Is There a Hold–up Problem?

Tore Ellingsen* Magnus Johannesson**

First version: September 30, 1999
This version: February 10, 2000

Stockholm School of Economics
Working Paper Series in Economics and Finance, No 357

Abstract

Most literature on the hold–up problem starts from the assumption that ex post bargaining outcomes are insensitive to prior investment costs. We argue that this approach is unsatisfactory. If the bargaining procedure is relatively symmetric, it typically admits multiple perfect equilibria, some of which give the investor a high enough payoff to sustain efficient investment. Even if the bargaining procedure is asymmetric and rigged against the investor, there may be investment if agents are driven by moral concerns or if communication creates commitment. Laboratory experiments indicate that communication is necessary and sufficient for agents to coordinate on efficient outcomes when the bargaining game is symmetric. When the bargaining game is rigged against the investor, the hold–up problem is mitigated, but not eliminated, by moral behavior. Communication is quite credible, and we find that promises are more believable than threats.

JEL classification: L14, C78.
Keywords: Specific investments, bargaining, fairness, communication.

---

*Address: Department of Economics, Stockholm School of Economics, Box 6501, S—113 83 Stockholm, Sweden. Email: gte@hhs.se.

**Address: Department of Economics, Stockholm School of Economics, Box 6501, S—113 83 Stockholm, Sweden. Email: hemj@hhs.se.

We are grateful to Magnus Allgulin, Mats Ekelund, Freddie Henriksson, Douglas Lundin, Arvid Nilsson, Joakim Ramsberg and Niklas Zethraeus for research assistance, and to the Swedish Council for Research in the Humanities and Social Sciences (Ellingsen) for financial support. Thanks to Oliver Hart, Bengt Holmström, David Kreps, Paul Milgrom, Torsten Persson and Jörgen Weibull for helpful discussions. The paper has also benefited from presentations at Stockholm School of Economics and Stockholm University.

1 Introduction

Many investments are only valuable to a particular trading partner. Well known examples of such relationship–specific investments include development of tailormade designs, plant location, and acquisition of firm–specific skills. Unless contracts are perfect, and often they are not, the specificity of these investments makes the investor vulnerable to ex post exploitation. This is the hold–up problem.

The hold–up problem is central to contract theory and to institutional economics. Workers are potentially held up by employers, lenders by borrowers, sellers by buyers, investors by governments—and the other way around. As a result, there is too little investment. This theme has inspired a vast recent literature, within which contracts and governance structures are seen as mitigating the hold–up problem by empowering the investor; see e.g. Williamson (1975,1985), Klein, Crawford and Alchian (1978), Grossman and Hart (1986), Hart and Moore (1990), North and Weingast (1989), and Shleifer and Vishny (1997).[1]

A fundamental premise for the whole literature on the hold–up problem is that investors cannot guarantee themselves a sufficient share of the return through unregulated ex post bargaining. Indeed, in all the formal models we are aware of it is simply assumed that the cost of the investment is irrelevant for the ex post bargaining outcome. The purpose of this paper is to argue theoretically, and to show empirically, that sunk costs might well affect bargaining outcomes. As a result, we need to revise and refine our analysis of the hold–up problem. The following is a brief overview of our argument.

The irrelevance of sunk costs is a cherished economic principle, but one which is only generally valid for non–strategic decisions. Technically speaking, the principle is violated in games with multiple equilibria, because in such games any past event can be used to select among equilibria in the remaining subgame. As we shall argue below, bargaining should indeed normally be thought of as a game with many equilibria. Hence investment levels might well matter for the bargaining outcome. Actually, there are sound theoretical reasons why observed bargaining outcomes should always be expected to reflect sunk costs.

For concreteness, consider a seller who can make a relationship specific investment of 60 which generates a gain from trade of 100 if the appropriate buyer agrees to trade. Suppose that it is impossible to contract in advance,

[1]The hold–up problem has also attracted the attention of macroeconomists. For example, Caballero and Hammour (1998) develop a general equilibrium model wherein the hold–up problem plays a key role.

on the investment as well as on the sharing of the gain.[2] If the seller does not invest, the game is over, and both parties get their reference utility of zero. If the seller invests, the two parties bargain over the terms of trade. If they cannot agree, the gain from trade is foregone. The seller then gets a total payoff of -60 and the buyer a total payoff of 0. If they agree, the payoffs are $-60+x_S$ and x_B, where $x_S + x_B \leq 100$. The standard assumption in the literature is that the gain is split according to the Nash bargaining solution, or just 50:50. If that is what the seller expects, the total expected payoff from investing is -10, and hence there is no investment in equilibrium. This modelling strategy mixes up strategies (the investment decision) with outcomes (the bargaining split), and ends up more or less assuming the result. The reason why sunk costs don't matter is that they are not allowed to.

