ECONOMIC RESEARCH REPORTS

Bargaining, Independence, and the Rationality of Fair Division

by **Giuseppe Lopomo**
and
Efe A. Ok

RR# 98-18 **May 1998**

C.V. STARR CENTER FOR APPLIED ECONOMICS

NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, NY 10003-6687

Bargaining, Interdependence, and the Rationality of Fair Division[1]

Giuseppe Lopomo[2]

and

Efe A. Ok[3]

May 1998

[1]The authors would like to thank J.P. Benoît, Gary Bolton, Kalyan Chatterjee, Faruk Gül, Levent Koçkesen, George Mailath, David Pearce, Ariel Rubinstein and Ennio Stacchetti for helpful comments. Of course, the usual disclaimer applies. Support from the C.V. Starr Center for Applied Economics at New York University is also gratefully acknowledged.

[2]Stern School of Business, New York University, and Department of Economics, University of Michigan. E-mail: glopomo@stern.nyu.edu; and glopomo@econ.lsa.umich.edu.

[3]Department of Economics, New York University, 269 Mercer St., New York NY 10003. E-mail: okefe@fasecon.econ.nyu.edu.

Abstract

We consider two person bargaining games with interdependent preferences, with and without bilateral incomplete information. We show that, both in the ultimatum game and in the two-stage alternating-offers game, our equilibrium predictions are fully consistent with all robust experimental regularities which falsify the standard game theoretic model: occurrence of disagreements, disadvantageous counteroffers, and outcomes that come close to the equal split of the pie. In the context of infinite horizon bargaining, the implications of the model pertaining to fair outcomes is even stronger. In particular, the Coase property in our case generates "almost" 50-50 splits of the pie, almost immediately. The present approach thus provides a positive theory for the frequently encountered phenomenon of the 50-50 division of the gains from trade. We also show that the potential interdependence of preferences entails the emergence of "near-fair" divisions in bargaining settlements.

Keywords: Ultimatum bargaining, fair division, Coase conjecture

Journal of Economic Literature **Classification number:** C78.

1 Introduction

Fair divisions of gains from trade are commonly observed in daily life. They occur even in the case of bargaining among two asymmetrically placed individuals one of whom holds a clear strategic advantage over another. A striking example of this phenomenon is provided by Guth, Schmittberger and Schwarz (1982) who studied experimentally the two-player *ultimatum game*, in which one player, A, proposes a division of some fixed monetary amount, and his opponent, B, either accepts or rejects. Guth et al. (1982) found that the average proposal by subjects in the role of player A was in the neighborhood of $(70\%, 30\%)$. Moreover, about 20 percent of proposals were rejected. They interpreted this evidence as follows:

> ... subjects often rely on what they consider a *fair* or justified result. Furthermore, the ultimatum aspect cannot be completely exploited, since subjects do not hesitate to punish if their opponents ask for "too much." [Guth et al. (1982), p. 389]

What is paradoxical about all this is that the unique subgame perfect equilibrium of this game has player A claiming the entire surplus for himself and player B agreeing to this. The findings of Guth et al. (1982) thus casts serious doubts on the ability of game-theoretic predictions to match observed behavior.

The anomalies encountered in the ultimatum bargaining experiments have motivated a large number of experimental and game-theoretical studies. (See Camerer and Thaler (1995) and Roth (1995) for excellent surveys of the related experimental literature.) Particularly related to our present inquiry are the important papers of Ochs and Roth (1989) and Bolton (1991), who ran large experimental sessions in which discount factors and subjects' experience were varied along with other structural parameters. Among others, these two studies uncovered three major robust empirical regularities in alternating-offers bargaining games. First, proposed divisions tend to move away from equilibrium divisions toward the 50-50 division; the actual outcomes are "more fair" than the usual prediction. Second, rejections, which should never be observed on the equilibrium path, occur in significant numbers. Third, more often than not, subjects who reject an offer make a *disadvantageous counteroffer*, that is, after rejecting a proposal that would leave them with x dollars, they propose a new division that spares them less than x dollars.[1] Again, all of these observations are in sharp contrast with the standard game-theoretic predictions.

One way of interpreting these findings is to argue that pure expected profit maximization cannot be the only criterion guiding the choices of the players in bargaining games. For instance,

[1] "Some plausible, but flawed" explanations of these regularities, based on subjets' confusion or lack of experience, or even accidental designers' influence, were considered and convincingly ruled out by Bolton (1991).

Ochs and Roth suggested that a notion of "fairness" may be influencing the subjects behavior: more precisely, "fairness" leads players to reject "insultingly low" offers, and the anticipation of this leads to proposing interior divisions.[2] Similarly, Bolton (1991) has proposed the idea that "fairness" does guide a subject's behavior, but *only* when he is getting *less* surplus than his opponent. To account for rejections, and disadvantageous counteroffers, Bolton also considered a model in which players have incomplete information about their opponents' preferences, and engage in a myopic learning algorithm about the distribution of their opponents' preferences. This model is, however, built on restrictive assumptions.[3]

In this paper we examine the implications of the possibility that each player's utility depends not only on her absolute level of earnings, but also on her *relative* share of the total surplus. In the resulting bargaining model, each player's private 'type' is referred to as "independent" if and only if his utility depends exclusively on her absolute earnings, and "(negatively) interdependent" otherwise. The latter kind of preferences relates to the time-honored relative income hypothesis, and is thus much studied in various branches of economic literature (see Frank (1987), Ok and Koçkesen (1997), and references cited therein). In the context of bargaining, on the other hand, a negatively interdependent individual may be thought of as a *competitive* player who does not wish to "lose the game" unless he is sufficiently compensated for it. Such preference structures are particularly appealing in the present context, for, as we shall demonstrate in the sequel, they square nicely with the presence of the "fear of punishment/rejection" a bargainer may have. As the above quotation from Guth et al. (1982) states, and as many other experimental studies demonstrate, much explanatory power may be gained by explicitly introducing the notion of "fear of rejection" to the bargaining models. One major objective of the present paper is to show that this can be accomplished by means of negatively interdependent preferences.[4]

Among the empirical regularities noted above, we are foremost interested in the one that pertains to the unexpectedly fair divisions of the pie. Apart from the experiments, the real life transactions that frequently result in the 50-50 split motivates our interest in this phenomenon. There is, of

[2] The idea that the players are simply trying to be "fair," did not pass the 'dictator game' test satisfactorily. In this game, one subject single-handedly decides how to divide a given surplus between himself and another subject. In the related experiments, the deciding subject has most often claimed the entire surplus for himself (Forsythe, Horowitz, Savin and Sefton, 1994).

[3] For instance, Bolton assumes that the true distribution of players' types is such that the game has a unique equilibrium, in which the initial offer is always accepted.

[4] Alternative models that perturb the individual utility functions are considered by Daughety (1994), Kirschteiger (1994) and Levine (1996). Several other psychological-economic approaches that provide compelling alternatives to the standard expected profit maximization paradigm are outlined in Rabin (1998).

course, a literature that studies the emergence of the equal division as a focal point, but few studies examined the conditions under which the 50-50 split is actually an equilibrium outcome of a bargaining game played among *rational* individuals.[5] In this paper, we aim to show that very egalitarian outcomes can indeed be sustained in bargaining equilibrium, provided that bargainers have *possibly* interdependent preferences that leads to a "fear of rejection" on the part of the players. When one models the bargaining environment in such a way that the level of competitiveness (interdependence) of the players are their private information, it also becomes possible to "explain" the occurrence of disagreements and disadvantageous counteroffers.

The paper is organized as follows. In Section 2 we specify the assumptions on the players' preferences and the bargaining environment, and briefly comment on the significance of the interdependence assumption both for bargaining theory and for other areas of economic analysis. In Section 3, we show that the ultimatum game in which players' preferences may be interdependent has a unique equilibrium, in which the proposed division can be rejected, and is always between the equilibrium with independent preferences and the 50-50 division.

In Section 4, we analyze two-stage alternating offer bargaining games, and prove the existence of perfect Bayesian equilibria in which both rejections and disadvantageous counteroffers are possible. This result provides a rational explanation for the occurrence of disadvantageous counteroffers, say, in the experiments of Ochs and Roth (1989). We also prove that disadvantageous counteroffers cannot occur in a model in which the monetary surplus remains constant across time periods, and the players' *utility* is discounted through time. We therefore claim that disadvantageous counteroffers can only be observed in the experiments in which, for practical reasons, it is *the size of the pie* that is discounted.

In Section 5, we focus on the implications of (possibly) interdependent preferences in bargaining, by examining certain scenarios with infinite horizon. It turns out that an interesting case in which players have interdependent preferences is essentially identical to the "gap case" of the much studied buyer-seller bargaining model, with risk neutral players. Therefore, all the results known in that context apply immediately to ours. In particular, the famous "Coase conjecture," established by Gul, Sonnenschein and Wilson (1986), implies the following in the present setting: as the frequency of offers increases arbitrarily, even the smallest doubt in the mind of the player who makes all the

[5] Most papers that examine the origin of fair outcomes in bargaining games are evolutionary in nature. Young (1993), for instance, studies an evolutionary model of the Nash demand game played between two populations who learn adaptively. Young shows that the equal split can be the unique stable division depending on the nature of the expected utility function. (See Ellingsen, 1997, for a similar analysis.) Bolton (1997), on the other hand, provides an alternative bargaining game at least one (limit) evolutionarily stable equilibrium of which results in the equal split. (A modification of the model ensures the uniqueness of this equilibrium.)

offers about his opponent's degree of interdependence being relatively high, induces him to propose a division relatively close to the 50-50 split, almost immediately. We believe that this result is but an important step towards providing a rational theory of equal division.

Finally, in Section 6, we adopt a mechanism design approach, and examine the conditions under which the potential interdependence of preferences would enforce the emergence of fair outcomes in bargaining settlements. The paper concludes with a brief discussion of directions for future research.

2 The Bargaining Environment and Interdependent Preferences

We consider a bargaining environment in which two players try to agree on how to divide a pie of size $2m$, where $m > 1$. In case of disagreement, each player receives $\varepsilon \in (0, m)$.[6] Thus, the size of the pie, net of the bargainers' holdings, is $2m - 2\varepsilon$. Equivalently, we can assume that the players have an initial level of wealth ε and are bargaining over a pie of size $2m - 2\varepsilon$, which is wasted if disagreement occurs. Without loss of generality, we let $\varepsilon = 1$, so that the set of all feasible divisions of the pie is

$$X \equiv \{(x_A, x_B) \geq (1, 1) : x_A + x_B \leq 2m\},$$

while the set of all efficient divisions is given by

$$Y \equiv \{(x_A, x_B) \in X : x_A + x_B = 2m\}.$$

In this paper, we examine the consequences of the *possibility* that the bargainers' welfare depend not only on the absolute gains that they may achieve through bargaining, but also on the *relative* sizes of the slices of the pie that they get. More precisely, we assume that the utility function of each player $i = A, B$ on X, denoted u_i, is of the following form:

$$u_i(x_A, x_B) \;=\; V_i\left(x_i, \frac{x_i}{\bar{x}}\right) \tag{1}$$

where V_i is a continuous and strictly increasing function on $\mathbb{R}^2_+$, and $\bar{x} \equiv (x_A + x_B)/2$. Thus, we model the preferences as (possibly) being *negatively interdependent*: player i cares not only about her share of the pie x_i, but also about how x_i compares with the *average* level of earnings $\bar{x}$.[7] For simplicity, we normalize V_i so that the disagreement outcome gives zero utility to each player: $V_i(1, 1) = 0$.

