denote the support of μ. We assume that $X \subset \{x \in \mathbb{R}^2 \mid x_1 x_2 \leq 0\}$, consistent with our interpretation of a project as conferring a benefit on one player and imposing a cost on the other. We also assume that $\mathbb{E}[x_i] > 0$ and $\mathbb{E}[\max\{x_i, 0\}] < +\infty$ for $i = 1, 2$, and that $\mu(\{x \mid x_1 x_2 = 0\}) = 0$.

The stage game has a (essentially unique) Nash equilibrium with value $(0, 0)$ in which all projects are rejected (except possibly those on the zero-measure set $\{x \mid x_1 x_2 = 0\}$, which motivates the "essentially unique").

III.2 The Repeated Game

We consider an infinitely repeated game Γ with perfect monitoring and public randomization. We model the latter by assuming that, in every period t, before the project x^t is realized, the players publicly observe the value ζ^t randomly drawn from a uniform distribution in $[0, 1]$. Appendix A.2 presents an example showing that we need the random device

in order to ensure the set of equilibrium payoffs is convex, and so we are more explicit about the device than is often the case.

The random project x^t is independent of ζ^t and $\{(\zeta^t, x^t)\}$ is a sequence of independent and identically distributed random variables. The actions $a^t = (a_1^t, a_2^t)$ chosen by the players in period t are publicly observed at the end of the period. Given a sequence of projects and actions $\{(x^t, a^t)\}_{t=0}^{\infty}$, player i's payoff is

$$(1 - \delta) \sum_{t=0}^{\infty} \delta^t x_i^t a_1^t a_2^t,$$

where $\delta \in (0, 1)$ is the common discount factor. Notice that each realization of ζ^0 leads to a distinct subgame Γ_{ζ^0}, but since ζ^0 has no effect on payoffs, the games Γ_{ζ^0} are identical.

A period-t *history* h^t includes the outcome of the public randomization device, the project, and the choices of the players in periods $0, \ldots t - 1$ (a period-0 history is empty). A behavior strategy σ_i specifies a probability $\sigma_i(h^t, \zeta^t, x^t)$ that player i accepts the period-t project x^t after observing history h^t and the period-t outcome ζ^t of the public randomization device. A pure strategy σ_i is such that $\sigma_i(h^t, \zeta^t, x^t) \in \{0, 1\}$ for each *interim history* (h^t, ζ^t, x^t). For any (behavior) strategy σ_i for player i and any period-t history h^t, we denote by $\sigma_i|_{h^t}$ the continuation strategy of σ_i after history h^t. Note that because the subgame after history h^t is identical to the original game, $\sigma_i|_{h^t}$ is itself a strategy for the original game.

For a strategy profile $\sigma = (\sigma_1, \sigma_2)$, let $u(\sigma) = (u_1(\sigma), u_2(\sigma))$ denote its expected value. That is,

$$u_i(\sigma) = \mathbb{E}_\sigma \Big[(1 - \delta) \sum_{t=0}^{\infty} \delta^t a_1^t a_2^t x_i^t \Big], \quad i = 1, 2.$$

With a small abuse of notation, if σ is a strategy for the subgame Γ_ζ, then we also denote its expected value by $u(\sigma)$. We also denote the history $(\emptyset, \zeta, x)$ simply by (ζ, x).

III.3 Equilibrium

Let $V \subset \mathbb{R}^2$ be the set of subgame perfect equilibrium payoffs of Γ. Notice that if σ is a subgame perfect equilibrium for Γ, then the continuation strategy σ_ζ must be a subgame perfect equilibrium for Γ_ζ for all $\zeta \in [0, 1]$.

Let σ^A denote the pure strategy where both players reject every project after every history. This strategy plays a stage-game Nash equilibrium in every period, and hence is trivially a subgame perfect equilibrium. It has payoffs $(0, 0)$. We refer to σ^A as the autarky equilibrium and say that the players have terminated the relationship whenever they follow the subgame perfect equilibrium σ^A. Each player can ensure a payoff of at least 0 by rejecting every project. Therefore, σ^A simultaneously delivers the worst equilibrium payoff of 0 to both players.

The existence of the autarchy equilibrium ensures that V is nonempty. For any strategy σ, we have

$$u(\sigma) = \int_0^1 u(\sigma|_\zeta)d\zeta.$$

This ensures the set of equilibrium payoffs V is the convex hull of the equilibrium payoffs of the (identical) games V_ζ, and hence is convex. Standard arguments ensure that V is compact. We thus immediately have the following:

Proposition 1. *The equilibrium value set V is nonempty, compact and convex.*

III.4 Analytical Tools

This section presents some preliminary results that simplify the analysis. We decompose strategies into their current actions and continuation payoffs, along the lines developed by Abreu, Pearce and Stacchetti (1990). We then show that attention can be restricted to strategies that are pure and simple (in a sense to be made precise). This allows us to characterize the set of subgame perfect equilibria.

III.4.1 Factorization

We first develop a convenient description of the conditions for an equilibrium. Let σ be a behavior strategy. For any outcome $\zeta \in [0, 1]$ of the public randomization device and any project $x \in X$ for period 0, let

$$\alpha_i(\zeta, x) \quad i = 1, 2,$$

denote respectively the probabilities that each player approves the project. For any pair of choices $a = (a_1, a_2) \in \{0, 1\}^2$, let

$$w^a(\zeta, x) = u(\sigma|_{(\zeta, x, a)})$$

denote the value of the continuation strategy after the players choose a in period 0. Hence, (α, w) describes the current actions and the continuation payoffs induced by the behavior strategy. We can then write the resulting payoffs as

$$u(\sigma) = \int_0^1 \int_X \left[\alpha_1 \alpha_2 [(1 - \delta)x + \delta w^{1,1}] + \delta[\alpha_1(1 - \alpha_2)w^{1,0} \right.$$
$$\left. + (1 - \alpha_1)\alpha_2 w^{0,1} + (1 - \alpha_1)(1 - \alpha_2)w^{0,0}] \right] d\mu(x)d\zeta,$$

where, for simplicity, the argument (ζ, x) for α_i $(i = 1, 2)$ and w^a $(a \in \{0, 1\}^2)$ is omitted here and below. We say that (α, w), where $w = (w^{0,0}, w^{0,1}, w^{1,0}, w^{1,1})$, *factorizes σ*.

A factorization of the strategy σ expresses the payoffs induced by σ as a sum of current and continuation payoffs. The gives us our description of the conditions for equilibrium— the behavior strategy σ factorized by (α, w) is a subgame perfect equilibrium if and only if for all $\zeta \in [0,1]$ and $x \in X$,

(i) $\sigma|_{(\zeta, x, a)}$ is a subgame perfect equilibrium for each $a \in \{0,1\}^2$,

$$(ii)\ (1-\delta)x_1 + \delta[\alpha_2 w_1^{1,1} + (1-\alpha_2)w_1^{1,0}] \left.\begin{cases} \geq \\ = \\ \leq \end{cases}\right\} \delta[\alpha_2 w_1^{0,1} + (1-\alpha_2)w_1^{0,0}] \qquad (IC_1)$$

$$(1-\delta)x_2 + \delta[\alpha_1 w_2^{1,1} + (1-\alpha_1)w_2^{0,1}] \left.\begin{cases} \geq \\ = \\ \leq \end{cases}\right\} \delta[\alpha_1 w_2^{1,0} + (1-\alpha_1)w_2^{0,0}] \qquad (IC_2)$$

where the relevant inequalities are

$$\begin{cases} \geq & \text{if } \alpha_i = 1 \\ = & \text{if } \alpha_i \in (0,1) \\ \leq & \text{if } \alpha_i = 0. \end{cases}$$

Condition (i) requires that continuation strategies are themselves equilibria, ensuring that continuation values w^a are in V (for all $a \in \{0,1\}^2$), while condition (ii) ensures that current actions are incentive compatible. Conversely, if $\alpha_i : [0,1] \times X \to [0,1]$ for $i = 1,2$ and $w^a : [0,1] \times X \to V$ for $a \in \{0,1\}^2$ satisfy (IC_i) for $i = 1,2$, then (α, w) is the *factorization* of a subgame perfect equilibrium.

III.4.2 Pure Strategies Suffice

The next proposition establishes that without loss of generality we can restrict our analysis to pure strategy subgame perfect equilibria. The idea behind this result is that, since each period begins with a draw from the public randomization device, no further mixtures are required. This is an expected but convenient result.

Let V_P be the set of values of all pure strategy subgame perfect equilibria.

Proposition 2. $V_P = V$, *i.e., any subgame perfect equilibrium payoff is the payoff of a pure strategy subgame perfect equilibrium.*

The proof (presented in Appendix A.1) considers a subgame perfect equilibrium σ such that after some interim history at least one player randomizes between accepting and rejecting the current project. For convenience, suppose in fact that this randomization occurs in the first period and let (α, w) be the factorization of the equilibrium σ. The proof first shows that we can replace the factorization (α, w) with a lottery over two pure factorizations that gives the same expected payoff. We choose these two pure factorizations to ensure that each satisfies the incentive constraints, so that each represents a subgame perfect equilibrium. This would complete the argument if, after every interim history,

we could appeal to a public random variable to mediate the lottery over the two pure factorizations that we have just attached to that interim history. The final step of the proof shows that we can accomplish the same by linking these factorizations to the draw of the public randomization device that occurs at the *beginning* of the period. This gives us our pure strategy.

III.4.3 Simple Strategies

We can further simplify the analysis of the set of equilibrium payoffs V by restricting attention to *simple* pure strategies. Let σ be a subgame perfect equilibrium in pure strategies. We can specify another pure strategy profile $\hat{\sigma}$ that agrees with σ on the equilibrium path and that plays the autarky equilibrium after every out-of-equilibrium history. More precisely, the strategy $\hat{\sigma}$ is obtained from σ as follows. After any history h^t that includes a deviation from the (random, given the public randomization device) outcome path of σ in period $t-1$, let $\hat{\sigma}|_{h^t} = \sigma^A$. For all other histories h^t (including the empty history), let $\hat{\sigma}(h^t, \zeta^t, x^t) = \sigma(h^t, \zeta^t, x^t)$ for all $\zeta^t \in [0,1]$ and $x^t \in X$. Since $\hat{\sigma}$ generates the same random outcome path as σ, we have $u(\hat{\sigma}) = u(\sigma)$. Moreover, $\hat{\sigma}$ is a subgame perfect equilibrium—intuitively, we have replaced any punishments in σ with the harshest punishment available (the autarky equilibrium).

Following Abreu (1986), we say that $\hat{\sigma}$ is simple because it only uses the autarky equilibrium as punishment. Though simple strategies are used to simplify the characterization of a subgame perfect equilibria, we extend it to include any pure strategy: we say that any pure strategy is simple if any deviation from its outcome path is followed by autarky.

III.4.4 Factorizing Simple, Pure Strategies

The factorization of a simple strategy is (as one would hope) simple. Let σ be a simple pure strategy for Γ_ζ. Let (α, w) be σ's factorization (here α and w are functions of the project x alone, since we have already conditioned on a draw of ζ). Since $\alpha(x) \in \{0,1\}^2$, the function α partitions X into two sets:

$$A = \{x \in X \mid \alpha(x) = (1,1)\} \quad \text{and} \quad R = \{x \in X \mid \alpha(x) \neq (1,1)\}.$$

The set A contains all projects that are accepted by σ in period 0 and R contains all projects that are rejected. Since σ is simple, $w^a(x) = (0,0)$ for all $x \in X$ and $a \neq \alpha(x)$. In particular, $w^a(x) = (0,0)$ for all $x \in A$ and $a \neq (1,1)$. Let $\hat{w}(x) = w^{\alpha(x)}(x)$ for all $x \in X$. The function $\hat{w}(x)$ specifies the continuation value when the players obey σ's recommendation in period 0. The triple $(A, R, \hat{w})$ provides a complete description of σ and thus it is an alternative representation of a *factorization* of σ. Hereafter, we drop the hat from w and simply write (A, R, w). The following characterization of equilibrium strategies is then immediate:

Lemma 1. *Let (A, R, w) be the factorization of a simple and pure strategy profile σ for Γ_ζ. Then*

$$u(\sigma) = \int_A [(1 - \delta)x + \delta w(x)]d\mu(x) + \int_R \delta w(x)d\mu(x).$$

Moreover, if σ is a subgame perfect equilibrium then $w : X \to V$ and

$$(1 - \delta)x_i + \delta w_i(x) \geq 0 \quad \text{for all } x \in A \text{ and } i = 1, 2. \tag{IC}$$

Conversely, if (A, R, w) satisfies (IC) and $w : X \to V$, there exists a subgame perfect equilibrium $\hat{\sigma}$ that has factorization (A, R, w).

The first equation again writes payoffs as a sum of current and continuation payoffs, while the second equation states the incentive constraints for current actions. Since $V \subset \mathbb{R}^2_+$, note that the incentive constraint (IC) is automatically satisfied for player i when $x \in X$ is such that $x_i \geq 0$; players are happy to accept benefits, and the incentive constraints revolve around getting them to incur costs. We let

$$\beta = (1 - \delta)/\delta,$$

and note that when the incentive constraint is binding for player i for some $x \in A$, then $w_i(x) = -\beta x_i$.

III.5 An Example

Let $X = \{(-1, 3), (3, -1)\}$, with $\mu(-1, 3) = \mu(3, -1)) = 1/2$, and $\delta = 1/2$. We characterize the equilibrium set for this example. As we have noted, there is an autarkic equilibrium featuring payoffs $(0, 0)$. Our task is to characterize the remainder of the equilibrium set.

There are four pure strategy profiles in the stage game, consisting of (i) accepting neither project, (ii) accepting only $(-1, 3)$, (iii) accepting only $(3, -1)$, and (iv) accepting both projects. The set of feasible stage-game payoffs is the convex hull of the corresponding payoff profiles, given by (remembering that each project is equally likely in each period)

$$\text{co} \{(0, 0), (-1/2, 3/2), (3/2, -1/2), (1, 1)\}.$$

Since the minmax payoff for each player is 0, the set of feasible and individually rational payoffs is (see Figure 1):

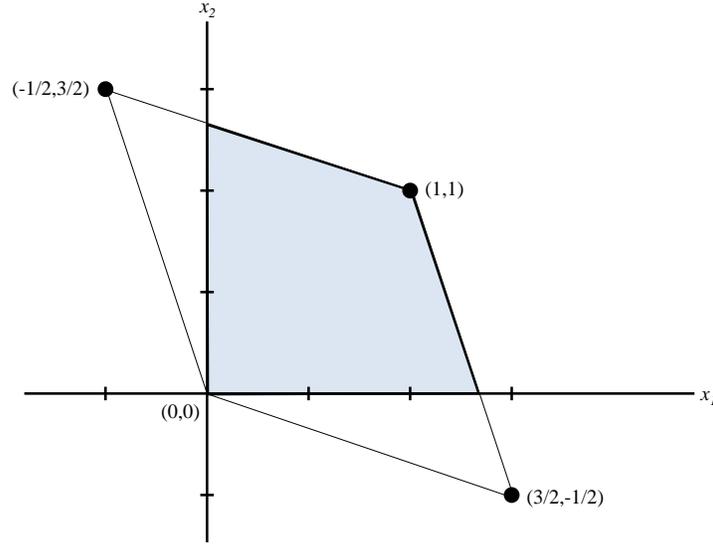$$\text{co} \{(0, 0), (4/3, 0), (0, 4/3), (1, 1)\}.$$

Figure 1: *Illustration of Example III.5. The set of feasible payoffs is the convex hull of $\{(0,0),(-1/2,3/2),(3/2,-1/2),(1,1)\}$, and the set of equilibrium payoffs is the shaded region, namely the convex hull of $\{(0,0),(4/3,0),(0,4/3),(1,1)\}$.*

There is an equilibrium in which every project is accepted, with any history exhibiting a rejection giving rise to the autarky equilibrium. The expected payoff from this equilibrium, and from every continuation equilibrium along the equilibrium path, is $(1,1)$. We need verify only one incentive constraint to confirm that this strategy profile is an equilibrium, ensuring that a player who earns an immediate payoff of -1 from accepting a project is willing to do so. The relevant inequality is

$$(1-\delta)(-1)+\delta \geq 0,$$

which is obviously satisfied (since $\delta = 1/2$). The left side is the payoff from accepting a project, coupling a current payoff of -1 with a continuation payoff of 1, while the right side is the payoff from rejecting, which couples a current payoff of 0 with a continuation payoff (from the autarchy equilibrium) of 0. This is the largest symmetric equilibrium payoff.

The equilibrium that maximizes player 2's payoff gives payoffs $(0,4/3)$, and features strategies in which (i) if $(-1,3)$ is drawn in the first period, then both players accept and the continuation payoff is $(1,1)$, with continuation payoff $(0,0)$ should either player reject, (ii) if $(3,-1)$ is drawn in the first period, then both players reject, and play continues in the next period just as it does after the null history. It is a straightfroward application of Lemma 1 (confirming that (IC) holds) that these strategies constitute an equilibrium. It is clear that player 1 receives a zero payoff in this equilibrium, since the first project to be accepted is of the form $(-1,3)$, followed by a continuation payoff of 1 for player 1. Player 2's payoff solves

$$v_2 = (1/2)[(1-\delta)3+\delta 1]+(1/2)[\delta v_2],$$

which we can solve for $v_2 = 4/3$.

The roles of player 1 and 2 in the preceding paragraph can obviously be reversed to obtain an equilibrium with payoffs $(4/3, 0)$. We thus have equilibria with payoffs $(0, 0)$, $(4/3, 0)$, $(0, 4/3)$, and $(1, 1)$. The set V is the convex hull of these points, which is the set of feasible, individually rational stage-game payoffs. Figure 1 illustrates.

III.6 The Pareto Frontier

This section characterizes the Pareto frontier ∂V of the set V of equilibrium payoffs. It will be useful to have a parametric representation of the Pareto frontier ∂V. We denote the maximum equilibrium payoff to player i as

$$\bar{v}_i = \max\{v_i \mid v \in V\} \quad i = 1, 2,$$

and then define a pair of functions $\varphi_j : [0, \bar{v}_i] \to [0, \bar{v}_j]$, $j = 1, 2$, describing the Pareto frontier: $(v_i, v_j) \in \partial V$ if and only if $v_i \in [0, \bar{v}_i]$ and $v_j = \varphi_j(v_i)$.

For $\lambda \in \mathbb{R}_+^2 \setminus \{(0, 0)\}$ define the set

$$\partial V^\lambda = \operatorname{argmax}\{\lambda \cdot v \mid v \in V\} \subset \partial V$$

of all equilibrium payoffs that maximize the weighted sum of payoffs $\lambda_1 v_1 + \lambda_2 v_2$. Any point on the frontier of efficient payoffs belongs to a set ∂V^λ for some λ. The set ∂V^λ may not be a singleton, and so we define $\bar{v}_i^\lambda = \max\{v_i \mid v \in \partial V^\lambda\}$, $i = 1, 2$.

Since $V = \operatorname{co}(V_\zeta)$, all extreme points of V are contained in V_ζ. Moreover, $\partial V^\lambda = \operatorname{co}(\partial V_\zeta^\lambda)$ for all $\lambda \in \mathbb{R}_+^2 \setminus \{(0, 0)\}$. Hence, one can find achieve the extreme points of the Pareto frontier with equilibria that do not use the public randomization device in the first period. Obviously, if ∂V^λ is a singleton $\{v\}$, then v is an extreme point of V and $\partial V_\zeta^\lambda = \{v\}$. In general, $(\bar{v}_1^\lambda, \varphi_2(\bar{v}_1^\lambda)), (\varphi_1(\bar{v}_2^\lambda), \bar{v}_2^\lambda) \in \partial V_\zeta^\lambda$ for all $\lambda \in \mathbb{R}_+^2 \setminus \{(0, 0)\}$. It is possible that $\partial V^\lambda \neq \partial V_\zeta^\lambda$ when ∂V^λ is a line segment, in which case, some $v \in \partial V^\lambda$ can only be attained with the use the public randomization device.

