DISCUSSION PAPER SERIES

No. 5495

BARGAINING WITH IMPERFECT ENFORCEMENT

Lucy White and Mark Williams

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP5495.asp**

ISSN 0265-8003

BARGAINING WITH IMPERFECT ENFORCEMENT

Lucy White, Harvard Business School and CEPR
Mark Williams, National Economic Research Associates (NERA)

Discussion Paper No. 5495
February 2006

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 5495

February 2006

ABSTRACT

Bargaining with Imperfect Enforcement*

The game-theoretic bargaining literature insists on non-cooperative bargaining procedure but allows 'cooperative' implementation of agreements. The effect of this is to allow free-reign of bargaining power with no check upon it. In reality, courts cannot implement agreements costlessly, and parties often prefer to use 'non-cooperative' implementation. We present a bargaining model which incorporates the idea that agreements may be enforced non-cooperatively. We show that this has a substantial impact in limiting the inequality of agreements, and results in a non-montonicity of the discount rate. The general need to maintain incentives for co-operation means it may appear that 'other-regarding' elements enter agents' utility functions. This helps us to understand why experimental subjects might begin negotiations anticipating 'fair' bargains. The model also explains why some parties may have incentives to deliberately write incomplete contracts which cannot be enforced in a court of law.

JEL Classification: C72, C78, C91 and D23
Keywords: enforcement, incomplete contracts, non-cooperative bargaining and strength in weakness

Lucy White
Harvard Business School
Soldiers Field Road
Boston, MA 02163
USA
Tel: (1 617) 495 0645
Fax: (1 617) 496 7357
Email: lwhite@hbs.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=148708

Mark Williams
Director
NERA
15 Stratford Place
LONDON
W1N 9AF
Tel: (44 20) 7629 6787
Fax: (44 20) 7493 5937
Email: mark.williams@nera.com

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=147631

*The authors would like to thank Michael Bacharach, Christopher Bliss, Paul Klemperer, Jon Levin, Jim Malcolmson and Jean Tirole for helpful comments and suggestions. White acknowledges financial support from the ESRC and from the European Commission.

Submitted 16 January 2006

"Here's the rule for bargains: 'Do[1] other men, for they would do you.' That's the true business precept." Jonah Chuzzlewit in Charles Dickens' 'Martin Chuzzlewit'.

I Introduction

Non-Cooperative bargaining theory contains a kind of contradiction. It is assumed that during negotiations, bargainers have *no* recourse to a third party, even though a third party might be useful in enforcing participation, or ensuring that the agreement has desirable properties such as symmetry, Pareto efficiency, equality and so on. The agreement will have these properties *only* if they emerge from strategic interaction in a well-defined game. Nevertheless, when the time comes to *implement* the agreement, the implicit assumption is invariably that players *do* make use of a third party who will *costlessly* enforce the agreement they have reached. In short, game-theorists have typically assumed that while bargaining is non-cooperative, implementation is cooperative. Yet if a third party is sufficiently well-informed to costlessly implement an agreement, one might suppose that he could be of some use in arbitrating the initial dispute between the two parties. That is, although the assumption of cooperative enforcement goes hand-in-hand with cooperative bargaining, non-cooperative bargaining matches more naturally with non-cooperative enforcement.

In practice, of course, such well-informed third parties are few and far between. In many countries in the developing world and in transition economies legal enforcement of contractual obligations is often far from perfect (see e.g. Johnson *et al* 1999). Even in the United States, the incomplete contracts literature has made clear that legal enforcement of contracts is used only as a last resort and that the best defence against someone not meeting his contractual obligations is to threaten not to do business with him again (MaCaulay 1963). Thus the typical bargained agreement has to be implemented non-cooperatively. This point is even more obvious when we consider international agreements between nations where no third party has any real enforcement power. A similar difficulty arises in enforcing agreements to divide output between colluding firms. The need for self-enforcement of agreements was recognised as long ago as Telser (1980), but its implications for bargaining theory have never been spelled out. In this paper we provide what to our knowledge is the first analysis of the impact of *non-cooperative implementation on non-cooperative bargaining outcomes*.

[1] To "do" someone is London slang for taking them for a ride, i.e. giving them a rotten deal.

The results are quite striking. The most important result is that non-cooperative implementation *places limits on the inequality of agreements*. The reasoning behind the result is quite simple. If an individual receives relatively little from an agreement, he has little incentive to cooperate in implementing it. Rather, he may find it more profitable to pay lip-service to the agreement, all the while planning to violate the agreement and steal all its profit for himself. In other contexts, game theory has supposed that stiff retaliatory punishments can be threatened to avoid such opportunistic behaviour. We employ a simple model of this type, but show that when bargaining strengths (and hence outcomes) are relatively unequal, even minimax punishment will be insufficient to sustain co-operation. At this point "the stick" becomes ineffective and "the carrot" must be employed: the lot of the weaker player must be improved in order that he finds continuing the agreement worthwhile. Fundamentally, agreements must be sufficiently even-handed that all parties have an interest in seeing them carried out.

This result is important for two reasons. Firstly, it is precisely in the case of unequal agreements where existing non-cooperative bargaining theory fails to predict well.[2] This can now be seen to be at least parly due to the unrealistic assumption of cooperative implementation, which does not match the institutional environment in which bargaining typically takes place. Existing theories for subjects' anomalous behaviour in experiments tend to be based on the idea that subjects care not only about their own pay-off, but that there are additional elements of altruism or envy in their utility functions (see Roth 1995, Bolton 1991). The difficulty of such an approach, as argued eloquently by Cole et al (1992) is that one must exercise caution: "If one is "allowed" to put [an arbitrary property] into agents' utility functions, then it is possible to explain anything. When agents care enough about some arbitrary property, then it should not be surprising that this property emerges." Thus we follow their pioneering approach in trying to explain the presence of other-regarding concepts in the utility function with reference to more fundamental aspects. We argue that though subjects may act as if they care about their opponent's payoff, their day to day experience with non-cooperatively implemented bargains will have taught them to do so: unequal agreements are generally not very profitable, because they are not implementable.[3]

---

[2] For example, Güth and Tietz (1988) write that "the consistency of experimental observations and game theoretic predictions observed by Binmore et al as well as by Fouraker and Siegel is solely due to the moderate relation of the equilibrium payoffs which makes the game theoretic solution socially more acceptable". In fact these authors reject strategic reasoning altogether and argue instead that what drives the result is the (social) need for an equal outcome.

[3] This is not to argue that bargaining agents do not have other-regarding elements in their utility functions; but rather to note that even a selfish player might wish to behave in what would appear to be a somewhat altruistic way. For a similar observation in a different context, see Fehr and Schmidt (2002).

Second, notice that the incomplete contracts literature in fact embodies the same sort of contradiction contained in non-cooperative bargaining theory itself. Ex ante, it is impossible to describe the object in a contract (otherwise the parties could set a fixed price which would in many cases ensure efficient investment), but ex post, the parties bargain and presumably reach an enforceable agreement as to the object to be exchanged and its price. In some cases - such as limited foresight of the parties involved - this is reasonable. But it is strange to assume that ex ante contractual incompleteness arises because it is impossible to write a contract which unspecialised courts will understand, while ex post this difficulty is suddenly removed. Clearly, if courts cannot verify performance at all, the only way object quality can be enforced is non-cooperatively. This calls for a model of non-cooperative bargaining and implementation such as that presented here.

The recognition that the outcome of ex post bargaining in incomplete contract models should, for logical consistency, be non-cooperatively enforceable leads to an important insight for that literaure. The need for implementation affects the bargaining outcome, so *it is possible for some parties to gain from a bargained contract being difficult to enforce in a court of law*. Thus we have a potential foundation for contractual incompleteness. A party which is weak at bargaining would gain a small share of a bargain which can be enforced by courts. But such a party would gain a larger share of an agreement which must be implemented non-cooperatively. Thus he has little incentive to see that the contract is written in a clear and unambiguous manner. It is true that there may be a cost to ambiguous contracts - the inefficient levels of ex ante investment which have been so much emphasised by the literature. But weak parties will trade-off the social costs from not writing a careful contract to induce efficient investment against private gains from writing a contract which does not stipulate their required performance in fine detail, and thus forces their partner to provide them with financial incentives to cooperate. This trade-off between attempts to increase the size of the pie and to increase one's share of it, which is familiar from management textbooks, has until now been absent from economic theory.[4]

The plan of the paper is as follows. The basic model is set out in Section II with the equilibrium derived in Section III. Section IV lists our main results, which are threefold. First, recognising the need

[4] In standard bargaining models, players' pay-offs from bargaining never decline in the size of the pie, so it is difficult to understand why managerial textbooks suggest that there is a trade-off. We thank Bob Wilson for this observation.

for agreements to be enforceable non-cooperatively has an important effect on a substantial proportion of bargained agreements. Second, as stressed above, the typical effect of imposing the requirement of non-cooperative enforcement is to make outcomes more egalitarian. Third, the comparative statics of players' discount factors are non-monotonic in our framework. This comes about because, whilst patience is helpful in negotiation for the standard reasons, a more impatient player can credibly demand a larger share to prevent him from acting opportunistically.

In Section V we single out the two applications which we believe to be of particular importance: Section VA discusses the implications of our model for the incomplete contracts literature, and Section VB its relation to the experimental bargaining literature. Our conclusions are presented in Section VI.

II The Model

Time is infinite and is split into discrete periods 1,2,...,t,... . We assume that players 1 and 2 have risk-neutral time-separable utility functions, with per-period discount parameters δ_1 and δ_2 respectively. In each period, players may either continue bargaining or, if an agreement has been reached, collect a pie. A pie of size one is available to be split between the two players in each period following agreement. The idea behind this is that the each player must take an action to produce the pie, and they will not be willing to do this until agreement has been reached.[5] The game can thus be split into two parts, the bargaining stage and the enforcement stage, as shown in figure 1.

In the bargaining stage, players bargain non-cooperatively to reach an agreement. A bargaining agreement reached at time t consists of a sharing rule (x,1-x) allocating shares x and 1-x of each period's pie to players 1 and 2 respectively. Thus we make the following assumption, which will facilitate comparisons with the existing literature:

A1 Players are restricted to making agreements in which they receive the same share of the pie in each period.[6]

---

[5] In order to keep the model simple and easily comparable with the standard Rubinstein model, we do not model this production process explicitly. It is, however, easy to interpret the stage-game pay-offs in a way consistent with this story – see footnote 7.

[6] One might justify such an assumption on the grounds that when contracts are unwritten, it is better that they are simple and 'focal'. It is in any case without loss of generality when one player is *marginal* (see below) - which we think is the interesting case. In general, efficiency considerations favour increasing the share of the more patient player over time, but co-operation considerations favour increasing the share of the less patient player over time in order to increase the

The precise model of bargaining used here is not important for our results; we choose the well-known infinite-horizon alternating-offer model of Rubinstein (1982).[7] Thus in period 1, player 1 makes an offer (x,1-x) to player 2; if player 2 accepts, the game passes to the enforcement stage; if player 2 rejects, players wait until the next period when player 2 can make a counter-offer, and so on. Then in the absence of enforcement considerations, we have the following result, which is familiar from Rubinstein (1982).

Lemma 1: Assume that any agreement can be enforced (co-operatively). In the absence of enforcement considerations, the immediate outcome of bargaining is (x,1-x) where $x=(1-\delta_2)/(1-\delta_1\delta_2)$. If player 2 moves first, the outcome is (y, 1-y) where $y=\delta_1(1-\delta_2)/(1-\delta_1\delta_2)$.