[6] We require $\varepsilon > 0$ in order to avoid the indeterminate form $0/0$.

[7] This particular utility representation of interdependent preferences is axiomatically characterized by Ok and Kockesen (1997).

A natural index for the degree to which player i is interdependent is the minimum share of the pie that she must be given in order for her to be at least as well off as she would be at the disagreement outcome. That is, the *reservation amount*, r_i, which is uniquely determined by the equation[8]

$$V_i\left(r_i, \frac{r_i}{m}\right) \equiv 0, \tag{2}$$

indicates how interdependent player i really is. We say that individuals A and B are *equally interdependent* (or, equally competitive) if $r_A = r_B$.

The presence of negative interdependence is the main feature that distinguishes this paper from the existing literature on bargaining. We contend that negative interdependence is a meaningful postulate, the significance of which extends well beyond bargaining theory. Indeed, there is a voluminous literature examining the reasons behind and the consequences of Duesenberry's *relative income hypothesis* which holds that economic agents care not only about their absolute level of earnings, but also about how their earnings compare with those of others. There is now abundant empirical and experimental evidence in support of this hypothesis; see Frank (1987a), Clark and Oswald (1996), and references cited therein.

In bargaining situations, the possibility that players may have interdependent preferences can also account for the *fear of rejection* that a bargainer may have in case she makes an offer that favors herself disproportionately. It is not unreasonable to think of a bargainer A who does not propose an allocation that spares the entire pie to herself as reasoning along the following lines: "If I offer an allocation that would leave B with a very small share of the pie, she may get upset and reject my offer. My share of the pie should then be still higher than hers, but not so large as to offend her." If we assume that the preferences of B are as in (1), it is easy to see that the gist of this reasoning will influence A's behavior when it is her turn to make a proposal. Introducing interdependent preferences into bargaining models may then be thought of as modeling (in reduced-form) the notion of "fear of rejection," which commands considerable experimental support (Bolton and Zwick, 1990).

There are only few studies that investigate how the main insights of the theory of bargaining would be affected by the presence of interdependent preferences. These studies, most notably Bolton (1991) and Kirschteiger (1996), focus only on specific formulations of interdependent preferences, which are tailored to the experimental bargaining games that they analyze.[9] Therefore, while

[8] Since $m > 1$, we have

$$V_i\left(1, \frac{1}{m}\right) < V_i(1,1) = 0 < V_i\left(m, 1\right).$$

Therefore, by strict monotonicity and continuity of V_i, there exists a unique $r_i \in (1, m)$ such that (2) is satisfied.

[9] For instance, the preferences considered by Bolton (1991) are not well-defined in an infinite horizon bargaining

these authors demonstrate nicely that the idea of interdependent preferences passes important preliminary tests in explaining experimental evidence, it is still not clear at this point what general insight can be gained by introducing such preferences into bargaining theory. As noted earlier, the aim of this paper is to make use of the general utility specification given in (1), and argue that the possibility of interdependent preferences can explain commonly observed regularities such as the occurrence of disagreements, disadvantageous counteroffers, and the outcomes that come very close to the equal split of the pie.

We proceed by reviewing some other rudiments of bargaining theory. The *bargaining problem* (in the sense of Nash) associated with the abstract setting specified above is $(\mathcal{U}, d)$ where the utility possibility set $\mathcal{U}$ is

$$\mathcal{U} \equiv \left\{ \left(V_A \left(x_A, \frac{x_A}{\bar{x}} \right), V_B \left(x_B, \frac{x_B}{\bar{x}} \right) \right) \;:\; x \in X \right\}$$

and the disagreement point d is $(0,0)$. The *bargaining set* of the problem $(\mathcal{U}, d)$, denoted $B(\mathcal{U})$, is defined as the set of all Pareto optimal and individually rational utility allocations in $\mathcal{U}$, that is,

$$B(\mathcal{U}) \equiv \left\{ \left(V_A \left(x_A, \frac{x_A}{\bar{x}} \right), V_B \left(x_B, \frac{x_B}{\bar{x}} \right) \right) \geq (0,0) : x \in Y \right\}.$$

It is important to note that the bargaining set of a situation with interdependent preferences is smaller than that of a usual bargaining situation with independent preferences. As Figure 1 illustrates, this is due to the individual rationality constraint.

For future reference, we note that $\mathcal{U}$ is compact and has nonempty interior in $\mathbb{R}^2_{++}$. Given only the assumptions on V that we have made so far (namely, continuity and monotonicity), it is not possible to say more on the structure of $\mathcal{U}$ and $B(\mathcal{U})$. In particular, $\mathcal{U}$ need not be convex (even when both u_A and u_B are concave in their first argument). If each u_i is a concave function, however, then $\mathcal{U}$ is convex. Moreover, it can be shown that $\mathcal{U}$ is, in general, not "too non-convex", in the sense that $B(\mathcal{U})$ is always a connected set.

3 The Ultimatum Game

In this section we study the classic ultimatum game, with interdependent preferences. The rules of the game are simple: player A proposes an allocation $x \in Y$, and player B either accepts or rejects A's offer. If player B accepts, the proposed allocation is realized. Otherwise, both players receive the disagreement outcome $(1,1)$.

---

model (with discounting) where the size of the pie remains *unchanged* through time. On the other hand, while the model of "envious preferences" of Kirchsteiger (1996) is more general (for it allows the discounting of utilities) and is closer to our model in spirit, it does not allow for bargaining games with incomplete information.

Before proceeding, we should note that the ultimatum game has played a central role in the recent experimental work on bargaining. This is largely due to the fact that the standard equilibrium prediction, based on the assumption that bargainers are money-maximizers, differs substantially from experimental findings. Indeed, if the players care only about their own earnings, the unique subgame perfect equilibrium of the game dictates that player A should claim the whole pie (net of what player B is guaranteed by design). However, in experiments it is consistently observed that subjects in the role of player A tend to propose divisions of the pie that spare their opponents significantly more than the feasible minimum. Moreover, second-movers (i.e., subjects in the role of player B) tend to reject offers which provide them with relatively small shares of the pie, even though these offers would leave them with strictly more than the minimum possible.[10] In other words, in experiments one typically observes that the players either disagree, or agree on a division of the pie which is "more fair" than what the theory predicts. The analysis in this section will show that these are precisely the predictions of game theory if one admits the possibility that the preferences of player B may exhibit some degree of interdependence.

3.1 Ultimatum Bargaining with Complete Information

In this subsection we assume that both player's utility functions are common knowledge. Part of this assumption is actually unnecessary, however: the discussion below only assumes that player A knows her opponent's utility function.

Let us begin by noting that there is a unique offer in the set Y which leaves player B indifferent between accepting and rejecting. Recalling (2), we observe that this offer is $(2m-r_B, r_B)$, where r_B is the reservation amount of player B. In any subgame perfect equilibrium then, player B accepts (rejects) any proposal $x \in Y$ with $x_B > (<)\ r_B$; and player A proposes $(2m - r_B, r_B)$, because $m > r_B$ implies that

$$V_A\left(2m - r_B, \frac{2m - r_B}{m}\right) > V_A(1,1) = 0$$

by strict monotonicity of V_A. Thus, we have:

Proposition 1. *The unique subgame perfect equilibrium outcome of the ultimatum bargaining game described above is: Player A offers $(2m - r_B, r_B)$ and player B accepts.*

Proposition 1 shows that if player A knows his opponent's utility function, he cannot receive as large a share as $2m - 1$ when player B has interdependent preferences. This is because A knows

[10] As noted in the introduction, these regularities we first noted by Guth, Schmittberger and Schwarze (1982). Camerer and Thaler (1995) and Roth (1995) provide detailed surveys of the related experimental literature.

that B strictly prefers the disagreement allocation to the allocation $(2m-1,1)$. Consequently, the presence of interdependent preferences induces deviations towards the "equal split" in the ultimatum game with complete information. The intuition is simply that, knowing that B has interdependent preferences, A fears rejection. As also noted by Bolton (1991) and Kirchsteiger (1994), the interdependent preferences could thus be thought of as an indirect way of modeling the presence of "fear of rejection" on the part of the first-movers.

3.2 Ultimatum Bargaining with Incomplete Information

In the ultimatum game with independent preferences, the assumption of complete information is innocuous, since the equilibrium does not depend on the specifics of the players' utility functions (as long as they are increasing in one's own income). With interdependent preferences, however, the complete information postulate is less readily acceptable. It is more realistic to assume instead that each player is unsure about the his opponent's degree of interdependence.[11] Thus, in this subsection we reformulate the ultimatum game with interdependent preferences as a game of incomplete information.

Take any $\Theta_i > 0$, and assume that $[0, \Theta_i]$ is the type space of each player $i = A, B$. We express the utility function of player i, $u_i : X \times [0, \Theta_i] \to \mathbb{R}$, as

$$u_i\left(x_A, x_B \mid \theta_i\right) \equiv V_i\left(x_i, \frac{x_i}{\bar{x}} \mid \theta_i\right), \tag{3}$$

where $V_i(1,1 \mid \theta_i)$ is normalized to zero for each $\theta_i \in [0, \Theta_i]$. All functions $V_i(\cdot, \cdot \mid \theta_i)$ for $\theta_i \in (0, \Theta_i]$ are continuous and strictly increasing in both arguments, while $V_i(\cdot, \cdot \mid 0)$ is assumed to be strictly increasing in its first argument, and independent of the second argument. Thus, type $\theta_i = 0$ corresponds to the type with independent preferences.

Recall that (2) defines player i's unique reservation amount for a given utility function V_i. Of course, different types θ_i have different reservation amounts. The equation $V_i\left(r_i, r_i/m \mid \theta_i\right) = 0$ defines a function $r_i : [0, \Theta_i] \to \mathbb{R}$, which specifies the reservation amount for each type $\theta_i \in [0, \Theta_i]$. Clearly, $r_i(0) = 1$, since $V_i(1, 1/m \mid 0) = V_i(1,1 \mid 0) = 0$. Moreover, $r_i(\theta_i) \in (1, m)$ for any $\theta_i \in (0, \Theta_i]$ since

$$V_i\left(1, \frac{1}{m} \mid \theta_i\right) < V_i\left(1,1 \mid \theta_i\right) = V_i\left(r_i\left(\theta_i\right), \frac{r_i\left(\theta_i\right)}{m} \mid \theta_i\right) = V_i\left(1,1 \mid \theta_i\right) < V_i\left(m,1 \mid \theta_i\right).$$

[11] As rightly stated by Bolton (1991, p. 1112), "... the marginal rate of substitution between absolute and relative money most likely varies by individual, making utility functions private information." A similar point was made also by Kennan and Wilson (1993), p. 93, who argue that the most bargaining experiments in the literature "can be interpreted, in effect, as involving bargaining with private information, as evidently most players did not know the preferences of the opposing bargainer."