We can characterize the set of efficient equilibria with the following decomposition:

Proposition 3. *Fix $\lambda \in \mathbb{R}_+^2 \setminus \{(0, 0)\}$ and $v^\lambda \in \partial V_\zeta^\lambda$, and assume σ is a pure and simple subgame perfect equilibrium such that $u(\sigma) = v^\lambda$. Let (A, R, w) be σ's factorization. For $i = 1, 2$, define*

$$U_i = \{x \in X \mid -\bar{v}_i^\lambda \leq \beta x_i < 0 \text{ and } \lambda \cdot x > 0\}$$

$$C_i = \{x \in X \mid -\bar{v}_i \leq \beta x_i < -\bar{v}_i^\lambda \text{ and } \beta \lambda_j x_j > \lambda \cdot v^\lambda - \lambda_j \varphi_j(-\beta x_i)\}$$

Then,
(a) $A = U_1 \cup U_2 \cup C_1 \cup C_2$ and $R = X \setminus A$.
(b) $\lambda \cdot w(x) = \lambda \cdot v^\lambda$ for all $x \in R \cup U$.

(c) $w(x) = (-\beta x_i, \varphi_j(-\beta x_i))$ *for all* $x \in C_i$, $i = 1, 2$.
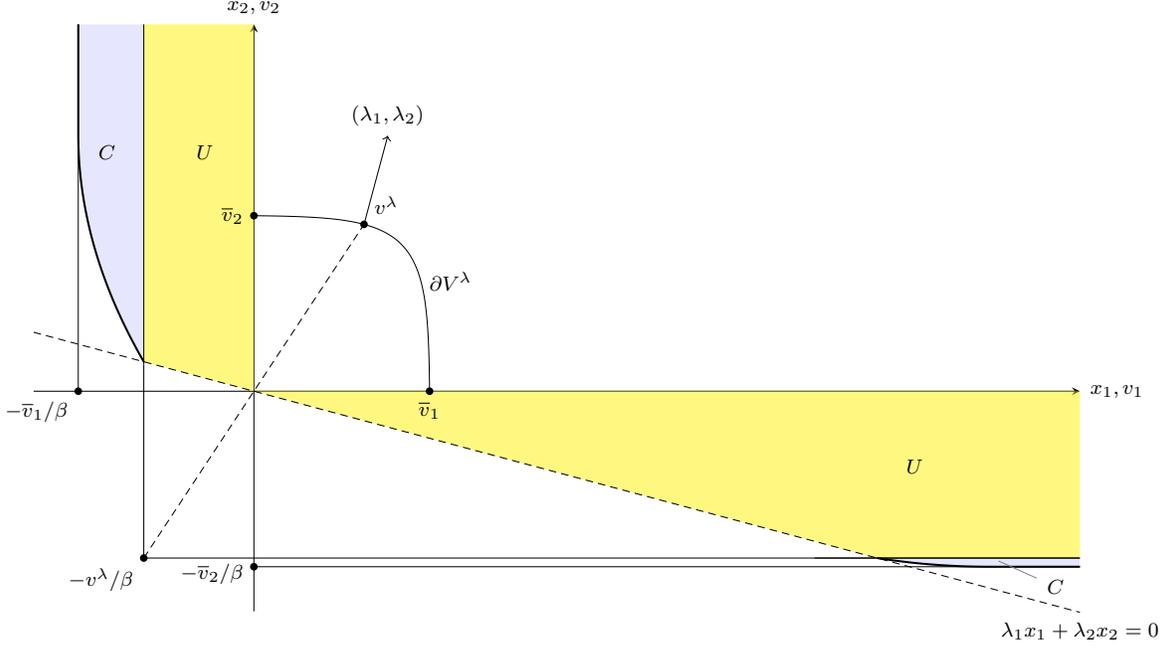

Figure 2: *Acceptance regions without transfers, for an equilibrium that maximizes the weighted sum of equilibrium payoffs $\lambda_1 v_1 + \lambda_2 v_2$. Projects in the unconstrained region U (combining U_1 and U_2) are accepted, with continuation value v^λ that maximizes $\lambda_1 v_1 + \lambda_2 v_2$. Projects in the constrained region C (combining C_1 and C_2) are accepted, but the continuation value is pushed away from v^λ in favor of the player incurring the cost of the project. Projects for which $\lambda_1 x_1 + \lambda_2 x_2 < 0$ are rejected, since they reduce the value of the objective. Projects for which $\beta x_i < -\overline{v}_i$ for some i are rejected, because continuation values are insufficient to motivate i to accept such projects. Some projects for which $\lambda_1 x_1 + \lambda_2 x_2 > 0$ and $\beta x_i > -\overline{v}_i$ are rejected, because the relatively small amount by which they increase the objective is overwhelmed by the loss in value incurred by the required adjustment of continuation values away from v^λ.*

    Figure 2 illustrates this result and Appendix A.3 contains the proof. To interpret this decomposition, recall that we are describing an equilibrium that maximizes (over the set of equilibrium payoffs) the weighted sum $\lambda_1 v_1 + \lambda_2 v_2$. As Figure 2 illustrates, projects for which $\lambda \cdot x < 0$ are rejected (lie outside $A = U_1 \cup U_2 \cup C_1 \cup C_2$) since they reduce the value of the objective $\lambda_1 v_1 + \lambda_2 v_2$. Item (b) indicates that after a project $x \in R$ is rejected, the continuation payoff is drawn from ∂V^λ. There are no incentive constraints in the event of a rejection, and so the optimal continuation payoff maximizes the weighted sum $\lambda \cdot w(x)$.

    A project for which $\lambda \cdot x > 0$ increases the value of the objective function, but it may still be optimal to reject it. A project x should be accepted only if, together with proper continuation payoffs $w(x)$ that ensure incentive compatibility, it would give a higher welfare than rejecting the project and continuing with continuation payoffs v^λ, i.e., only if

$$\lambda \cdot [(1 - \delta)x + \delta w(x)] > \delta \lambda \cdot v^\lambda. \tag{1}$$

The sets U_1 and U_2 ("Unconstrained") identify projects with the property that the incentive constraints can be satisfied with a continuation payoff drawn from ∂V^λ. Item (b) indicates that, in this case, it is optimal to choose a continuation payoff from ∂V^λ. When $w(x) \in \partial V^\lambda$, inequality (1) becomes

$$\lambda_1 x_1 + \lambda_2 x_2 > 0.$$

The sets C_1 and C_2 ("Constrained") identify projects with the property that the incentive constraint for player i to accept the project can be satisfied only if i's continuation payoff exceeds v_i^λ. Item (c) indicates that in this case, it is optimal to set a continuation payoff to exactly satisfy the incentive constraint for player i. In particular, pushing the continuation payoff away from v^λ reduces welfare, and so the strategy calls for the smallest departure from v^λ consistent with satisfying the incentive constraints. We thus have $w(x) = ((-\beta x_i, \varphi_j(-\beta x_i)))$, and accepting the project has a higher welfare than rejecting it only if

$$\beta \lambda_j x_j > \lambda \cdot v^\lambda - \lambda_j \varphi_j(-\beta x_i).$$

The sets U_i and C_i combine to comprise the acceptance set.

The function $\varphi_j(-\beta x_i)$ describing the Pareto frontier is concave and decreasing in $x_i \in [-\bar{v}_i, 0]$, and so $\lambda \cdot v^\lambda - \lambda_j \varphi_j(-\beta x_i)$ is increasing and convex function of x_i. Hence the acceptance set A is convex.

III.7 Interpretation

To interpret these equilibria, we can think of the agents as entering each period with an efficient continuation value $v \in \partial V_\zeta^\lambda$ (after public randomization) for some welfare weights λ. The welfare weights determine the projects that are accepted in the current period. For projects involving a sufficiently small cost (i.e., $x \in U$), the project is accepted and the welfare weights remain unchanged. Incentives are relatively easy to create here, and the threat of autarky suffices, without adjusting continuation payoffs. For projects imposing larger costs, incentives are created by rewarding the person incurring the cost of the project with a larger continuation value. For example, player 2 is induced to approve a particularly costly project by a shift in continuation values in favor of player 2. The corresponding welfare weights associated with the new efficient continuation value give rise to a larger set of projects in the second orthant that player 1 is prescribed to accept next period, and a smaller set of projects in the fourth orthant that player 2 is prescribed to accept next period, pushing continuation payoffs in favor of player 2. Play then continues with this new behavioral prescription, with continuation payoffs unchanged as long as projects are not too costly, and with the next especially costly project shifting the continuation payoffs in the favor the agent incurring the cost.

Projects with small costs thus have no effect on continuation play. The players keep track only of large costs, with each sufficiently costly project prompting a shift of the acceptable projects toward the person bearing the cost, with this shift remaining until the next large cost is incurred.

IV. Transfers

We now modify the previous game by allowing the players to make transfers.

IV.1 The Game

In every period, after a project x has been either adopted or rejected, simultaneously each player i makes a voluntary transfer $m_i \geq 0$ to player j. A transfer m_j from player j to player i increases i's utility by m_j. Player i's total payoff for the period is then $x_i - m_i + m_j$ when the project is accepted, and $-m_i + m_j$ when a project is rejected.

With transfers, the autarky equilibrium σ^A prescribes that after every history both players reject every project and make no transfers. We say that a strategy σ is simple if in any period where the players do not follow its voting recommendation, the players then make no transfers, and after any history that involves a deviation, the players revert to autarky.

We again restrict attention to simple strategies. The factorization of a simple strategy σ is now represented by a tuple (α, m, w), where $\alpha(\zeta, x) \in [0,1]^2$ specifies players' accept/reject decisions, $m = (m^{0,0}, m^{0,1}, m^{1,0}, m^{1,1})$ specifies the transfers and $w = (w^{0,0}, w^{0,1}, w^{1,0}, w^{1,1})$ specifies the continuation values. Hence, for each voting outcome $a \in \{0,1\}^2$, the vector $m^a(\zeta, x) \in \mathbb{R}_+^2$ specifies the pair of transfers after the players choose a, and $w^a(\zeta, x) = u(\sigma|_{(\zeta, x, a, m^a(\zeta, x))})$ is the continuation value after the players choose a and make the proper transfers. Since σ is simple, the factorization only specifies continuation values after the proper transfers $m^a(\zeta, x)$ are made. By definition, when the transfers are different from $m^a(\zeta, x)$, the players revert to autarky and the continuation value is $(0,0)$. Moreover, if $a \in \{0,1\}^2$ is not given positive probability by $\sigma(\zeta, x)$, then $m^a(\zeta, x) = (0,0)$ and $\sigma|_{(\zeta, x, a, \hat{m})} = \sigma^A$ for all $\hat{m} \in \mathbb{R}_+^2$, i.e. transfers are zero and continuation play is given by autarky.

This factorization implicitly assumes that transfers are deterministic. More generally, for an arbitrary behavioral strategy σ, the transfers after (ζ, x, a) could be drawn randomly, in which case $m^a(\zeta, x)$ would be a probability distribution over $\mathbb{R}_+^2$ instead of just a vector in $\mathbb{R}_+^2$. But, as we explain in the following subsection, it will suffice to consider behavioral strategies with deterministic transfers only.

Remark 1 [Complementarity Condition]: Without loss of generality, we can restrict attention to strategies σ that satisfy a *complementarity condition* that in every period and for any history, at most one player makes a positive transfer. In particular, let σ be a subgame perfect equilibrium with factorization (α, m, w). Suppose that for a given (ζ, x, a), $m_i^a(\zeta, x) > 0$ for both $i = 1, 2$. Since each player i only cares about the net transfer $m_j^a(\zeta, x) - m_i^a(\zeta, x)$, we can reduce the players' transfers while keeping constant the net transfers. If $m_j^a(\zeta, x) - m_i^a(\zeta, x) > 0$, for example, then we let $\hat{m}_i^a(\zeta, x) = 0$ and $\hat{m}_j^a(\zeta, x) = m_j^a(\zeta, x) - m_i^a(\zeta, x)$. Reducing the transfers relaxes incentive constraints and hence $(a, \hat{m}, w)$ also represents a subgame perfect equilibrium. Hereafter we will

assume that every subgame perfect equilibrium satisfies the complementarity condition $m_1^a(\zeta, x) \cdot m_2^a(\zeta, x) = 0$ for all (ζ, x, a).

IV.2 Pure Strategies Suffice

Once again, it sacrifices no generality to restrict attention to pure strategies. An argument paralleling that of Proposition 2 (presented in Appendix A.4) establishes:

Proposition 4. *With transfers, $V_P = V$, i.e., any subgame perfect equilibrium payoff is the payoff of a pure subgame perfect equilibrium.*

For a simple pure strategy σ, we adopt again the simpler factorization representation (A, R, m, w), where $A, R \subset X$ are respectively the sets of projects accepted or rejected in the first period, and m and w specify the transfers and continuation values (on the outcome path) in the first period.

IV.3 The Importance of Transfers

We illustrate the potential importance of transfers in this section. We illustrate how allowing transfers can enlarge the set of equilibrium payoffs, while also showing how the resulting set of payoffs can fall short of an intuitive upper bound. In addition to serving as a comparison for our examination of the effect of transfers, this upper bound may be useful as a point of departure for an iterative self-generation algorithm designed to calculate the set of equilibrium payoffs.

IV.3.1 An Upper Bound

We first identify an upper bound on the set V of equilibrium payoffs for the game *with* or *without* transfers. Our argument is reminiscent of that developed by Fudenberg and Levine (1994), in that it identifies a half-plane from which continuation payoffs must be drawn, and then asks what equilibrium payoffs can be constructed if we allow *any* continuation payoffs drawn from this half-plane.

Let $T^* = \max \{v_1 + v_2 \mid v \in V\}$, where V is the relevant equilibrium value set (again, for the game with or without transfers). Then it is obvious (though not yet obviously helpful, given the self-reference) that $V \subset V_{T^*}$, where for any $T > 0$, we define

$$V_T = \{(v_1, v_2) \in \mathbb{R}_+^2 \mid v_1 + v_2 \leq T\}.$$

Hence, $V \subset V_{T^*}$ is simply the statement that no equilibrium can give a sum of payoffs larger than the maximum sum of equilibrium payoffs.

We next argue that, for the game with or without transfers, the set of projects A that are accepted in equilibrium (in any period) is bounded by $A \subset A_{T^*}$, where for any $T > 0$ we define

$$A_T = \{x \in X \mid x_1 + x_2 \geq 0 \text{ and } \beta x_i \geq -T \text{ for } i = 1, 2\}.$$

The set A_{T^*} is the set of projects whose cost does not exceed the maximum (discounted) sum of equilibrium payoffs. To see that $A \subset A_{T^*}$, let (A, R, w) be the factorization of a subgame perfect equilibrium for the game *without* transfers. Let $x \in A$ be such that $x_i < 0$. Then $(1 - \delta)x_i + \delta w_i(x) \geq 0$. Since $w(x) \in V \subset V_{T^*}$, we have $w_i(x) \leq T^*$ and hence $\beta x_i \geq -w_i(x) \geq -T^*$. Analogously, let (A, R, m, w) be the factorization of a subgame perfect equilibrium for the game *with* transfers. If $x \in A$ is such that $x_i < 0$, then $(1 - \delta)(x_i + m_j(x)) + \delta w_i(x) \geq 0$. Also $-(1 - \delta)m_j(x) + \delta w_j(x) \geq 0$. Adding these two inequalities, we conclude that $(1 - \delta)x_i + \delta(w_1(x) + w_2(x)) \geq 0$, which again implies that $\beta x_i \geq -T^*$.

To construct our upper bound, we now define the function

$$S(T) = \int_{A_T} (x_1 + x_2)d\mu(x).$$

Then $S(T)$ is the expected payoff of accepting all projects in $A(T)$. If T is a fixed point of S, then accepting the projects in the set A_T gives a total expected payoff of T. Clearly S is a weakly increasing function, $S(0) = 0$, $S(\infty)$ is finite, and so S must have a largest fixed point, $\bar{T}$, and this gives us our upper bound:

Lemma 2. *With our without transfers, the set of equilibrium payoffs V is a subset of $V_{\bar{T}}$.*

Proof: As we have noted above, it is immediate from the definition that $V \subset V_{T^*}$. Next, A_{T^*} is an upper bound on the set of profitable projects that are accepted in any period and in any equilibrium, and hence, we have $S(T^*) \geq T^*$, i.e., accepting all of the projects in A_{T^*} must give a sum of payoffs at least as large as the largest equilibrium sum of payoffs. But the inequality $S(T^*) \geq T^*$ implies that the largest fixed point $\bar{T}$ of S is larger than T^*. We thus have $\bar{T} \geq T^*$, which in turn gives $V \subset V_{T^*} \subset V_{\bar{T}}$. $\qquad \square$

The upper bound on the sum of equilibrium payoffs is thus the largest sum $\bar{T}$ that can be generated by having the players accept all projects with discounted costs no greater than $\bar{T}$.

IV.3.2 Achieving the Upper Bound

In this section, we show that if the distribution of projects is symmetric, then the upper bound $\bar{T}$ on the sum of equilibrium payoffs can be achieved.

The distribution μ is symmetric if, for any $B \subset X$, we have $\mu(B) = \mu(B^T)$, where

$$B^T = \{(x_1, x_2)|(x_2, x_1) \in B\}.$$

If μ is symmetric and transfers can be made, then the payoff provile $(\bar{T}/2, \bar{T}/2)$ is an equilibrium payoff. The equilibrium strategy is to accept any project that increases the joint surplus and whose cost is no greater than $\bar{T}/\beta$, with continuation values always given by $(\bar{T}/2, \bar{T}/2)$. Notice that this would not be part of an equilibrium strategy if there were no transfers, since then we could not support a project whose current cost is $\bar{T}/\beta$ with a continuation value of $(\bar{T}/2, \bar{T}/2)$. However, with transfers, one player make a transfer of up to $\bar{T}/2$ to the other player. Hence, for any project whose cost is no greater than $\bar{T}/\beta$, we can effectively split the cost between the two players. Appendix A.5 proves:

Lemma 3. *If μ is symmetric and transfers can be made, then $(\bar{T}/2, \bar{T}/2) \in V$ and $T^* = \bar{T}$.*

One might conjecture that under the conditions of Lemma 3, the set of equilibrium payoffs is $V_{\bar{T}} = \{(v_1, v_2) \in \mathbb{R}_+^2 | v_1 + v_2 \leq \bar{T}\}$, because transfers will allow the total surplus of $\bar{T}$ to be divided arbitrarily among the two players. The following example shows that instead, pushing the payoff profile away from $(\bar{T}/2, \bar{T}/2)$ can have efficiency costs.

IV.3.2 An Example

This example shows that the Pareto frontier with transfers can lie strictly above that without, and that the set of equilibrium payoffs with transfers can be a strict subset of $V_{\bar{T}}$. We retain the symmetry of μ in this example, so $(\bar{T}/2, \bar{T}/2)$ remains an equilibrium payoff.

Let $\delta = (1 - \delta) = 1/2$ and let μ attach probability $1/2$ to project $(-1, 2)$ and probability $1/2$ to project $(2, -1)$. Here $\bar{T} = 1$ and

$$V_{\bar{T}} = \{v \in \mathbb{R}_+^2 | v_1 + v_2 \leq 1\}.$$

Lemma 3 immediately implies that there exists an equilibrium in the game with transfers with payoffs $(1/2, 1/2)$.

We first argue that without transfers, $V = \{(0, 0)\}$. Hence, without transfers, there is no equilibrium in which any project is accepted, leaving only the autarky equilibrium. Coupled with the previous observation, this implies that equilibrium value set of the relationship with transfers is strictly larger than the equilibrium value set of the relationship without transfers. The argument has two steps.

Our first step is to show that $(1, 0) \notin V$ (and by symmetry, $(0, 1) \notin V$). Assume σ is a subgame perfect equilibrium that attains the expected value $(1, 0)$. Then σ must reject the project $x' = (-1, 2)$ since accepting this project assures player 2 a positive payoff. Similarly, project $x = (2, -1)$ must be accepted, since otherwise no subsequent pattern of acceptances can generate a payoff for player 1 equal to 1. Constrained to continuation values from $V_{\bar{T}}$, project x is accepted in period 0 by player 2 if and only if this is followed with the continuation value $w(x) = (0, 1)$. We also need to choose $w(x')$ so that we achieve the target equilibrium payoffs:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2} \left[(1 - \delta) \begin{pmatrix} 2 \\ -1 \end{pmatrix} + \delta \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] + \frac{1}{2} \delta w(x').$$

This implies that $w(x') = (2, 0)$. But $(2, 0) \notin V_{\bar{T}}$, so $(2, 0) \notin V$, and hence $(1, 0) \notin V$ (since, by Lemma 2, $V \subset V_{\bar{T}}$).