Once we model the bargaining stage as a game, it becomes clear both that sunk costs might matter, and that there might be equilibria with investment. For example, consider bargaining according to the Nash demand game. Both parties simultaneously make a claim; if claims are compatible, each gets the claimed share; if claims ar incompatible, there is no trade. Observe that it is quite unreasonable for the buyer ever to claim as much as 50 in this game. To see this, note that the seller should never invest and then claim less than the investment cost–this is a strictly dominated strategy. Eliminating all strictly dominated strategies for the seller, it then becomes a (weakly) dominated strategy for the buyer to claim more than 40 (if he has observed the seller investing). Hence, the assumption that bargaining would be insensitive to sunk costs is contradicted by the assumption that agents are able to perform two rounds of elimination of dominated strategies! On the other hand there are a host of efficient subgame perfect equilibria of this game with the property that the seller invests and makes a claim $c_S \geq 60$ in the negotiation and that the buyer claims $c_B = 100 - c_S$.

Multiplicity of equilibria is a feature of many bargaining games. Even the famous alternating offer game, associated with Ståhl (1972) and Rubinstein (1982), has multiple subgame perfect equilibria under plausible assumptions about the set of offers (finite), the rate of discounting (low), and the speed of communication (high); see van Damme et al. (1990).[3] Thus, the possibility that sunk costs can serve as a coordination device is not confined to a narrow

[2]The literature provides various reasons why contracting might be ruled out; (i) use of legal enforcement might be too costly, (ii) the identity of the appropriate buyer might not be known in advance (Pitchford and Snyder, 1999), (iii) the investment is too complicated to specify (Hart and Moore, 1999, Segal, 1999). We do not here wish to take a stand on whether these reasons are convincing or not.

[3]When agents bargain over a stream of services rather than a single trade, the alternating offers procedure gives rise to multiple equilibria even when the set of offers is continuous;

class of bargaining games.

A class of bargaining games which do not usually admit multiple subgame perfect equilibria are those which vests the power to propose with one of the parties exclusively. For example, if the buyer has the ability to commit to a take–it–or–leave–it offer, and offers can not be made before investment decisions, conventional wisdom says that the only equilibrium entails a very meager offer that the seller accepts. Theory thus appears to say that if the seller has a weak bargaining position, there is necessarily a hold–up problem. However, even this conclusion needs qualification. If the buyer feels bad about exploiting the seller, or if the seller is willing to take a loss in order to punish greedy behavior by the buyer, the game may nonetheless have subgame perfect equilibria in which the seller invests. These are examples of social (or interdependent) preferences.[4]

At first sight, social attitudes such as inequity aversion and altruism would seem to be unambiguously beneficial for the solution of the hold–up problem. This is only the case if preferences are common knowledge or the bargaining game is rigged against the investor. Indeed, as we indicated above, there are quite plausible circumstances under which narrowly self–interested agents would be able to coordinate on efficient outcomes. However, if an agent knows that there is a possibility that the opponent has social preferences, the basis for coordination is weaker. The introduction of informational asymmetry about preferences might therefore, as we shall show, ruin the possibilities for attaining coordination—at least if parties cannot communicate.

We do not deny that people are boundedly rational, and that behavior is sometimes driven by erroneous beliefs. For example, it may happen that the buyer makes a generous offer in the belief that the seller would punish a meager offer. However, this case is observationally equivalent to the case in which the buyer is inequity averse, presenting an example of a general difficulty in distinguishing bounded rationality from social preferences.[5] In the theoretical part of this paper, we shall thus maintain the assumption that people are rational, in the sense that they maximize expected utility and apply backward induction, but we shall allow the possibility that some agents have social preferences in a well–defined sense.[6]

---

see Haller and Holden (1990) and Fernandez and Glazer (1991).

[4]For literature on social preferences in bargaining games, see the surveys by Roth (1995) and Camerer (1998). See also Fehr and Schmidt (1999).

[5]For a recent attempt to distinguish bounded rationality from social preferences within the context of simple bargaining games, see Binmore et al. (1999).

[6]A different approach, pursued in a companion paper by Ellingsen and Robles (2000), is to apply evolutionary game theory, which do not impose maximizing behavior at all.