Table 1: Payoffs in the Stage-game of the Enforcement Stage

	Payoffs (P1,P2)	Player 2's Action *Cooperate*	 *Defect*
Player 1's Action	*Cooperate*	x,1-x	-ε,1
	Defect	1,-ε	0,0

We now turn to the need to enforce agreements. In the enforcement stage, players must decide in each period whether to go ahead in implementing the agreement that they have just reached. The problem is that rather than sticking to the agreement, they can "take advantage" and try to "steal" more than their agreed share from an unsuspecting partner. For example, after signing an agreement with a mine-mouth electricity generator, a coal mining company may reduce the quality of the coal it provides. We model this decision in the simplest possible way. If players agree on a sharing rule (x,1-x) at time t, a pie of size one becomes available to them in this and each subsequent period t, t+1, t+2, ..., ∞. In each period t, t+1, t+2, ..., ∞, each player must decide whether to "co-operate" in implementing the agreement, or to "defect" and steal the whole pie. If both players co-operate, then the agreement is implemented successfully that period, and each gets his agreed share (x for player 1 and (1-x) for

importance which he attaches to the future. We show later that when one player is marginal, these two effects exactly cancel, so shares must be constant over time.

[7] The sense in which this is true will become obvious below. Lemma 2 shows that the need to enforce agreements will always tend to limit inequality, regardless of the bargaining game; this will be especially true in any bargaining game in which impatience is a disadvantage, where the impact of enforcement will tend to make the discount rate act non-monotonically on share.

player 2); if one player defects, he successfully steals the whole pie (1) from his opponent, who incurs a cost ε;[8] if both players try to defect, neither receives anything in that round.[9] The stage-game payoffs are thus as given in Table 1 above. Regardless of the players' choices, the game then passes to the next period,[10] in which the stage game is repeated, and so on ad infinitum. The enforcement game is thus essentially a repeated prisoners' dilemma. Notice that once the bargaining has ended, the co-operative payoffs (x,1-x) are fixed; there is no opportunity to reopen the bargaining.[11] Despite this there will of course be many possible equilibria in the enforcement stage. To simplify matters we will proceed under the following assumption:

A2 Players enforce co-operation in the enforcement stage (if this is possible) by using trigger (or grim) strategies.[12]

The reader should note at the outset that A1 and A2 are made largely in the interests of expositional clarity. We will show below they are in fact without loss of generality for the central case where the need to ensure co-operation restricts the full exploitation of bargaining power. The play of trigger

---

[8] This cost may be arbitrarily close to zero - its only purpose is to improve the analogy to the repeated prisoners' dilemma by ensuring that (C,D) is worse than (D,D) for the cooperating player. On the other hand, if one thinks about the enforcement stage as a game of joint production, it is quite plausible that each cooperating player bears a real non-negligible cost of production in cooperating, whereas if a player fails to carry out his part of the agreement, he bears no such cost.

[9] One can imagine more complicated variations of the same story. For example, players could choose to 'steal' only part of the pie, i.e. if the agreed division is (x,1-x) player 1 could violate the agreement by taking some amount x' strictly between his agreed share x and the whole pie, with player 2 receiving 1-x'-ε. However, it is not difficult to see that if all of these actions are punished in the same way, taking the entire pie will be the optimal form of defection. Allowing different punishments for different degrees of violation of agreements would provide an interesting area of research (see Holmstrom and Kreps 1995 for an interesting start on the problem), but is tangential to the main points we want to make here, so this is a reasonable simplification.

[10] In order to simplify matters, we assume that the time period between successive offers in the bargaining stage has the same length as the time between stage-games in the enforcement stage. This makes it appropriate to use the same discount rates in each case, and thus simplifies the notation. In reality, of course, one might expect offers to be made more frequently in the bargaining stage than the pie is received in the enforcement stage. By effectively making players more impatient in the enforcement stage, this modification would tend to make enforcement considerations even more pressing, so making our point even more strongly.

[11] We take this to be a feature of the bargaining and pie-production technologies: players must meet to thrash out an agreement but may need to be elsewhere in order to produce the pie each period. But one might also argue, on psychological and legal grounds that it is not "reasonable" for players to reopen negotiations when no new information has arrived: see Hart and Moore (2004).

[12] As in footnote 2, allowing players to use less severe punishment strategies would make enforcement of Rubinstein outcomes more difficult, and thus make our point more strongly. Indeed, in the simple model presented here, enforcement may become impossible. This would not be the case in a complex model, where, for example, players cannot extract the whole pie through defection, but rather than complicate the analysis in this way, we restrict ourselves to this simple punishment strategy.

strategies and the use of constant shares over time will then instead arise endogenously through equilibrium refinement.

III The Equilibrium of the Game

We denote the set of possible bargaining outcomes by $B=\{(x_1,x_2): x_1\geq 0, x_2\geq 0, x_1+x_2\leq 1\}$. We denote the set of "enforceable agreements"- those after which cooperation is feasible in equilibrium - by *A*. Notice that players can receive positive payments in equilibrium only if they reach an agreement in the set *A*. The following lemma characterizes this set:

Lemma 2: For all $a\in A$, player 1 receives $x\geq(1-\delta_1)$, and player 2 receives at least $1-x\geq (1-\delta_2)$. *A* is non-empty as long as $\delta_1+\delta_2\geq 1$.

Proof: This is obvious if players use trigger strategies to punish players who defect from implementing the agreement. (By infinite co-operation player 1 can receive x each period, whereas by defection player 1 receives the whole pie, 1, now, and nothing thereafter.) But playing trigger strategies is the worst threat that can be made in the enforcement game, since this holds a player to his minimax payoff following the period of defection. If it is impossible to sustain co-operation through this most severe threat of punishment, then it must be impossible to sustain any other threatened punishment threat. Thus *A* includes only those agreements which have $x\geq(1-\delta_1)$, and $1-x\geq(1-\delta_2)$, so $A\neq\emptyset$ as long as $\delta_2+\delta_1\geq 1$; otherwise the players are so impatient that no agreement can be enforced. □

Thus the effect of recognising that bargains must be enforced non-cooperatively is to impose a lower-bound on each player's payoff from the bargaining game. We will call these lower-bound payoffs the *enforceability constraints* (ECs), since they constrain the players' ability to make low offers. The effect is illustrated in figure 2; whereas in the standard Rubinstein game the whole of the unit square contains feasible, mutually advantageous offers, in our modified Rubinstein game only those bargains inside the inner square (shaded and labelled *A*) are potential equilibrium points. We will define a player whose enforceability constraint just binds his opponent's offer to be a *marginal player*. Player 1 is marginal if he receives exactly $1-\delta_1$ in equilibrium, and Player 2 is marginal if he receives $1-\delta_2$.

Notice from lemma 2 that the lower bound is inversely related to the discount parameter δ; whereas low discount parameters typically receive low payoffs in bargaining games (lemma 1), so the lower-bound on payoffs is highest precisely when the payoffs would otherwise be lowest. As we shall see shortly, this fact means that the enforceability constraint will frequently bind and leads to non-monotonicity in the discount factor. In order to show this, however, we first need to show that the equilibrium of the modified bargaining game has the same qualitative properties as the Rubinstein game. Here we encounter a somewhat technical difficulty first raised by Muthoo (1990) (see also Fernandez and Glazer (1991)).

The Rubinstein game is unusual among infinite horizon games in having (under suitable assumptions on the utility function) a unique equilibrium. The reason why it has a unique equilibrium is that the acceptance of an offer terminates the game. Thus in this instance a player does not have the opportunity to punish his opponent for making or accepting an out-of-equilibrium offer. Once this opportunity is afforded, multiple equilibria again arise in the fashion typical of supergames. Muthoo demonstrates this in the context of a Rubinstein game in which the offerer can withdraw his offer and choose to continue bargaining even after his offer has been accepted. Provided that the players are patient enough, any shares between zero and one can be an equilibrium. This is because the proposer can punish his opponent for accepting an out-of-equilibrium offer (either one that is too generous or one that is too mean) by switching to his opponent's punishment phase.

Essentially the same point can be made even more directly in our model: to support a putative equilibrium agreement (x,1-x) in the set of enforceable bargains *A*, one simply has to suppose that the players expect immediate defection should any other agreement be reached. This provides a very effective punishment to deter players from moving off the equilibrium path, leading to a continuum of subgame-perfect equilibria.

Lemma 3: Any enforceable agreement $a \in A$ reached at any time t can be sustained as a subgame perfect equilibrium of the game.

Proof: Consider a putative equilibrium (x,1-x)$\in A$. Suppose that in the enforcement stage the players play strategies of the following form: play a trigger strategy if the agreement reached is (x,1-x) at time t; otherwise, always defect. Such strategies are clearly best responses to each other and therefore Nash equilibria of the enforcement subgame, and thus SPE of the whole game. Turning back to the bargaining stage, defection is anticipated in every agreement except (x,1-x), so neither player has an incentive to make any serious offer or accept any agreement other than this one at any time except for the specified time t. An agreement other than the norm (x,1-x) at t is expected to result in zero payoffs, whereas (x,1-x) will result in positive payoffs provided it is enforceable. □

The majority of such equilibria are unattractive, however, because they rely on the threat to defect should some agreement other than the "norm" agreement be reached. This threat is not credible because if the players reached an agreement $a \in A$ then they should simply recoordinate to a co-operative equilibrium. Given their discount rates, this could be enforced using suitable punishments if initial co-operation were expected. To formalise this argument, we require our solution to be renegotiation-proof, according to the definition of Pearce (1987) (see also Abreu, Pearce and Stachetti (1993)).[13] The idea behind this concept is roughly that if players find themselves in a punishment phase enduring a payoff *w* when there are subgame perfect equilibria of the game which never require a payoff as low as *w*, then the players will renegotiate to the best of these "better" equilibria. Intuitively, one of the players can complain that the punishment is unnecessarily severe, and that higher payoffs could be enjoyed by renegotiating. However, the proposed deviation from the punishment must itself be credible: if they renegotiate when payoffs are as low as *w*, players must recognise that they could do so again. Thus the proposed deviation must not be a deviation to an SPE which requires punishments of *w* or worse to sustain it.

We now apply this idea to our model. First we show that our assumption A2 was without loss of generality for the case of marginal players.

Lemma 4: Suppose one player is marginal. Then any renegotiation-proof equilibrium of the enforcement stage involves the play of a trigger strategy against the marginal player at the enforcement stage.

Proof: Consider an instance where the marginal player has deviated. Then his opponent will punish him by infinite defection, resulting in zero continuation payoffs. Players might then renegotiate to a "better" equilibrium, in which payoffs as low as zero are never required as punishment. But by

[13] Readers should be aware of an alternative definition of (weak) renegotiation-proofness in the literature, due to Farrell and Maskin (1989). Very roughly, this definition is more concerned with internal rather than external Pareto perfection: an equilibrium is renegotiation-proof if none of its continuation equilibria Pareto dominate one another. Whilst this definition would also serve to eliminate arbitrary social norms, it is too strong for our purposes here, because by limiting punishment options it precludes co-operation in this particular game. (See Pearce 1987 and Abreu and Pearce 1989 for a general argument that the Farrell-Maskin concept, by being relatively inflexible about available punishment, precludes too much

assumption, the defecting player is marginal, so the only equilibrium in which he co-operates is when he is threatened with a trigger strategy: a less severe punishment would result in defection, with zero payoffs. Thus there is no credible deviation from trigger punishment against the defecting player when he is marginal. □

Having imposed renegotiation-proofness on the equilibrium, we can now afford to drop assumption A2. Of course, dropping A2 will result in a different punishment for non-marginal players; the punishment of non-marginal players will not be trigger but rather the least severe punishment available subject to their still finding it optimal to co-operate. This is without consequence for our results regarding the influence of enforcement on bargaining outcomes, however. It is only the need to ensure the co-operation of *marginal* players that impinges on the bargaining outcome. If one wanted to offer a non-marginal player less in a bargaining agreement, then one could simply threaten to punish him more severely subsequently, so bargaining is not restricted. But in the case of marginal players who are already threatened with their minimax payoffs, offering them a smaller share when they co-operate results in defection and so is not an option.

Before turning to the equilibrium of the model, we note that assumption A1, that players agree to receive the same share in each time period, is without loss of generality when enforcement considerations bind:

Lemma 5: When at least one player is marginal, in equilibrium the players agree to receive the same share in each time period.
Proof: See Appendix.

The essential reason is that a marginal player's consumption cannot be brought forward because, by definition, he will then no longer co-operate; and since it is always the most impatient player who is marginal, it is inefficient to delay his consumption.