The second step returns to the (just noted) fact that in equilibrium, project x can be accepted only if the players are offered the continuation value $w(x) = (0, 1)$. Since $(0, 1)$ is not available, project x cannot be accepted by *any* subgame perfect equilibrium. By symmetry, project x' cannot be accepted either. Hence, autarky is the unique equilibrium of the relationship model without transfers and $V = \{(0, 0)\}$.

We next argue that the equilibrium value set with transfers, while strictly larger than the equilibrium value set without transfers, is a strict subset of $V_{\bar{T}}$. Suppose we seek the highest equilibrium payoff for player 2. We consider two possibilities. First, suppose that the project $(2, -1)$ is drawn in the first period. Accepting gives player 2 a payoff of $(1/2)(-1 + m_1 + w_2)$, where player 1's incentive constraint requires $-m_1 + w_1 \geq 0$ and feasibility requires $w_1 + w_2 \leq 1$. These last two inequalities imply that $m_1 + w_2 \leq 1$, which in turn implies that the maximum payoff for player 2, conditional on $(2, -1)$ being drawn and accepted, is zero. If this project is rejected, then player 1 can make a transfer m_1 to player 2, where incentive compatibility requires $-m_1 + w_1 \geq 0$. Player 2's maximum payoff if $(2, -1)$ is rejected is then $w_1 + w_2$, which is at most 1. Hence, project $(2, -1)$ must be rejected.
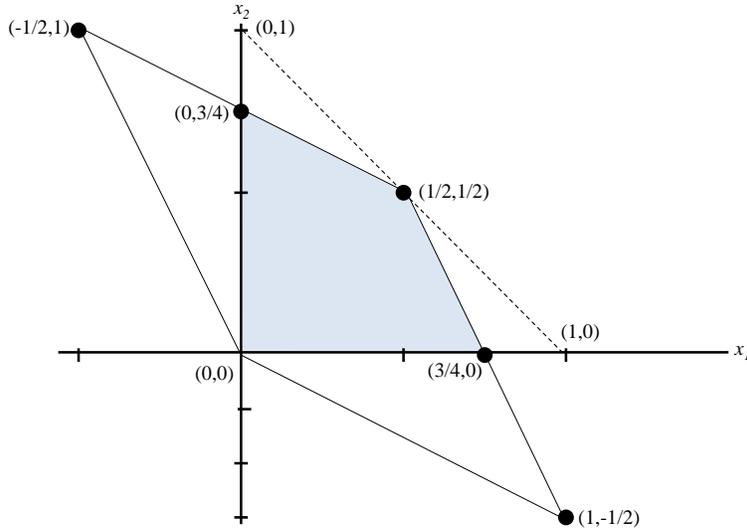

Figure 3: *Illustration of Example IV.3.2. The set of feasible payoffs in the relationship without transfers is the convex hull of $\{(0,0), (-1/2, 1), (1, -1/2), (1/2, 1/2)\}$ (solid lines). The set $V_{\bar{T}}$ is given by the convex hull of $\{(0,0), (0,1), (1,0)\}$ (bounded by the dotted line and the axes). Without transfers, the set of equilibrium payoffs is the singleton $\{(0,0)\}$, corresponding to the autarky equilibrium—no projects are accepted. With transfers, the set of equilibrium payoffs includes the convex hull of $\{(0,0), (3/4, 0), (0, 3/4), (1/2, 1/2)\}$ (shaded region). The largest equilibrium payoff for either player is $3/4$, ensuring that the set of equilibria with transfers is a strict subset of $V_{\bar{T}}$.*

Obviously, project $(-1, 2)$ must be accepted. In this case, the maximum payoff for player 2 is then $2 - m_2 + w_2$. Incentive compatibility for player 1 requires that $-1 + m_2 + w_1 \geq 0$, which (using $w_1 + w_2 \leq 1$) implies that 2's payoff in this case is maximized by setting $m_2 = w_2$, and hence is given by 2.

Putting these results together, the largest equilibrium payoff for player 2 is given by $(1/4)2 + (1/4)1 = 3/4$. This ensures that V is a strict subset of $V_{\bar{T}}$.

The analysis in the preceding paragraph in fact identifies a subgame perfect equilibrium that delivers the equilibrium value $v = (0, 3/4)$. The strategy specifies that in the first period, if project $(2, -1)$ arises, it is rejected, player 1 transfers $1/2$ to player 2, and continuation payoffs are $(1/2, 1/2)$. If project $x(-1, 2)$ arises, it is accepted, player 2 transfers $1/2$ to player 1, and continuation payoffs are $(1/2, 1/2)$.

By symmetry, $(3/4, 0) \in V$ as well. We have thus established that the set of equilibrium payoffs V contains at least the convex hull of the equilibrium payoffs $(0, 0)$, $(0, 3/4)$, $(3/4, 0)$ and $(1/2, 1/2)$. Proposition 5 ensures that it can contain no other points. Therefore, as figure 3 illustrates,

$$V = \text{co} \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3/4 \end{pmatrix}, \begin{pmatrix} 3/4 \\ 0 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \right\}.$$

IV.4 The Pareto Frontier

We now characterize the Pareto frontier of the game with transfers. Let w^* be a vector of continuation payoffs that achieve the (maximal) sum of equilibrium payoffs T^*.

Fix a project x. If the project x is to be accepted, followed by transfers $m(x)$, and continuation payoffs $w(x)$, then the following incentive constraints must hold:

$$(1 - \delta)[x_1 - n] + \delta w_1 \geq 0$$
$$-(1 - \delta)m_1 + \delta w_1 \geq 0$$
$$(1 - \delta)[x_2 + n] + \delta w_2 \geq 0$$
$$-(1 - \delta)m_2 + \delta w_2 \geq 0,$$

where $n = m_1 - m_2$ is the *net* transfer from player 1 to player 2 (cf. Remark 1), and the argument x in transfers and continuation values has been omitted. The first and third constraints ensure that the project is undertaken, and the second and fourth ensure that transfers are made.

Each point on the Pareto frontier maximizes (over the set of equilibrium payoffs) a weighted sum of the players' welfare. For a given $\lambda \in \mathbb{R}^2_+ \setminus \{(0, 0)\}$, the welfare to be maximized is

$$\lambda \cdot [(1 - \delta)x + \delta w] + (1 - \delta)(\lambda_2 - \lambda_1)n. \tag{2}$$

We assume for this presentation that $\lambda_2 > \lambda_1$ (the case $\lambda_2 < \lambda_1$ is analogous), so that the welfare function puts more weight on 2's payoffs, and characterize the frontier for such values. Notice that in this case, any transfer from player 1 to player 2 increases welfare, and everything else being equal, such transfers should be maximized.

We further divide the analysis into two cases. We first consider projects in which $x_2 < 0$, so that player 2 bears the cost of the project. In this case, since $\lambda_2 > \lambda_1$, it is optimal to set $m_2 = 0$ and to maximize m_1. Hence, the second incentive constraint binds and

$$(1 - \delta)m_1 = \delta w_1.$$

Since $x_1 \geq 0$ and $m_2 = 0$, the first and last constraints are slack, and after we substitute the optimal value of n $(= m_1)$, the third constraint becomes

$$(1 - \delta)x_2 + \delta(w_1 + w_2) \geq 0.$$

This constraint may be impossible to satisfy if $|x_2|$ is too large. Finally, substituting the (binding) second constraint into (2), the welfare to be maximized becomes

$$(1 - \delta)\lambda \cdot x + \delta\lambda_2[w_1 + w_2]. \tag{3}$$

If the project is rejected, then the welfare to be maximized (from (2)) is

$$(1 - \delta)(\lambda_2 - \lambda_1)n + \delta\lambda \cdot w.$$

In this case, we optimally set $m_2 = 0$ and maximize m_1. The constraint on m_1 is that $-(1 - \delta)m_1 + \delta w_1 \geq 0$, so $(1 - \delta)m_1 = \delta w_1$ and the welfare upon rejecting x is

$$\delta\lambda_2(w_1 + w_2). \tag{4}$$

It is clear from (3)–(4) that whether the project is accepted or rejected, the continuation value should be chosen so as to maximize $w_1 + w_2$. The project should then be accepted if and only if $\lambda \cdot x > 0$ and $\beta x_2 \geq -T^*$. Notice that player 1's payoff in this case is given by x_1, since the binding incentive constraint $(1 - \delta)m_1 = \delta w_1$ ensures that 1's payoff in the continuation game starting with the period-1 transfer is 0.

Let us now consider the case in which $\lambda_2 > \lambda_1$, as before, but $x_1 < 0$. Transfers now serve a dual role. Transfers from 1 to 2 again increase welfare, but now transfers from 2 to 1 may be necessary to create the incentives for 1 to endorse the project.

Because $\lambda_2 > \lambda_1$, if the project x is accepted, at an optimum the first incentive constraint is satisfied with equality, so

$$(1 - \delta)n = (1 - \delta)x_1 + \delta w_1,$$

and since $x_1 < 0$, the second constraint is slack (if $m_1 = 0$ the second constraint is clearly satisfied, and if $m_1 > 0$ then $n = m_1$). If $m_2 = 0$ the last two constraints are satisfied, and if $m_2 > 0$ then $m_2 = -n$ and the last constraint is more demanding than the third (because $x_2 \geq 0$) and can be written as

$$(1 - \delta)x_1 + \delta(w_1 + w_2) \geq 0.$$

Again, substituting the constraints into (2), we see that whether the project is accepted or rejected, $w_1 + w_2$ should be maximized, and the project should be accepted if and only if $x_1 + x_2 > 0$ and $\beta x_1 \geq -T^*$. In this case, player 1's payoff is 0.
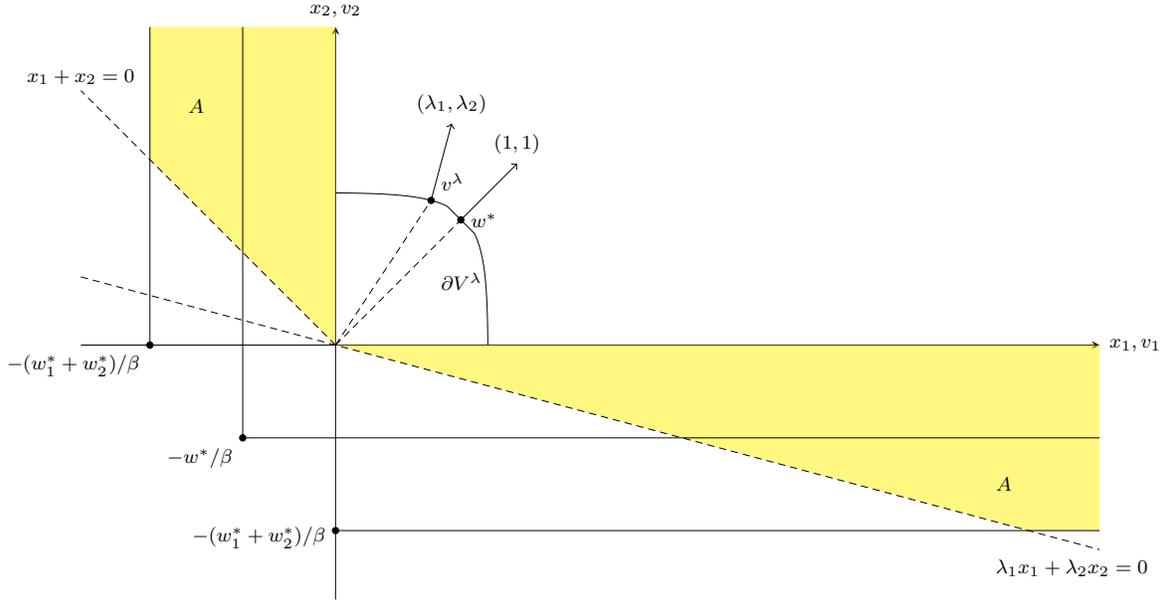
Figure 4: *Acceptance regions with transfers, for an equilibrium that maximizes the weighted sum of equilibrium payoffs $\lambda_1 v_1 + \lambda_2 v_2$, with $\lambda_2 > \lambda_1$. Regardless of the project accepted, continuation values maximize the unweighted sum $w_1 + w_2$. Projects are accepted only if the cost is less than the maximized (and normalized) value $(w_1^* + w_2^*)/\beta$ of this unweighted sum. Within this constraint, projects with $x_2 < 0$ are accepted if and only if $\lambda_1 x_1 + \lambda_2 x_2 > 0$, i.e., if and only if they increase the value of the objective. Projects with $x_1 > 0$ are less valuable because they require directing surplus to player 1 in order to motivate him to accept the project, whereas the objective favors transferring surplus to player 2 (since $\lambda_2 > \lambda_1$), and these projects are accepted only if $x_1 + x_2 > 0$.*

Notice that there is an asymmetry here between the case $x_1 < 0$ and $x_2 < 0$. The asymmetry concerns the criterion for accepting the current project. When $x_2 < 0$, current projects are accepted if and only if they are incentive compatible and increase the value $\lambda_1 x_1 + \lambda_2 x_2$. When $x_1 < 0$, then current projects are accepted if and only if they are incentive compatible and increase the sum $x_1 + x_2$. Why the difference? Player 2 gets a larger weight in the welfare criterion, and there are two ways of increasing welfare. One is to make payments from player 1 to player 2, and the other is to accept current projects. When $x_2 < 0$, accepting a current project places no restrictions on the amount of money 1 can transfer to 2, and so the criterion is to accept if and only if the current project increases the objective $\lambda_1 x_1 + \lambda_2 x_2$. When $x_1 < 0$, accepting the current project reduces the net transfer that 1 can make to 2. As a result, the current project is accepted only if it increases the sum $x_1 + x_2$.

Using the fact that a similar analysis applies to the case $\lambda_1 < \lambda_2$, we have established:

Proposition 5. *Fix $\lambda \in \mathbb{R}_+^2 \setminus \{(0,0\}$ with $\lambda_2 > \lambda_1$ (with the reverse inequality analogous). Assume σ is a subgame perfect equilibrium such that $u(\sigma) = v^\lambda$. Let (A, R, w) be σ's factorization. Then,*

(a) $w(x) = w^*$ *for all* $x \in X$.

(b) $R = X \backslash A$, *where* A *consists of those* $x \in X$ *satisfying*

$$(1-\delta)x_i + \delta T^* \geq 0 \quad i = 1, 2, \quad and \quad \begin{cases} \lambda_1 x_1 + \lambda_2 x_2 \geq 0 & if \ x_2 < 0 \\ x_1 + x_2 \geq 0 & if \ x_1 < 0. \end{cases}$$

Figure 4 illustrates this characterization of equilibrium strategies.

Once the first project is accepted in this equilibrium, the continuation equilibrium maximizes the sum of payoffs, and continues to do so after every subsequent history. Hence, there is no memory of past projects. Any asymmetry in payoffs is accomplished by adjusting the acceptance region for the initial project and the transfers accompanying that project. Subsequently, the acceptance regions are symmetric and transfers are just enough to create the current incentives to undertake projects.

We can illustrate this result by returning to the example illustrated in Figure 3. Suppose $(\lambda_1, \lambda_2) = (1, 2)$. The line $\lambda_1 x_1 + \lambda_2 x_2 = 0$ includes $(0,0)$ and the project $(2,-1)$. Maximizing the sum $\lambda_1 w_1 + \lambda_2 w_2$ thus calls for project $(-1,2)$ to be accepted in the first period, and allows project $(2,-1)$ to be either accepted or rejected. Any weighted average that puts more weight on 2's payoff would accept only the project $(-1,2)$ in the first period, while any weighted sum that places lower weight on player 2 (but still favors player 2) must accept both projects. In any of these cases, continuation payoffs are $(1/2, 1/2)$. Notice that, because the sum $x_1 + 2x_2$ is maximized both by an equilibrium that accepts both projects in the first period (giving payoffs $(1/2, 1/2)$) and an equilibrium accepting only $(-1,2)$ (giving payoffs $(0, 3/4)$), there can be no equilibrium payoffs above the shaded area in Figure 3, as indicated in Section IV.3.2.

An interesting aspect of Proposition 5 is that there are multiple equilibria consistent with the equilibrium payoffs. When projects arise so large that they cannot otherwise be carried out, then transfers must be made. These are projects that cost player i more than w_i^*/β (falling in the lower or leftward part of the A region in Figure 4). Transfers can also be made after small projects (that cost less than w_i^*/β)), but need not be made, as long as the transfer conventions are arranged so as to have the same expected payoffs. We could thus have equilibria in which transfers are always made, but there also equilibria in which transfers only rarely undertaken. This freedom in making transfers gives rise to the flat segment on the payoff frontier in Figure 4. Different points along this flat segment correspond to equilibria with the same accept/reject decisions, but different transfer conventions. In one of these equilibria, transfers are made only in the event of a project so large that it otherwise would not be carried out.

IV.5 Comparison

Without transfers, projects of moderate size are accepted without any effect on continuation values, while larger projects are accepted only at the cost of pushing continuation values away from the maximizer of the relevant weighted sum of continuation payoffs. In

the presence of transfers, we focus on equilibria in which projects of moderate size are analogously accepted without transfers, while larger projects requires transfers. What is the difference in the two situations, other than that "transfers" has been substituted for "continuation values?"

In the absence of transfers, every variation in continuation payoffs required to create incentives entails a sacrifice in the welfare value of continuation payoffs. In contrast, transfers involve no such sacrifice. As a result, relationships with transfers can typically support higher payoffs than those that do not (as in Example IV.3.2). However, this very ability to support larger payoffs pushes outward the boundary between "moderate" projects, that can be supported without transfers, and "large" projects that requires transfers. The effectiveness of transfers in creating incentives ensures that we can expect to see less evidence of these very incentives. Instead, it will more often be the case that the ability to continue the relationship unaltered will suffice for the required incentives.

IV.6 Interpretation

The equilibrium described in Proposition 5 captures the basic features of the even-up strategy examined by Ellickson (1991) and summarized in Section I.1. First, under ordinary circumstances, no transfers are made and the continuation payoffs that lie along the equilibrium path do not respond to current actions. Actions and continuation payoffs are chosen to maximize the sum of the players' payoffs. This behavior continues until the relationship is thrown "out of balance" by the appearance of a project that is particularly costly for one player. Such a project pushes the burden imposed on that player beyond the bounds that can be sustained without recourse to either a transfer or an adjustment in continuation payoffs. Depending on the circumstances, this may entail a player who incurs exceptional cost in granting a favor to another, or it may involve a player who captures a benefit only by imposing a particulary large cost on the other. The equilibrium then gives us the second even-up characteristic, namely a transfer that (perhaps only partly) compensates the player incurring the cost. We then continue, with no further transfers and with continuation play continuing to maximize the sum of payoffs, until the next particularly costly project. Third, the incentives that lie behind these equilibrium actions arise out of the fact that, should one party fail in their equilibrium action, whether by failing to undertake a project or by failing to make a transfer, then the relationship enters a punishment phase. We have simplified the equilibrium analysis by considering simple strategies, exhibiting the harshest punishments possible. Ellickson (1991) notes that punishments are typically measured, being only as severe as called for by the relevant transaction, and typically are temporary. As is familiar from work in repeated games, our harsh punishments are a convenience only, and could easily be replaced by measured and temporary punishments.

V. Extensions

We sketch here some of the many possible extensions of this model.