In what follows we wish to think of higher types as more competitive players. Therefore, we invoke the following assumption that associates higher types θ_i with higher levels of reservation amounts:

Assumption [A]: r_i is strictly increasing on $[0, \Theta_i]$, $i = A, B$.

This assumption relates the magnitude of each player's type to her level of interdependence in a natural way: if type θ_i of player i prefers the disagreement outcome $(1, 1)$ to an outcome $x = (x_A, x_B)$, then any type $\theta_i' > \theta_i$ strictly prefers $(1, 1)$ to x.

We assume that θ_B is privately known by player B. It will become clear shortly that the equilibrium in the ultimatum game does not depend on whether player A's type is private information or not. Player A's beliefs about θ_B are represented by a continuous cumulative distribution function $F : [0, \Theta_B] \to [0, 1]$.[12]

To determine the subgame perfect equilibria of the ultimatum game, we define the function τ_B as

$$\tau_B(x_B) \equiv r_B^{-1}(x_B) \qquad \text{for all } x_B \in [1, r_B(\Theta_B)] .$$

This function specifies a critical level of interdependence $\tau_B(x_B)$ below which player B accepts and above which he rejects any given offer $(2m - x_B, x_B)$. Indeed, by [A], $\theta_B \lesseqgtr \tau_B(x_B)$ implies $r_B(\theta_B) \lesseqgtr x_B$ so that

$$0 = V_B\left(r_B(\theta_B), \frac{r_B(\theta_B)}{m} \mid \theta_B\right) \lesseqgtr V_B\left(x_B, \frac{x_B}{m} \mid \theta_B\right).$$

Consequently, in any subgame perfect equilibrium, player A of type θ_A chooses $x_B \in [1, m)$ in order to maximize the following objective function:

$$V_A\left(2m - x_B, \frac{2m - x_B}{m} \mid \theta_A\right) F(\tau_B(x_B)). \tag{4}$$

We thus obtain the following result.

Proposition 2. *Under the assumption [A], any perfect Bayesian equilibrium outcome of the ultimatum bargaining game with interdependent preferences has the following structure: Player A proposes a division $(2m - x_B(\theta_A), x_B(\theta_A))$ where $x_B(\theta_A) \in [1, m)$ maximizes the expression in (4). Player B accepts if her true type θ_B is strictly lower than the critical threshold $\tau_B(x_B(\theta_A))$, and rejects if $\tau_B(x_B(\theta_A)) < \theta_B$.*

[12] Continuity of F does not really play a significant role here. It is easy to modify the following analysis to account for discrete distributions.

Proposition 2 establishes that the two major "puzzling" regularities observed in ultimatum bargaining experiments, namely interior offers and rejections, are fully consistent with game theoretic equilibrium behavior. Once one accepts the idea that the competitive nature of bargainers plays a role in ultimatum experiments, it is realistic to assume that the exact degree of a player's interdependence is her private information. In this case, Proposition 2 shows that, for generic probability distributions representing player A's beliefs about his opponent's preferences, all equilibrium offers must indeed be "more fair" than $(2m-1,1)$, but less fair than (m,m), and rejections occur with positive probability.

4 Two-Period Bargaining

In this section we investigate the two-period alternating-offer bargaining game where players have interdependent preferences, as defined in Subsection 2.1. The game is played in two periods. In the first period, player A proposes an allocation in Y which player B either accepts or rejects. In the former case, player A's proposal is realized, while in the latter case the game advances to the second period, in which the players switch roles: player B proposes an allocation (x_A, x_B), and player A either accepts or rejects, with rejection resulting in disagreement.

The standard Stahl (1972) model posits that, if B's proposal is accepted in the second period, player $i = A, B$ achieves a utility level of $\delta_i u_i(x_A, x_B)$, where $\delta_i \in (0,1)$ is player i's *discount factor*. However, this particular representation of the players' time preferences makes it difficult to understand the implications of our results for the bargaining experiments. This is because, in experiments, it is rather the *size* of the pie that is discounted as the game moves from one period to the next. Thus, the utility of player i at the allocation (x_A, x_B) in period 2 is $u_i(\delta_A x_A, \delta_B x_B)$ (cf. Bolton, 1991). As we shall see in Subsection 4.2, it is this particular feature of the experimental design that generates disadvantageous counteroffers in the presence of interdependent preferences and incomplete information.

Before turning to the experiments however, we shall first briefly comment on how interdependence of preferences alters the subgame perfect equilibrium strategies in the standard Stahl bargaining model with complete information. While this model is not applicable to the two-period bargaining experiments, it is useful in identifying the basic interplay between the roles of the discount factors and the reservation amounts of the individuals. It also paves the way towards the analysis of the Rubinstein bargaining model presented in the next section. In Subsection 4.2, we shall turn our attention to bargaining experiments, and relate the implications of the present model (with incomplete information and alternating offers) to the findings of Ochs and Roth (1989) and

Bolton (1991).

4.1 Alternating-Offer Bargaining with Complete Information

Throughout this subsection, we assume that, in case player B's proposal (x_A, x_B) is realized in the second period, the utility of player i is $\delta_i u_i(x_A, x_B)$. Of course, if player A rejects B's offer, the game ends in disagreement.

Just as in the standard case (in which players are assumed to have independent preferences) the unique subgame perfect equilibrium of the game is obtained by backward induction in a straightforward manner: if the second period is reached, B proposes the allocation $(r_A, 2m - r_A)$ where $r_A \in (1, m)$ is A's reservation amount, and A accepts. Therefore, player A in the first period proposes the allocation that leaves player B indifferent between accepting and rejecting: that is, she proposes $(2m - x_B(\delta_B), x_B(\delta_B))$, where $x_B(\delta_B)$ is uniquely determined by

$$V_B\left(x_B(\delta_B), \frac{x_B(\delta_B)}{m}\right) = \delta_B V_B\left(2m - r_A, \frac{2m - r_A}{m}\right). \tag{5}$$

The structure of the equilibrium strategies readily from this equation: player A proposes $(2m - x_B(\delta_B), x_B(\delta_B))$ in the first period, and accepts in the second period a division like $(\xi_A, 2m - \xi_A)$ if and only if $\xi_A \geq r_A$; player B accepts the allocation $(2m - \xi_B, \xi_B)$ in the first period if and only if $\xi_B \geq x_B(\delta_B)$, and proposes $(r_A, 2m - r_A)$ in the second period.

Owing to the continuity and strict monotonicity of V_B, the function $x_B : (0, 1) \to (r_B, 2m - r_A)$ is well-defined and strictly increasing. Thus, the introduction of interdependence does not modify the basic insights of the standard two-period alternating offer bargaining game: the equilibrium payoff of player B increases as her degree of patience increases, and/or the amount of time elapsing between the first and second period decreases. As δ_B becomes arbitrarily close to zero, say because the second period is pushed far into the future, the two-stage game collapses into the ultimatum game that we have analyzed in the previous section. Thus player B is better off in the two-stage game, relative to the ultimatum game, for any positive value of his discount factor. At the other extreme, for δ_B close to one, the two-stage game is essentially equivalent to the ultimatum game in which player B proposes the division $(r_A, 2m - r_A)$ and player A accepts.

Clearly, the 50-50 division (m, m) is obtained in equilibrium if and only if $\delta_B = x_B^{-1}(m)$; thus, in the two-stage game, player A enjoys a first-mover advantage if and only if her discount factor exceeds the critical value $x_B^{-1}(m)$. Yet, given its generality, the present model has little to say about the magnitude of $x_B^{-1}(m)$. Whether there is a first-mover advantage or not depends both on how patient B is and how competitive both players are. It turns out, however, that if the utility function of player B is (strictly) concave, then the critical value satisfies $x_B^{-1}(m) \geq (>)\ 1/2$, provided that the players are *equally interdependent*, that is, when $r_A = r_B$. Therefore, if V_B is concave and $\delta_B \in [0, 1/2)$, a first-mover advantage obtains even when the players are equally competitive.[13] (This observation parallels Proposition 2 of Bolton, 1991.)

We summarize the above discussion in the following

Proposition 3. *The unique subgame perfect equilibrium outcome of the two-period alternating-offer bargaining game described above has the following structure: player A proposes* $(2m - x_B(\delta_B),$ $x_B(\delta_B))$ *in the first period, where* $x_B(\delta_B)$ *is defined by (5), and player B accepts. Moreover, we have:*

(a) $r_B < \xi_B(\delta_B) < 2m - r_A$ *for all* $\delta_B \in (0, 1)$*;*

(b) the function $\xi_B(\delta_B)$ *is a continuous and strictly increasing;*

(c) if $r_A \geq r_B$ *and* V_B *is concave, then* $\xi_B(\delta_B) < m$ *for all* $\delta_B \in [0, 1/2)$*.*

Proposition 3 does not say much about the fairness of the equilibrium divisions, that is, about their closeness to the 50-50 split (m, m). Given the generality of the utility functions we consider here and the effect of the discount factor, which increases the bargaining power of player B, it is difficult to see if the presence of interdependent preferences causes a deviation towards the 50-50 division in general.[14] The following, example, however, illustrates clearly that this would be precisely the prediction of our model, had we confined our attention to a linear preference specification. We shall in fact make extensive use of this specification in Section 5.

Example. Let $\theta \in [0, 1)$ and assume that

$$u_i(x_A, x_B) = V\left(x_i, \frac{x_i}{\bar{x}}\right) = (1 - \theta)\, x_i + \theta \frac{x_i}{\bar{x}} - 1$$

for all $(x_A, x_B) \in X$, $i = A, B$. Note that the two players are equally interdependent, and $V(1, 1) = 0$. The reservation amount of each individual, as a function of θ, is found via (2) as

$$r(\theta) = \left(1 - \theta + \frac{\theta}{m}\right)^{-1}.$$

Hence, by using (5) and Proposition 3, we determine that in equilibrium player A offers $(2m - x_B(\delta_B; \theta), x_B(\delta_B; \theta))$ where

$$x_B(\delta_B; \theta) \equiv r(\theta)\ [2\delta_B\ (1 - \theta)\ (m - 1) + 1] \quad \text{for all } \theta, \delta_B \in [0, 1). \tag{6}$$

[13] The condition $\delta_B < 1/2$ is sufficient but not necessary for the presence of a first-mover advantage when u_i is concave. For instance, if $u_i(x) = \sqrt{x_i} + \sqrt{x_i/\bar{x}}$ for all $x \in X$, we have $m > x(\delta_B)$ for all $\delta_B < 0.513$.

[14] Once again, this question is meaningful only when players are equally competitive, i.e., $r_A = r_B$.