For readers uneasy about our use of this particular equilibrium refinement, we offer the following two considerations. Firstly, one may think of this invocation of renegotiation-proofness as imposing a

---

co-operation.) Changing the payoffs slightly would allow us to use the Farrell-Maskin concept (which is perhaps the more generally accepted) instead, at the cost of some extra complexity.

certain history-independence on the game. This seems particularly appropriate since we are in effect trying to *explain* the social norm of relatively equal agreements. We might have argued that when agreements must be enforced non-cooperatively, *any* division of the pie can be supported as a social norm; that 50-50 agreements are in fact the norm; and that this explains why so many agreements give exactly equal divisions. However, such an argument would beg several questions: why a social norm should arise at all; why 50-50 division should be selected as the norm; and why many agreements give only relatively (and not exactly) equal shares. Instead we prefer to *explain* the existence of a "social norm" rather than supposing it from the outset. Consequently, we refine equilibria in such a way that arbitrary social norms are essentially ruled out. If players renegotiate when they come to an inefficient outcome, this is essentially the same thing as having them regard all outcomes as in some sense "equally possible" (that is, ruling out any prior expectations about a norm agreement that must be reached, or that defection will ensue if an agreement other than the norm is reached).

Secondly, notice that by adopting Pearce's (1987) renegotiation-proofness as our refinement concept, we are essentially subordinating punishment considerations as far as possible, since we allow the degree of punishment to adapt to the bargaining outcome, treating the outcome of bargaining as a bygone. It is therefore perhaps surprising that enforcement considerations take such a prominent role (see section III). Naturally, if the available punishment did not become more severe as the bargaining outcome demanded, enforcement considerations would become still *more* important. For example, if it were not possible to threaten marginal players with such a severe punishment (or if for some reason the punishment meted out to weak bargainers was less severe) then enforcement considerations would play a greater part. We therefore suspect that our central point - that bargaining models need to take account of the need to enforce agreements - is robust to changes in the equilibrium concept (and indeed to changes in the bargaining model), though the particular refinement used here is very convenient in isolating a unique equilibrium. It is to this that we now turn.

Proposition 1:

Suppose that enforcement is possible, i.e. $\delta_1+\delta_2\geq 1$ so $A\neq\emptyset$. Then the unique renegotiation-proof equilibrium of the Rubinstein game with enforcement consists of a pair of offers (x,y) made by Players 1 and 2 respectively such that:

$x=1-\max(1-\delta_2, \delta_2(1-y))$ and $y=\max(1-\delta_1, \delta_1 x)$

Player 2 accepts any offer x such that $\delta_1 \geq 1-x \geq \max(1-\delta_2, \delta_2(1-y))$, and Player 1 accepts any offer y such that $\delta_2 \geq y \geq \max(1-\delta_1, \delta_1 x)$. In equilibrium, agreement is reached immediately. Co-operation by marginal players is enforced through the use of trigger strategies against them.

Proof: As in the standard Rubinstein game, the best response of the offering player to the recipient's strategy is to offer as little as possible subject to eliciting acceptance of the offer. The best response of the recipient is to accept any offer which gives him at least as much as he can expect in the continuation game starting with his own counteroffer. In the Rubinstein game this implies a pair of offers such that $1-x=\delta_2(1-y)$ and $y=\delta_1 x$. Here we have the same except that from lemma 2 we have the additional constraint that players must offer deals that are enforceable.

From lemma 4 we know that if a player is marginal, then the use of a trigger strategy against him is not subject to renegotiation. So to induce player 2 to accept, player 1 must make an offer which is not only greater than player 2's Rubinstein continuation payoff (1-y), but also greater than player 2's enforceability constraint $(1-\delta_2)$. (The assumption that $\delta_1+\delta_2 \geq 1$ ensures that player 1 still leaves himself enough $(1-\delta_1)$ that he too will co-operate.) If the EC is the larger, then player 2 is marginal, and can be credibly threatened with a trigger strategy.

Player 2 faces analogous considerations when he makes an offer. The remainder of the proof (which is given in the Appendix) runs along familiar lines. The reader is referred to Rubinstein (1982), Shaked and Sutton (1984), and Osborne and Rubinstein (1990) for a more detailed discussion.□

IV Results

A. The Importance of Enforcement Constraints

From Proposition 1, we can see that five types of outcome are possible, depending on the relative values of the players' discount rates:

(i) $\delta_1+\delta_2 \geq 1$, $\max(1-\delta_1, \delta_1 x)=\delta_1 x$ and $\max(1-\delta_2, \delta_2(1-y))=\delta_2(1-y)$: Neither enforceability constraint binds: the bargained outcome is exactly that described by Rubinstein (1982).

(ii) $\delta_1+\delta_2 \geq 1$, $\max(1-\delta_2, \delta_2(1-y))=1-\delta_2$ and $\max(1-\delta_1, \delta_1 x)=\delta_1 x$: player 2's enforceability constraint binds only player 1's offer.

(iii) $\delta_1+\delta_2 \geq 1$, $\max(1-\delta_1, \delta_1 x)=1-\delta_1$ and $\max(1-\delta_2, \delta_2(1-y))=\delta_2(1-y)$: player 1's enforceability constraint binds only player 2's offer.

(iv) $\delta_1+\delta_2\geq 1$, $\max(1-\delta_2, \delta_2(1-y))=1-\delta_2$ and $\max(1-\delta_1, \delta_1 x)=1-\delta_1$: player 2's enforceability constraint binds player 1's offer and player 1's enforceability constraint binds player 2's offer.

(v) $\delta_1+\delta_2<1$, no enforceable agreement is possible. This happens when a player's own EC binds his own offer, since (in the case of player 1 for example) this implies that $\delta_1<1-\delta_2$ so $(1-\delta_1)+(1-\delta_2)>1$.

Table 2 summarises when each of these possibilities arises. One can get an idea of the discount factors for which these different possibilities arise by finding for each possible value for δ_2, the critical value for δ_1, which we denote δ_1^*, at which player 1's enforceability constraint just binds player 2's Rubinstein offer. That is, $\delta_1^*(\delta_2)$ solves the equation:

$$1-\delta_1^*=(1-\delta_2)/(1-\delta_1^*\delta_2) \quad (1)$$

The resulting function δ_1^* is illustrated in figure 3. $\delta_2^*(\delta_1)$ can be found similarly as the value of δ_2 for which player 2's enforceability constraint just binds player 1 to make the Rubinstein offer $\delta_2(1-\delta_1)/(1-\delta_1\delta_2)$, and is also illustrated in the figure. It is evident that if players are too impatient, no enforceable bargain can be struck. What is perhaps more surprising is that even among discount rates where enforcement is possible, the "Rubinstein zone" in which neither player is constrained by enforcement considerations is relatively small.[14] Quite naturally, it consists in the area where players are relatively patient (so that enforcement considerations are relatively unimportant), *and* have reasonably similar discount rates (so that the Rubinstein shares are relatively equal and no player can make a very large gain by defecting and stealing the whole pie). We thus have:

Result 1: Introducing the need to enforce bargains non-cooperatively has an important impact on bargaining outcomes. The set of outcomes which are unaffected by it (a) is relatively small and (b) consists of outcomes which are already rather equal.

Of course, the probability of different combinations of discount factors arising in reality is largely an empirical matter, so that figure 3 can tell us nothing directly about the *empirical* importance of

[14] Integrating the area between the two curves indicates that the Rubinstein zone R takes up only about 7% of the unit square (or 14% of enforceable bargains). The areas EC_{12}^* and EC_{21}^* zones occupy about 21% of the unit square each and the region C* where both players are constrained in their ability to make Rubinstein offers occupies about 1.5%. (In fact, because the figure considers only the case where players' Rubinstein offers are constrained, the area C where players'

enforcement considerations. However, we offer three considerations on this issue. Firstly, notice that enforcement considerations remain important *even when both players are relatively patient*. (This is because a given difference between players' discount rates becomes more critical in the Rubinstein game as both players become more patient, as can be seen from table 3a and figure 4a below.) Thus the need to agree an equal bargain to ensure co-operation is not limited to players who meet infrequently to bargain or collect their pie.

Secondly, the importance of enforcement considerations is underlined by considering circumstances where the choice of discount rates might be endogenous. In particular, suppose that a principal P_i hires an agent A_i to reach and implement a bargained agreement with his opponent P_j. In order to provide the agent with incentives, the principal will allocate a share β of the profits from the agreement to the agent, keeping a fraction $1-\beta$ for himself.[15] Suppose that P_i can choose how patient an agent to hire, and that he can really commit to employing an agent with preferences different to his own. (The difficulties in doing so are familiar from Katz 1991). Then the weak best response of P_i to P_j's discount factor δ_j is to choose $\delta_i=1-\delta_j$.[16] In other words, when a player's discount factor is modelled as endogenous, in equilibrium that player's enforcement constraint will bind.[17]

Thirdly, the available experimental evidence suggests that subjects act as if enforcement considerations are relevant even when they have been "artificially" removed by the experimenter (see Section V part B). This suggests that in the everyday experience which shapes subjects' "rules of thumb", enforcement considerations are indeed important.

B. The Effect of Enforcement Constraints

---

actual offers are constrained is somewhat larger that C*, at the expense of EC_{12} and EC_{21} being somewhat smaller; the important point, however, is that R is unaffected by this and relatively small.)

[15] For simplicity we treat β as exogenously given and independent of the agent's discount rate. We assume that the agent cannot afford to buy the project from the principal, so that $\beta\neq 1$. We also assume that the agent continues to receive the fraction β of profits if he defects, so that the size of β does not affect incentives when parties are risk-neutral.

[16] The reason for this is the following. Since the P_i shares profits with A_i, whatever P_i's discount rate, he will always choose an agent with a discount factor which maximises the stage-game payoff. This is achieved by holding his bargaining partner to the minimum payoff that still ensures enforcement. So if his partner has discount factor δ_j, in equilibrium the principal will choose an agent with discount factor δ_i such that $\delta_i=1-\delta_j$. An alternative (and equally profitable) strategy, if available, would be to choose $\delta_i=1$; but if both principals are able to delegate, $\delta_i=\delta_j=1$ is not a Nash equilibrium, whereas $\delta_i=1-\delta_j$ is.

[17] We could alternatively have allowed the principal to delegate each of the bargaining and implementation functions to two independent agents. In this case it is obvious that the principal will try to arrange things so that enforcement constraints bind. But what is more surprising is that when the principal has to delegate both bargaining and implementation to the same

At this stage, it might seem obvious to the reader that, because (a) enforcement constraints place a lower-bound on the amount that each player can receive, and (b) this lower bound is often effective, the effect of enforcement constraints is to make agreements "more equal" than the standard Rubinstein game would predict. In fact such a claim is non-trivial. The reason for this is that it is possible for the Rubinstein game to yield relatively equal outcomes where player 1 is somewhat more impatient than player 2, since then the former's first mover advantage is offset by the latter's greater patience. In this case, introducing enforcement constraints reduces the extent to which player 2 can gain from his relative patience, and so can result in a more unequal division.[18] However, it is true that the "typical" effect of imposing enforceability is to reduce the inequality of stage-game pay-offs, in the sense of making the division closer to 50-50 (compare tables 3a and 3b). The tendency towards equality is even greater when one compares total inter-temporal pay-offs, since the imposition of enforcement constraints invariably (weakly) advantages the least patient player (see tables 4a and b). We summarise our results in the following:

Result 2: Requiring the non-cooperative enforcement of bargains typically restricts the "injustice" that can be done to the less patient player (or to the player moving second), thus making outcomes more equal.

In fact, enforcement considerations provide a second, distinct, reason why the outcome of real-world (as opposed to experimental) bargaining is likely to be equal. Unequal outcomes can result from first-mover advantage, which is greater when players are impatient. But when enforceability is imposed, only players who are together relatively patient ($\delta_1+\delta_2\geq 1$) will be able to form agreements at all. The (impatient) players who would receive the most inequitable offers are excluded from bargains because it is known that they will not stick to them. This selection process means that the average agreement reached is more equitable than might otherwise be supposed. The operation of such a selection mechanism would also tend to reinforce subjects' initial expectation that agreements reached in experiments will be relatively equal.