V.1 Imperfectly Transferable Utility

Let $f : \mathbb{R} \to \mathbb{R}$ be a strictly decreasing, concave function with $f(0) = 0$. The interpretation is that if m is a transfer from player 2 to player 1, then $f(m)$ is the cost of this transfer to player 2. The function f takes the entire real line as its domain. If $m < 0$, money is flowing from 1 to 2 with benefit to 2 of $f(m) > 0$. The previous sections examined the case of perfectly transferable utility, or $f(m) = -m$. When f is strictly concave, we have imperfectly transferable utility.
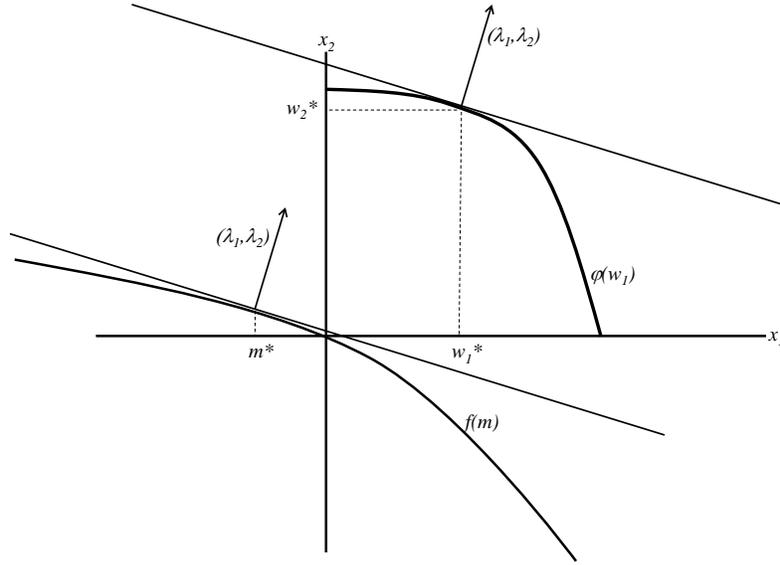

Figure 5: *Configuration (m^*, w_1^*, w_2^*) that maximizes the expected value $\lambda_1[(1 - \delta)m + \delta w_1] + \lambda_2[(1 - \delta)f(m) + \delta w_2]$.*

To characterize the efficient frontier, we consider welfare weights $\lambda \in \mathbb{R}_+^2 \setminus \{(0,0)\}$. Suppose that a transfer m and continuation payoffs (w_1, w_2) are to be chosen so as to maximize the weighted sum

$$\lambda_1[(1 - \delta)m + \delta w_1] + \lambda_2[(1 - \delta)f(m) + \delta w_2].$$

The solution satisfies (unless no such m^* exists, in which case we set $\varphi_2'(w_1^*) = -\lambda_1/\lambda_2$ and $m^* = -\infty$)

$$f'(m^*) = \varphi_2'(w_1^*) = -\lambda_1/\lambda_2.$$

Figure 5 illustrates this solution. The transfers $(m^*, f(m^*))$ and the continuations payoff $(w_1^*, \varphi_2(w_2))$ are chosen from the curve f and the frontier φ_2 at points where the slopes equal $-\lambda_1/\lambda_2$, maximizing (respectively) $\lambda_1 m + \lambda_2 f(m)$ and $\lambda_1 w_1 + \lambda_2 w_2$.

To interpret this result, suppose that $m^* < 0$, so that a transfer is made from player 1 to player 2. This will be the case if $f'(0) < -\lambda_1/\lambda_2$, as illustrated in Figure 5. Then the transfer is set so that the player-1 and player-2 payoffs trade off at the same rate via transfers as they do via continuation payoffs.

Now, continuing with our choice of λ, consider a subgame perfect equilibrium σ such that $u(\sigma) \in \partial V^\lambda$. Suppose a project (x_1, x_2) has been drawn and equilibrium actions are taken, and we must determine the corresponding transfer and continuation values (m, w_1, w_2) provided by σ. These values will be given by (m^*, w_1^*, w_2^*), as illustrated in Figure 5, unless a constraint intervenes. How could a constraint intervene? We consider two possibilities. The first includes the following related scenarios:

- Project (x_1, x_2) is (in equilibrium) rejected and $m^* + w_1^* < 0$.
- Project (x_1, x_2) is (in equilibrium) accepted, $x_2 < 0$, and $m^* + w_1^* < 0$.
- Project (x_1, x_2) is (in equilibrium) accepted, $x_1 < 0$, and $x_1 + m^* + w_1^* < 0$.

In these scenarios, the incentives for player 1 to either make a transfer as large as $|m^*|$ or to accept x_1 if $x_1 < 0$ are insufficient. In equilibrium, the transfer from player 1 to player 2 is necessarily smaller than $|m^*|$, that is $m > m^*$, and accordingly, the continuation value w_1 satisfies (as long as w_1 remains interior) $\varphi_2'(w_1) = f'(m) < -\lambda_1/\lambda_2$, so $w_1 > w_1^*$ and $w_2 = \varphi_2(w_1) < w_2^*$. Alternatively, we may have:

- Project (x_1, x_2) is (in equilibrium) accepted, $x_2 < 0$, and $x_2 + f(m^*) + w_2^* < 0$.

In this case, the configuration (m^*, w_1^*, w_2^*) provides inadequate incentives for player 2 to accept the project. To strengthen these incentives, the equilibrium increases the transfer from player 1 to player 2 and decreases the continuation value for player 1 (increases the continuation value for player 2) until $x_2 + f(m) + \varphi_2(w_1) = 0$ and (as long as $\varphi_2(w_1)$ remains interior) $f'(m) = \varphi_2'(w_1) \ (> -\lambda_1/\lambda_2)$. Note that under either possibility, barring corner solutions the equality $f'(m) = \varphi_2'(w_1)$ must hold.

The relevant observation is that in general, every period will exhibit a transfer. Fix a pair of welfare weights (λ_1, λ_2). If this pair of weights does not satisfy $f'(0) = -\lambda_1/\lambda_2$, then a transfer occurs, and indeed we can expect (barring knife-edge cases) a transfer in every subsequent period. Hence, perhaps paradoxically, under imperfectly transferrable utility, transfers are used pervasively, despite being relatively inefficient in transferring utility from one player to another. Under perfectly transferable utility, transfers are used more rarely. Behind this difference lies the fact that under imperfectly transferrable utility, the utility exchange rate between the two players deteriorates as the size of the transfer grows (because f is strictly concave). There is then an incentive to break any transfer into a succession of smaller transfers in subsequent periods. The players will then constantly make transfers, in order to avoid any single especially large transfer. No such force arises under perfectly transferrable utility. Hence, an observations of a relationship in which transfers are rare is an indication that utility is reasonably modeled as transferable in that relationship.

V.2 Imperfect Information

We now suppose that there is uncertainty concerning the measure μ that governs the distribution of projects. Players 1 and 2 begin the game with a common prior over the measure μ, and update their beliefs as they observe the series of projects drawn by Nature.

Their beliefs remain symmetric throughout. We confine our analysis to an example, and keep things simple by assuming there are no transfers.

V.2.1 An Example

Suppose there are four possible projects, given by $(-1, 10)$, $(-6, 22)$, $(10, -1)$ and $(22, -6)$. There is an underlying state of the world that is either high, in which case the probability distribution over projects is

$$f((-1, 10)) = f((10, -1)) = \frac{1}{2}(1 - p)$$
$$f((-6, 22)) = f((22, -6)) = \frac{p}{2}$$

or low, in which case the distribution is

$$f((-1, 10)) = f((10, -1)) = \frac{p}{2}$$
$$f((-6, 22)) = f((22, -6)) = \frac{1}{2}(1 - p).$$

To match the names given to the states, we assume $p > 1/2$. Hence, the outcomes $(-1, 10)$ and $(10, -1)$ are equally likely in either state, as are the outcomes $(-6, 22)$ and $(22, -6)$, but the latter are more likely in the high state. Let us refer to the former as the low pair and the latter as the high pair. For our calculations, we take $p = 3/4$. To keep the calculations simple, we assume $\delta = 1/2$.

Let q_0 be the prior probability that the state is high. Then the posterior probability that the state is high depends only on the number of projects that have been drawn from the low pair and the number that have been drawn from the high pair. If the difference in the number of draws from the high pair and the low pair is given by k, then we denote the posterior probability of the high state by q_k, where

$$q_k = \frac{q_0}{q_0 + (1 - q_0)((1 - p)/p)^k}$$

We will refer to k as the state of the relationship.

There is obviously an autarky equilibrium in which no projects are accepted. There is also an equilibrium outcome in which the low pair of projects is always accepted, and the high pair always rejected. To confirm that this is an equilibrium outcome, we need to show that for any state k,

$$(1 - \delta) \leq \delta \underline{V}(k - 1),$$

where the left side is the cost that one of the players pays for accepting a low project, and we define $\underline{V}(k - 1)$ to be the continuation value of an equilibrium in which (only) low projects are accepted, given that the state is $k - 1$. The continuation value $\underline{V}(k - 1)$ is

lower the higher is k (and hence q_k), indicating that in this equilibrium it is bad news to find that the state is high, since low projects are then less plentiful. To confirm that we have an equilibrium, it then suffices to verify that this inequality holds in the limit as k becomes arbitrarily large. As $k \to \infty$ and hence $q_{k-1} \to 1$, $\underline{V}_{k-1}$ approaches $(1-p)(9/2)$, and (recalling that $1 - \delta = \delta = 1/2$ and $p = 3/4$) the desired inequality holds.

We now note that it is *not* an equilibrium outcome for players to accept every project in every state. If this were to be an equilibrium outcome, we need to show that for any state k,

$$(1 - \delta)6 \le \delta \overline{V}(k + 1),$$

where the left side is the cost that one of the players pays for accepting a low project, and we define $\overline{V}(k + 1)$ to be the continuation value of an equilibrium in which all projects are accepted, given that the state is $k + 1$. The continuation value $\overline{V}(k + 1)$ is lower the lower is $k + 1$ (and hence q_{k+1}), since high projects are then less plentiful. To confirm that this is not an equilibrium, it suffices to note that As $k \to -\infty$, $\overline{V}_{k+1}$ approaches $p(9/2) + (1-p)(16/2)$, and hence (again recalling that $1 - \delta = \delta = 1/2$ and $p = 3/4$) the required inequality fails.

We now argue that there exists an equilibrium featuring a value k^* with the property that only low projects are accepted for values $k < k^*$, while all projects are accepted for values $k \ge k^*$. Indeed, there are many such equilibria, with varying values of k^*. However, the fact that it is not an equilibrium to always accept all projects ensures that there is a lower bound of the value of k^*.

Let $V(k, k^*)$ be the value of this equilibrium in state k. Then, for $k \ge k^*$,

$$V(k,k^*)=\frac{1}{2}\Big[[q_k(1-p)+(1-q_k)p]\big[\frac{9}{2}+V(k-1,k^*)\big]+[q_kp+(1-q_k)(1-p)]\big[\frac{16}{2}+V(k+1,k^*)\big]\Big].$$

It is clear that q_k is increasing in k and that $V(k, k^*)$ is increasing in k for $k \ge k^*$. Moreover, $V(k, k^*) \ge 9/2$ for all k. It is immediate that in the proposed equilibrium the players prefer to accept low projects, and to reject high projects for values $k < k^*$. We need only verify that the players prefer to accept high projects in every state $k \ge k^*$. That is,

$$(1 - \delta)6 \le \delta V(k + 1, k^*) \quad \text{for all } k \ge k^*.$$

Since $V(k, k^*)$ is increasing in k, it suffices to verify that $6 \le V(k^* + 1, k^*)$. We now show that if k^* is chosen large enough, this is the case.

Since $V(k, k^*) \ge 9/2$ and $V(k+1, k^*) > V(k, k^*)$ for $k \ge k^*$, we have that for all $k \ge k^*$

$$V(k,k^*) > [q_k(1 - p) + (1 - q_k)p]\frac{9}{2} + \frac{1}{2}[q_kp + (1 - q_k)(1 - p)][\frac{16}{2} + V(k,k^*)].$$

This implies that

$$V(k,k^*) > \frac{35 - 2q_k}{7 - 2q_k} \quad \text{for all } k \ge k^*,$$

and hence $V(k^* + 1, k^*) \geq 6$ if $q_{k^*+1} \geq 7/10$. Since $q_k \to 1$, there exists k^* such that $q_{k^*+1} \geq 7/10$, and for this k^* the proposed strategy is indeed an equilibrium. Our analysis also shows that if the proposed strategy is an equilibrium for a particular threshold k^*, then a similar strategy with a higher threshold k^{**} is also an equilibrium.

V.2.3 Starting Small

An interesting aspect of the preceding example is that for relatively small priors, the relationship has a "starting small" pattern. Initially, the relationship is small in the sense that only relatively small projects (in the example, $(-1, 10)$ and $(10, -1)$) are accepted. If the state is low, the posterior probability of a high state will tend to remain low, and the relationship will remain small. If the state is high, the posterior will grow until larger projects are also accepted.

It is intuitive that relationships should start small and build as the participants build trust. One may be happy to loan a new neighbor one's power tool, but the loan of a car or the use of one's house may come only after the relationship has developed. A new employee in an investment firm may initially manage small accounts and face trading limits, before being given a no-limits access to the largest funds. Countries may begin with cultural exchanges, working up to trade agreements, political cooperation, joint military exercises and finally a demilitarized border.

Watson (1999, 2002) offers one interesting perspective on this "starting small" phenomenon, based on incomplete information about the rate at which players discount future interactions. In our case, the players undertake larger (i.e., more costly) projects as they receive encouraging information about the distribution μ. It is no surprise that players might find it optimal to adjust the size of a relationship as they learn valuable information about that relationship. However, it appears as if this would cause relationships to shrink as readily as grow, whereas the starting small literature emphasizes growth. One possibility, appearing in Watson (1999, 2002), is that declining relationships may fall below a viability threshold and disappear, so that what we primarily observe are growing relationships. If these is some cost of maintaining a relationship, perhaps because it requires monitoring, then this would be the case in our setting. However, the example of the preceding section shows that the uncertainty concerning μ may be such that relatively small projects are profitable no matter what, while encouraging news is required to support large projects. Then the typical observed behavior will be that relationships start small, as indicated by the scale of projects that are undertaken, and (if anything) grow.

V.3 Imperfect Monitoring

We now consider a game with imperfect monitoring and (linear) transfers. At the beginning of each period t, one of the two players is chosen randomly, with probability $1/2$ each. If player i is chosen, the project $x^t = (x_i^t, x_j^t)$ is drawn randomly from a distribution μ_i with support in the $X^i \subset [0, \infty) \times (-\infty, 0]$. That is, in project $x^t \in X^i$, player i is the beneficiary and player j is the benefactor. The benefit $x_i^t > 0$ is publicly observed,

but the cost $x_j^t \leq 0$ is only observed by player j. However, after observing his cost, player j voluntarily reports a cost to player i. Of course, player j can lie and report $\hat{x}_j^t \neq x_j^t$. Moreover, the true cost x_j^t is never revealed to player i (neither this period, nor later). As in the perfect monitoring case, after j sends his report, the players simultaneously vote whether to accept or reject the project. If the project is adopted, the players privately collect their current period payoffs (x_1^t, x_2^t). Then, whether the project is adopted or not, at the end of the period the players simultaneously send each other voluntary (linear) transfers.

We study subgame perfect equilibria that do not make use of the public randomization device. They will suffice to characterize the extreme points of the Pareto efficient frontier of the equilibrium value set. Without loss of generality we restrict attention to subgame perfect equilibria where benefactors always report their costs honestly. As before, let V denote the equilibrium value set.

Let σ be a pure and simple subgame perfect equilibrium without public randomization. Then, its factorization is given by (α, m, w), where $\alpha(i, x)$ represents the vote in period 0 when the beneficiary is i and the observed/reported project is x, and for each vote outcome a, $m^a(i, x)$ and $w^a(i, x)$ represent the corresponding transfers and continuations value.

Suppose player i is the beneficiary and the project is x. Since the benefactor's cost x_j is never observed, $m^a(i, x)$ and $w^a(i, x)$ can depend on x_j only in irrelevant ways. Suppose for example that for some $\hat{x}_j \neq x_j$, $\alpha(i, x) = \alpha(i, x_i, \hat{x}_j) = (1, 1)$ and

$$(1 - \delta)[m_i^{1,1}(i, x) - m_j^{1,1}(i, x)] + \delta w_j^{1,1}(i, x)$$
$$< (1 - \delta)[m_i^{1,1}(i, x_i, \hat{x}_j) - m_j^{1,1}(i, x_i, \hat{x}_j)] + \delta w_j^{1,1}(i, x_i, \hat{x}_j).$$

Then player j would rather report $\hat{x}_j$ even if his true cost were x_j. We will also see that player i (the beneficiary) always prefers to accept the project. Thus, we can restrict attention to "delegation" subgame perfect equilibra, where player i approves all projects (x_i, x_j). That is, $\alpha_i(i, x_i, x_j) = 1$ for all (x_i, x_j). This gives player j (the benefactor) veto power to decide whether a project is accepted or rejected. To simplify notation, the factorization of a delegation subgame perfect equilibrium σ is given by (α, m, w), where now $m = (m^R, m^A)$ and $w = (w^R, w^A)$: $m^R(i, x_i)$ and $w^R(i, x_i)$ are respectively the transfers and continuation value when player j rejects the project, and $m^A(i, x_i)$ and $w^A(i, x_i)$ are the corresponding vectors when player j accepts the project. If player i rejects the project, by assumption this triggers the punishment where no transfers are made followed by autarky.

It is easy to see that if player j accepts project (x_i, x_j) in equilibrium, then he should also accept any project (x_i, x_j') with $0 \geq x_j' > x_j$. Hence, for each x_i, there exists $\bar{x}_j(x_i)$ such that player j accepts a project with benefit x_i if and only if his payoff is greater or equal to $\bar{x}_j(x_i)$.

As usual, we assume that transfers satisfy the complementarity condition $m_1^A(i, x_i) \cdot m_2^A(i, x_i) = 0$ and $m_1^R(i, x_i) \cdot m_2^R(i, x_i) = 0$. Hereafter the arguments (i, x_i) in transfers and continuation values will be omitted. Continuation values $w^A(i, x_i)$ and $w^R(i, x_i)$ must

belong to V, and together with the transfers $m = (m^A, m^R)$ must satisfy the incentive constraints

$$-(1-\delta)[m_i^A - m_j^A] + \delta w_i^A \geq 0 \quad \text{and} \quad -(1-\delta)[m_i^R - m_j^R] + \delta w_i^R \geq 0 \quad (TIC_i)$$
$$-(1-\delta)[m_j^A - m_i^A] + \delta w_j^A \geq 0 \quad \text{and} \quad -(1-\delta)[m_j^R - m_i^R] + \delta w_j^R \geq 0 \quad (TIC_j)$$

since transfers are voluntary and simultaneous. Let $\Delta^A = m_i^A - m_j^A$ and $\Delta^R = m_i^R - m_j^R$ be the net transfers from i to j (these net transfers can be positive or negative). Then (TIC) are equivalent to

$$-w_j^A \leq \beta\Delta^A \leq w_i^A \quad \text{and} \quad -w_j^R \leq \beta\Delta^R \leq w_i^R. \qquad (TIC)$$

Let $F_j(\bar{x}_j|x_i) = \mathbb{P}[x_j \leq \bar{x}_j \mid x_i]$. For player i to accept the project after he learns the benefit x_i the following incentive constraint must be satisfied:

$$\bar{F}_j(\bar{x}_j(x_i)|x_i)[-(1-\delta)\Delta^R + \delta w_i^R] + (1 - \bar{F}_j(\bar{x}_j(x_i)|x_i))[(1-\delta)(x_1 - \Delta^A) + \delta w_i^A] \geq 0.$$

But this incentive constraint is superfluous since it is implied by the upper bounds for Δ^A and Δ^R included in (TIC). This shows that player i always prefers to accept the project, as claimed earlier.

Finally, player j will vote to approve a project with cost x_j if and only if

$$x_j \geq [\Delta^R + w_j^R/\beta] - [\Delta^A + w_j^A/\beta] = \bar{x}_j(x_i). \qquad (VIC)$$

That is, $-\bar{x}_j(x_i)$ is the largest cost that player 2 will accept.