(Notice that by choosing $\theta = 0$, we obtain the equilibrium share of player B in the standard case in which all players are independent.) There are two things that are interesting about (6). First, it shows that $x_B(1/2;\theta) = m$ for all $\theta \in [0,1)$. Thus, a first-mover advantage exists in this game if *and only if* $\delta_B < 1/2$; the bound found in Proposition 3(c) turns out to be tight in this example, owing to the fact that V is an affine function. Second, (6) shows that *the more interdependent players are, the more fair the equilibrium outcomes will be*. To see this, observe that

$$\frac{\partial x_B(\delta_B;\theta)}{\partial \theta} = \left(1 - \frac{1}{m}\right)(1 - 2\delta_B)\, r(\theta)^2 \gtrless 0 \quad \text{whenever} \quad \delta_B \lessgtr 1/2.$$

Thus, if $\delta_B < (>)\ 1/2$, there is first (second)-mover advantage and the equilibrium share of the second (first)-mover is strictly increasing in θ.

4.2 Alternating-Offer Bargaining with Incomplete Information

Arguably, the most paradoxical regularity emerging from the experimental data is that "a substantial percentage of rejected offers were followed by disadvantageous counterproposals" (Ochs and Roth, 1988, p. 376). Neither Stahl's original model with independent preferences nor our variation with interdependent preferences outlined in the previous subsection, with or without incomplete information, can reconcile this evidence with equilibrium behavior. The reason is that a necessary condition for B to reject A's initial offer in equilibrium is

$$u_B(x_A, x_B) = V_B(x_B, \frac{x_B}{m}) \leq \delta_B V_B(y_B, \frac{y_B}{m}) = \delta_B\, u_B(y_A, y_B), \tag{7}$$

where $(y_A, y_B) \in Y$ is B's second period equilibrium offer. But since sequential rationality ensures $y_B \geq r_B$, i.e. $u_B(y_A, y_B) \geq 0$, the monotonicity of V_B and (7) jointly imply that $y_B \geq x_B$ for all $\delta_B \in [0,1]$.

This argument, however, depends crucially on the assumption that individuals' utilities are discounted through periods. But in the experiments it is the size of the *pie* that is discounted when the game moves from one stage to the next, while the entire game lasts only a short amount of time; hence we can reasonably assume that the subjects incur no cost due to the passage of real time: that is, if the game ends in agreement at the second stage, and the pie shrinks from size k to size δk, the utility of each player $i = A, B$ is $u_i(x_A, \delta_i k - x_A)$ instead of $\delta_i u_i(x_A, k - x_A)$.

This difference between the standard theoretical model and the game played in the experiments, while inessential under the assumption that the players preferences are known to be independent, is crucial for the presence of disadvantageous counteroffers in equilibrium once player A admits the possibility that his opponent may have interdependent preferences. To see why, consider the following model which captures the essential features of the games played in the experiments, while conforming to our assumptions on the players' preferences. First, player A proposes an allocation in Y, and player B either accepts, thus ending the game, or rejects. If B rejects, the game enters a second stage, in which the *net* pie[15] shrinks from $2m-2$ to $\delta\,(2m-2)$, and B proposes a division in the set

$$Y(\delta) \equiv \{(y_A, y_B) \geq (1,1) : y_A + y_B = 2 + \delta(2m-2)\},$$

where $\delta \in (0,1]$ stands for the *common* rate at which the pie is discounted.

We assume that the player i is either of type 0 or of type $\Theta_i > 0$, $i = A, B$,[16] and, as in Subsection 3.2, the utility functions satisfy (3). Type 0 has independent preferences, hence her reservation amount is $r_i\,(0) = 1$, while type Θ_i has (strictly) interdependent preferences, hence her reservation amount $r_i(\Theta_i)$ is strictly higher than 1. If the game ends in the first stage, with agreement on division $(x_A, x_B) \in Y$, the utility of player i is $V_i(x_i, x_i/m \mid \theta_i)$. If the game enters the second stage, and B proposes $(y_A, y_B) \in Y(\delta)$, the utility of player i would be $V(y_i, \frac{y_i}{1+\delta(m-1)} \mid \theta_i)$ if A accepts, and zero otherwise. Finally, the beliefs of player i about player j are represented by $\pi_j \in [0,1]$, with $\pi_j \equiv \text{Prob}[\theta_j = \Theta_j]$; that is, π_A is the probability that B assigns to the event that A is interdependent, and similarly for π_B.

Of course, if $\pi_A = \pi_B = 0$, then the model collapses into a complete information bargaining game with independent preferences. It is readily verified that, in the unique subgame perfect equilibrium of this game, player A proposes $(1 + (1-\delta)\,(2m-2)\,, 1 + \delta\,(2m-2))$ in the first period, and player B accepts. Following Bolton (1991), we shall refer to this allocation as the *pecuniary equilibrium* from now on.

When $\pi_A, \pi_B > 0$, the model carries the basic features of a signaling game, thus admitting great many perfect Bayesian equilibria, the full characterization of which falls beyond our present scope. Our aim in this subsection is simply to show that all the experimental regularities, and in particular the occurrence of disadvantageous counteroffers, are consistent with equilibrium behavior.

Before stating the formal results, we provide a heuristic argument that shows how disadvantageous counteroffers may arise in equilibrium. Consider Figure 2 in which the bargaining sets Y and $Y(\delta)$ are plotted with some indifference curves for both players. The independent type of each player does not care about his opponent's share of the surplus: hence A's indifference curves are horizontal if $\theta_A = 0$, and B's curves are vertical if $\theta_B = 0$. The indifference curves of the

[15] Recall that each player must receive at least one unit. Hence the surplus on which the players bargain is effectively $2m-2$.

[16] The analysis of the model and the results reported below would remain true with inessential modifications, if we allowed for different discount rates and more than two types. We invoke these two assumptions here only to simplify the exposition.

interdependent type of each player instead are upward sloping, since their utility decreases if their opponent's share increases. For instance, type Θ_A of player A is indifferent between receiving the disagreement outcome $(1, 1)$ and the allocation y. Similarly, type Θ_B of B is indifferent between the allocations y and z^1.

Now consider an allocation like x in Figure 2, and suppose that both types of A offer x in the first period. If B is pessimistic enough, i.e., π_A is high enough, then, in any perfect Bayesian equilibrium, he will propose the division $y \in Y(\delta)$ in the second stage, whatever his level of interdependence, and both types of A would accept. Therefore, it will be optimal for type 0 of B to accept the proposal x.

But is it optimal for type Θ_B of B to accept this offer when π_A is high? The answer is no, for if B rejects x and offers y in the second period, she is certain that both types of A will accept her offer. Since type Θ_B of B strictly prefers y to x (y is on a higher indifference curve than is x), type Θ_B will indeed reject A's offer x, and counteroffer y in the second stage. But notice that y is a disadvantageous counteroffer for B; that is $x_B > y_B$!

The above argument suggests that disadvantageous counteroffers should not be considered paradoxical when we allow for potentially interdependent preferences. The main result of this section is stated next.

Proposition 4. *There exist $\pi_A, \pi_B \in (0, 1)$ such that, for all $\delta \in (0, 1)$, the two-period alternating-offer bargaining game described above admits a continuum many pure strategy perfect Bayesian equilibria that have the following features:*

(i) both types of A make the same initial offer;

(ii) type 0 of B accepts while type Θ_B of B rejects and makes a disadvantageous counteroffer in the second period, which both types of A accept;

(iii) there exists $\bar{\delta} \in [0, 1)$ such that A's initial offer deviates from the pecuniary equilibrium in the direction of 50-50 division for all $\delta \geq \bar{\delta}$;

(iv) we have

$$\lim_{\delta \to 1} x(\delta, \theta) = \{(r_A(\Theta_A), 2m - r_A(\Theta_A))\} \qquad \text{for all } \theta \in \{0, \Theta_A\} \times \{0, \Theta_B\},$$

where $x(\delta, \theta) \subset X$ denotes the set of all equilibrium allocations at state θ.

Proposition 4 is interesting in that it shows that the bargaining model at hand is capable of predicting the rejection of first-period offers that are followed by disadvantageous counteroffers in equilibrium.[17] Moreover, this result shows that one should not be surprised to see opening

[17] As for belief-based refinement properties of these equilibria, we should note that they satisfy the intuitive criterion of Cho and Kreps (1987).

offers in the experiments that deviate towards the 50-50 division when δ is sufficiently high. Of course, not all equilibria possess these properties; there are many other equilibria of the game than those mentioned in Proposition 4. The point is that when we model the bargaining experiments by using interdependent preferences with the degree of interdependence of players being private information, what seems like paradoxical plays become perfectly reasonable equilibrium behavior. In a qualitative sense, then, we would like to argue that the present model fits the data fairly well.[18]

In passing, we note that the outcome-fairness properties of the equilibria mentioned in proposition 4 are in line with the main thesis of the paper. Since the presence of discounting again shadows the potential of making comparisons with the pecuniary equilibrium, we choose to make this comparison for large δ. In this case, all equilibria in Proposition 4 envisage a deviation of both the opening offers and the equilibrium allocations towards the 50-50 division. This follows from part (iv) of the proposition and the fact that the pecuniary equilibrium converges to $(1, 2m-1)$ as $\delta \to 1$.

5 Infinite Horizon Bargaining Models

The findings of the previous sections suggest that enriching the bargaining games with the potential presence of interdependent preferences can be useful. However, while these findings allow us to compare the predictions of the proposed models with those of the experiments, they do not apply to more realistic bargaining scenarios, with a potentially infinite time duration. It is important to study infinite horizon bargaining games because in many situations that involve substantial gains from trade, it would be unrealistic to postulate that bargainers could credibly commit themselves not to trade just because a certain amount of time has elapsed. This is in fact the main reason why an extensive part of the bargaining theory is devoted to infinite horizon games.

Consequently, in this section we study a number of infinite horizon bargaining models with interdependent preferences. We begin with the classical Rubinstein model, and then consider bargaining games with incomplete information. Our aim is again to understand the implications of (possibly) interdependent preferences with regard to the fairness (i.e. closeness to the 50-50 split) of the equilibrium outcomes. We wish to show that the insight of our earlier findings remain valid in infinite horizon models as well: admitting the possibility of interdependent preferences generates results that are consistent with the high frequency of "fair divisions" commonly observed in daily

[18] Another oft-quoted regularity observed in the related experiments is the considerable effect that the discount factor of A has on the equilibrium outcome. It is easy to see that this phenomenon would be completely accounted for in our model, had we allowed for different discount rates for players. This is again primarily due to the discounting of the pie as opposed to utilities.

life. In fact, we shall observe that in some cases even the slightest possibility of interdependence is enough to guarantee such predictions.

5.1 The Rubinstein Bargaining Game

In Rubinstein's model, player A (player B) proposes an allocation in Y in period 1 (period 2, resp.), and at every odd (even, resp.) period, until one of the offers is accepted. If no offer is ever accepted, then the outcome of the game is the disagreement outcome, i.e. $(1,1)$. If a proposal $x \in Y$ is accepted in period t, the utility of player i is $\delta^{t-1}u_i(x_A, x_B)$, where $\delta \in (0,1)$ is the discount rate.[19] Both players obtain zero utility in the case of disagreement.