C. The Comparative Statics of the Discount Factor

---

agent, he will tend to choose a relatively impatient one; in other words that enforcement rather than bargaining considerations will predominate.

Further inspection of Tables 3b and 4b yield the observation that when the Rubinstein game is modified by the requirement of enforceability, a player's payoffs are no longer monotonic in his own or his opponent's discount rate. In the Rubinstein zone R of figure 3, players benefit from being patient and from their opponent's being impatient, as one would expect. In all other regions, however, at least one player's discount rate has perverse effects. In Zone EC_{12}, player 1's enforceability constraint binds player 2's offer; therefore player 1's stage-game payoff is *decreasing* in his own discount factor. In this region player 1's total payoff between now and infinity is invariant to his discount factor, since in equilibrium player 2 ensures that player 1 receives just enough to prevent him defecting. But because of this effect, player 2's total and stage-game payoffs are both higher if his opponent is *more* rather than less patient locally.

Similarly, in Zone EC_{21}, player 2's stage payoff is increasing as he becomes locally more impatient; player 1 is made worse off by this. In Zone C, where both players' offers are constrained by the need to ensure enforcement, both players' payoffs are increasing in their own impatience. In summary, and in contrast to the usual results about the effect of discount factors in bargaining games, we have:

Result 3: When player *i* is constrained in the offer he can make to player *j* by the need for enforcement, player *i* is better off if player *j* becomes more patient. Player *j* increases his stage-game payoff if he becomes more impatient. The reverse holds when player *j* is sufficiently patient that player *i* is not constrained, so the effect of player *j*'s discount factor on player *j*'s payoff is non-monotonic.

V Some Applications

A. Contractual Incompleteness

The typical incomplete contracts model runs as follows (see Hart 1995). In Stage 0, parties choose non-contractible investments. It is not possible to contract on a price for the widget at this stage because the widget cannot be described in a contract. In Stage 1, investments determine the size of the pie to be divided between parties; ex post bargaining determines this division of surplus. The process of ex post bargaining is usually black-boxed. To the extent that it is discussed, it is usually assumed that Nash or Rubinstein bargaining takes place. But as we have shown above, this will not generally be the case, because for consistency it should also be impossible to contract on widget production ex post.

[18] Notice that for some parameters the division does become more equal in the sense of *moving in the direction of equality* by benefiting the player who is initially worse off, even though the eventual division of stage-game pay-offs is further

(It is not obvious why courts should suddenly become better equipped to deal with the subtleties of the situation.[19]) This being the case, having reached an ex post bargained agreement, the only way to ensure performance is if the parties are in a repeated relationship. This is possible since both parties know what constitutes appropriate behaviour according to their agreement even though they cannot explain this to a court: they can interpret delivery (by the seller) and payment of the agreed amount (by the buyer) as cooperative strategies and non-delivery or non-payment (except under what according to their informal contract are extenuating circumstances) as "cheating". The lack of recourse to courts means that cheats are best dealt with by refusing to deal with them in future.[20]

This means that the appropriate model of ex post bargaining for the incomplete contracts literature is one with ex post non-cooperative enforcement, such as the one presented here. Up until now, the model of ex post bargaining has been seen as a relatively unimportant feature of incomplete contracts models. But the important point that arises from our model is that *contractual incompleteness leads to a shift in bargaining power*. Thus it may often be *in the interest of at least one of the parties to have an incomplete contract.* [21] By abstracting from the details of ex post bargaining, economists have previously overlooked this point, and thus an important foundation for contractual incompleteness has gone unrecognised.

To investigate this issue further, we relax our assumption that defecting players can simply "steal" the whole pie with impunity in the enforcement stage should they decide to do so, and assume instead that by writing a "better" contract, players can limit the fraction of the pie that can be extracted by a defecting player to θ ($0 \leq \theta \leq 1$).[22] This yields the pay-off matrix for the stage-game of the enforcement

---

away from 50-50 in the other direction.

[19] One important exception to this arises if the incompleteness is due not to contracting difficulties, but to parties' limited foresight as to what type of widget they might want (see Segal 1999).

[20] The work of Baker, Gibbons and Murphy on relational contracting is concerned with the implications of this fact for the design of incentive contracts within and between firms; for a survey, see Baker et al (2001).

[21] This possibility can never arise in an incomplete contracts model where bargaining is over surpluses and it is assumed that bargaining power is the same ex post and ex ante. Then parties will receive $\alpha V(I_1*,I_2*)$ and $(1-\alpha)\ V(I_1*,I_2*)$ with a complete contract, and $\alpha V(I_1,I_2)$, $(1-\alpha)\ V(I_1,I_2)$ with an incomplete contract, where surplus is higher when investment is effficient: $V(I_1*,I_2*) > V(I_1,I_2)$. Then both parties want a complete contract. This will not be generally true when bargaining powers differ between the two situations, as - we have argued - will certainly be the case. Note that bargaining powers might also differ ex post and ex ante (rather than with complete and incomplete contracts) if investment cost functions are asymmetric. This might give another reason for preferring incompleteness.

[22] θ can be given a wide variety of interpretations within this framework: it might represent a deterministic fraction of the pie that can be extracted without cost; the expected gain from defection net of possible legal punishments; the probability of one's opponent not finding it worthwhile to go to court to get the stolen pie back; or assuming that court action is taken, the probability of being found innocent and not forced to repay the stolen pie. All of these possibilities can be thought of as

stage given in table 5 below. The degree to which it is possible for defecting agents to gain from defection most likely depends upon how carefully the contract agreed by the bargaining parties is written. For example, it may be very difficult to hold someone to a verbal contract; the more "loopholes" that can be eliminated from a written contract, the smaller the fraction of the pie that players can steal without being punished. If the contract agreed is very detailed, covers most eventualities and is written in relatively unambiguous language, then the expected costs of getting a court to rule on whether the contract has been breached will be relatively low, as will the uncertainties associated with the likely decision.

Table 5: Payoffs in the Stage game of the Enforcement Stage with degree of incompleteness θ

	Payoffs (P1,P2)	Player 2's Action: *Cooperate*	*Defect*
Player 1's Action	*Cooperate*	$x,1-x$	$-\varepsilon,\theta$
	Defect	$\theta,-\varepsilon$	$0,0$

The effect of varying the degree of incompleteness θ is felt on the enforceability constraints, which become slacker: cheating is less attractive if less pie can be stolen. Thus the discount factor $\delta_1^*(\delta_2,\theta)$ at which player 1's enforceability constraint is just binding on Player 2's Rubinstein offer is now given by:

$$\theta(1-\delta_1^*)=(1-\delta_2)/(1-\delta_1^*\delta_2) \qquad (2)$$

As anticipated, increasing the degree of contractual completeness reduces the tendency towards equal outcomes and the other results noted in section IV. However, though enforcement becomes easier as contracts become more complete, the Rubinstein zone remains relatively small compared to the areas where at least one of the enforceability constraints is binding for quite low θ. Figures 4a-c illustrate the cases where θ=0.75, 0.5 and 0.25 respectively, and show that the results given above are in no way special to the case θ=1. It is rather the Rubinstein game - where bargaining power is given free reign - that represents a polar case.

dependent on what we might loosely call the degree of contractual incompleteness. The case θ=1 dealt with up until now then represents a wholly incomplete (or "non-existent") contract; the case θ=0 a wholly complete contract, where players cannot gain from defection. An alternative interpretation of θ is the (inverse) "cost of breaking your word", which might be either psychological or reputational.

In general, the more that contracts are incomplete, the more players must "trust" each other not to defect, and the more equal outcomes must be. When the courts can make defection unprofitable, outcomes can be much more unequal; strong bargainers can push their advantage to a greater extent.[23] To address the issue from another perspective, players who are relatively impatient will prefer to bargain in situations where contracts are incomplete and legal enforcement is difficult, provided that they can find a willing partner.

Since increasing the degree of completeness slackens the enforceability constraints, it is the players whose enforceability constraints bind (raising their payoff above what they could expect in a Rubinstein framework) who will object to having a more complete contract written. Thus, although one might think that incompleteness would allow the strong to exploit the weak, quite the contrary is true: contractual incompleteness benefits the weaker bargainer. Even if writing a complete contract is entirely costless (as we have assumed here), the weaker bargainer would prefer to have a more informal agreement instead. In fact in our model, the weaker bargainer, given the unilateral choice of θ, will generally wish to choose the most incomplete contract that is feasible, subject to an enforceable agreement still being possible.

Of course, the stronger bargainer on the other hand will often have every incentive to make sure that a detailed contract drawn up. However, contract writing is rarely a unilateral process. If the stronger bargainer is to eliminate loopholes through which the weaker bargainer can defect, he may need the latter's assistance in doing so, and the latter may have little incentive to provide the requisite information (because after it is provided he can be restricted to the Rubinstein pay-off).[24] There will also be cases where both players are marginal and thus neither wishes to take a contractual approach. In fact, it is this last situation that seems to correspond most closely to the situation which MaCaulay (1963) describes.[25]

---

[23] It is tempting to regard this as corresponding to a distinction made by the general public between hard-nosed "big business" (which suffers disapproval because of its contractual approach), and small, more informally organised firms.

[24] Our argument here is essentially that it is extremely difficult to write contracts for the transfer of information. If the weak player W could guarantee to provide the strong player with the correct information to write into the contract to ensure that a widget is accurately described, and the strong player S could guarantee to pay a lump sum for the information to compensate the weak player for his lost bargaining power, then there would be a Pareto gain. But precisely because it is difficult to contract on the description of a widget in the first place, so it is difficult or impossible for a court to rule on whether W indeed provided S with the correct information; and thus for S to promise to make a payment contingent on this provision.

[25] In general each party will have an incentive to limit the ability of his partner to defect, whilst retaining as many loopholes as possible for himself. This is reminiscent of MaCaulay's description of "the battle of the forms".

The foundation we have provided for contractual incompleteness has several testable predictions. One should expect to see strong bargainers attempting to take a more contractual approach than weak bargainers. Impatient bargainers prefer informality. Other things being equal, agreements which are more legalistic will tend to involve more unequal shares than informal agreements. It is also probable that more contractual solutions will be chosen when the pie to be divided is larger.[26]

One may object to this foundation for contractual incompleteness by arguing that we have here neglected the costs of incomplete contracts: the inefficient investment levels which after all are the focus of the literature. In a more general model including ex ante investment, greater incompleteness may lead to reduced investment incentives and thus a smaller pie to be divided. So the weaker party would choose θ to trade-off his private bargaining gain from incompleteness against any loss due to inefficient investment. But if investment is not too elastic, it is entirely possible that he will choose to have a larger share of a smaller pie. Notice moreover that once the endogeneity of bargaining power is recognised, it is no longer obvious that increased incompleteness (from $\theta>0$) does indeed reduce investment incentives: if the party whose investment is more important is also the weaker bargainer, investment incentives may be improved by contractual imcompleteness, because he will extract a larger share of ex post surplus.[27] Thus we have a sort of theory of the second-best: if contracts must be a little incomplete, it might be best for them to be very incomplete. We also note in passing that the vertical integration sometimes advocated by the incomplete contracts literature as a solution to the under-investment problem is a very blunt instrument for allocating bargaining power to improve investment incentives; it is easy to imagine that varying the degree of contractual incompleteness θ may be much less costly.