Let $T^R = w_1^R + w_2^R$, $T^A = w_1^A + w_2^A$ and $T^* = \max\{v_1 + v_2 \mid v \in V\}$, and for any $T > 0$ define
$$D(T) = \{v \in \mathbb{R}_+^2 \mid v_1 + v_2 = \delta T\}.$$

It is easy to see that the set of feasible interim continuation values that can be attained with transfers that satisfy (TIC) are

$$\{(v_i^R, v_j^R) = (1-\delta)(-\Delta^R, \Delta^R) + \delta(w_i^R, w_j^R) \mid -w_j^R \leq \beta\Delta^R \leq w_i^R\} = D(T^R)$$
$$\{(v_i^A, v_j^A) = (1-\delta)(-\Delta^A, \Delta^A) + \delta(w_i^A, w_j^A) \mid -w_j^A \leq \beta\Delta^A \leq w_i^A\} = D(T^A).$$

Clearly, given (w^R, w^A), there is a one-to-one map between (v^A, v^R) and (Δ^R, Δ^A). Moreover, (VIC) is equivalent to $x_j \geq v_2^R - v_2^A = (1-\delta)\bar{x}_2(x_1)$.

For any welfare weights $\lambda \in \mathbb{R}_{++}^2$ (avoiding some straightforward special cases by requiring both weights to be positive), the conditional expected welfare of σ given $x_i > 0$ is

$$W = F_j(\bar{x}_j|x_i)[\lambda \cdot v^R] + (1 - F_j(\bar{x}_j|x_i)[\lambda \cdot v^A] + (1-\delta)\int_{\bar{x}_j}^0 \lambda \cdot x \, dF_j(x_j|x_i).$$

The following proposition characterizes equilibria for given acceptance regions. Appendix A.6 proves:

Proposition 6. *For any pair of functions $\bar{x}_1, \bar{x}_2 : (0, \infty) \to [-T^*/\beta, 0]$ there exists a simple delegation subgame perfect equilibrium σ such that a project (i, x) is accepted in period 0 if and only if $x_j \geq \bar{x}_j(x_i)$. Moreover, let $\lambda \in \mathbb{R}^2_{++}$ and σ be such a delegation subgame perfect equilibrium with factorization (α, m, w) and acceptance regions given by $\bar{x}_1$ and $\bar{x}_2$. Then, σ maximizes $\lambda \cdot u(\sigma)$ among all such subgame perfect equilibria if and only if almost surely for $x_i > 0$ ($i = 1, 2$), $w^R(i, x_i), w^A(i, x_i) \in V^{(1,1)}$ and*

$$\Delta^R = \begin{cases} -w_j^R/\beta & \text{if } \lambda_i > \lambda_j \\ \bar{x}_j(x_i) + w_i^R/\beta & \text{if } \lambda_i < \lambda_j \end{cases}$$

$$\Delta^A = \begin{cases} -\bar{x}_j(x_i) - w_j^A/\beta & \text{if } \lambda_i > \lambda_j \\ w_i^A/\beta & \text{if } \lambda_i < \lambda_j. \end{cases}$$

Two familiar features emerge from this proposition. First, regardless of the point on the Pareto frontier being characterized, and the corresponding weighted sum of payoffs to be maximized, continuation values maximize the sum of payoffs $w_1 + w_2$ (i.e., are drawn from $V^{(1,1)}$). Second, the net transfers Δ^R and Δ^A are spread far enough apart to create the appropriate accept/reject decision, and consistent with this difference, are pushed as far as possible in the direction of transferring surplus to the player with the larger welfare weight (the displayed equations). Note that by complementarity (m^R, m^A) are uniquely defined by (Δ^R, Δ^A): if $\Delta^A > 0$ then $(m_i^A, m_j^A) = (\Delta^A, 0)$ and otherwise $(m_i^A, m_j^A) = (0, -\Delta^A)$. Similarly, if $\Delta^R > 0$ then $(m_i^R, m_j^R) = (\Delta^R, 0)$ and otherwise $(m_i^R, m_j^R) = (0, -\Delta^R)$.

Proposition 6 takes the acceptance regions (that is, the functions $\bar{x}_1$ and $\bar{x}_2$) as fixed and fills in the remaining details of the equilibrium. It remains to characterize the acceptance regions for an efficient subgame perfect equilibrium σ. Now let $\lambda \in \mathbb{R}_+ \backslash \{(0, 0)\}$ and suppose that σ is efficient: $u(\sigma) \in \partial V^\lambda$. In order to maximize total welfare, we would like to induce player j to accept all projects (i, x) such that $\lambda \cdot x > 0$. That is, we would like to set $\bar{x}_j(x_i) = -\frac{\lambda_i}{\lambda_j} x_i$. But this is possible only if x_i is not too large. In particular, let

$$x_i^* = \frac{\lambda_j}{\lambda_i} T^*/\beta \quad i = 1, 2. \tag{5}$$

If $x_i > x_i^*$, there are no feasible transfers that satisfy (VIC) for $\bar{x}_j(x_i) = -\frac{\lambda_i}{\lambda_j} x_i < -T^*/\beta$.

By Proposition 6, after every equilibrium history except the null history, the continuation strategies $\hat{\sigma}$ maximize the total surplus, i.e., $u(\hat{\sigma}) \in V^{(1,1)}$. The next Proposition characterizes the acceptance regions for any such subgame perfect equilibrium, showing that current projects are accepted if and only if they increase the total surplus, subject to the identified in (5).

Proposition 7. *Let $\lambda_1 = \lambda_2 = 1$ and assume that σ is an subgame perfect equilibrium with value $u(\sigma) \in V^{(1,1)}$. Then, its acceptance regions are defined by the functions $\bar{x}_j$, $j = 1, 2$, where*

$$\bar{x}_j(x_i) = \begin{cases} -x_i & \text{if } x_i < x_i^* \\ -T^*/\beta & \text{if } x_i \geq x_i^* \end{cases}$$

Proof: Let (α, m, w) be σ's factorization. By Proposition 6, $w^R, w^A \in V^{(1,1)}$ and $w_1^R + w_2^R = w_1^A + w_2^A = T^*$. Assume $(i, x_i) = (1, x_1)$. When $\lambda_1 = \lambda_2 = 1$, the conditional expected welfare of σ given x_1 becomes

$$W = (1 - \delta) \int_{\bar{x}_2}^0 (x_1 + x_2) dF_2(x_2|x_1) + \delta T^*.$$

We want to maximize W subject to the constraint that $-(1 - \delta)\bar{x}_2 = v^A - v^R$ for some $v^R, v^A \in D(T^*)$. If $x_1 < x_1^* = T^*/\beta$, the optimum is not constrained and W is maximized when $\bar{x}_2 = -x_1$. If $x_1 \geq x_1^*$, the optimum is constrained and W is maximized when $\bar{x}_2 = -T^*/\beta$ (and $v^R = (0, \delta T^*)$ and $v^A = (\delta T^*, 0)$). $\qquad\square$

Propositions 6 and 7 fully characterize σ, including the acceptance regions for all periods except for period 0. We now examine the acceptance regions for σ in period 0.

Let $\lambda \in \mathbb{R}_+^2 \backslash \{(0,0)\}$ with $\lambda_1 \neq \lambda_2$ and σ be a subgame perfect equilibrium such that $u(\sigma) \in \partial V^\lambda$. We organize the presentation by considering projects (x_i, x_j) with $x_i > 0$ and $x_j < 0$, so that i is the beneficiary and j the benefactor. We then divide the analysis into two cases.

First, suppose $\lambda_j > \lambda_i > 0$, so it is beneficial to transfer surplus to the benefactor. From Proposition 6, when $\lambda_j > \lambda_i > 0$, the optimal interim continuation values are $(v_i^R, v_j^R) = (-(1 - \delta)\bar{x}_j, \delta T^* + (1 - \delta)\bar{x}_j)$ and $(v_i^A, v_j^A) = (0, \delta T^*)$. Therefore, $v^R - v^A = (-(1 - \delta)\bar{x}_j, (1 - \delta)\bar{x}_j)$ and

$$W = (1 - \delta) F_j(\bar{x}_j|x_i)(\lambda_j - \lambda_i)\bar{x}_j + \delta \lambda_j T^* + (1 - \delta) \int_{\bar{x}_j}^0 \lambda \cdot x dF_j(x_j|x_i).$$

Hence,

$$\frac{\partial W}{\partial \bar{x}_j} = (1 - \delta) \Big[(\lambda_j - \lambda_i)[f_j(\bar{x}_j|x_i)\bar{x}_j + F_j(\bar{x}_j|x_i)] - (\lambda_i x_i + \lambda_j \bar{x}_j) f_j(\bar{x}_j|x_i) \Big]$$

$$= (1 - \delta) \Big[(\lambda_j - \lambda_i) F_j(\bar{x}_j|x_i) - \lambda_i (x_i + \bar{x}_j) f_j(\bar{x}_j|x_i) \Big].$$

When $\lambda_i = \lambda_j$, the first term in the brackets is zero and this derivative is negative as long as $x_i + \bar{x}_j > 0$. Thus, optimally, $\bar{x}_j$ should be decreased until $\bar{x}_j$ reaches $-x_i$ or $-\delta T^*$, as was already established in Proposition 7. In this case there are no gains from transferring surplus from one agent to the other (reflected in the zero first term), and so projects are accepted if and only if (within the bounds set by incentive compatibility) they increase the total surplus.

When $\lambda_j > \lambda_i$, the objective of increasing surplus competes with the objective of transferring surplus to the benefactor, player j. The first term in brackets is positive and bounded away from zero. As a result, there exists $x_i^0 > 0$ such that the derivative $\frac{\partial W}{\partial \bar{x}_j}$ is positive for all $x_i \in (0, x_i^0)$ and hence optimally $\bar{x}_j = 0$. Reducing $\bar{x}_j$ reduces the amount

that can be transferred to player j after a rejection, in order to create incentives to accept projects, which overwhelms the modest surplus gains from accepting projects, for values of $x_i < x_i^0$. At the other end, there exists $x_i^1 > x_i^0$ such that for all $x_i \geq x_i^1$, optimally $\bar{x}_j = -\lambda_j/\lambda_i[T^*/\beta]$. Here, projects are sufficiently valuable that they accepted up to the limits of the ability to create the requisite incentives, despite sacrificing the opportunity to make transfers to agent j. For values of x_i in between x_1^0 and x_i^1, the optimal value $\bar{x}_j$ is greater than 0 but less than $\max \{-x_i, -\lambda_j/\lambda_i[T^*/\beta]\}$. Here, x_i is sufficiently large that W is maximized by accepting some projects, but the ability to increase transfers to player j by reducing the incentives to accept projets still pushes the acceptance boundary above the level $\max \{-x_i, -\lambda_j/\lambda_i[T^*/\beta]\}$.

Now we continue to maintain that $x_i > 0$ and $x_j < 0$, but assume that $\lambda_i > \lambda_j > 0$. When $\lambda_i > \lambda_j > 0$, the optimal interim continuation values are

$$\begin{pmatrix} v_i^A \\ v_j^A \end{pmatrix} = \begin{pmatrix} \delta T^* + (1-\delta)\bar{x}_j \\ -(1-\delta)\bar{x}_j \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} v_i^R \\ v_j^R \end{pmatrix} = \begin{pmatrix} \delta T^* \\ 0 \end{pmatrix}.$$

Therefore, $v^R - v^A = (-(1-\delta)\bar{x}_j, (1-\delta)\bar{x}_j)$,

$$W = (1-\delta)(1 - F_j(\bar{x}_j|x_i))(\lambda_i - \lambda_j)\bar{x}_j + \delta\lambda_i T^* + (1-\delta)\int_{\bar{x}_j}^0 \lambda \cdot x dF_j(x_j|x_i),$$

and

$$\frac{\partial W}{\partial \bar{x}_j} = (1-\delta)\Big[(\lambda_i - \lambda_j)(1 - F_j(\bar{x}_j|x_i)) - \bar{x}_j f_j(\bar{x}_j|x_i)] - (\lambda_i x_i + \lambda_j \bar{x}_j)f_j(\bar{x}_j|x_i)\Big].$$

$$= (1-\delta)\Big[(\lambda_i - \lambda_j)(1 - F_j(\bar{x}_j|x_i)) - \lambda_i(x_i + \bar{x}_j)f_j(\bar{x}_j|x_i)\Big].$$

Once again, if $\lambda_i = \lambda_j$, this derivative reduces to $-(x_i + \bar{x}_j)f_j(\bar{x}_j|x_i)$, and a project will be accepted only if $x_i + x_j > 0$. When $\lambda_i > \lambda_j$, the first term in brackets in the derivative above is positive, capturing the objective of transferring surplus to player i. Creating incentives for j to accept projects reduces the effectiveness of such transfers. Hence, we again have $x_i^1 > x_1^0 > 0$ such that $\bar{x}_j = 0$ (no projects are accepted) when $x_i < x_i^0$ and $\bar{x}_j = -\lambda_j/\lambda_i[T^*/\beta]$ (all projects for which we can create the requisite incentives) are accepted for $x_i > x_i^1$. For values of x_i in between x_1^0 and x_i^1, the optimal value $\bar{x}_j$ is greater than 0 but less than $\max \{-x_i, -\lambda_j/\lambda_i[T^*/\beta]\}$.

V.4. Two-Sided Imperfect Monitoring

In this section we assume that whenever a project is generated, neither player can observe the value of the project to the other player. Hence, the players may know (for example) that the project confers a benefit on player 1 and imposes a cost on player 2, but neither players knows the magnitude of the effect on the other players.

V.4.1 Efficient Continuation Payoffs?

A key result running through all of our previous analysis is that in the presence of transfers, the continuation values of an efficient equilibrium must invariably maximize total surplus. We construct an example showing that this is not the case with two-sided imperfect monitoring, where we find that continuation equilibria may be strictly Pareto dominated.

Let $\delta = 1/2$, to simplify the calculations. With probability half, each player is chosen to be the beneficiary of the project. The beneficiary's benefit is equally likely to be 2 or 9 and independently the benefactor's cost is equally likely to be 2 or 9. Thus, conditional on player 1 being the beneficiary, the project is equally likely to be one of

$$(2, -2), \ (9, -2), \ (2, -9), \ (9, -9),$$

and conditional on player 2 being the beneficiary, the project is equally likely to be one of

$$(-2, 2), \ (-2, 9), \ (-9, 2), \ (-9, 9).$$

We first construct an equilibrium σ in which only projects $(9, -2)$ and $(-2, 9)$ are accepted in the first period. Using notation developed in the previous section, we describe the equilibrium in terms of the factorization (α, m, w). Here, $\alpha(i, x) = (\alpha_1(i, x_1), \alpha_2(i, x_2))$ specifies how the players vote when player i is the beneficiary and the current project is x. This notation reflects the fact that, by assumption, player j only observes x_j, $j = 1, 2$. Let $m^a(i) = (m_1^a(i), m_2^a(i))$ and $w^a(i) = (w_1^a(i), w_2^a(i))$ be the transfers and continuation values when player i is the beneficiary and $a \in \{0, 1\}^2$ is the vote outcome in the first period. Also let $v^a(i) = (1 - \delta)m^a(i) + \delta w^a(i)$ be the interim continuation values.

Our candidate equilibrium calls for the beneficiary to accept the project if the benefit is 9 and reject if it is 2, while the benefactor accepts the project if the cost is 2 and reject it if the cost is 9. For the players to vote this way, if player 1 is the beneficiary the following incentives constraints must be satisfied:

$$\frac{1}{2}[(1 - \delta)2 + v_1^{1,1}(1)] + \frac{1}{2}v_1^{1,0}(1) \leq \frac{1}{2}[v_1^{0,1}(1) + v_1^{0,0}(1)]$$

$$\leq \frac{1}{2}[(1 - \delta)17 + v_1^{1,1}(1)] + \frac{1}{2}v_1^{1,0}(1)$$

$$\frac{1}{2}[-(1 - \delta)17 + v_2^{1,1}(1)] + \frac{1}{2}v_2^{0,1}(1) \leq \frac{1}{2}[v_2^{1,0}(1) + v_2^{0,0}(1)]$$

$$\leq \frac{1}{2}[-(1 - \delta)2 + v_2^{1,1}(1)] + \frac{1}{2}v_2^{0,1}(1).$$

Let $T = u_1(\sigma) + u_2(\sigma)$ and let the continuation values be given by

$$v^{0,0}(1) = \begin{pmatrix} T/2 \\ 0 \end{pmatrix}, \ v^{0,1}(1) = \begin{pmatrix} T/4 \\ T/4 \end{pmatrix}, \ v^{1,0}(1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ v^{1,1}(1) = \begin{pmatrix} 0 \\ T/2 \end{pmatrix}.$$

Similar incentive constraints must be satisfied when player 2 is the beneficiary, and in that case we symmetrically set

$$v^{0,0}(2) = \begin{pmatrix} 0 \\ T/2 \end{pmatrix}, \ v^{0,1}(2) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ v^{1,0}(2) = \begin{pmatrix} T/4 \\ T/4 \end{pmatrix}, \ v^{1,1}(2) = \begin{pmatrix} T/2 \\ 0 \end{pmatrix}.$$

In this factorization, with probability $1/4$ the players capture a total surplus of 9-2=7 in the first period and with probability $3/4$ the total value $v_1^a(i) + v_2^a(i)$ of the interim continuation value is $T/2$. Therefore

$$T = u_1(\sigma) + u_2(\sigma) = \frac{1}{4}\frac{7}{2} + \frac{3}{4}\frac{T}{2}.$$

Thus, $T = 7/5$ and the incentive constraints above are indeed satisfied.

Conditional on player 1 being the beneficiary, the interim continuation values above are delivered by the following transfers and continuation values:

$$m^{0,0}(1) = \begin{pmatrix} 0 \\ T/2 \end{pmatrix}, \ m^{0,1}(1) = m^{1,0}(1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ m^{1,1}(1) = \begin{pmatrix} T/2 \\ 0 \end{pmatrix}$$

$$w^{0,0}(1) = w^{0,1}(1) = w^{1,1}(1) = \begin{pmatrix} T/2 \\ T/2 \end{pmatrix}, \ w^{1,0}(1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The corresponding transfers and continuation values when player 2 is the beneficiary are symmetrically constructed.

In equilibrium σ, the players cooperate until a period in which either player 1 is the beneficiary and the vote outcome is $a = (1,0)$ or player 2 is the beneficiary and the vote outcome is 1 or $a = (0,1)$. While the players are cooperating, the benefactor accepts a project if the cost is 2 and rejects it if the cost is 9 while the beneficiary accepts a project if the benefit is 9 and rejects it if the benefit is 2, and a transfer of $T/2$ is made by the benefactor if both players reject the project and by the beneficiary if both players accept the project.

We thus have an equilibrium σ with payoffs $u(\sigma) = (T/2, T/2) = (7/10, 7/10)$. However, one of the continuation payoffs in this equilibrium does not maximize total surplus, since continuation payoffs following an outcome in which the beneficiary accepts and the benefactor rejects are $(0,0)$.

We now show that this example is not pathological—the surplus-maximizing equilibrium necessarily contains some continuation values that do not maximize surplus. Let σ^* be an equilibrium such that $u(\sigma^*) \in \partial V^{(1,1)}$. That is, σ^* is an equilibrium that maximizes total surplus: $u_1(\sigma^*) + u_2(\sigma^*) = T^*$. Let (α, m, w) be its corresponding factorization. We show that not all continuation values $w^a(i)$ can maximize total surplus. That is, for some (i, a) it must be that $w^a(i) \notin \partial V^{(1,1)}$.

By definition, $w_1^a(i) + w_2^a(i) \leq T^*$ (so $v_1^A(i) + v_2^A(i) \leq \delta T^* = T^*/2$) and the maximal expected total surplus in period 1 is $7/4$ (which is attained when only projects $(9, -2)$ and $(-2, 9)$ are adopted in the first period). Therefore an upper bound for T^* is given by

$$T^* = u_1(\sigma^*) + u_2(\sigma^*) \leq (1 - \delta)\frac{7}{4} + \delta T^* = \frac{7}{8} + \frac{T^*}{2}.$$

That is, $T^* \leq 7/4$.