Throughout this section, we shall adopt the following simple one-parameter, specification of the utility function u_i:

$$u_i\left(x_A, x_B \mid \theta_i\right) \equiv \left(1-\theta_i\right)\, x_i\, +\, \theta_i \frac{x_i}{\left(x_A+x_B\right)/2} - 1, \qquad i = A, B, \tag{8}$$

where $(x_A, x_B) \in X$ and $0 \leq \theta_i < 1$. There are two important advantages of this specification. First, it guarantees the existence of a unique subgame perfect equilibrium of the game at hand.[20] Second, its simplicity allows us to perform a meaningful comparative statics analysis. This is important, since our main objective here is to compare the equilibrium outcomes of Rubinstein's bargaining model, with and without interdependent preferences.

The equilibrium is of course invariant under linear transformations of the individual utility function u_i. Therefore, we may assume that the utility function of individual i of type θ_i is given by

$$v_i\left(x_A, x_B \mid \theta_i\right) \equiv r\left(\theta_i\right)\, u_i\left(x_A, x_B \mid \theta_i\right),$$

where

$$r(\theta_i) \equiv (1-\theta_i+\theta_i/m)^{-1} \in (1, m)$$

is the reservation amount of individual i of type θ_i. It follows from (8) that

$$v_i\left(x_A, x_B \mid \theta_i\right) \; = \; x_i - r\left(\theta_i\right) \qquad \text{for all } \; (x_A, x_B) \in Y.$$

This shows that, given (8), the present model is essentially identical to the standard buyer-seller bargaining model. Indeed, if we think of player A as the owner of an indivisible object with

[19] Again we assume that the players have the same discount factor for simplicity. Modifying the analysis in the case of distinct discount factors is straightforward.

[20] In the general case where u_i is defined as in (1), the equilibrium need not be unique since the bargaining set $\mathcal{U}$ can then be non-convex (cf. Binmore, 1987, and Herrero, 1989). However, note that the existence of equilibrium is guaranteed even at this level of generality since $B(\mathcal{U})$ is a connected set.

value $r(\theta_A)$, and of player B as the potential 'buyer' with value $2m - r(\theta_B)$, we can interpret any division $(x_A, 2m - x_A)$ in our model as a 'sale' at price x_A. Consequently, we may readily apply the standard analysis of the Rubinstein bargaining game, and conclude that, for any fixed $\theta_i \in [0,1)$, the subgame perfect equilibrium of the game is unique and is of the following form: player A always proposes $x(\theta_A, \theta_B, \delta) \in Y$ and accepts any proposal in which she is offered at least $y_B(\theta_A, \theta_B, \delta)$; player B always proposes $y(\theta_A, \theta_B, \delta) \in Y$ and accepts any proposal in which she is offered at least $x_B(\theta_A, \theta_B, \delta)$, where $\delta v_A(x(\theta_A, \theta_B, \delta) \mid \theta_A) = v_A(y(\theta_A, \theta_B, \delta) \mid \theta_A)$ and $\delta v_B(y(\theta_A, \theta_B, \delta) \mid \theta_B) = v_B(x(\theta_A, \theta_B, \delta) \mid \theta_B)$. By solving these two equations, we find that the following allocation is reached in equilibrium with no delay:

$$x(\theta_A, \theta_B, \delta) = \left(r(\theta_A) + \frac{1}{1+\delta}(2m - r(\theta_A) - r(\theta_B)),\ r(\theta_B) + \frac{\delta}{1+\delta}(2m - r(\theta_A) - r(\theta_B)) \right).$$

That is, in equilibrium, each player receives her Rubinstein share of the *net* pie $2m - r(\theta_A) - r(\theta_B)$, in addition to her reservation amount $r(\theta_i)$. Clearly, by setting $\theta_A = \theta_B = 0$, we recover the equilibrium allocation of the standard Rubinstein game with *linear* individual utility functions.

This result shows that the essential feature of this model, common to both the buyer-seller and our interpretation, is that the players are trying to share a 'pie' whose size varies with their characteristics, i.e. their 'types'. In our context, the players' types have to do with their degree of interdependence: the more interdependent any player is, the smaller the pie is. Therefore, while the players' discount factor δ still determines the size of the first mover advantage, the equilibrium division also depends on both players' interdependence levels. In particular, as the discount factor δ approaches 1, (i.e., as the interval between offers approaches zero) the equilibrium division converges to

$$\left(\frac{1}{2}[2m - r(\theta_A) - r(\theta_B)],\ \frac{1}{2}[2m - r(\theta_B) - r(\theta_A)] \right).$$

The structure of the equilibrium allocation $x(\theta_A, \theta_B, \delta)$ also shows that the notion of "being more interdependent" can be interpreted as "being less eager to agree" to any given division x in Y, hence as having a smaller marginal rate of substitution between money and time.[21] Consequently, the degree of interdependence of a bargainer plays a similar role here to the one played by the discount factor in the standard model. When the bargainers are equally interdependent, however, the extent of the first-mover advantage is smaller here than it is in the standard Rubinstein model.

[21] The marginal rate of substitution between money and time for player i is here found as

$$\frac{\mathbf{D}_t(\delta^{t-1} v_i(x(\theta_A, \theta_B, \delta) \mid \theta_i))}{\mathbf{D}_{x_i}(\delta^{t-1} v_i(x(\theta_A, \theta_B, \delta) \mid \theta_i))} = [x_i - r(\theta_i)] \ln \delta$$

which is strictly increasing in θ_i for any $\delta \in (0,1)$. That is, the higher θ_i, the flatter the indifference curve of player i in the (t, x_i) space which means that time is less important for player i.

Indeed, it is easily checked that $x_A(\theta, \theta, \delta)$ is strictly decreasing in θ on $[0, 1)$ so that the model predicts in this case a more equal division of the pie than the usual for *any* $\delta \in (0, 1)$. Thus, the more interdependent players are, the more fair the equilibrium outcomes will be. When the individuals are extremely competitive, therefore, one is likely to observe very fair outcomes:

$$\lim_{\theta \to 1} x(\theta, \theta, \delta) = (m, m) \quad \text{for all } \delta \in (0, 1).$$

The following proposition summarizes these findings.

Proposition 5. *In the unique subgame perfect equilibrium of the Rubinstein bargaining game described above, player A offers $x(\theta_A, \theta_B, \delta)$ in the first period and player B accepts. If $\theta_A = \theta_B$, then, for any $\delta \in (0, 1)$,*

(a) *there is a first-mover advantage*;

(b) $x(\theta_A, \theta_B, \delta)$ *is closer to the 50-50 division than the standard equilibrium allocation* $x(0, 0, \delta)$;

(c) *the more interdependent players are, the closer the equilibrium is to the 50-50 division.*

The equilibrium division $x(\theta_A, \theta_B, \delta)$ has the property that, for *any* $\delta \in (0, 1)$, each player's equilibrium share increases with his degree of interdependence. Therefore, in games with incomplete information, a player with a low degree of interdependence can only benefit from his opponent assigning positive probability to the event of facing a highly interdependent type. In some cases, small degrees of uncertainty may in fact generate substantial gains. This issue is explored next.

5.2 Infinite Horizon Bargaining Games with Incomplete Information

In this section, we shall study the following *concession game*: player A, whose type θ_A is common knowledge, makes all the offers, while player B, who has private information about his type θ_B, can only accept or reject. Player A's beliefs about B's type are represented by a distribution function $F : [0, \Theta_B] \to [0, 1]$, where $0 < \Theta_B < 1$. We assume that F is common knowledge, and posit the following technical condition:

Assumption [T]: $\liminf_{\theta \to \Theta_B} \dfrac{1 - F(\theta)}{\Theta_B - \theta} > 0.$

This assumption is a weak regularity condition that requires the slope of F to be bounded away from zero near Θ_B. For instance, if F is left-differentiable at Θ_B and $F'_-(\Theta_B) > 0$, or if F has a mass point on Θ_B, then it satisfies [T]. Intuitively speaking, all [T] says is that player A assigns strictly positive probability to the event that player B is arbitrarily close to being the most interdependent type possible.

Given assumption [T], it can be shown that a perfect Bayesian equilibrium of the game at hand exists. In fact, under this assumption, one equilibrium may differ from another only with respect to the first-period proposal of player A; the equilibrium is *generically* unique. What is more, as the time interval between offers gets very small (i.e., as the offers take place very quickly), all equilibrium first-period offers converge to a unique allocation:

Proposition 6. *Assume* [T], *consider any perfect Bayesian equilibrium of the concession game defined above, and denote the corresponding equilibrium sequence of offers made by player A (conditional on the discount rate and the type space of B) by $\{x^t(\delta, \Theta_B)\}_{t=1}^{\infty} \in Y^{\infty}$. We have:*

$$\lim_{\delta \to 1} x^1(\delta, \Theta_B) = (2m - r(\Theta_B), r(\Theta_B)).$$

Proposition 6 says that if player A assigns positive probability, however small, to the event that B is maximally interdependent, then, as the interval between offers becomes small, his strategy becomes close to the strategy that he would use were he certain that B's reservation amount is near $r(\Theta_B)$: that is, his first offer becomes close to allocation that will be accepted even by the maximally interdependent type, and the probability that the game ends immediately converges to one. In particular, if player A thinks that B may be the most competitive type who is *almost* indifferent between 50-50 division and the disagreement outcome (i.e., if Θ_B is close to 1), then the equilibrium outcome is (almost) the 50-50 split, no matter how unlikely A may think this event really is:

$$\lim_{\Theta_B \to 1} \lim_{\delta \to 1} x^1(\delta, \Theta_B) = (m, m).$$

Very fair outcomes may thus be rational, after all.

Proposition 6 is essentially a corollary of a well-known result in bargaining theory. Indeed, the concession game described above is isomorphic to the single-sale bargaining model with one-sided incomplete information. The limit result stated in Proposition 6 is essentially identical to the famous *Coase conjecture* which states that the seller's expected gain from trade tends to its lowest possible value when the frequency of price offers becomes arbitrarily large. Moreover, since here it is common knowledge that mutually beneficial agreements exist, i.e., $r(\theta_A) + r(\Theta_B) < 2m$ for all $\theta_A \in [0,1]$, our model corresponds to the so-called "gap" case of the single-sale model (see Fudenberg and Tirole, 1991, Chapter 10). Consequently, the related results of Fudenberg, Levine and Tirole (1985) and Gul, Sonnenschein and Wilson (1986) that establish the validity of the Coase conjecture in the gap case of the single-sale model almost immediately entail our Proposition 6.[22] (The details of this claim are found in the appendix.)

[22] We owe the generic uniqueness of the equilibrium again to Fudenberg, Levine and Tirole (1985) and Gul, Sonnenschein and Wilson (1986).