B. Discussion of Experimental Evidence

Experimental work on non-cooperative bargaining has shown that subjects do not behave as ruthlessly in making offers as the standard theory predicts: pie-division is often "too equal" compared to the

[26] If writing a more complete contract is costly, then the stronger bargainer will trade-off the cost of reducing θ against the cost of giving his opponent a larger share in order to induce co-operation. Supposing, as seems plausible, that the cost of drafting a contract rises less than in proportion to the total value of the agreement, when only small sums are involved it will be relatively cheaper to simply give away a larger share of the pie; more contractual solutions will be chosen when the pie is larger.

perfect equilibrium. Moreover, unequal offers are frequently rejected. This has led some experimenters to call for a complete rejection of game theory as a modelling tool (Guth and Tietz 1990, p440). Others take the less extreme view that whilst subjects do act strategically and aim to maximise their utility, the traditional view of economic man as entirely self-interested needs to be modified (Ochs and Roth 1989, Bolton 1991, Fehr and Fischbacher 2002).[28]

However, many economists remain wedded to the idea of rational (strategic) self-interested agents, and thus it is important to see how far we can get using this notion.[29] As argued forcefully by Gale *et al* (1995) and Cole *et al* (1992), we should be hesitant in inserting additional elements into the utility function, because in itself this may add little to our understanding. We need to be very careful to avoid tautology in explaining equitable agreements by arguing that agents gain utility from equitable agreements.[30] Here we put forward an alternative explanation. We argue that one reason that current game-theoretic models may fail is that they omit a crucial feature of the institutional environment which is present in virtually all real-world bargaining games, and which boundedly rational subjects may therefore incorporate into their "rules of thumb" for playing bargaining experiments. This feature is the fact that bargained agreements need to be self-enforcing. Once included in game-theoretic bargaining models, their predictions come much more into line with the available evidence.

---

[27] This property is likely to be model-specific however, depending for example on whether investment is "selfish" or "cooperative". See Che and Hausch (1999) for a definition of these terms, and an explanation of how investment may be higher without an explicit contract.

[28] In his survey of the literature on experimental bargaining, Roth (1995) sums up the current consensus as follows: "the evidence from all these sequential bargaining games suggests that some of the away-from-equilibrium behaviour...results from bargainers' preferences that concern not only their own income but also their relative share. At the same time, the evidence suggests that much of the away-from-equilibrium behaviour does not have such a simple cause, but results instead from strategic considerations...". Our model shows that there is in fact no tension between self-interested strategic motivation and the apparent desire that responders should not receive too little. For more recent surveys of reciprocal preferences in a more general set of games see Fehr and Schmidt (2002).

[29] See Fehr and Schmidt (1999) for an interesting theory which incorporates both reciprocal and self-interested agents and studies the interactions between them. See also Fehr and Schmidt (2000) which uses the existence of these preferences to try to explain why contracts may be endogenously incomplete. Our theory can be viewed complementary to theirs in that it can also help understand why, if agents begin with inequity-averse preferences, these do not die out. If contracts are incomplete, we show that equitable bargains are more profitable than inequitable ones, so such preferences may be evolutionarily stable in such an environment.

[30] There may also be other problems with treating fairness as a primitive concept. For example, Binmore et al (1991) show that subjects seem to modify their views about what is fair according the strategic situation as well as what actually occurs. Moreover, in some bargaining situations it seems entirely implausible that there is any altruism or envy between the parties (see e.g. the discussion of Eckel and Grossman 1995). In addition, Ochs and Roth (1989) note that their evidence is consistent with the idea that proposers were self-interested and trying to maximise their earnings; it is just that the perverse behaviour of responders meant that the way to achieve this was to make relatively equal offers.

In surveying previous experimental bargaining studies[31] and reconciling them with their own results, Ochs and Roth (1989) draw attention to five particular regularities in the data which are not consistent with standard non-cooperative bargaining theory, and which they and subsequent authors have sought to explain:

- *R1* A consistent first mover advantage was observed in all cells, regardless of the value of δ_2 (i.e. regardless of whether there should be a first mover advantage).
- *R2* The discount factor of player 1 was observed to influence the outcome even in two-period games.
- *R3* A substantial percentage of first offers were rejected.
- *R4* The observed mean agreements deviate from the equilibrium predictions in the direction of equal division.
- *R5* A substantial proportion of rejected offers were followed by disadvantageous counter-offers.

These results can be explained by adopting the hypothesis of this paper: that experimental subjects - and indeed economic agents in general - habitually encounter negotiation situations where the agreement subsequently has to be enforced *non-cooperatively*. Comparing tables 3a and 3b, there is a consistent first-mover advantage (R1) in the Rubinstein game with non-cooperative enforcement since on the occasions when player 1's relative impatience would normally offset his first-mover advantage, his enforceability constraint prevents him receiving too little. Since the same factor operates in favour of player 2, the hypothesis predicts that agreements will differ from the expected ones in the direction of equal division (R4). Indeed, in this respect the results of the model are very much as Spiegel *et al* (1994) remark of their experimental investigation,

> "First players who have a great advantage in the sense that the Stahl/Rubinstein division would give them a large share of the initial pie demand more than half of the pie, but do not fully exploit their advantage. First players who are at a disadvantage demand equal divisions."

Moreover, the influence of enforceability constraints in a two-period game would make player 1's discount rate relevant, thus explaining (R3). The fact that offers are rejected in favour of making disadvantageous counter-offers (R5) can also be explained in this framework. Whether an agreement is enforceable or not depends not on the size of the pie, but on the relative share of the pie that each

[31] Güth *et al* (1982), Binmore *et al* (1985), Güth and Tietz (1988), and Neelin *et al* (1988).

player receives, for in our model it is the share that influences the incentive to defect. Thus it would be rational to reject an unenforceable agreement on a large pie in favour of an enforceable agreement on a small pie.

Of course, when they play the bargaining games designed by experimenters, subjects do not actually face problems of enforceability. So to the extent that one retains the assumption of complete rationality on the part of experimental subjects, the model does not explain the data. However, it is more than plausible that subjects initially try to solve these new problems based on their previous experience, especially since the stakes are usually quite small. And their experience almost certainly involves everyday agreements, e.g. with other students, in which tasks are shared for mutual benefit. Since such agreements are only enforced by the threat of termination, rewards cannot be too unequal if the agreement is to be profitable. Thus our model suggests that subjects' everyday experience leads them to regard agreements that are too one-sided with suspicion; so that quite naturally they do not begin by suggesting such agreements. Even if they themselves know that the usual problems with unequal bargains have no weight in the laboratory context, they may be uncertain as to whether their opponent has fully recognised this fact. Indeed, the fact that unequal offers are sometimes made and rejected (R3) can be attributed to the lack of common knowledge on this point.[32]

The experimental evidence can thus be explained solely in terms of fully self-interested - albeit boundedly rational - individuals. But why, if behaviour in experiments is a matter of bounded rationality, as compared to a matter of genuine preference, do experienced subjects not learn and "correct" their behaviour? The answer is that learning in experiments is a very subtle matter, as has only recently been fully realised. Roth and Erev (1995) show that the way in which an experiment will proceed is very sensitive to the initial conditions, so that if subjects *begin* by offering relatively equal amounts and rejecting unequal ones, the experiment will not necessarily converge to an unequal equilibrium.[33] The model presented here explains why subjects *do* begin by making equal offers.[34] In a

---

[32] For a discussion of how lack of common knowledge increases the frequency of disagreement in another bargaining context see Roth and Murnighan (1982).

[33] The intuition for their result is that proposers who have unequal offers rejected lose much more than the responders who reject them; thus the former are likely to learn more quickly not to make unequal offers than the latter not to reject them. On this basis, they argue that the experimental data on the Ultimatum game can be understood as a set of medium run observations of a learning process; Gale et al (1995) argue that the persistence of dispersed proposals and rejected offers tend to imply that this is indeed the case.

sense what we have set out are some ‘microfoundations’ for a theory of gift exchange (Akerlof 1982) or reciprocity (Fehr and Gächter 2000): one party gives the other more than he is entitled to on “pure” bargaining considerations. However, he does so not for altruistic reasons but because this reflects his own best interests in inducing co-operation in pursuit of a mutual pie.

VI Conclusion

Game-theoretic bargaining theory has until now implicitly assumed that whilst agreements must be reached *non-cooperatively*, their implementation takes place entirely *cooperatively*. Yet it is well known that in reality when an agreement is signed many loopholes remain through which a reluctant party can escape its obligations. As has been made clear by the incomplete contracts literature, in the vast majority of cases legal solutions are sufficiently uncertain and expensive that they are simply not worth pursuing; the only way to enforce co-operation is to refuse to do any further business with the offender (MaCauley 1963).

Therefore it is important to model the use of non-cooperative punishments in implementing bargaining agreements, and this paper provides the first attempt to do so. Our results are threefold. First, we have shown that this modification constrains the degree to which bargaining advantages can be pressed home. This constraint is frequently binding, even when players are relatively patient. Second, pay-offs are non-monotonic in players’ discount rates: impatient players may receive a larger share of the pie by becoming *more* impatient. Third, and most important, outcomes are typically more equal than predicted by standard bargaining models (with co-operative implementation). Non-cooperative implementation requires that each party has a large enough stake in the agreement to find implementation worthwhile. Thus real-world outcomes are egalitarian, not simply due to some innate altruism, but because ‘unfair’ agreements are frequently unprofitable.

These results may help us to understand two further phenomena. Firstly, when subgame perfect equilibrium would entail inequitable division, laboratory subjects often fail to play SPE strategies. This may be because they other-regarding elements (altruism or envy) enter their utility functions (Ochs and Roth (1989), Bolton (1991), Fehr and Fischbacher 2002). But it is important to explore under what conditions boundedly rational self-interested individuals would exhibit the same behaviour. It is

[34] Gale *et al* (1995) explain the starting point by suggesting that in the real world bargaining power is usually symmetrically distributed. Our argument is stronger: even if bargaining power were asymmetrically distributed, this could

known that if self-interested subjects start out with egalitarian expectations then plausible learning processes may fail to produce subgame-perfect outcomes (Roth and Erev (1995), Gale *et al* (1995)). Using the results of this paper, we are able to explain why subjects may begin with egalitarian expectations or preferences. *In the real world rational self-interested economic agents will not reach very inegalitarian agreements, even if bargaining power is distributed very asymmetrically, because bargained agreements need to be self-enforcing.*

Secondly, our result can be used to provide foundations for the incomplete contracts literature. It will generally *not* be in the interest of the weaker (more impatient) party to write a contract which is easily interpreted and enforced by the courts. The weaker party may prefer to sign an *incomplete* contract, since this must be non-cooperatively enforced, and so must provide him with the financial incentive to behave well. A full model of negotiation of the form of the contract itself is outside the scope of this paper, but remains an important area for future research (see also Fehr and Schmidt 2000). We hope that - by recognising the links between bargaining, implementation and contractual form - this paper may provide some first steps towards understanding the prevalence of simple, fair contracts in the real world.

---

not be exploited in unequal bargains because they are not enforceable.

Appendix

I Proof of Proposition 1:

The pattern for the proof is drawn from Osborne and Rubinstein (1990). Assume that A is non-empty, i.e., $\delta_1+\delta_2\geq 1$.

Claim: If A is non-empty, the following is a subgame perfect equilibrium:

P1 always offers x* such that $1\text{-}x=\max(1\text{-}\delta_2, \delta_2(1\text{-}y))$, P2 always offers y* such that $y=\max(1\text{-}\delta_1, \delta_1 x)$. P2 accepts any offer x* such that $\delta_1\geq 1\text{-}x\geq\max(1\text{-}\delta_2, \delta_2(1\text{-}y))$, and P1 accepts any offer y* such that $\delta_2\geq y\geq\max(1\text{-}\delta_1, \delta_1 x)$.

Proof:

1) Is P1's offering strategy a best response to P2's acceptance strategy?

- for equilibrium reasons, if P2 can get 1-y next round, P1 must offer at least $1\text{-}x = \delta_2(1\text{-}y)$
- for enforcement reasons, player 1 must give P2 at least $1\text{-}x=1\text{-}\delta_2$
- for enforcement reasons, player 1 can give P2 no more than $1\text{-}x=\delta_1$

P1 must choose $\delta_1\geq 1\text{-}x\geq\max(1\text{-}\delta_2, \delta_2(1\text{-}y))$, or else P2 will reject. Obviously since P2 will accept any offer in this region, if P1 wants to get P2 to accept, he does best by choosing the best acceptable offer for himself in this region, so he minimises 1-x, so he sets $1\text{-}x=\max(1\text{-}\delta_2, \delta_2(1\text{-}y))$.