Since $T^* > 0$ (since we have exhibited an equilibrium with a positive payoff), it must be that σ^* approves projects $(9, -2)$ and $(-2, 9)$ in the first period. Moreover, it must also reject all other projects. To see the latter, suppose (for example) that the beneficiary accepts all projects. Then, when the beneficiary is player 1, for player 2 to accept a project with cost 2 it must be that

$$v_2^{1,0}(1) \leq -2(1 - \delta) + v_2^{1,1}(1).$$

But even if we choose $v_2^{1,0}(1) = 0$ and $v_2^{1,1}(1) = T^*/2 \leq 7/8$, the constraint is violated. That is, if the beneficiary accepts all projects, it is not possible to provide incentives for the benefactor to accept a project with cost 2 (and hence surely not to accept a project with cost 9). In a similar fashion, one can easily see it is not possible to provide incentives for player 2 to accept a project with cost 9.

Assume player 1 is the beneficiary. For player 1 to reject a project with benefit 2 and for player 2 to accept a project with cost 2, the following incentives must be satisfied:

$$
\begin{aligned}
(1 - \delta)2 + v_1^{1,1}(1) + v_1^{1,0}(1) &\leq v_1^{0,1}(1) + v_1^{0,0}(1) \\
v_2^{1,0}(1) + v_2^{0,0}(1) &\leq -(1 - \delta)2 + v_2^{1,1}(1) + v_2^{0,1}(1).
\end{aligned}
$$

Clearly setting $v^{1,1}(1) = (0, T^*/2)$ and $v^{0,0}(1) = (T^*/2, 0)$ helps with incentives and the objective of maximizing total surplus. With these choices, $v^{1,0}(1)$ and $v^{0,1}(1)$ must satisfy the constraints

$$v_1^{0,1}(1) - v_1^{1,0}(1) \geq 1 - \frac{T^*}{2} \geq \frac{1}{8} \quad \text{and} \quad v_1^{0,1}(1) - v_1^{1,0}(1) \geq 1 - \frac{T^*}{2} \geq \frac{1}{8}.$$

But then, $v^{0,1}(1) - v^{1,0}(1) \geq (1/8, 1/8)$, which implies that $w^{1,0}(1) \notin \partial V^{(1,1)}$. Continuation payoffs after the benefactor rejects a project and the beneficiary accepts are strictly Pareto dominated.

V.4.2 Sufficient Conditions for Efficient Continuation Payoffs

We now establish sufficient conditions under which all continuation values maximize expected surplus (i.e., all continuation values satisfy $w^a(i) \in V^{(1,1)}$), for any $\lambda \in \mathbb{R}_+^2 \setminus \{(0,0)\}$ and any equilibrium σ with $u(\sigma) \in \partial V^\lambda$. We depart from our previous analysis in requiring the players to be sufficiently patient, explaining at the end of this section how this exposes a fundamental effect of two-sided imperfect monitoring.

In each period, with probability half each player is chose to be the beneficiary. The benefit is then drawn from a distribution F in $[0, \infty)$ and independently, the cost is drawn from a distribution G in $(-\infty, 0]$. Assume F and G are absolutely continuous with densities f and g, respectively.

Fix $\lambda \in \mathbb{R}^2_{++}$ (again avoiding special cases by requiring both to weights to be positive). Let σ be an equilibrium with factorization (α, m, w) such that $u(\sigma) \in \partial V^\lambda$. The vote policy α for the first period corresponds to two threshold vectors $\bar{x}(i) = (\bar{x}_1(i), \bar{x}_2(i))$, $i = 1, 2$, such that when player i is the beneficiary, player j accepts a project x if and only $x_j \geq \bar{x}_j(i)$, $j = 1, 2$.

We first identify the threshold vectors that maximize the contribution of the current period to the weighted welfare objective. When player 1 is the beneficiary, the expected contribution of the current project to the objective is

$$H = \int_{\bar{x}_1(1)}^{\infty} \int_{\bar{x}_2(1)}^{0} \lambda \cdot x f(x_1) g(x_2) dx_2 dx_1.$$

The thresholds that maximize H satisfy the first-order conditions

$$\frac{\partial H}{\partial \bar{x}_2(1)} = -\int_{\bar{x}_1(1)}^{\infty} (\lambda_1 x_1 + \lambda_2 \bar{x}_2(1)) g(\bar{x}_2(1)) f(x_1) dx_1 = 0$$

$$\frac{\partial H}{\partial \bar{x}_1(1)} = -\int_{\bar{x}_2(1)}^{0} (\lambda_1 \bar{x}_1(1) + \lambda_2 x_2) f(\bar{x}_1(1)) g(x_2) dx_2 = 0$$

which we can solve for

$$\bar{x}_1(1) = -\frac{\lambda_2}{\lambda_1} \mathbb{E}_G[x_2 | x_2 \geq \bar{x}_2(1)] \quad \text{and} \quad \bar{x}_2(1) = -\frac{\lambda_1}{\lambda_2} \mathbb{E}_F[x_1 | x_1 \geq \bar{x}_1(1)].$$

These conditions are just what one would expect, and have an intuitive interpretation. Player 2 should accept the project if and only if $\lambda_2 \cdot$[the cost] is less then $\lambda_1 \cdot$[expected benefit to player 1]. Similarly, player 1 should accept the project if and only if $\lambda_1 \cdot$[the benefit] is larger than $\lambda_2 \cdot$[the expected cost to player 2]. It is straightforward that the solution to the problem of maximizing H is interior. The thresholds that maximize the expected welfare contribution of the first period when player 2 is the beneficiary are similarly

$$\bar{x}_1(2) = -\frac{\lambda_2}{\lambda_1} \mathbb{E}_F[x_2 | x_2 \geq \bar{x}_2(2)] \quad \text{and} \quad \bar{x}_2(2) = -\frac{\lambda_1}{\lambda_2} \mathbb{E}_G[x_1 | x_1 \geq \bar{x}_1(2)].$$

Next, consider the special case $\lambda = (1, 1)$ and let $(\bar{x}^*(1), \bar{x}^*(2))$ be the corresponding thresholds. In this case, $\bar{x}_1^*(2) = \bar{x}_2^*(1)$ and $\bar{x}_2^*(2) = \bar{x}_1^*(1)$. The total expected surplus if, in every period, a project is implemented if and only if the benefit is more than $\bar{x}_1^*(1)$ and the cost is less than $\bar{x}_2^*(1)$, is given by

$$T^* = (1 - \bar{F}^*)(1 - \bar{G}^*)[\mathbb{E}_F[x_1 | x_1 \geq \bar{x}_1^*(1)] + \mathbb{E}_G[x_2 | x_2 \geq \bar{x}_2^*(1)]] = -(1 - \bar{F}^*)(1 - \bar{G}^*)[\bar{x}_2^* + \bar{x}_1^*],$$

where $\bar{F}^* = F(\bar{x}_1^*(1))$ and $\bar{G}^* = G(\bar{x}_2^*(1))$. This value T^* is an upper bound on the total expected surplus one could expect an equilibrium to generate. We show below that if the

players are sufficiently patient, this upper bound is achieved, i.e., $(T^*/2, T^*/2) \in \partial V^{(1,1)}$ for δ sufficiently large.

We now argue that in equilibrium σ, where $u(\sigma) \in \partial V^\lambda$ for an arbitrary $\lambda \in \mathbb{R}^2_+ \backslash \{(0,0)\}$, a project x is implemented in the first period if and only if $x_j \geq \bar{x}_j(i)$ for $j = 1, 2$ when the beneficiary is player i ($i = 1, 2$). Moreover, $w^a(i) = (T^*/2, T^*/2)$ for all $a \in \{0, 1\}^2$ and $i = 1, 2$. Hence, first-period decisions are arranged to maximize the expected value of the objective $\lambda_1 x_1 + \lambda_2 x_2$, while continuation payoffs invariably maximize the social surplus. The argument proceeds in three steps.

First, the argument that continuation payoffs are drawn from $V^{(1,1)}$, and hence maximize the social surplus, is analogous to that given in Proposition 6 for the one-sided imperfect monitoring case, and we will not repeat it here. Second, taking this as given, we show that we can create the required incentives in the first period. As usual, we will work with the interim value $v^a(i) = (1 - \delta)m^a(i) + \delta w^a(i)$. Conditional on player 1 being 1 the beneficiary, the voting incentive constraints are

$$\bar{G}v_1^{0,0}(1) + (1 - \bar{G})v_1^{0,1}(1) = \bar{G}v_1^{1,0}(1) + (1 - \bar{G})[(1 - \delta)\bar{x}_1 + v_1^{1,1}(1)]$$
$$\bar{F}v_2^{0,0}(1) + (1 - \bar{F})v_2^{1,0}(1) = \bar{F}v_2^{0,1}(1) + (1 - \bar{F})[(1 - \delta)\bar{x}_2 + v_2^{1,1}(1)],$$

where $\bar{F} = F(\bar{x}_1(1))$ and $\bar{G} = G(\bar{x}_2(1))$. Since $w_1^a(i) + w_2^a(i) = T^*$, we must have that $v_1^a(i) + v_2^a(i) = \delta T^*$ for all $a \in \{0, 1\}^2$ and $i = 1, 2$. Substituting $v_1^a(1) = \delta T^* - v_2^a(1)$ in the first constraint we get

$$\bar{G}v_2^{0,0}(1) + (1 - \bar{G})v_2^{0,1}(1) = \bar{G}v_2^{1,0}(1) + (1 - \bar{G})[-(1 - \delta)\bar{x}_1 + v_2^{1,1}(1)]$$
$$\bar{F}v_2^{0,0}(1) + (1 - \bar{F})v_2^{1,0}(1) = \bar{F}v_2^{0,1}(1) + (1 - \bar{F})[(1 - \delta)\bar{x}_2 + v_2^{1,1}(1)].$$

These constraints can be rewritten as

$$A \cdot z = \frac{1 - \delta}{\delta} \begin{bmatrix} (1 - \bar{G})\bar{x}_1(1)/T^* \\ (1 - \bar{F})\bar{x}_2(1)/T^* \end{bmatrix} \equiv \begin{bmatrix} 1 - \delta \\ \delta \end{bmatrix} b,$$

where

$$A = \begin{bmatrix} -\bar{G} & -(1 - \bar{G}) & \bar{G} & (1 - \bar{G}) \\ \bar{F} & -\bar{F} & (1 - \bar{F}) & -(1 - \bar{F}) \end{bmatrix} \quad \text{and} \quad z = \frac{1}{\delta T^*} \begin{bmatrix} v_2^{0,0}(1) \\ v_2^{0,1}(1) \\ v_2^{1,0}(1) \\ v_2^{1,1}(1) \end{bmatrix}.$$

Thus, σ is an equilibrium if there exists $z \in [0, 1]^4$ that solves the linear system of equations. Now, $C = A([0,1]^4)$ is a convex polyhedron in $\mathbb{R}^2$ that contains the four points corresponding to the four columns of A (since, for example, $(-\bar{G}, \bar{F}) = A(1, 0, 0, 0)$). Hence $(0, 0)$ is an interior point of C. Therefore, no matter what are the values of $\bar{x}_1(1)$ and $\bar{x}_2(1)$, there exists $\bar{\delta} \in (0, 1)$ such that $[(1 - \delta)/\delta]b \in C$ for all $\delta \in [\bar{\delta}, 1)$. A symmetric analysis

applies to the thresholds $(\bar{x}_1(2), \bar{x}_2(2))$. This establishes that σ is an equilibrium and that $u(\sigma) \in \partial V^\lambda$ for δ sufficiently high.

Third, we need to show that $(T^*/2, T^*/2) \in V^{(1.1)}$ for δ sufficiently high is proved. The argument here is analogous to that of the second step. One proceeds by considering the strategy σ^* with factorization (α^*, m, w) where α^* corresponds to the thresholds $\bar{x}^*(1)$ and $\bar{x}^*(2)$, and $w^a(i) = (T^*/2, T^*/2)$ for all (a, i).

We have thus established conditions under which every continuation path in an efficient subgame perfect equilibrium maximizes the expected surplus. The key to this result is patience. In contrast, we have not made a similar appeal patience in the previous settings we have considered, because doing so is ineffective in those settings. With no information on either side as to the size of the project, equilibrium play in the current setting is quite simple, revolving around a pair of cutoff values. Moreover, there is an upper bound on the range of such values that one might want to implement, arising out of the technology that generates projects and requiring only that this technology generate finite expected values. We then need only make agents sufficiently patient that this upper bound can be implemented. In all of our previous settings, the value of the project on at least one side is known, and we must implement a cutoff on the other side. However, the value of the cutoff to be implemented is conditioned on the realized value on the observed side, and may be arbitrarily large. No amount of patience will then suffice to ensure that we do not sometimes encounter a desired cutoff that is simply too large to implement. Putting these observations together, we can achieve a folk-theorem type of result only if monitoring is sufficiently (i.e., two-sided) imperfect.

VI. Discussion

VI.1 Maintaining Relationships: Summary

Section IV presents our workhorse model, featuring transfers and complete information. This gives us a relationship in which the players routinely adopt projects, to the cost of one and the benefit of the other. Projects that are not out of the ordinary require no adjustment in continuation play and need not be remembered, with the players simply confirming that the relationship is on track. Larger projects prompt transfers to push the relationship back on track. Depending on the distribution of projects, transfers may seldom be observed, and indeed may be rarer than would be adjustments in continuation values in the absence of transfers, and yet may also be important in supporting the equilibrium.

VI.2 Why Transfer?

An important feature of the transfers in our model is that they are voluntary. At the end of each period, player i is welcome to make a transfer to player j, but need not do so, and any incentives to do so must be created by the appropriate conditioning of continuation payoffs on current actions.

Three approaches to transfers appear in the literature. First, and at one extreme, many applications in game theory assume that transfers are simply impossible. Von Neumann and Morgenstern (1944) take this as their point of departure for the study of noncooperative games, commenting that they will begin with "no mention of coalitions between players and the compensations they can pay to each other" (Von Neumann and Morgenstern (1944, p. 46)). In this view, any exchange involving a transfer effectively requires some enforcement mechanism, since otherwise there is no way to ensure the transfer will be made, and hence transfers are out of place in a noncooperative theory.

This argument may motivate one to omit transfers from many static models, but it is less obvious that transfers should be omitted from the study of repeated games. Here, the prospect of future interactions may well create incentives for making transfers. Indeed, some work on repeated games stresses that continuation payoffs can often be viewed as transfers in a current mechanism design problem (e.g., Athey and Bagwell (2001), Athey, Bagwell and Sanchirico (2004)). Then why do repeated games typically preclude transfers? One possibility is that much of the motivation for work in repeated games come from models of repeated oligopoly (e.g., Friedman (1971)), where transfers presumably give rise to potential legal problems.

Second, and in contrast, it is standard in many economic models to simply assume that transfers are mandatory and will be made, with scant mention of the incentives to do so. We typically assume that firms can hire inputs in exchange for appropriate remuneration, but we do not consider the possibility that the firm consumes the input and then declines payment. This is the case even in models that otherwise view the firms as engaged in an noncooperative game with one another. In models of competitive markets we typically assume that a consumer can either purchase a good at the market price or choose not to purchase the good, but we do not allow the possibility that the consumer can take delivery of the good and decline the corresponding transfer. Once again, this often occurs in the context of an otherwise noncooperative model.

Why not simplify our model by similarly making transfers mandatory? We believe that at the basic level, all transfers are voluntary, and will be made only if people find it in their best interests to make them. In many cases, however, it is sufficiently obvious that in equilibrium the transfers will be made that we can conveniently omit the mechanism creating the attendant incentives from our model. Firms pay their workers because failure to pay risks a lawsuit. People pay for their purchase after it has been packaged and handed to them in the store because they do not relish the thought of running through a crowded mall with their arms full of packages and mall security at their heels. Little is typically lost by omitting the legal system and mall security from the resulting models and simply assuming transfers are mandatory.

In other circumstances the enforcement mechanism behind transfers is perhaps less obvious. Suppose one comes to the end of a taxi journey to find no one around. Why pay the fare? The driver has no opportunity to sue if you do not. One possibility is that if you do not, the driver will pursue you and compel payment, much like the mall security guards of the preceding paragraph. But if the driver can do this, why doesn't he do so

after you have paid, extracting a second fare? And if you could successfully resist this second attempt, why not resist the first? Indeed, if either side has the ability to force an outcome, it is not clear why this side does not always exercise this ability, and even exercise it repeatedly.

Basu (1983) explores these questions (using the case of a taxi ride as his motivating example), and concludes that conventional models can explain outcomes in which no payment is made, or outcomes in which multiple payments are extorted, but not an outcome in which a single payment is made. Basu (1983) then hypothesizes that people pay the fare at the end of the taxi journey because they *prefer* to do so. He extends this point to argue that a similar preference to pay plays an essential role in making economic transactions possible. Far from the stereotypical complaint that economists assume that people are inherently selfish, Basu argues that it is a pervasive desire to part with money under the appropriate circumstances that makes economic activity possible.

We find this appeal to preferences unconvincing as an explanation for transfers. At the end of the taxi journey, the driver can attempt to compel payment and the passenger can attempt to avoid payment. We think of the resulting interaction as a war of attrition. The war of attrition has an inefficient symmetric equilibrium, and two efficient asymmetric equilibria, one in which the passenger makes the payment without a fuss, and one in which the driver foregoes payment, again without a fuss. The circumstances of the interaction allow the participants to coordinate on one of these asymmetric equilibria. At the end of the ride, convention selects the equilibrium in which the passenger pays. Should the driver attempt to extort a second payment, convention selects the equilibrium in which the passenger does not pay. We thus have no difficulty accounting for an interaction that successfully elicits one and only one payment. We believe that there are many similar situations in which the mechanism for enforcing transfers is hidden, but is nonetheless effective, and in which it is reasonable to assume that transfers are essentially mandatory, while omitting the underlying war of attrition from the model.

In the current setting, however, we are uncomfortable with simply assuming that transfers are mandatory, enforced by either formal or informal means. A theme of the descriptive studies motivating our research is that the participants deliberately eschew an appeal to contracts, legal measures, or formal institutions as a means of governing their interactions. We regard the war of attrition as a compelling explanation for many of the more mundane transactions in which people engage, but less convincing in many other circumstances.

This brings us to the third approach, commonly gathered under the heading of relational contracts and incorporated in our model. Interactions are repeated, the players can make transfers, but the equilibrium strategies must ensure the incentives to do so.

VI.3 What to Transfer?

We have said little about the form of transfers. Ellickson (1991) reports that the subjects of his study exhibited a very strong preference to make transfers in kind rather than in cash. A farmer whose crops had been damaged by another's trespassing cattle might be invited to "come down and take some hay" (Ellickson, 1991, p. 56), or someone who has financed

a disproportionate share of joint fencing may claim the use of his neighbor's bulldozer (Ellickson, 1991, p. 80), but money would typically not change hands. Our model is silent on the form taken by transfers. Transfers are less efficient the more curvature there is in the function f in Section IV.1, and so we would expect transfers to be made in a currency for which this function does not exhibit too much curvature. However, we have no reason to believe that hay is inferior to money in this respect.

The aversion to monetary transfers described by Ellickson (1991) appears to go beyond a simple efficiency comparison. Instead, the decision to make transfers in cash or in kind appears to contain a signaling component absent from our model. Interactions in relationships are carried out in kind, while arms-length interactions are carried out in cash. Offering hay to someone whose crops your cattle have damaged is a confirmation that your relationship is intact; while offering cash is an indication that incentives are to be created and resolved contemporaneously rather than intertemporally, putting the relationship at risk.