In passing, we note that extending the present model to allow for two-sided offers and/or two-sided incomplete information introduces a vast multiplicity of equilibria. For instance, in any alternating offers game with incomplete information, many allocations can be sustained as equilibrium outcomes, simply by making each player's beliefs concentrated on the independent type of his opponent off the equilibrium path. (Under these beliefs, the only constraint acting on the equilibrium strategies is the sequential rationality on the equilibrium path.) Nevertheless, by suitably refining the sequential equilibria, one may identify those equilibria that have the property that A's initial offer converges to $(2m - r(\Theta_B), r(\Theta_B))$ as $\delta \to 1$. Indeed, both in the alternating offers game with one-sided private information (Gul and Sonnenschein, 1988), and in the concession game with two-sided private information (Cho, 1990), all equilibria that satisfy certain monotonicity and stationarity properties exhibit the Coase property.

6 A Mechanism Design Approach

In this section, we depart from the "positive" analysis of bargaining among interdependent players and adopt a normative approach. This is of interest, because in many instances of real-world bargaining, negotiations are brought to an end by a settlement that is guided by a third party. The main question we ask then is this: if an arbitrator realizes that players may possess interdependent preferences, in which manner should she resolve the pie division problem? More specifically, we wish to see if there is any reason to suspect that the potential interdependence of preferences would enforce the emergence of egalitarian outcomes in bargaining settlements. To address this issue, we adopt a mechanism design approach.

Let $\mathcal{J}_i$ denote the set of all utility functions on X that can be written as in (1) for some continuous and strictly increasing V_i with $V_i(1,1) = 0$, $i = A, B$. Take any $\mathcal{T}_i \subseteq \mathcal{J}_i$, and interpret $\mathcal{T}_i$ as the set of all preferences that the arbitrator conceives as admissible for individual i; that is, $\mathcal{T}_i$ is the *type space* of player i. The product $\mathcal{T} \equiv \mathcal{T}_A \times \mathcal{T}_B$ then corresponds to the set of all *states of nature.*

A social choice function is any function that assigns to a state of nature a particular division of the pie. Formally, we define a *social choice function* (SCF) on $\mathcal{T}$ as any function $f : \mathcal{T} \to Y$. Thus, $f(u_A, u_B)$ stands for the (efficient) division of the pie that the arbitrator would choose, had she known that the true utility (type) of player i was u_i.[23] As is usual, we say that f is *individually*

[23] Alternatively, we may define a SCF on $\mathcal{T}$ as mapping a type profile (u_A, u_B) to a 3-tuple $(x, t) \in Y \times \mathbf{Z}_+$, where (x, t) denotes the allocation awarded to players at period t. In this case, our definition of a SCF would postulate implicitly the property of *ex post efficiency.*

rational, if f does not allocate to any type a share that is strictly below its reservation amount, i.e., if

$$u_i(f(u_A, u_B)) \geq 0 \quad \text{for all } u_i \in \mathcal{T}_i,\ i = A, B,$$

where $r(u_i)$ is defined through the equation $V_i(r(u_i), r(u_i)/m) = 0$. Needless to say, individual rationality is a participation constraint that needs to be satisfied by any reasonable settlement.

We now ask if it is possible here to implement a SCF in dominant strategies.[24] While it is well-known that this is in general impossible (the Gibbard-Satterthwaite theorem), it is easily seen that the restricted domain we consider here admits individually rational SCFs that are dominant strategy implementable. For instance, the social choice function that assigns to *any* type profile the equal split of the pie is individually rational and dominant strategy implementable. Our next result provides a characterization of all such SCFs.

Proposition 7. *A SCF f on $\mathcal{T}$ is individually rational and dominant strategy implementable if, and only if,*

$$f(u_A, u_B) \geq \left(\sup_{u_A \in \mathcal{T}_A} r(u_A),\ \sup_{u_B \in \mathcal{T}_B} r(u_B) \right) \quad \textit{for all } (u_A, u_B) \in \mathcal{T},$$

$r(u_i)$ is defined through the equation $u_i(r(u_i), 2m - r(u_i)) = 0$, $i = A, B$.

Proposition 7 shows that if the domain of a SCF includes highly interdependent preferences, then that SCF must choose highly egalitarian outcomes at all states of nature. The following corollary of Proposition 7 will drive this point home.

Corollary 8. *A SCF f on $\mathcal{J}_A \times \mathcal{J}_B$ is individually rational and dominant strategy implementable if, and only if,*

$$f(u_A, u_B) = (m, m) \quad \textit{for all } (u_A, u_B) \in \mathcal{T}.$$

This result tells us that if the class of all interdependent preferences considered by the arbitrator is sufficiently rich, then the only possible SCF is the one that assigns the 50-50 division *regardless of the state of nature.*[25] This observation provides a rigorous normative rationale for the equal-split solution the applications of which arise abundantly in daily life. For instance, consider a parent who has to divide the last slice of the pie among two siblings. It would not be unreasonable to

[24] A *mechanism* is any list $(\{Z_A, Z_B\}, h)$ where Z_i is an arbitrary message space and $h : Z_A \times Z_B \to Y$ is an arbitrary outcome function. Such a mechanism is said to *implement* the SCF f on $\mathcal{T}$ *in dominant strategies* iff the 2-person normal form game $(\{Z_i, u_i \circ h\}_{i=A,B})$ has a dominant strategy equilibrium $z(u_A, u_B) \in Z_A \times Z_B$ such that $h(z(u_A, u_B)) = f(u_A, u_B)$, for all $(u_A, u_B) \in \mathcal{T}$.

[25] One certainly does not need the type space of player i to consist of all members of $\mathcal{J}_i$ for this result to hold. If, for instance, $\mathcal{T}_i$ contains all of the affine utility functions considered in Section 5, then the result goes through.

expect that the parent will divide the pie equally (to the best of her abilities) to avoid any possible conflict between the siblings that may arise due to envious feelings. Corollary 8 suggests a rigorous foundation precisely for this decision rule.

7 Conclusion

Admitting the possibility of interdependent preferences may not be the only way in which one can modify standard bargaining models in order to obtain predictions that are consistent with experimental evidence. To the best of our knowledge, however, the combination of interdependent preferences and private information yields the first model to date capable of accommodating the observed experimental regularities that are in direct conflict with the predictions of standard bargaining theory. The model has the additional merit of providing a rational theory of fair outcomes. Moreover, the idea of interdependent preferences, which is already in use in other areas of economics, seems particularly appropriate in bilateral bargaining contexts. Indeed, it is not unreasonable to think that competitive feelings may arise and influence the players' choices in such contexts.

There are several directions for future research. First, it will be interesting to design experiments to identify the predictive limitations of the finite horizon bargaining models examined here. Second, the infinite horizon model with nonlinear interdependent utility functions remains to be investigated. This case is difficult since, unlike the model of Section 5.2, it is not necessarily isomorphic to the standard buyer-seller bargaining model. Finally, the analysis of multilateral bargaining with possibly interdependent preferences is left for future research.

Appendix

Proof of Proposition 3. **(a)** Define $r_B \in (1,m)$ by $V_B(r_B, r_B/m) = V_B(1,1) = 0$. Since V_B is strictly increasing, $x_B(\delta_B)$ is determined by (5), and $2m - r_A > m > r_B$, we have

$$V_B\left(x_B(\delta_B), \frac{x_B(\delta_B)}{m}\right) > \delta_B V_B\left(r_B, \frac{r_B}{m}\right) = 0,$$

hence $x_B(\delta_B) > r_B$ for all $\delta_B \in (0,1)$.

(b) This follows immediately from (5).

(c) Since V_B is concave, $r_A \geq r_B$, and $V_B(r_B, r_B/m) = 0$, we have

$$\begin{aligned} V_B(m,1) &\geq \frac{1}{2}V_B(r_A, \frac{r_A}{m}) + \frac{1}{2}V_B\left(2m - r_A, \frac{2m - r_A}{m}\right) \\ &\geq \frac{1}{2}V_B(r_B, \frac{r_B}{m}) + \frac{1}{2}V_B\left(2m - r_A, \frac{2m - r_A}{m}\right) \\ &= \frac{1}{2}V_B\left(2m - r_A, \frac{2m - r_A}{m}\right). \end{aligned}$$

By (5) and monotonicity of V_B, therefore, we have

$$V_B(m,1) \geq V_B\left(x_B(1/2), \frac{x_B(1/2)}{m}\right) > V_B\left(x_B(\delta_B), \frac{x_B(\delta_B)}{m}\right),$$

hence $m > x_B(\delta_B)$ for all $0 \leq \delta_B < 1/2$.

Proof of Proposition 4. Let k denote the size of the pie net of the bargainers holdings, that is, set $k \equiv 2m - 2$. Since V is strictly increasing and continuous, (3) and $V(1,1 \mid \Theta_A) = 0$ imply that there exists a unique $\alpha \in (0,1)$ such that

$$u_A(1 + (1-\alpha)\delta k, 1 + \alpha\delta k \mid \Theta_A) = 0.$$

Clearly, α is the maximum share for B that both types of A will accept. Throughout this proof we shall let

$$(y_A, y_B) \equiv (1 + (1-\alpha)\delta k, 1 + \alpha\delta k).$$

We define next the functions $\omega : \{0, \Theta_B\} \to (0,1)$ and $\hat{\omega} : \{0, \Theta_B\} \to (0,1)$ by

$$u_B(1 + (1 - \omega(\theta_B))k, 1 + \omega(\theta_B)\ k \mid \theta_B) = u_B(y_A, y_B \mid \theta_B) \tag{9}$$

and

$$u_B(1 + (1 - \hat{\omega}(\theta_B))k, 1 + \hat{\omega}(\theta_B)k \mid \theta_B) = u_B(1, 1 + \delta k \mid \theta_B), \tag{10}$$

respectively. One can think of $\omega(\theta_B)$ as the minimum share for B that type θ_B will accept in stage 1, provided that B has sufficiently pessimistic beliefs about A's type (i.e. π_A is high enough). In contrast, $\hat{\omega}(\theta_B)$ is the minimum share for B that type θ_B will accept in stage 1, provided that B is sufficiently optimistic about A's type.

Claim 1. $\alpha\delta = \omega(0) < \omega(\Theta_B)$ and $\delta = \hat{\omega}(0) < \hat{\omega}(\Theta_B)$.