If instead P1 makes an offer outside this region and P2 rejects, P1 will get one of (a) the infinite disagreement outcome (b) $y=\max(1\text{-}\delta_1, \delta_1 x)$ if he accepts P2's offer next period, or worse if he accepts 2's offer at some later date. (c) Some agreement inside the region $\delta_1\geq 1\text{-}x\geq\max(1\text{-}\delta_2, \delta_2(1\text{-}y))$ if he decides to make an acceptable offer at a later date. Since pie is valued by the players, (a) is worse for P1. The payoff to (b) is y next period (or later) where $y=\max(1\text{-}\delta_1, \delta_1 x)$. Now if $y=\max(1\text{-}\delta_1, \delta_1 x)=\delta_1 x$, this is clearly worse than x now. Now suppose $y=\max(1\text{-}\delta_1, \delta_1 x)=1\text{-}\delta_1$. We know that $x\geq 1\text{-}\delta_1$ since $\delta_1\geq 1\text{-}x\geq\max(1\text{-}\delta_2, \delta_2(1\text{-}y))$. So making the offer x is better than waiting for $1\text{-}\delta_1$ next period. If P1 follows strategy (c) he can only get $x\leq x^*$ at some later date: waiting to make an acceptable offer is not profitable, it's better to make one now, since the set of acceptable offers doesn't change with time when the game is stationary.

2) Is P2's Acceptance Strategy a best response to P1's offering strategy?

Suppose P1 makes an offer x*: $\delta_1\geq 1\text{-}x\geq\max(1\text{-}\delta_2, \delta_2(1\text{-}y))$. Can P2 do better by rejecting such an offer? His payoff would be:

(a) $\delta_2(1\text{-}y)=\delta_2(1\text{-}\max(1\text{-}\delta_1,\ \delta_1 x))$, if P2 rejects the offer and makes the best acceptable counter-offer next period. Since $1\text{-}x \geq \delta_2(1\text{-}y)$, P2 does not gain from this strategy.

(b) 1-x in two periods (or later) from now if he delays accepting P1's offer: this is clearly worse: if you're going to accept an offer, do it now.

(c) infinite disagreement.

Thus P2's acceptance strategy is a best response to P1's offering strategy. The arguments for P2's offering strategy and P1's acceptance strategy are exactly analogous.

Claim:

If there is a unique pair of solutions to the simultaneous equations:

$1\text{-}x=\max(1\text{-}\delta_2,\ \delta_2(1\text{-}y))$

$y=\max(1\text{-}\delta_1,\ \delta_1 x)$

Then the equilibrium described above is the unique SPE of the game.

Because the game is stationary, all subgames starting with Player i's offer are isomorphic; let G_i be such a subgame. Define M_i as the supremum of the expected discounted payoffs that can be attained by player i in any SPE of G_i. Define m_i as the corresponding infimum. If we can show that the following condition holds:

Condition 1: $M_1=m_1=x_1^*$ and $M_2=m_2=y_2^*$

(where x_i^* denotes the value to player i of agreement x*, etc.,) so that the present value to Player 1 of any SPE outcome of G_1 is x_1^*, and similarly for Player 2 the present value of any SPE of G_2 is y_2^*, this is sufficient to prove that the have a unique SPE. The reason for this is as follows. If condition 1 holds, in any SPE the first offer is accepted. If Player 1's first offer x* is rejected, the players must follow an SPE of G_2. The present value of this to Player 1 is only $y_1^*<x_1^*$. But $x_1=m_1$, the least that Player 1 can get in any SPE of G_1. Hence under condition 1, there is no SPE in which Player 1's offer of x_1^* is rejected. (Otherwise we would have a contradiction).

Similarly, when Player 2 offers y*, he must be accepted or else we would contradict condition 1. Thus it must be the case that in any SPE of G_1, Player 1 always offers x* and is accepted by Player 2, and Player 2 always offers y* and is accepted by Player 1. Therefore this is the unique SPE of the game starting with Player 1's offer.

It remains to establish Condition 1. This involves 2 steps.

Firstly we establish that $m_2 \geq 1-\delta_1 M_1$. Suppose Player 2 proposes the agreement z with $z_1 \geq Max(1-\delta_1, \delta_1 M_1)$. If player 1 accepts, he receives z_1 now. If he rejects, he receives at most $\delta_1 M_1$ i.e. the best possible payoff following the game in which he offers, which is by assumption worse than z_1 now. So, since $z_1 \geq 1-\delta_1$ ensures that the proposal is enforceable, Player 1 certainly accepts the proposal z. Player 2 should therefore offer the lowest possible z. Player 2's payoff when he offers therefore cannot be any lower than $z_2 = 1-z_1$ now so that $m_2 \geq 1-Max(1-\delta_1, \delta_1 M_1)$.

Secondly, it must be the case that $M_1 \leq 1-Max(1-\delta_2, \delta_2 m_2)$. If Player 2 rejects Player 1's offer, he can get at least m_2 next period. So Player 2 will reject any offer x* such that $1-x<\delta_2 m_2$. Player 1 will also reject proposals if they are not enforceable, i.e. if they give him $1-x<1-\delta_2$. The most Player 1 can get if agreement is reached this period is therefore $M_1 \leq 1-Max(1-\delta_2, \delta_2 m_2)$. The most he can get if agreement is delayed is $\delta_1(1-m_2) \leq 1-m_2 \leq 1-\delta_2 m_2$, i.e. no more than he can get if agreement is reached now. Hence the most Player 1 can get in any SPE of G_1 is $M_1 \leq 1- Max(1-\delta_2, \delta_2 m_2)$.

The same arguments, with the roles of the players reversed, would establish $m_1 \geq 1-Max(1-\delta_2, \delta_2 M_2)$ and $M_2 \leq 1-Max(1-\delta_1, \delta_1 m_1)$. Thus we have the following inequalities:

$m_2 \geq 1-Max(1-\delta_1, \delta_1 M_1)$ (1)

$M_1 \leq 1-Max(1-\delta_2, \delta_2 m_2)$ (2)

$m_1 \geq 1-Max(1-\delta_2, \delta_2 M_2)$ (3)

$M_2 \leq 1-Max(1-\delta_1, \delta_1 m_1)$ (4)

Since we know that the intersection of the two curves:

$1-x=Max(1-\delta_2, \delta_2(1-y))$ or $x=1- Max(1-\delta_2, \delta_2(1-y))$

$y=Max(1-\delta_1, \delta_1 x)$ or $1-y=1- Max(1-\delta_1, \delta_1 x)$

constitutes an equilibrium, we will consider the relation of these inequalities to these two curves.

Eq(1) implies (i) that the point $(m_2, 1-M_1)$ lies to the right of the line $x=1- Max(1-\delta_2, \delta_2(1-y))$, and (ii) that the point $(M_1, 1-m_2)$ lies below the line $y=Max(1-\delta_1, \delta_1 x)$.

Eq(2) implies (i) that the point (M_1, 1-m_2) lies to the left of the line x=1- Max(1-δ_2, δ_2(1-y)) in (x,y) space. It also implies (ii) that the point (m_2, 1-M_1) lies above the line y=Max(1-δ_1, δ_1x).

Eq(3) implies (i) that the point (m_1, 1-M_2) lies to the right of the line x=1- Max(1-δ_2, δ_2(1-y)) and (ii) that the point (M_2,1-m_1) lies below the line y=Max(1-δ_1, δ_1x).

Eq(4) implies (i) that the point (M_2, 1-m_1) lies to the left of the line x=1- Max(1-δ_2, δ_2(1-y)) and (ii) that the point (m_1, 1-M_2) lies above the line y=Max(1-δ_1, δ_1x).

The regions indicated for the point (M_1, 1-m_2) by (1)(ii) and (2)(i) intersect in the Region Q only below and to the left of the point (x_1*,y_1*). But we have already shown that (x*,y*) constitutes an SPE and thus it must be the case that $M_1 \geq x_1$*. Hence (M_1, 1-m_2)=(x_1*,y_1*).

The regions indicated for the point (m_1, 1-M_2) by (3)(i) and (4)(ii) intersect in the region S above and to the right of the point (x_1*,y_1*). But we have already shown that (x*,y*) constitutes an SPE and thus it must be the case that $M_2 \geq 1-y_1$*, i.e. 1-$M_2 \leq y_1$*. Hence (m_1, 1-M_2)=(x_1*,y_1*). Thus $M_1=m_1$=x* and $M_2=m_2=1-y_1*=y_2*$, establishing condition 1.

It remains to show that the two curves 1-x=Max(1-δ_2, δ_2(1-y)) and y=Max(1-δ_1, δ_1x) do indeed have a unique intersection. This will be the case when $\delta_1+\delta_2>1$, (otherwise there will be no intersection); probably the simplest way to see this is to plot the curves.

II Proof of Lemma 5

Notice first that if only one player is marginal, it is always the least patient player. This is true almost by definition of marginality. A marginal player is the player whose EC constraint binds his opponent's offer. Thus in considering whether a player is marginal, one is considering him as the *recipient* of a proposal, regardless of whether in fact he will make the first or the second offer. Thus considerations of first-mover advantage do not enter into the consideration of who is marginal, though they will of course affect the final outcome, so it is simply a question of who is in the weaker position. One can show this formally by comparing the equilibrium equations:

Denote by v the offer made by *i* to *j*, and w the offer made by *j* to *i*:

If player *i* is marginal, then w=Max $(1-\delta_i, \delta_i(1-v))= 1-\delta_i$

If player *j* is not marginal then v=Max $(1-\delta_j, \delta_j(1-w))= \delta(1-w)$

Therefore $v=\delta_j\delta_i$. By rearranging, one can deduce from this that $\delta_i<\delta_j$.

Of course, if both players are marginal, then the least patient player must be marginal. So now consider the least patient player, supposing that he is marginal. We will call him player *i* and his opponent *j*, as above. We consider whether there is any gain to offering player *i* an agreement in which his share is not constant over time. Since he is least patient, efficiency favours moving some of his payoff backwards in time, so that he receives more now and less later. Consider reducing his payoff next period by ε/δ_i, and increasing his payoff this period by ε. If this were possible, his opponent *j* would lose ε this period and gain $\varepsilon\delta_j/\delta_i$, next period, which is profitable since $\delta_j>\delta_i$. But this is not a feasible change, since *i*'s no-defection constraint will be violated next period (since it holds with equality in the absence of a change). The same is true for any attempt to make *i*'s payoff fall over time. Now consider making *i*'s payoff rise over time. Adding ε/δ_i to *i*'s payoff next period and subtracting ε from his payoff this period would leave him just indifferent to cheating, but would reduce *j*'s total surplus by $\varepsilon-\varepsilon\delta_j/\delta_i<0$, and so is an unprofitable change for *j*, who is already giving up more surplus than he would like to *i*. Consequently, when one player is marginal, he will receive a constant share over time (and hence his opponent must do so also, since the surplus is constant over time).

Bibliography

Abreu, Dilip, and Pearce, David (1989)“A Perspective on Renegotiation in Repeated Games” in “Game Equilibrium Models II” (Reinhard Selten, ed.), Springer-Verlag, Berlin 1991

Abreu, Dilip, Pearce, David and Stacchetti, Ennio “Renegotiation and Symmetry in Repeated Games” *Journal of Economic Theory*, 1993, *60*, 217-240

Akerlof, George A. “Labour Contracts as Partial Gift Exchange” *Quarterly Journal of Economics*, 1982, *97*, 543-569

Baker, George, Gibbons, Robert and Murhy, Kevin J. “Bringing the Market Inside the Firm?” *American Economic Review* 2001, 91, 212-218

Binmore, Ken, Shaked, Avner and Sutton, John “Testing Non-Cooperative Bargaining Theory: A Preliminary Study” *American Economic Review,* 1985, *75*, 1178-1180

Binmore, Ken, Morgan, Peter, Shaked, Avner and Sutton, John “Do People Exploit their Bargaining Power? An Experimental Study” *Games and Economic Behavior*, 1991, *3*, 295-322

Bolton, Gary E. “A Comparative Model of Bargaining: Theory and Evidence” *American Economic Review*, 1991, *81*, 1096-1136

Che, Yeon-Koo and Donald B. Hausch (1999) “Cooperative Investments and the Value of Contracting” *American Economic Review*, 1999, 89, 125-147

Cole, Harold L., Mailath, George J. and Postlewaite, Andrew “Social Norms, Savings Behavior and Growth” *Journal of Political Economy,* 1992, *100*, 1092-1125

Eckel, Catherine C. and Grossman, Philip J. “Altruism in Anonymous Dictator Games” *Games and Economic Behavior*, 1996, *16*, 181-191

Farrell, Joseph and Maskin, Eric “Renegotiation in Repeated Games” *Games and Economic Behavior* 1989, *1*, 295-326

Fehr, Ernst and Fischbacher, Urs “Why Social Preferences Matter: The Impact of Non-Selfish Motives on Competition, Cooperation and Incentives” *Economic Journal* 2002, *112*, c1-c33

Fehr, Ernst and Gächter, Simon “Fairness and Retaliation: The Economics of Reciprocity” *Journal of Economic Perspectives*, 2000, *14*, 159-181.