A. Appendix

A.1 Proof of Proposition 2

We begin with a preliminary observation. Let V_ζ denote the equilibrium value set of the subgame Γ_ζ, i.e. the subgame after a draw has been taken from the public randomization device. Notice that V_ζ does not depend on ζ, and that if σ is a subgame perfect equilibrium for Γ, then the continuation strategy σ_ζ must be a subgame perfect equilibrium for Γ_ζ for all $\zeta \in [0, 1]$. Since $V_\zeta \subset \mathbb{R}^2$, by Caratheodory's theorem (Aliprantis and Border (2006, P. 184)), any value in V can be expressed as the convex combination of at most three points in V_ζ. That is, for any $v \in V$ there exist $v^k \in V_\zeta$ and $\lambda_k \in [0, 1]$, $k = 1, 2, 3$, such that

$$ v = \sum_{k-1}^{3} \lambda_k v^k \quad \text{and} \quad \sum_{k-1}^{3} \lambda_k = 1. $$

We can then construct a subgame perfect equilibrium with value v as follows. Let σ^k be a subgame perfect equilibrium for Γ_ζ such that $v^k = u(\sigma^k)$, $k = 1, 2, 3$. Then, partition the interval $(0, 1)$ into the subintervals $I_1 = (0, \lambda_1)$, $I_2 = [\lambda_1, \lambda_1 + \lambda_2]$ and $I_3 = (1 - \lambda_3, 1)$, and let $\sigma|_\zeta = \sigma^k$ for all $\zeta \in I_k$, $k = 1, 2, 3$.

Proof of Proposition 2: Let σ be a subgame perfect equilibrium in behavior strategies such that for some history h^t and random outcomes (ζ^t, x^t), at least one player strictly randomizes between accepting and rejecting the project x^t. We show how to modify the continuation strategy $\sigma|_{h^t}$ to obtain another subgame perfect equilibrium σ^P such that $\sigma^P(\zeta^t, x^t) \in \{0, 1\}^2$ (that is, σ^P makes only deterministic choices in period t) and $u(\sigma^P) = u(\sigma|_{h^t})$. By inductively repeating this process, we can "purify" strategy σ in every period to obtain a pure strategy subgame perfect equilibrium.

Without loss of generality, assume $h^t = h^0$, that σ randomizes in period 0 after the empty history and after some draw (ζ, x). Let $\Lambda = \{\lambda \in \mathbb{R}^3_+ \mid \sum \lambda_k = 1\}$. As we have noted, by Caratheodory's theorem, we can assume that there is $(\lambda_1, \lambda_2, \lambda_3) \in \Lambda$ with a corresponding collection of intervals $\{I_1, I_2, I_3\}$ and continuation strategies $\{\sigma^1, \sigma^2, \sigma^3\}$ such that $\sigma|_\zeta = \sigma^k$ for all $\zeta \in I_k$, $k = 1, 2, 3$.

Let (α, w) be the factorization of σ. Pick (ζ, x) such that $\sigma_i(\zeta, x) \notin \{0, 1\}$ for at least one player i (so i is strictly randomizing his choice after (ζ, x)). Note that

$$u(\sigma|_{(\zeta,x)}) = (1 - \delta)\alpha_1 \alpha_2 x + \delta[\alpha_1 \alpha_2 w^{1,1} + \alpha_1(1 - \alpha_2)w^{1,0}$$
$$+ (1 - \alpha_1)\alpha_2 w^{0,1} + (1 - \alpha_1)(1 - \alpha_2)w^{0,0}],$$

where the arguments (ζ, x) have been omitted in α_i $(i = 1, 2)$ and in w^a $(a \in \{0, 1\}^2)$. If $\alpha_1 \alpha_2 = 0$, the randomization is irrelevant; the project x is rejected for sure. In this case, replace $(\alpha(\zeta, x), w(\zeta, x))$ by $(\hat{\alpha}, \hat{w})$, where $\hat{\alpha} = (0, 0)$ (that is, make both players reject project x for sure) and

$$\hat{w}^{0,0} = \alpha_1(1 - \alpha_2)w^{1,0} + (1 - \alpha_1)\alpha_2 w^{0,1} + (1 - \alpha_1)(1 - \alpha_2)w^{0,0}$$
$$\hat{w}^\alpha = (0, 0) \quad \text{for } \alpha = (0, 1), (1, 0), (1, 1),$$

where again, the arguments (ζ, x) have been omitted in α_i and w^α. By construction, $\delta\hat{w}^{0,0} = u(\sigma|_{(\zeta,x)}) \geq (0, 0)$ and $(\hat{\alpha}, \hat{w})$ satisfies (IC_1) and (IC_2) at x. Moreover, $\hat{w}^{0,0} \in V$ as it is a convex combination of continuation values in V, and $\hat{w}^\alpha = (0, 0) \in V$ for $\alpha = (0, 1), (1, 0), (1, 1)$.

Now, suppose that $\alpha_1 \alpha_2 > 0$. This implies that $\alpha_i \in (0, 1]$ for $i = 1, 2$ and is interior for at least one player (that is, both players accept with positive probability after (ζ, x), with at least one player accepting with probability less than one). Assume that $x_1 \geq 0$ and $x_2 \leq 0$ (the case $x_1 \leq 0$ and $x_2 \geq 0$ is symmetric). Now we replace the factorization (α, w) at (ζ, x) with a lottery over two factorizations. We first specify the lottery and establish the resulting implications, and then return to how the lottery is accomplished.

With probability $\alpha_1 \alpha_2$, we replace $(\alpha(\zeta, x), w(\zeta, x))$ by the tuple $(\hat{\alpha}, \hat{w})$, where

$$\hat{\alpha} = (1, 1), \; \hat{w}^{1,1} = \alpha_1 w^{1,1} + (1 - \alpha_1)w^{0,1}, \text{ and } \hat{w}^\alpha = (0, 0) \text{ for } \alpha = (0, 0), (0, 1), (1, 0),$$

and with probability $1 - \alpha_1 \alpha_2$, replace $(\alpha(\zeta, x), w(\zeta, x))$ by the tuple $(\tilde{\alpha}, \tilde{w})$, where

$$\tilde{\alpha} = (0, 0), \quad \tilde{w}^{0,0} = \sum_\alpha \beta^\alpha w^\alpha, \quad \text{and} \quad \tilde{w}^\alpha = (0, 0) \text{ for } \alpha = (0, 1), (1, 0), (1, 1),$$

and

$$\beta^{1,1} = \frac{\alpha_2 \alpha_1(1 - \alpha_1)}{1 - \alpha_1 \alpha_2^*}, \; \beta^{0,1} = \frac{\alpha_2(1 - \alpha_1)^2}{1 - \alpha_1 \alpha_2}, \; \beta^{1,0} = \frac{(1 - \alpha_2)\alpha_1}{1 - \alpha_1 \alpha_2}, \; \beta^{0,0} = \frac{(1 - \alpha_2)(1 - \alpha_1)}{1 - \alpha_1 \alpha_2}.$$

Since $\tilde{w}^{0,0} \geq 0$, it is easy to see that $(\tilde{\alpha}, \tilde{w})$ satisfies (IC_1) and (IC_2). Since $x_1 \geq 0$, it is also clear $(\hat{\alpha}, \hat{w})$ satisfies (IC_1). Finally, $\alpha_i \in (0, 1]$ for $i = 1, 2$, implies that $(\alpha(\zeta, x), w(\zeta, x))$ satisfies (IC_2). Since $w_2^{1,0} \geq 0$ and $w_2^{0,0} \geq 0$, (IC_2) implies

$$(1 - \delta)x_2 + \delta\hat{w}_2^{1,1} = (1 - \delta)x_2 + \delta[\alpha_1 w_2^{1,1} + (1 - \alpha_1)w_2^{0,1}] \geq \delta[\alpha_1 w_2^{1,0} + (1 - \alpha_1)w_2^{0,0}] \geq 0.$$

Hence, $(\hat{\alpha}, \hat{w})$ also satisfies (IC_2).

The substitution of $(\alpha(\zeta, x), w(\zeta, x))$ by the lottery of $(\hat{\alpha}, \hat{w})$ with probability $\alpha_1 \alpha_2$ and $(\tilde{\alpha}, \tilde{w})$ with probability $1 - \alpha_1 \alpha_2$ preserves the expected value:

$$\alpha_1\alpha_2[(1 - \delta)x + \delta w^{1,1}] + \delta[\alpha_1(1 - \alpha_2)w^{1,0} + (1 - \alpha_1)\alpha_2 w^{0,1} + (1 - \alpha_1)(1 - \alpha_2)w^{0,0}]$$
$$= \alpha_1\alpha_2[(1 - \delta)x + \delta\hat{w}^{1,1}] + (1 - \alpha_1\alpha_2)\delta\tilde{w}^{0,0}.$$

So far, for each (ζ, x) we have been able to replace the continuation strategy $\sigma|_{(\zeta,x)}$ by a lottery between two continuation strategies $\hat{\sigma}|_{(\zeta,x)}$ and $\tilde{\sigma}|_{(\zeta,x)}$ (represented by their factorizations $(\hat{\alpha}(\zeta, x), \hat{w}(\zeta, x))$ and $(\tilde{\alpha}(\zeta, x), \tilde{w}(\zeta, x))$). We now construct σ^P that implements these lotteries.

Recall the definition of $(\lambda_1, \lambda_2, \lambda_3)$ and $\{I_1, I_2, I_3\}$ for period 0. For a fixed x and $k \in \{1, 2, 3\}$, note that $\alpha^k(x) = \alpha_1(\zeta, x)\alpha_2(\zeta, x)$ is constant for $\zeta \in I_k$. Pick any $\zeta_k \in I_k$, $k = 1, 2, 3$. For simplicity, let ℓ_k and h_k be such that $I_k = (\ell_k, h_k)$.

Construct the factorization (α^P, w^P) as follows:

$$(\alpha^P(\zeta, x), w^P(\zeta, x)) = \begin{cases} ((1, 1), \hat{w}(\zeta^k, x)) & \text{if} \quad \ell_k < \zeta \leq \ell_k + \alpha^k(x)(h_k - \ell_k) \\ ((0, 0), \tilde{w}(\zeta^k, x)) & \text{if} \quad \ell_k + \alpha^k(x)(h_k - \ell_k) < \zeta < h_k \end{cases}$$

The factorization (α^P, w^P) represents a subgame perfect equilibrium σ^P that duplicates the payoffs of σ and that makes only deterministic choices in period 0.   □

A.2 The Importance of Public Randomization, Section III.5

This section continues with the Example from Section III.5, exhibiting an equilibrium payoff in the set V that cannot be obtained without public randomization. The payoff in question is $(1/2, 7/6)$. This payoff is the half-half convex combination of $(1, 1)$ and $(0, 4/3)$, and hence can be achieved with the help of the public randomization device. We show that it cannot be achieved otherwise.

For the first step of the argument, we suppose that if the first project drawn is $(3, -1)$, then player 2 accepts with positive probability. Then an upper bound on 2's value must be

$$v_2 \leq (1/4)(-1 + (4/3)) + (1/4)(3 + 1) = 13/12.$$

The first term corresponds to the case in which $(3, -1)$ is the first project, and assumes that after accepting, player 2 gets the maximum continuation payoff. The second term corresponds to the case in which the first project is $(3, -1)$, in which case the highest payoff

for player 2 is generated by having player 1 accept (for an immediate payoff to player 2 of 3), but then player 1 must have a continuation payoff of at least 1, in order to satisfy the incentive constraint, and hence 2's continuation payoff can be at most 1. This gives an upper bound on 2's payoff of $13/12$, which falls short of $7/6$. It must then be that if the first project is $(3, -1)$, it is rejected.

For the second step, suppose that project $(-1, 3)$ is rejected in the first period. Then all projects are rejected in the first period, and the largest payoff available to player 2 is the discounted value of a second-period continuation value of $4/3$, or $(1/2)(4/3) = 2/3 < 7/6$, a contradiction.

For the third step, we now know the following about the strategies. If the first project is $(3, -1)$, it is rejected, with continuation payoffs that we will denote (z_1, z_2). If the first project is $(-1, 3)$ is accepted with positive probability $p_1 \in (0, 1]$, with acceptance giving rise to continuation payoffs (w_1, w_2) and rejection giving rise to continuation payoffs (y_1, y_2) for some $y_1 \leq w_1 - 1$ and y_2, with the payoff $y_1 \leq w_1 - 1$ ensuring that player 1 is willing to at least mix in accepting $(-1, 3)$. We can then construct an upper bound on player 1's equilibrium payoff, given by letting $z_1 = 4/3$ and deriving the inequality from the fact that $w_1 \leq 4/3$,

$$v_1 \leq (1/4)(w_1 - 1) + (1/4)(4/3) \leq 5/12,$$

a contradiction.

As a result, the payoff $(1/2, 7/6)$ is achievable only with public randomization.

A.3 Proof of Proposition 3

Proof: For simplicity, we offer the argument for the case in which $\partial V^\lambda = \{v^\lambda\}$. By self-generation,

$$v^\lambda = \max \int_A \lambda \cdot [(1 - \delta)x + \delta w(x)]d\mu(x) + \int_R \delta\lambda \cdot w(x)d\mu(x)$$

$$\text{s.t.} \quad (A, R) \text{ is a partition of } X, \ w : X \to V$$

$$(1 - \delta)x_i + \delta w_i(x) \geq 0 \quad \text{for all } x \in A \text{ and } i = 1, 2 \qquad (IC)$$

Since there is no incentive constraint for $x \in R$, to maximize the objective function it is optimal to set $w(x) = v^\lambda$ for all $x \in R$. Similarly, we want to choose $w(x)$ to maximize $\lambda \cdot w(x)$ when $x \in A$. When $x \in U$, we can optimally set $w(x) = v^\lambda$, but when $x \in C_i$, $w(x) = v^\lambda$ is not feasible and the choice of $w(x)$ is restricted by the (IC) constraint for player i. The (IC) is satisfied with equality when $w_i(x) = -\beta x_i$ and the optimal choice in this case is $w(x) = (-\beta x_i, \varphi_j(-\beta x_i))$.

Finally we turn to the optimal choice of partition (A, R). For any $x \in X$, the players can always reject the project, set $w(x) = v^\lambda$, and obtain a total welfare contribution of $\delta\lambda \cdot v^\lambda$ for this project. Hence, to improve total welfare, the players should accept the project only if $\lambda \cdot [(1 - \delta)x + \delta w(x)] \geq \delta\lambda \cdot v^\lambda$ for some feasible continuation value $w(x)$. Let

$$X_i^+ = \{x \in X \mid -v_i^\lambda \leq \beta x_i < 0\} \quad \text{and} \quad X_i^- = \{x \in X \mid -\bar{v}_i \leq \beta x_i < -v_i^\lambda\}.$$

A project $x \in X_i^+$ requires the consent of player i and his (IC) constraint can be satisfied by choosing $w(x) = v^\lambda$. Hence the project should be adopted if and only if $\lambda \cdot [(1-\delta)x + \delta v^\lambda] > \delta\lambda \cdot v^\lambda$, or equivalently, if $\lambda \cdot x > 0$ (in case $\lambda \cdot x = 0$ both decisions to adopt and reject the project are optimal). Similarly, a project $x \in X_i^-$ is adopted if and only if $\lambda \cdot [(1-\delta)x + \delta(-\beta x_i, \varphi_j(-\beta x_i))] > \delta\lambda \cdot v^\lambda$, or $\beta\lambda_j x_j > \lambda \cdot v^\lambda - \lambda_j\varphi_j(-\beta x_i)$. $\qquad\square$

A.4 Proof of Proposition 4

Proof: Let σ be a subgame perfect equilibrium in behavior strategies such that for some history h^t and outcomes (ζ^t, x^t), at least one player strictly randomizes between accepting and rejecting the project x^t. As we did in Proposition 2, we show how to modify the continuation strategy $\sigma|_{h^t}$ to obtain another subgame perfect equilibrium σ^P such that $\sigma^P(\zeta^t, x^t) \in \{0,1\}^2$ and $u(\sigma^P) = u(\sigma|_{h^t})$.

Without loss of generality, assume $h^t = h^0$, that σ randomizes in period 0, after the empty history and after some draw (ζ, x). Assume that there is a partition $(\lambda_1, \lambda_2, \lambda_3)$ and corresponding collection of intervals $\{I_1, I_2, I_3\}$ and continuation strategies $\{\sigma^1, \sigma^2, \sigma^3\}$ such that $\sigma|_\zeta = \sigma^k$ for all $\zeta \in I_k$, $k = 1, 2, 3$.

Let (α, m, w) be the factorization of σ. Since payoffs are linear in transfers, without loss of generality we can assume that m is deterministic. If m^a specified random transfers, we could replace m^a by its expected transfers without affecting any incentives or continuation values. Note that

$$
\begin{aligned}
u_1(\sigma|_{(\zeta,x)}) = (1-\delta)\big[&\alpha_1\alpha_2(x_1 + m_2^{1,1} - m_1^{1,1}) + \alpha_1(1-\alpha_2)(x_1 + m_2^{1,0} - m_1^{1,0}) \\
+ &(1-\alpha_1)\alpha_2(x_1 + m_2^{0,1} - m_1^{0,1}) + (1-\alpha_1)(1-\alpha_2)(x_1 + m_2^{1,1} - m_1^{1,1})\big] \\
+ \delta\big[&\alpha_1\alpha_2 w^{1,1} + \alpha_1(1-\alpha_2)w^{1,0} + (1-\alpha_1)\alpha_2 w^{0,1} + (1-\alpha_1)(1-\alpha_2)w^{0,0}\big]
\end{aligned}
$$

where the arguments (ζ, x) have been omitted in α_i, m_i^a and w_i^a for $i = 1, 2$ and $a \in \{0,1\}^2$. Since σ is a subgame perfect equilibrium, the players must be willing to make the required transfers at the end of the first period, that is

$$
(1-\delta)[m_i^a(\zeta, x) - m_j^a(\zeta, x)] \le \delta w_i^a(\zeta, x) \quad \text{for } i = 1, 2, \text{ and for all } a \in \{0,1\}^2. \quad (IC_T)
$$

If $\alpha_1\alpha_2 = 0$, the randomization is irrelevant; the project x is rejected for sure. In this case, replace $(\alpha(\zeta, x), m(\zeta, x), w(\zeta, x))$ by $(\hat\alpha, \hat m, \hat w)$, where $\hat\alpha = (0, 0)$ (that is, make both players reject project x for sure) and

$$
\begin{aligned}
\hat m^{0,0} &= \alpha_1(1-\alpha_2)m^{1,0} + (1-\alpha_1)\alpha_2 m^{0,1} + (1-\alpha_1)(1-\alpha_2)m^{0,0} \\
\hat w^{0,0} &= \alpha_1(1-\alpha_2)w^{1,0} + (1-\alpha_1)\alpha_2 w^{0,1} + (1-\alpha_1)(1-\alpha_2)w^{0,0} \\
\hat m^a &= \hat w^a = (0, 0) \quad \text{for} \quad a = (0, 1), (1, 0), (1, 1),
\end{aligned}
$$

By construction, $\delta\hat w^{0,0} \ge (0, 0)$ and thus each player prefers to reject x. Clearly $(\hat\alpha, \hat m, \hat w)$ preserves the original total expected value of $(\alpha(\zeta, x), m(\zeta, x), w(\zeta, x))$.

Moreover, $\hat{w}^{0,0} \in V$ as it is a convex combination of continuation values in V, and $\hat{w}^\alpha = (0,0) \in V$ for $\alpha = (0,1), (1,0), (1,1)$. Finally, (IC_T) implies that

$$(1-\delta)[\hat{m}_i^{0,0} - \hat{m}_j^{0,0}] \le \delta \hat{w}_i^{0,0} \quad i = 1, 2,$$

so the players are willing to make the transfers $\hat{m}^{0,0}$ at the end of the period.

Now, suppose that $\alpha_1 \alpha_2 > 0$. This implies that $\alpha_i \in (0,1]$ for $i = 1, 2$ and is interior for at least one player (that is, both players accept with positive probability after (ζ, x), with at least one player accepting with probability less than one). Assume that $x_1 > 0$ and $x_2 \le 0$ (the case $x_1 \le 0$ and $x_2 > 0$ is symmetric). Now we replace the factorization (α, m, w) at (ζ, x) with a lottery over two factorizations. We first specify the lottery and establish the resulting implications, and then return to how the lottery is accomplished.