Proof of Claim 1. Since $u_B(\cdot, \cdot \mid 0)$ is independent of its first argument, (9) readily yields $\omega(0) = \alpha\delta$. But then, by (3) and (9),

$$V\left(1 + \omega(\Theta_B)k, \frac{1}{m}(1 + \omega(\Theta_B)\ k) \mid \Theta_B\right) = V\left(1 + \omega(0)k, \frac{1}{1 + \delta(m-1)}(1 + \omega(0)k) \mid \Theta_B\right).$$

Since $m > 1 + \delta(m-1)$ for all $m > 1$ and $\delta \in [0, 1)$, by strict monotonicity of V, we must then have $\omega(\Theta_B) > \omega(0)$. The second part of the claim is proved similarly. ‖

Claim 2. There exists an $\omega^* \in (\omega(0), \min\{\omega(\Theta_B), \hat{\omega}(0)\})$ such that

$$u_A(1 + (1 - \omega^*)k, 1 + \omega^* k \mid \Theta_A) > 0.$$

Proof of Claim 2. Since $1 - \alpha\delta > (1 - \alpha)\delta$ and $\omega(0) = \alpha\delta$, by definition of α, we have

$$\begin{aligned} u_A(1 + (1 - \omega(0))k, 1 + \omega(0)k \mid \Theta_A) &= u_A(1 + (1 - \alpha\delta)k, 1 + \alpha\delta k \mid \Theta_A) \\ &> u_A(1 + (1 - \alpha)\delta k, 1 + \alpha\delta k \mid \Theta_A) \\ &= 0. \end{aligned}$$

So, since $\min\{\omega(\Theta_B), \hat{\omega}(0)\} > \omega(0)$ by Claim 1, the result follows by continuity. ‖

Take any $\omega_0 \in (\omega(0), \omega^*)$, where ω^* is as found in Claim 2. We propose the following assessment as a candidate for a pure strategy perfect Bayesian equilibrium.

Strategy of A: Both types of A offer

$$(x_A, x_B) \equiv (1 + (1 - \omega_0)k, 1 + \omega_0 k) \in Y$$

in stage 1. In case of rejection, each type of A responds optimally to B's offer in stage 2, that is, type 0 accepts any feasible offer $z \in Y(\delta)$, while type Θ_A accepts $z = (z_A, z_B) \in Y(\delta)$ if $z_A \geq r_A(\Theta_A)$ and rejects if $z_A < r_A(\Theta_A)$.

Strategy of B: If A offers $(1 + (1 - \omega)k, 1 + \omega k)$ in stage 1, then type 0 of B

$$\begin{cases} \text{accepts,} & \text{if } \omega = \omega_0, \text{ or if } \omega_0 \neq \omega > \hat{\omega}(0), \\ \text{rejects and proposes } (1, 1 + \delta k), & \text{if } \omega_0 \neq \omega \leq \hat{\omega}(0); \end{cases}$$

and type Θ_B of B

$$\begin{cases} \text{rejects and proposes } (y_A, y_B), & \text{if } \omega = \omega_0 \\ \text{accepts}, & \text{if } \omega_0 \neq \omega > \hat{\omega}(\Theta_B) \\ \text{rejects and proposes } (1, 1+\delta k), & \text{if } \omega_0 \neq \omega \leq \hat{\omega}(\Theta_B). \end{cases}$$

Beliefs: After any off-equilibrium offer $z \neq (x_{A,} x_B)$, the beliefs of both types of B are degenerate on $\theta_A = 0$.[26] On the equilibrium path, i.e., after A offers $(x_{A,} x_B)$, Bayes' rule applies, hence B still believes that $\theta_A = \Theta_A$ with probability π_A.

In what follows we shall show that there exist $\pi_A, \pi_B \in (0,1)$ such that the above assessment is a perfect Bayesian equilibrium. It is readily verified that this will complete the proof of Proposition 4. In particular, notice that the above assessment specifies a disadvantageous counteroffer for B on its equilibrium path since $\alpha\delta = \omega(0) < \omega_0$.

To establish sequential rationality, take any $\pi_A \in (0,1)$ such that

$$\pi_A \ > \ 1 \ - \max_{\theta_B \in \{0, \Theta_B\}} \frac{u_B\left(y_A, y_B \mid \theta_B\right)}{u_B\left(1, 1+\delta k \mid \theta_B\right)}. \tag{11}$$

Given the strategy of player A in the second stage, the decision problem of type θ_B of B in the second stage is

$$\underset{\gamma \in [\alpha, 1]}{\text{Max}} \begin{cases} (1-\pi_A)\ u_B\left(1+(1-\gamma)\delta k, 1+\gamma\delta k \mid \theta_B\right), & \text{if } \gamma \in (\alpha, 1] \\ u_B\left(y_A, y_B \mid \theta_B\right), & \text{if } \gamma = \alpha. \end{cases}$$

Therefore, (11) ensures that the optimal offer of both types of B is $(y_A, y_B) \in Y(\delta)$ in the second stage.

Sequential rationality for B: Since $\alpha\delta = \omega(0) < \omega_0$, we have

$$u_B\left(1+(1-\alpha)\delta k, 1+\alpha\delta k \mid 0\right) < u_B\left(1+(1-\omega_0)k, 1+\omega_0 k \mid 0\right).$$

Thus, accepting A's offer is optimal for type 0 of B. On the other hand, by (9) and since $\omega_0 < \omega(\Theta_B)$,

$$u_B\left(y_A, y_B\right) = u_B\left(1+\left(1-\omega\left(\Theta_B\right)\right)k, 1+\omega(\Theta_B)k \mid \Theta_B\right) > u_B\left(1+\left(1-\omega_0\right)k, 1+\omega_0 k \mid \Theta_B\right)$$

so that it is optimal for type Θ_B of B to reject A's offer and to counteroffer (y_A, y_B) in the second stage.

It is also clear that B's strategy is sequentially rational for both types $\theta_B \in \{0, \Theta_B\}$ after any off-equilibrium offer $z \neq (x_{A,} x_B)$, since the beliefs about A's type become degenerate on $\theta_A = 0$.

[26] This degenerate specification of beliefs is not necessary, but convenient. All we need to require here is that B's off equilibrium beliefs be sufficiently optimistic.

Sequential rationality for A: We need to show that offering $(x_A, x_B) \equiv (1 + (1 - \omega_0)\, k, 1 + \omega_0 k)$ in stage 1 is optimal for both types of A. Suppose A proposes $(1 + (1 - \omega')k, 1 + \omega' k)$ where $\omega_0 \neq \omega' \in [0, 1]$. Then, given the belief structure, if $\omega' < \hat{\omega}(0)$ the offer is rejected and A is offered $(1, 1 + \delta k)$ in stage 2. Thus, both types of A earn zero utility in this case. But, since $\omega_0 < \omega^*$, by Claim 2 above we have $u_A\,(x_A, x_B \mid \theta_A) > 0$ for any $\theta_A \in \{0, \Theta_A\}$; moreover $u_A\,(y_A, y_B \mid \theta_A) \geq 0$ for any $\theta_A \in \{0, \Theta_A\}$. Thus both types of A obtain strictly positive expected utility on the equilibrium path.

Suppose now that $\omega' \in [\hat{\omega}(0), \hat{\omega}(\Theta_B))$. In this case type 0 of player B accepts A's offer, while type Θ_B rejects and counteroffers $(1, 1 + \delta k)$ in the second stage; hence A's expected utility from offering $\omega' \in [\hat{\omega}(0), \hat{\omega}(\Theta_B))$ is $(1 - \pi_B)\; u_A\,(1 + (1 - \omega')\, k, 1 + \omega' k \mid \theta_A)$. Since $\omega_0 < \omega^* < \hat{\omega}(0) \leq \omega'$ and $u_A\,(y_A, y_B \mid \theta_A) \geq 0$, we have $(1 - \pi_B)\; u_A\,(x_A, x_B \mid \theta_A) + \pi_B\, u_A\,(y_A, y_B \mid \theta_A) > (1 - \pi_B)$

$$u_A \left(1 + \left(1 - \omega'\right) k, 1 + \omega' k \mid \theta_A\right)$$

for any $\theta_A \in \{0, \Theta_A\}$. Hence, given the belief structure, offering $(1 + (1 - \omega')\, k, 1 + \omega' k)$ is not a profitable deviation for either type of A in this case.

Finally, assume that $\omega' \geq \hat{\omega}(\Theta_B)$. In this case both types of B accept A's offer. Therefore, to complete the proof we need to show that

$$(1 - \pi_B)\; u_A\,(x_A, x_B \mid \theta_A) + \pi_B\, u_A\,(y_A, y_B \mid \theta_A) \geq u_A \left(1 + \left(1 - \omega'\right) k, 1 + \omega' k \mid \theta_A\right) \qquad (12)$$

holds for any $\theta_A \in \{0, \Theta_A\}$. Since $\omega' > \omega_0$, we again have

$$u_A\,(x_A, x_B \mid \theta_A) > u_A \left(1 + (1 - \omega')k, 1 + \omega' k \mid \theta_A\right).$$

Thus, if $(1 - \alpha)\, \delta \geq 1 - \omega'$, then $u_A\,(y_A, y_B \mid \theta_A) \geq u_A\,(1 + (1 - \omega')\, k, 1 + \omega' k \mid \theta_A)$ and (12) must hold for all $\theta_A \in \{0, \Theta_A\}$.[27] Assume then that $(1 - \alpha)\delta < 1 - \omega'$, so that $u_A\,(y_A, y_B \mid 0) < u_A\,(1 + (1 - \omega')\, k, 1 + \omega' k \mid 0)$; and there are two possibilities to consider: (i) $u_A\,(y_A, y_B \mid \Theta_A) < u_A\,(1 + (1 - \omega')k, 1 + \omega' k \mid \Theta_A)$; and (ii) otherwise. In case (i), (12) is established for all θ_A by choosing any $\pi_B \in (0, 1)$ such that

$$\pi_B \leq \min_{\theta_A \in \{0, \Theta_A\}} \frac{u_A(x_A, x_B \mid \theta_A) - u_A(1 + (1 - \hat{\omega}(\Theta_B))k, 1 + \hat{\omega}(\Theta_B)k \mid \theta_A)}{u_A(x_A, x_B \mid \theta_A) - u_A(y_A, y_B \mid \theta_A)}$$

since $\omega' \geq \hat{\omega}(\Theta_B)$. In case (ii), on the other hand, proof is completed upon choosing any $\pi_B \in (0, 1)$ such that

$$\pi_B \leq \frac{u_A(x_A, x_B \mid 0) - u_A(1 + (1 - \hat{\omega}(\Theta_B))k, 1 + \hat{\omega}(\Theta_B)k \mid 0)}{u_A(x_A, x_B \mid 0) - u_A(y_A, y_B \mid 0)}.$$

[27] Therefore, for high δ we do not have to put any restriction on the beliefs of A. In particular, for such δ, we may let $\pi_A = \pi_B$.