Fehr, Ernst and Schmidt, Klaus “A Theory of Fairness, Competition and Cooperation”, *Quarterly Journal of Economics* 1999, *114*, 817-868

Fehr, Ernst and Schmidt, Klaus “Fairness, Incentives and Contractual Choices” *European Economic Review*, 2000, *44*, 393-421

Fehr, Ernst and Schmidt, Klaus "Theories of Fairness and Reciprocity – Evidence and Economic Applications" M. Dewatripont, L. Hansen and St. Turnovsky (Eds.), *Advances in Economics and Econometrics - 8th World Congress Econometric Society Monographs*, Cambridge, Cambridge University Press 2002

Fernandez, Raquel and Glazer, Jacob "Striking for a Bargain between Two Completely Informed Agents" *American Economic Review* 1991, *81*, 240-252

Gale, John, Binmore, Kenneth G and Samuelson, Larry "Learning to be Imperfect: The Ultimatum Game" *Games and Economic Behaviour,* 1995, *8*, 56-90

Güth, Werner, Schmittberger, Rolf and Schwarze, Bernd "An Experimental Analysis of Ultimatum Bargaining" *Journal of Economic Behaviour and Organisation*, 1982, *3*, 367-388

Güth, Werner, and Tietz, Reinhard "Ultimatum Bargaining for a Shrinking Cake - An Experimental Analysis" In "Bounded Rational Behavior in Experimental Games and Markets" (R. Tietz, W. Albers, R. Selten eds.) Springer, Berlin 1988

Güth, Werner, and Tietz, Reinhard "Ultimatum Bargaining Behaviour: A Survey and Comparison of Experimental Results" *Journal of Economic Psychology*, 1990, *11*, 417-449

Hart, Oliver "Firms, Contracts and Financial Structure" 1995, Oxford: Clarendon Press

Hart, Oliver and Moore, John "Agreeing Now to Agree Later: Contracts that Rule Out but do not Rule In" 2004, mimeo, London School of Economics.

Holmström, Bengt and Kreps, David "Examples for and Questions about: An Economic (?) Theory of Promises", mimeo 1995

Johnson, Simon, McMillan, John and Woodruff, Christopher "Contract Enforcement in Transition" mimeo, MIT Sloan School

Katz, Michael L. "Game-Playing Agents: Unobservable Contracts as Precommitments" *Rand Journal of Economics*, Autumn 1991, *22*, 307-328

MaCaulay, Stewart "Non-Contractual Relations in Business", *American Sociological Review*, 1963, *28*, 55-70

Muthoo, Abhinay "Bargaining Without Commitment" *Games and Economic Behavior*, 1990, *2*, 291-7

Neelin, Janet, Sonnenschein, Hugo, and Spiegel, Matthew "A Further Test of Non-Cooperative Bargaining Theory" *American Economic Review*, 1988, *78*, 824-836

Ochs, Jack and Roth, Alvin E. "An Experimental Study of Sequential Bargaining" *American Economic Review*, June 1989, *79*, 355-84

Osborne, Martin J. and Rubinstein, Ariel "Bargaining and Markets" 1990, San Diego, CA: Academic Press

Pearce, David "Renegotiation-Proof Equilibria: Collective Rationality and Intertemporal Cooperation" Cowles Foundation Discussion Paper, Yale University, 1987

Roth, Alvin E. "Bargaining Experiments" in *Handbook of Experimental Economics* (J. Kagel and A.E. Roth eds.) 1995, pp253-348 Princeton, NJ: Princeton University Press

Roth, Alvin E. and Erev, Ido "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term" *Games and Economic Behavior*, 1995, *8*, 164-212

Roth, Alvin E. and Murnighan, J. Keith "The Role of Information in Bargaining: An Experimental Study" *Econometrica*, 1982, *50*, 1123-42

Rubinstein, Ariel "Perfect Equilibrium in a Bargaining Model" *Econometrica*, February 1982, *50*, 97-109

Segal, Ilya "Complexity and Renegotiation: A Foundation for Incomplete Contracts" *Review of Economic Studies,* 1999, *66*, 57-82

Shaked, Avner and Sutton, John "Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Model" *Econometrica*, 1984, *52*, 1351-64

Spiegel, Matthew, Currie, Janet, Sonnenschein, Hugo, and Sen, Arunava "Understanding When Agents are Fairmen or Gamesmen" *Games and Economic Behaviour*, 1994, *7*, 104-115

Telser, Lester "A Theory of Self-Enforcing Agreements" *Journal of Business*, 1980, *53*, 27-44

TABLES

Table 1: Payoffs in the Stage-game of the Enforcement Stage

	Payoffs (P1,P2)	Player 2's Action *Cooperate*	*Defect*
Player 1's Action	*Cooperate*	x,1-x	$-\varepsilon$,1
	Defect	1,$-\varepsilon$	0,0

Table 2: Regimes in the Enforcement Game:

U = No bargain can be enforced
EC_{12} = Player 1's EC constrains Player 2's offer
EC_{21} = Player 2's EC constrains Player 1's offer
C = Both players constrained by each others' EC's
R = Neither player constrained by the other's EC: Rubinstein Outcome

			Player 2's Offer= Max(1-δ_1,δ_1x)	
		δ_2 <max(1-δ_1, δ_1x)	1-δ_1	δ_1x
	δ_1<max(1-δ_2,δ_2(1-y))	U	U	U
Player 1's Offer=1-max(1-δ_2, δ_2(1-y))	1-δ_2	U	C	EC_{21}
	δ_2(1-y)	U	EC_{12}	R

[Tables 3-4 see overleaf]

Table 5: Payoffs in the Stage game of the Enforcement Stage with degree of incompleteness θ

	Payoffs (P1,P2)	Player 2's *Cooperate*	Action *Defect*
Player 1's Action	*Cooperate*	x,1-x	$-\varepsilon$,θ
	Defect	θ,$-\varepsilon$	0,0

Table 3a: Player 1's Rubinstein Stage-Game Pay-Off (to four decimal places)

		delta two										
		0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1-
delta one	0	1.0000	0.9000	0.8000	0.7000	0.6000	0.5000	0.4000	0.3000	0.2000	0.1000	0.0000
	0.1	1.0000	0.9091	0.8163	0.7216	0.6250	0.5263	0.4255	0.3226	0.2174	0.1099	0.0000
	0.2	1.0000	0.9184	0.8333	0.7447	0.6522	0.5556	0.4545	0.3488	0.2381	0.1220	0.0000
	0.3	1.0000	0.9278	0.8511	0.7692	0.6818	0.5882	0.4878	0.3797	0.2632	0.1370	0.0000
	0.4	1.0000	0.9375	0.8696	0.7955	0.7143	0.6250	0.5263	0.4167	0.2941	0.1563	0.0000
	0.5	1.0000	0.9474	0.8889	0.8235	0.7500	0.6667	0.5714	0.4615	0.3333	0.1818	0.0000
	0.6	1.0000	0.9574	0.9091	0.8537	0.7895	0.7143	0.6250	0.5172	0.3846	0.2174	0.0000
	0.7	1.0000	0.9677	0.9302	0.8861	0.8333	0.7692	0.6897	0.5882	0.4545	0.2703	0.0000
	0.8	1.0000	0.9783	0.9524	0.9211	0.8824	0.8333	0.7692	0.6818	0.5556	0.3571	0.0000
	0.9	1.0000	0.9890	0.9756	0.9589	0.9375	0.9091	0.8696	0.8108	0.7143	0.5263	0.0000
	1-	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.5000

Table 3b: Player 1's Enforceability-Constrained Stage-Game Pay-Off (to four decimal places)

		delta two										
		0+	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1-
delta one	0+	0	0	0	0	0	0	0	0	0	0	1-
	0.1	0	0	0	0	0	0	0	0	0	0.9	0.9
	0.2	0	0	0	0	0	0	0	0	0.8	0.82	0.8
	0.3	0	0	0	0	0	0	0	0.7	0.76	0.73	0.7
	0.4	0	0	0	0	0	0	0.6	0.7	0.68	0.64	0.6
	0.5	0	0	0	0	0	0.5	0.6	0.65	0.6	0.55	0.5
	0.6	0	0	0	0	0.4	0.5	0.6	0.58	0.48	0.46	0.4
	0.7	0	0	0	0.3	0.4	0.5	0.6	0.5882	0.4545	0.37	0.3
	0.8	0	0	0.2	0.3	0.4	0.5	0.6	0.6818	0.5556	0.3571	0.2
	0.9	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.7143	0.5263	0.1
	1-	1-	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.5

Key

0 Outcome not enforceable

Outcome more equal than Rubinstein Outcome

Outcome the same as Rubinstein Outcome

Outcome less equal than Rubinstein Outcome

Table 4a: Lifetime Pay-Offs in the Unconstrainted Rubinstein Game (To two decimal places)

Player 1's Utility, Player 2's Utility — delta two

delta one	0+	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1-
0+	1, 0+	0.9, 0.11	0.8, 0.25	0.7, 0.43	0.6, 0.67	0.5, 1	0.4, 1.5	0.3, 2.33	0.2, 4	0.1, 9	0+, inf
0.1	1.11, 0+	1.01, 0.10	0.91, 0.23	0.80, 0.40	0.69, 0.63	0.58, 0.95	0.47, 1.44	0.36, 2.26	0.24, 3.91	0.12, 8.90	0+, inf
0.2	1.25, 0+	1.15, 0.09	1.04, 0.21	0.93, 0.36	0.82, 0.58	0.69, 0.89	0.57, 1.36	0.44, 2.17	0.30, 3.81	0.15, 8.78	0+, inf
0.3	1.43, 0+	1.33, 0.08	1.22, 0.19	1.10, 0.33	0.97, 0.53	0.84, 0.82	0.70, 1.28	0.54, 2.07	0.38, 3.68	0.20, 8.63	0+, inf
0.4	1.67, 0+	1.56, 0.07	1.45, 0.16	1.33, 0.29	1.19, 0.48	1.04, 0.75	0.88, 1.18	0.69, 1.94	0.49, 3.53	0.26, 8.44	0+, inf
0.5	2, 0+	1.89, 0.06	1.78, 0.14	1.65, 0.25	1.5, 0.42	1.33, 0.67	1.14, 1.07	0.92, 1.79	0.67, 3.33	0.36, 8.18	0+, inf
0.6	2.5, 0+	2.39, 0.05	2.27, 0.11	2.13, 0.21	1.97, 0.35	1.79, 0.57	1.56, 0.94	1.29, 1.61	0.96, 3.08	0.54, 7.83	0+, inf
0.7	3.33, 0+	3.23, 0.04	3.10, 0.09	2.95, 0.16	2.78, 0.28	2.56, 0.46	2.30, 0.78	1.96, 1.37	1.52, 2.73	0.90, 7.30	0+, inf
0.8	5, 0+	4.89, 0.02	4.76, 0.06	4.60, 0.11	4.41, 0.20	4.17, 0.33	3.85, 0.58	3.41, 1.06	2.78, 2.22	1.79, 6.43	0+, inf
0.9	10, 0+	9.89, 0.01	9.76, 0.03	9.59, 0.06	9.38, 0.10	9.09, 0.18	8.70, 0.33	8.11, 0.63	7.14, 1.43	5.26, 4.74	0+, inf
1-	inf, 0+	inf, 0+	inf, 0+	inf, 0+	inf, 0+	inf, 0+	inf, 0+	inf, 0+	inf, 0+	inf, 0+	inf, inf