With probability $\alpha_1 \alpha_2$, we replace $(\alpha(\zeta, x), m(\zeta, x), w(\zeta, x))$ by the tuple $(\hat{\alpha}, \hat{m}, \hat{w})$, where

$$\hat{\alpha} = (1,1),$$
$$\hat{m}^{1,1} = \alpha_1 m^{1,1} + (1-\alpha_1)m^{0,1}, \quad \text{and} \quad \hat{m}^a = (0,0) \text{ for } a = (0,0), (0,1), (1,0),$$
$$\hat{w}^{1,1} = \alpha_1 w^{1,1} + (1-\alpha_1)w^{0,1}, \quad \text{and} \quad \hat{w}^a = (0,0) \text{ for } a = (0,0), (0,1), (1,0),$$

and with probability $1 - \alpha_1 \alpha_2$, replace $(\alpha(\zeta, x), m(\zeta, x), w(\zeta, x))$ by the tuple $(\tilde{\alpha}, \tilde{m}, \tilde{w})$, where

$$\tilde{\alpha} = (0,0)$$
$$\tilde{m}^{0,0} = \sum_\alpha \beta_\alpha w^\alpha, \quad \text{and} \quad \tilde{m}^\alpha = (0,0) \text{ for } \alpha = (0,1), (1,0), (1,1)$$
$$\tilde{w}^{0,0} = \sum_\alpha \beta_\alpha w^\alpha, \quad \text{and} \quad \tilde{w}^\alpha = (0,0) \text{ for } \alpha = (0,1), (1,0), (1,1)$$

and

$$\beta_{1,1} = \frac{\alpha_2 \alpha_1 (1-\alpha_1)}{1 - \alpha_1 \alpha_2}, \ \beta_{0,1} = \frac{\alpha_2 (1-\alpha_1)^2}{1 - \alpha_1 \alpha_2}, \ \beta_{1,0} = \frac{(1-\alpha_2)\alpha_1}{1 - \alpha_1 \alpha_2}, \ \beta_{0,0} = \frac{(1-\alpha_2)(1-\alpha_1)}{1 - \alpha_1 \alpha_2}.$$

Since $\alpha_i(\zeta, x) > 0$ for $i = 1, 2$, player i is willing to accept x in $(\alpha(\zeta, x), m(\zeta, x), w(\zeta, x))$. That is,

$$(1-\delta)[x_1 + \alpha_2(m_2^{1,1} - m_1^{1,1}) + (1-\alpha_2)(m_2^{1,0} - m_1^{1,0})] + \delta[\alpha_2 w_1^{1,1} + (1-\alpha_2)w_1^{1,0}]$$
$$\ge (1-\delta)[\alpha_2(m_2^{0,1} - m_1^{0,1}) + (1-\alpha_2)(m_2^{0,0} - m_1^{0,0})] + \delta[\alpha_2 w_1^{0,1} + (1-\alpha_2)w_1^{0,0}] \quad (IC_1)$$

$$(1-\delta)[x_2 + \alpha_1(m_1^{1,1} - m_2^{1,1}) + (1-\alpha_1)(m_1^{0,1} - m_2^{0,1})] + \delta[\alpha_1 w_2^{1,1} + (1-\alpha_1)w_2^{0,1}]$$
$$\ge (1-\delta)[\alpha_1(m_1^{1,0} - m_2^{1,0}) + (1-\alpha_1)(m_1^{0,0} - m_2^{0,0})] + \delta[\alpha_1 w_2^{1,0} + (1-\alpha_1)w_2^{0,0}]. \quad (IC_2)$$

Also, observe that (IC_T) implies that

$$(1-\delta)[\hat{m}_i^{1,1} - \hat{m}_j^{1,1}] \le \delta \hat{w}_i^{1,1} \quad \text{and} \quad (1-\delta)[\tilde{m}_i^{0,0} - \tilde{m}_j^{0,0}] \le \delta \tilde{w}_i^{0,0} \quad i = 1, 2.$$

That is, the players are willing to make the transfers at the end of the first period in $(\hat{\alpha}, \hat{m}, \hat{w})$ and $(\tilde{\alpha}, \tilde{m}, \tilde{w})$ respectively (the other transfers are all $(0,0)$).

We now verify that the players are willing to accept x in $(\hat{\alpha}, \hat{m}, \hat{w})$ and to reject x in $(\tilde{\alpha}, \tilde{m}, \tilde{w})$. Note that when the players choose $a \neq (1,1)$ in $(\hat{\alpha}, \hat{m}, \hat{w})$ or $a \neq (0,0)$ in $(\tilde{\alpha}, \tilde{m}, \tilde{w})$, they get a continuation value of $(0,0)$ (no transfers followed by autarky). The (IC_T) constraint for $(\tilde{\alpha}, \tilde{m}, \tilde{w})$ also implies that the players (weakly) prefer to reject project x. Now, $x_1 \geq 0$, (IC_T) and (IC_2) imply that

$$(1-\delta)[x_1 + \hat{m}_2^{1,1} - \hat{m}_1^{1,1}] + \delta \hat{w}_1^{1,1}$$
$$= \alpha_1[(1-\delta)(x_1 + m_2^{1,1} - m_1^{1,1}) + \delta w_1^{1,1}] + (1-\alpha_1)[(1-\delta)(m_2^{0,1} - m_1^{0,1}) + \delta w_1^{0,1}]$$
$$\geq \alpha_1(1-\delta_1)x_1 \geq 0 \quad \text{and}$$

$$(1-\delta)[x_2 + \hat{m}_1^{1,1} - \hat{m}_2^{1,1}] + \delta \hat{w}_2^{1,1}$$
$$\geq (1-\delta_1)[\alpha_1(m_1^{1,0} - m_2^{1,0}) + (1-\alpha_1)(m_1^{0,0} - m_2^{0,0})] + \delta[\alpha_1 w_2^{1,0} + (1-\alpha_1)w_2^{0,0}] \geq 0$$

Hence, players 1 and 2 (weakly) prefer to accept x with $(\hat{\alpha}, \hat{m}, \hat{w})$.

The substitution of $(\alpha(\zeta, x), m(\zeta, x), w(\zeta, x))$ by the lottery of $(\hat{\alpha}, \hat{m}, \hat{w})$ with probability $\alpha_1 \alpha_2$ and $(\tilde{\alpha}, \tilde{m}, \tilde{w})$ with probability $1 - \alpha_1 \alpha_2$ preserves the expected value. So far, for each (ζ, x) we have been able to replace the continuation strategy $\sigma|_{(\zeta,x)}$ by a lottery between two continuation strategies $\hat{\sigma}|_{(\zeta,x)}$ and $\tilde{\sigma}|_{(\zeta,x)}$ (represented by their factorizations $(\hat{\alpha}, \hat{m}, \hat{w})$ and $(\tilde{\alpha}, \tilde{m}, \tilde{w})$). We then construct σ^P as in Proposition 2. $\square$

A.5 Proof of Lemma 3

We now provide the factorization of a pure strategy (without public randomization) subgame perfect equilibrium σ with value $v^* = (\bar{T}/2, \bar{T}/2)$. The factorization of σ is specified by (A, R, m, w). In σ, transfers are only made by player i when $x \in A$ and $x_i > 0$. That is, $m(x) = (0,0)$ for all $x \in R$, and $m_i(x) = 0$ for $x \in A$ if $x_i \leq 0$.

Let $A = A_{\bar{T}}$, $R = X \backslash A$, and

$$m_i(x) = \max\left\{0, -\left[x_j + \frac{\bar{T}}{2\beta}\right]\right\} \text{ if } x \in A \text{ and } x_i > 0$$
$$w(x) = v^* \text{ for all } x \in X.$$

Assume $x \in A$ with $x_i > 0$ and $x_j < 0$. If $(1-\delta)x_j + \delta\bar{T}/2 \geq 0$, then $m_i(x) = 0$ and since $x_i > 0$ and $\bar{T} \geq 0$, $(1-\delta)x_i + \delta\bar{T}/2 \geq 0$ and both players prefer to accept project x. If $(1-\delta)x_j + \delta\bar{T}/2 < 0$, then $m_i(x) > 0$ and by definition,

$$(1-\delta)[x_j + m_i(x)] + \delta\bar{T}/2 = 0,$$

so player j weakly prefers to accept project x. Also, since $\beta x_j \geq -\bar{T}$,

$$-(1-\delta)m_i(x) + \delta\bar{T}/2 = (1-\delta)[x_j + \bar{T}/(2\beta)] + \delta\bar{T}/2 = (1-\delta)x_j + \delta\bar{T} \geq 0,$$

and player i is willing to make the transfer $m_i(x)$ after he has collected x_i. This also implies that i strictly prefers to accept the project because $x_i > 0$:

$$(1 - \delta)[x_i - m_i(x)] + \delta\bar{T}/2 > 0.$$

All this establishes that (A, R, m, w) is incentive compatible.

Finally, we show that the expected value of the factorization (A, R, m, w) is indeed v^*. The total expected surplus of the factorization is

$$(1 - \delta) \int_A (x_1 + x_2)d\mu(x) + \delta\bar{T} = \bar{T},$$

and by symmetry, each player gets $\bar{T}/2$. $\qquad\square$

A.6 Proof of Proposition 6

Pick an arbitrary pair of functions $\bar{x}_1, \bar{x}_2 : (0, \infty) \to [-T^*/\beta, 0]$. The existence of a subgame perfect equilibrium σ that accepts projects (i, x) if and only if $x_j \geq \bar{x}_j(x_1)$, $i = 1, 2$, is implied by the construction of an optimal subgame perfect equilibrium we present below.

Fix $\lambda \in \mathbb{R}^2_{++}$. Assume σ is a subgame perfect equilibrium with factorization (α, m, w) that satisfies the proposed project acceptance rule. Suppose, for example, that $(i, x_i) = (1, x_1)$ and assume that $w^A(1, x_1) \notin V^{(1,1)}$. We now construct another factorization associated with a subgame perfect equilibrium that strictly increases the expected conditional welfare given x_1. As usual, we drop the variables (i, x_i) from relevant functions below. Let $\hat{w}^R = w^R$, $\hat{v}^R = v^R$, $\hat{w}^A \in V^{(1,1)}$, and $\hat{v}^A = (v_1^A + S, v_2^A) \in D(\hat{T}^A) = D(T^*)$, where $S = \delta[(\hat{w}_1^A + \hat{w}_2^A) - (w_1^A + w_2^A)] > 0$ (so $\hat{\Delta}^R = [\hat{v}_2^R - \delta\hat{w}_2^R]/(1 - \delta)$ and $\hat{\Delta}^A = [\hat{v}_2^A - \delta\hat{w}_2^A]/(1 - \delta)$). Clearly $(\alpha, \hat{m}, \hat{w})$ is the factorization of another subgame perfect equilibrium $\hat{\sigma}$ (it satisfies all the necessary incentive constraints: $\hat{v}^R \in D(\hat{T}^R)$, $\hat{v}^A \in D(\hat{T}^A)$, and $\hat{v}_2^A - \hat{v}_2^R = -(1 - \delta)\bar{x}_2(x_1)$), and increases the conditional expected welfare after x_1 by $\lambda_1(1 - F_2(\bar{x}_2|x_1))S$. A similar welfare improving modification can be made if $w^R(i, x_i) \notin V^{(1,1)}$. Thus, if $w^A(i, x_i) \notin V^{(1,1)}$ or $w^R(i, x_i) \notin V^{(1,1)}$ for x_i in a set of positive measure, the strategy σ could be modified accordingly in the same set, obtaining a subgame perfect equilibrium $\hat{\sigma}$ that has a strictly higher welfare value $\lambda \cdot u(\hat{\sigma})$. Therefore, σ maximizes $\lambda \cdot u(\sigma)$ among all subgame perfect equilibria with an acceptance rule $(\bar{x}_1, \bar{x}_2)$ only if $w^R(i, x_i) \in V^{(1,1)}$ and $w^A(i, x_i) \in V^{(1,1)}$ almost surely for $x_i > 0$, $i = 1, 2$.

Let α be such that $\alpha_i(i, x) \equiv 1$ and $\alpha_j(i, x) = 1$ if $x_j \geq \bar{x}_2(x_1)$ and 0 otherwise. Pick any $w = (w^R, w^A)$ such that $w^R(i, x_i), w^A(i, x_i) \in V^{(1,1)}$ for all (i, x_i). Hence $w_1^R(i, x_i) + w_2^R(i, x_i) = w_1^A(i, x_i) + w_2^A(i, x_i) = T^*$ for all (i, x_i). We now construct optimal transfers $m = (m^R, m^A)$ such that (α, m, w) is the factorization of a subgame perfect equilibrium σ that maximizes $\lambda \cdot u(\sigma)$.[1] Fix (i, x_i). As argued above, transfers (m^R, m^A) are uniquely

---

[1] Any continuation values $w^R(i, x_i), w^A(i, x_i) \in V^{(1,1)}$ are possible. If $V^{(1,1)}$ is not a singleton, all such continuation values lead to the same interim continuation value sets $D(T^*)$. So while different choices lead to different optimal subgame perfect equilibria (because they provide different continuation values), they will all share the same interim continuation values and therefore the same total expected value.

defined by continuation values (v^R, v^A) and conversely. To satisfy (VIC) we need to choose $v^R, v^A \in D(T^*)$ such that $v_j^A - v_j^R = -(1 - \delta)\bar{x}_j(x_i)$. When $-\beta\bar{x}_j(x_i) = T^*$, $(v_i^R, v_j^R) = (0, \delta T^*)$ and $(v_i^A, v_j^A) = (\delta T^*, 0)$ are the only feasible choices. These choices correspond to $\Delta^R = -w_j^R/\beta$ and $\Delta^A = w_i^A/\beta$, which in turn correspond to $(m_i^R, m_j^R) = (0, w_j^R/\beta)$ and $(m_i^A, m_j^A) = (w_i^A/\beta, 0)$. When $-\beta\bar{x}_j(x_i) < T^*$, there is a range of possible values for v^R and v^A such that $v_j^A - v_j^R = -(1 - \delta)\bar{x}_j(x_i)$. If $\lambda_i > \lambda_j$, these values should be chosen to maximize player i's expected value. In this case, $(v_i^A, v_j^A) = (\delta T^* + (1 - \delta)\bar{x}_j(x_i), -(1 - \delta)\bar{x}_j(x_i))$ and $(v_i^R, v_j^R) = (\delta T^*, 0)$, so $\Delta^A = -w_j^A/\beta - \bar{x}_j(x_i)$ and $\Delta^R = -w_j^R/\beta$. If $\lambda_i < \lambda_j$, these values should be chosen to maximize player j's expected value, and $(v_i^A, v_j^A) = (0, \delta T^*)$ and $(v_i^R, v_j^R) = (-(1 - \delta)\bar{x}_j(x_i)), \delta T^* + (1 - \delta)\bar{x}_j(x_i))$. $\square$

References

Abreu, Dilip and David Pearce and Ennio Stacchetti (1990), "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica* **58**, 1041–1064.

Aliprantis, Charalambos D. and Kim C. Border (2006), *Infinite Dimensional Analysis*, Third edition, Springer.

Athey, Susan and Kyle Bagwell (2001), "Optimal Collusion with Private Information," *RAND Journal of Economics* **32**(2), 428–465.

Athey, Susan, Kyle Bagwell and Chris Sanchirico (2004), "Collusion and Price Rigidity," *Review of Economic Studies* **71**(2), 317–349.

Abdulkadiroğlu, Atila and Kyle Bagwell (2013), "Trust, Reciprocity and Favors in Cooperative Relationships," *American Economic Journal: Microeconomics* **5**(2), 213–259.

Baker, George, Robert Gibbons and Kevin J. Murphy (2002), "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics* **117**(1), 39–84.

Bernstein, Lisa (1992), "Opting Out of the Legal System: Extralegal Contractual Relations in the Diamond Industry," *Journal of Legal Studies* **21**(1), 115-157.

Bramoullé, Yann and Sanjeev Goyal (2013), "Favoritism," Aix-Marseille University and University of Cambridge.

Cosmides, Leda and John Tooby (1992a), "The Psychological Foundations of Culture," in *The Adapted Mind*, edited by Jerome H. Barkow, Leda Cosmides and John Tooby, Oxford University Press, 19–136.

Cosmides, Leda and John Tooby (1992a), "Cognitive Adaptations for Social Exchange," in *The Adapted Mind*, edited by Jerome H. Barkow, Leda Cosmides and John Tooby, Oxford University Press, 163–228.

Basu, Kuashik (1983), "On Why We Do Not Try to Walk Off without Paying After a Taxi Ride," *Economic and Political Weekly* **18**, 2011–2012.

Bowen, T. Rene, David M. Kreps and Andrzej Skrzypacz (2013), "Rules with Discretion and Local Information," *Quarterly Journal of Economics* **128**, 1273–1320.

Ellickson, Robert C. (1991), *Order without Law: How Neighbors Settle Disputes*, Harvard University Press, Cambridge, Massachusetts.

Fiske, Alan Page (1992), "The Four Elementary Forms of Sociality: Framework for a Unified Theory of Social Relations," *Psychological Review* **99**(4), 689–723.

Friedman, James W. (1971), "A Noncooperative Equilibrium for Supergames," *Review of Economic Studies* **38**(1), 1–12.

Fudenberg, Drew and David Levine (1994). "Efficiency and Observability with Long-run and Short-run Players," *Journal of Economic Theory* **62**(1), 103–135.

George, Christine (2009), *The Psychological Contract*, Open University Press.

Greif, Avner (1997), "Microtheory and Recent Developments in the Study of Economic Institutions through Economic History," in *Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress*, edited by David M. Kreps and Kenneth F. Wallis, Cambridge University Press, 79–113.

Greif, Avner (2005), *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*, Cambridge University Press.

Greif, Avner, Paul Milgrom and Barry R. Weingast (1994), "Coordination, Commitment, and Enforcement: The Case of the Merchant Guild," *Journal of Political Economy* **102**(4), 745–776.

Hauser, Christine and Hugo A Hopenhayn (2005), "Trading Favors: Optimal Exchange and Forgiveness," University of Rochester and University of California, Los Angeles.

Leo, Gregory C. (2014), "Taking Turns," University of California Santa Barbara.

Levin, Jonathan (2003), "Relational Incentive Contracts," *American Economic Review* **93**(3), 835–857.

Macaulay, Stewart (1963), "Non-Contractual Relations in Business: A Preliminary Study," *American Sociological Review* **28**(1), 55–67.

Mailath, George J. and Larry Samuelson (2006), *Repeated Games and Reputations: Long-Run Relationships*, Oxford University Press.

Mailath, George J. and Larry Samuelson (20013), "Reputations in Repeated Games," in *Handbook of Game Theory, Volume 4*, edited by H. Peyton Young and Shmuel Zamir, Elsevier Press, 165–238.

Miller, David A. (2009), "Attainable Payoffs in Repeated Games with Interdependent Private Information," University of Michigan.

Pearce, David G. and Ennio Stacchetti (1998), "The Interaction of Implicit and Explicit Contracts in Repeated Agency," *Games and Economic Behavior* **23**(1), 75–96.

Mobius, Markus (2001), "Trading Favors."

Paley, Thomas M. (1985), "Avoiding Regulatory Constraints: Contracting Safeguards and the Role of Informal Agreements," *Journal of Law, Economics and Organization* **1(1)**, 155–175.

Pasquale, Denise and Edward L. Glaeser (1998), "Incentives and Social Capital: Are Homeowners Better Citizens?," *Journal of Urban Economics* **45**(2), 354–384.

Putnam, Robert D. (1995), *Bowling Alone*, Simon and Schuster.

Putnam, Robert D., editor (2002), *Democracies in Flux: The Evolution of Social Capital in Contemporary Society*, Oxford University Press.

Richman, Barak D. (2006), "How Communities Create Economic Advantage: Jewish Diamond Merchants in New York," *Law and Social Inquiry* **31**(2), 383–420.

Rousseau, Denise M. (1995), *Psychological Contracts in Organizations*, Sage Publications.

Samuelson, Larry (2006), "The Economics of Relationships," *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress, volume 1*, edited by Richard Blundell and Torsten Person, Cambridge University Press, 136–185.

Von Neumann, John and Oskar Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton University Press.

Watson, Joel (1999), "Starting Small and Renegotiation," *Journal of Economic Theory* **85**(1), 52–90.

Watson, Joel (2002), "Starting Small and Commitment" *Games and Economic Behavior* **38**(1), 176–199.