Proof of Proposition 6. Let $r(\theta) \equiv (1-\theta+\theta/m)^{-1}$ and $b(\theta) \equiv 2m - r(\theta)$ for all $\theta \in [0, \Theta_B)$. Define

$$v_A(x \mid \theta_A) \equiv r(\theta_A)\, u_A(x \mid \theta_A) \quad \text{and} \quad v_B(x \mid \theta_B) \equiv b(\theta_B) u_B(x \mid \theta_B)$$

for all $x \in X$. It follows from (8) that

$$v_A(x \mid \theta_A) = x_A - r(\theta_A) \quad \text{and} \quad v_B(x \mid \theta_B) \equiv b(\theta_B) - x_B \quad \text{for all } x \in Y.$$

Let $\underline{b} \equiv b(\Theta_B) > m$ and $\overline{b} \equiv b(0) = 2m - 1 > \underline{b}$. Define next the distribution function $P : [\underline{b}, \overline{b}] \to [0,1]$ as

$$P \equiv 1 - (F \circ b^{-1}).$$

P is the distribution function of the random variable $b(\cdot)$:

$$P(b) = \text{Prob }(b(\theta) \leq b) = 1 - F(b^{-1}(b)), \quad \underline{b} \leq b \leq \overline{b}.$$

The equilibrium outcomes of our model are, therefore, equivalent to those of the standard single-sale model in which the seller with the known cost $r(\theta_A)$ makes (all) offers to the buyer whose valuation b is distributed on the interval $[\underline{b}, \overline{b}]$ according to the c.d.f. P (cf. Fudenberg and Tirole, 1991, Chapter 10). Moreover, since $\Theta_B < 1$, this bargaining model corresponds to the *gap* case: for all $\theta_A \in [0, 1]$,

$$r(\theta_A) \leq m < 2m - r(\theta_B) = \underline{b}.$$

Consequently, if we can show that P^{-1} is Lipschitz continuous at 0, we may then apply Theorem 3 (and Remark 6.2) of Gul et al. (1986) to complete the proof. To see this, we use assumption (T) to find a $\theta^* \in (0, \Theta_B)$ and $K_0 > 0$ such that $1 - F(\theta) > K_0(\Theta_B - \theta)$ for all $\theta \in [\theta^*, \Theta_B)$. Since r is easily checked to be convex, we thus have

$$1 - F(\theta) > K_0(\Theta_B - \theta) = K_1 r'(\Theta_B)(\Theta_B - \theta) \geq K_1(r(\Theta_B) - r(\theta)), \quad \theta^* \leq \theta < \Theta_B,$$

where $K_1 \equiv K_0/r'(\Theta_B)$. Therefore,

$$1 - F(b^{-1}(b(\theta))) > K_1(b(\theta) - b(\Theta_B)), \quad \theta^* \leq \theta < \Theta_B$$

that is,

$$P(b) > K_1(b - \underline{b}), \quad \underline{b} < b \leq b^*$$

where $b^* \equiv b(\theta^*)$. Since $P(\underline{b}) = 0$, we have $q > K_1(P^{-1}(q) - P^{-1}(0))$ for all $q \in (0, P(b^*)]$. Letting $q^* \equiv P(b^*)$ and $K \equiv 1/K_1$, therefore, we find

$$\left|P^{-1}(0) - P^{-1}(q)\right| < Kq, \quad 0 < q \leq q^*,$$

that is, P^{-1} is Lipschitz continuous at 0.

Proof of Proposition 7. Suppose that f on $\mathcal{T}$ is individually rational and implementable in dominant strategies. By the revelation principle, there must then exist a direct mechanism that truthfully implements f. This implies that we must have

$$u_i(f(u_A, u_B)) \geq u_i(f(u_i^*, u_{-i})) \quad \text{for all } u_i, u_i^* \in \mathcal{T}_i \text{ and } u_{-i} \in \mathcal{T}_{-i}$$

for any $i = A, B$. By (1), therefore,

$$V_i\left(f_i(u_A, u_B), \frac{f_i(u_A, u_B)}{m}\right) \geq V_i\left(f_i(u_i^*, u_{-i}), \frac{f_i(u_i^*, u_{-i})}{m}\right)$$

for all $u_i, u_i^* \in \mathcal{T}_i$ and $u_{-i} \in \mathcal{T}_{-i}$. Thus, by monotonicity of V_i,

$$f_i(u_A, u_B) \geq f_i(u_i^*, u_{-i}) \text{ for all } u_i, u_i^* \in \mathcal{T}_i \text{ and } u_{-i} \in \mathcal{T}_{-i}.$$

Notice that f_i must then be independent of its i-th component, that is, $f_i(v, u_{-i}) = f_i(w, u_{-i})$ for any $v, w \in \mathcal{T}_i$ and any $u_{-i} \in \mathcal{T}_{-i}$. Consequently, we may write

$$f_i(u_A, u_B) = \varphi_i(u_{-i}) \text{ for all } (u_A, u_B) \in \mathcal{T}$$

for some function $\varphi_i : \mathcal{T}_{-i} \to Y$. But by individual rationality, we must have $\varphi_i(u_{-i}) \geq r(u_i)$ for all $(u_A, u_B) \in \mathcal{T}$ so that

$$f_i(u_A, u_B) = \varphi_i(u_{-i}) \ \geq \sup_{u_i \in \mathcal{T}_i} r(u_i) \text{ for all } (u_A, u_B) \in \mathcal{T},$$

which completes the proof of the "only if" part. The validity of the "if" part of the proposition is self-evident.

Proof of Corollary 8. Suppose that f is individually rational and implementable in dominant strategies. Let $\mathcal{V}_i \equiv \{u_i : X \to \mathbf{R} : u_i(x) = (1-\theta)x_i + \theta x_i/\bar{x} - 1,\ 0 < \theta < 1\}$, $i = A, B$. Applying Proposition 7, we have

$$f_i(u_A, u_B) \geq \sup_{u_i \in \mathcal{J}_i} r(u_i) \geq \sup_{u_i \in \mathcal{V}_i} r(u_i) = \lim_{\theta \to 1}\left(1 - \theta + \frac{\theta}{m}\right)^{-1} = m$$

for all $(u_A, u_B) \in \mathcal{J}_A \times \mathcal{J}_B$. The converse claim is readily verified.

References

Binmore, K., 1987, "Perfect equilibria in bargaining models," in Binmore, K. and P. Dasgupta, eds., *The Economics of Bargaining*, Basil Blackwell.

Bolton, G., 1991, "A comparative model of bargaining: theory and evidence," *American Economic Review* 81: 1096-136.

Bolton, G., and R. Zwick, 1995, "Anonymity versus punishment in ultimatum bargaining," *Games and Economic Behavior* 10: 95-121.

Bolton, G., 1997, "The rationality of splitting equally," *Journal of Economic Behavior and Organization* 32: 365-81.

Camerer, C., and R. Thaler, 1995, "Anomalies: Ultimatums, Dictators and Manners,"*Journal of Economic Perspectives*, 9: 209-219.

Cho, I.-K., 1990, "Uncertainty and Delay in Bargaining," *Review of Economic Studies*, 57: 575-96.

Cho, I.-K., and D. Kreps, 1987, "Signalling games and stable equilibria," *Quarterly Journal of Economics*, 102: 179-221.

Clark, A., and A. Oswald, 1996, "Satisfaction and comparison of income," *Journal of Public Economics*, 61: 359-81.

Daughety, A., 1994, "Socially-influenced choice: Equity considerations in models of consumer choice and in games," mimeo, University of Iowa.

Ellingsen, T., 1997, "The Evolution of Bargaining Behavior," *Quarterly Journal of Economics,* 113: 581-602.

Forsythe, R., J. Horowitz, N. Savin, M. Sefton, 1994, "Fairness in simple bargaining experiments," *Games and Economic Behavior*, 6, 347-69.

Frank, R., 1987, *Choosing the Right Pond*, New York: Oxford University Press.

Fudenberg. D., and J. Tirole, 1991, *Game Theory*, MIT Press.

Fudenberg. D., D. Levine, and J. Tirole, 1985, "Infinite-horizon models of bargaining with one-sided incomplete information," in Roth, A., ed., *Game-Theoretic Models of Bargaining*, Cambridge University Press.

Gul, F., and H. Sonnenschein, 1988, "On Delay in Bargaining with One-Sided Uncertainty," *Econometrica* 56: 601-11.

Gul, F., H. Sonnenschein, and R. Wilson, 1986, "Foundations of Dynamic Monopoly and the Coase Conjecture," *Journal of Economic Theory* 39: 155-90.

Guth, W., R. Schmittberger, and B. Schwarze, 1982, An experimental analysis of ultimatum bargaining, *Journal of Economic Behavior and Organization* 3: 367–88.

Herrero, M., 1989, "The Nash program: non-convex bargaining problems," *Journal of Economic Theory*, 49, 266-77.

Kennan, J. and R. Wilson, 1993, "Bargaining with Private Information," *Journal of Economic Literature*, 31, 45–104.

Kirchsteiger, G., 1994, "The role of envy in ultimatum games," *Journal of Economic Behavior and Organization* 25: 373-89.

Levine, D., 1996, "Modeling altruism and spitefulness in experiments," mimeo, UCLA.

Ochs, J., and A. Roth, 1989, "An experimental study of sequential bargaining," *American Economic Review,* 79: 355-84.

Ok, E. A., and L. Koçkesen, 1997, "Negatively interdependent preferences," C.V. Starr Center working paper 97-02, New York University.

Rabin, M., 1998, "Psychology and Economics," *Journal of Economic Literature*, 36: 11-46.

Roth, A., 1995, "Bargaining experiments," in Kagel, J. and A. Roth, eds., *Handbook of Experimental Economics*, Princeton: Princeton University Press.

Rubinstein, A., 1982, "Perfect equilibrium in a bargaining model," *Econometrica*, 50: 97-109.

Young, P., 1993, "An evolutionary model of bargaining," *Journal of Economic Theory*, 59: 145–68.

Figure 1

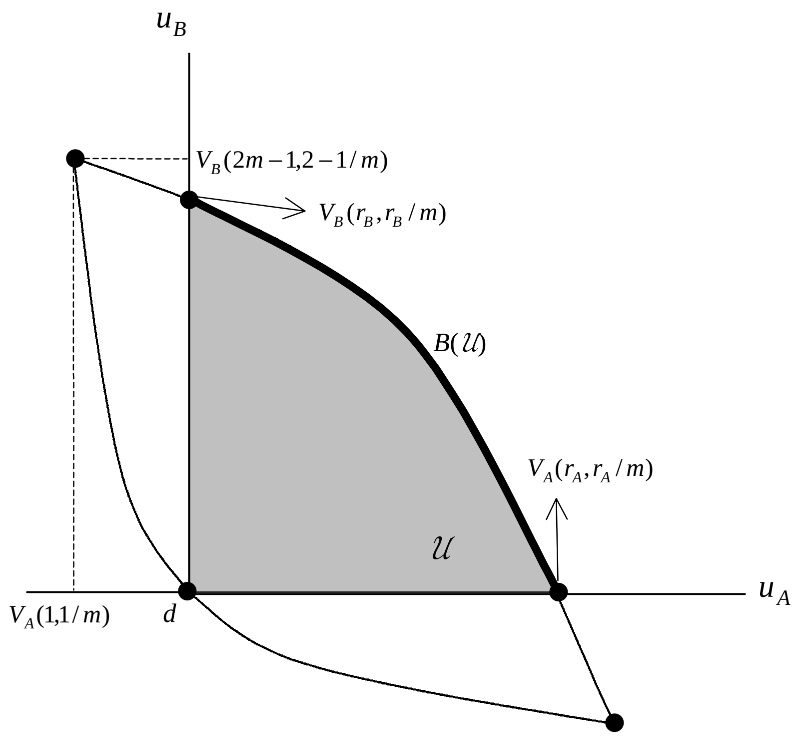

Figure 2