Table 4b: Lifetime Pay-Offs in the Enforcement-Constrained Game (To two decimal places)

Player1's Utility, Player 2's Utility — delta two

delta one	0+	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1-
0+	0	0	0	0	0	0	0	0	0	0	1,inf
0.1	0	0	0	0	0	0	0	0	0	1,1	1,inf
0.2	0	0	0	0	0	0	0	0	1, 1	1.03, 1.8	1,inf
0.3	0	0	0	0	0	0	0	1,1	1.09, 1.2	1.04, 2.7	1,inf
0.4	0	0	0	0	0	0	1, 1	1.17, 1	1.13, 1.6	1.07, 3.6	1,inf
0.5	0	0	0	0	0	1, 1	1.2, 1	1.3, 1.17	1.2, 2	1.1, 4.5	1,inf
0.6	0	0	0	0	1, 1	1.25, 1	1.5, 1	1.45, 1.4	1.2, 2.6	1.15, 5.4	1,inf
0.7	0	0	0	1, 1	1.33, 1	1.67, 1	2, 1	1.96, 1.37	1.52, 2.73	1.23, 6.3	1,inf
0.8	0	0	1, 1	1.5, 1	2, 1	2.5, 1	3, 1	3.41,1.06	2.78, 2.22	1.79, 6.43	1,inf
0.9	0	1,1	2, 1	3, 1	4, 1	5, 1	6, 1	7, 1	7.14, 1.43	5.26, 4.74	1,inf
1-	inf, 1	inf, 1	inf, 1	inf, 1	inf, 1	inf, 1	inf, 1	inf, 1	inf, 1	inf, 1	inf, 1

Key

- 0 Outcome not enforceable
- (light stipple) Outcome more equal than Rubinstein Outcome
- (dark grey) Outcome the same as Rubinstein Outcome
- (hatched) Outcome less equal than Rubinstein Outcome

Figure 1: The Game Tree

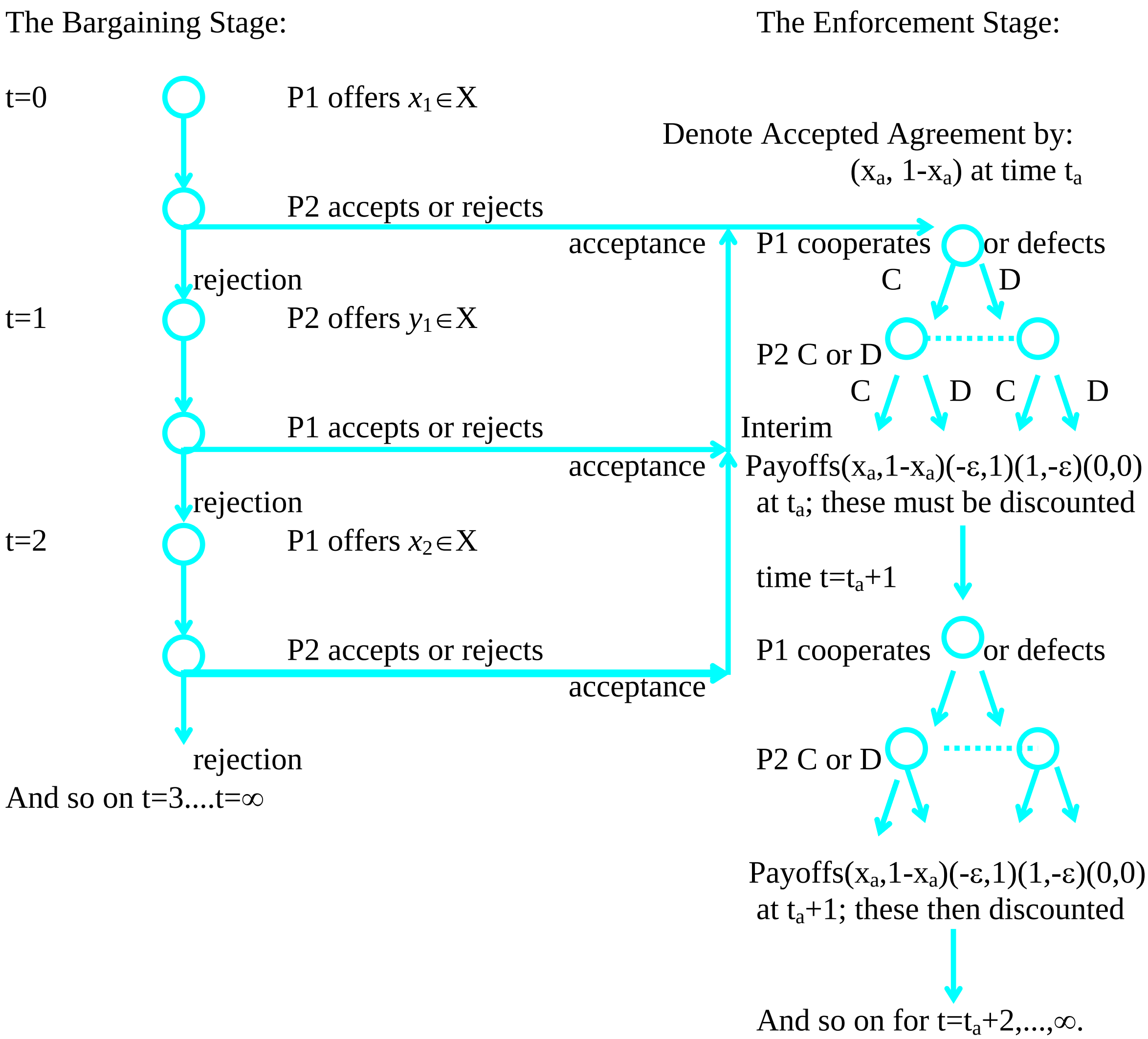

Figure 2(a): The bargaining game when enforcement constraints do not bind.

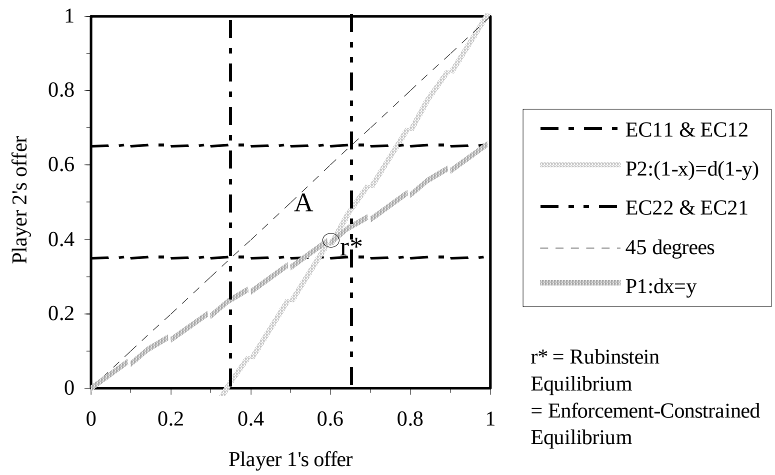

Figure 2(b): The bargaining game when enforcement constraints bind.

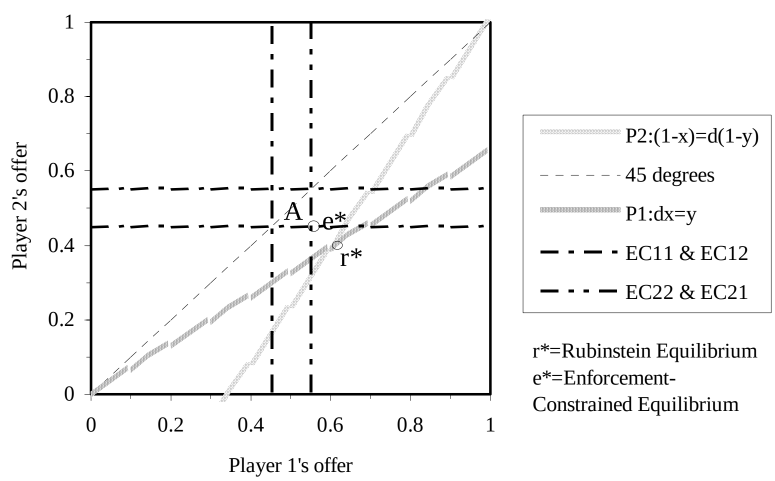

Figure 3: The Rubinstein Zone

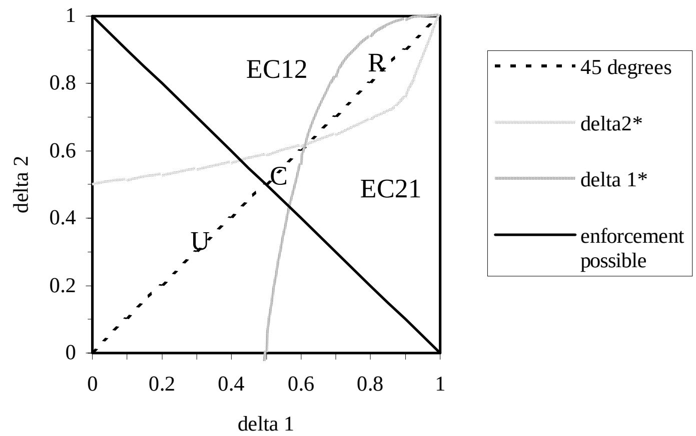

Figure 4(a): The Rubinstein Zone with degree of Contractual Incompleteness theta = 0.75

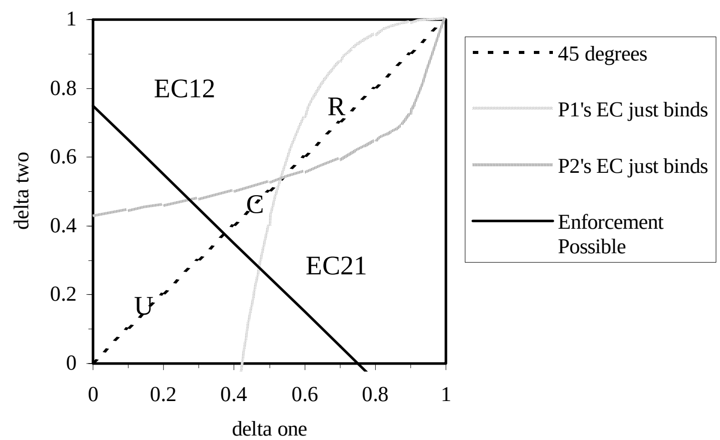

Figure 4(b): The Rubinstein Zone with degree of Contractual Incompleteness theta = 0.5

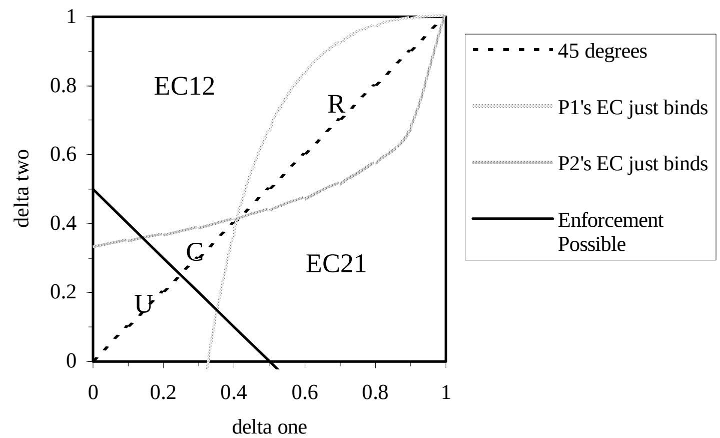

Figure 4(c): The Rubinstein Zone with degree of Contractual Incompleteness = 0.25