The final issue that we discuss is the role of communication. We argue that, in theory, cheap talk is enough to restore coordination on efficient outcomes when the bargaining procedure admits multiple equilibria, but that it is ineffective otherwise. However, psychologists have persistently argued that there is more to communication than just cheap talk. In particiular, non–binding promises to cooperate in social dilemmas are often credible, as demonstrated already by Loomis (1959). Kerr and Kaufman–Gilliland (1994) and Sally (1995) survey experimental results along these lines.

To probe empirically the theoretical issues raised above, we have conducted a sequence of experiments. In all the experiments, one party had the opportunity to invest 60 Swedish crowns (about 8 US dollars) to create a benefit of 100, which would only be reaped upon succesful negotiation with another party. Interaction was anonymous. As bargaining procedures we used either the Nash demand game or the ultimatum game with the buyer as proposer. Both bargaining treatments were carried out with (i) no communication, (ii) investor communication, and (iii) trading partner communication, yielding a total of six experiments. Communication took the form of a single written message, which was "cheap talk" in the sense that it did not formally commit the sender to a particular course of action. The idea was that authority, in the form of a communication monopoly, might allow improved coordination at the bargaining stage, which in turn should affect investment. However, there was no restrictions on (or guidelines for) the messages. When conducted by the trading partner, communication preceded the investment decision. Investor communication was simultaneous with the investment decision.

The results overwhelmingly reject the notion that sunk costs are irrelevant. For example, in the Nash demand game without communication, no buyer ever claimed more than the difference between the benefit and the investment cost. Despite this, versions of the hold–up problem prevail to a considerable extent. When parties could not communicate, only about a quarter of the subjects chose to invest under each bargaining protocol. The violation appears to be caused by sellers' (justified) fear that buyers may claim 40 or more. We argue that this coordination failure can be explained by even a tiny heterogeneity in people's preferences.

When the investor could send a message, the rate of investment increased substantially. This is predicted by theory in the Nash demand treatment, since communication allows the agents to coordinate on a specific subgame perfect equilibrium, thus removing strategic uncertainty. In the ultimatum game treatment, cheap talk should not have any effect according to conventional game theory, since there is a unique subgame perfect outcome. The experimental finding that communication nonetheless increases investment is quite remarkable, although it is quite consistent with psychological research on the ability of communication to solve social dilemmas. In both bargaining treatments, the investor tended to do somewhat better when the trading partner conducted the communication. This is perhaps our most surprising result. In the Nash demand game, theory predicts that the party which sends a message can choose its favored equilibrium, and hence that it is better to be sending a message than receiving it. However, in the experiment, the splits are about the same irrespectively of who proposes them. We suggest that the seller is better off under buyer communication because a buyer message reveals whether or not the buyer is "reasonable," whereas a seller message is never certain to convince the buyer to behave reasonably.

In the ultimatum game, messages should not matter, and hence in theory it should not matter who sends them either. However, we find that promises made by the buyer are more credible than threats made by the seller, and this is what makes the seller better off in a position as receiver than sender. The superior credibility of promises over threats is a puzzle which we hope will get more attention in the future.

In the experiments, there is a clear tendency to coordinate on an 80:20 split, especially when communication is allowed. This outcome gives both parties the same net payoff. The well–known tendency of agents to divide 50:50 when the gain from trade is exogenously given thus carries over to the case of endogenous investment, with the important qualification that it is the *net gain* that is split equally, not the gross gain as has been assumed in the hold–up literature.

The paper is organized as follows. The next section formally describes and analyses the two games we consider. Section 3 describes the experiments and the results. Section 4 discusses related literature and identifies some avenues for further research. Section 5 concludes.

2 Model

Consider a seller S and a buyer B. At stage 1, the seller can make a fixed, non–contractible investment at cost F. This decision is given by the indicator variable I, which is 1 if the seller invests and zero otherwise. At stage 2, there is a potential gain from trade $g(I)$, where

$$g = \begin{cases} G & \text{if } I = 1; \\ 0 & \text{if } I = 0. \end{cases}$$

In other words, there is a potential gain from trade if and only if the seller invests. To make the problem interesting, we assume that $G > F$. The gain from trade is divided through some explicit bargaining procedure.

As a convention, when we talk about equilibria below, without explicit mention of some refinement, we have in mind the concept of Nash equilibrium.

We now turn to the first bargaining procedure.

2.1 The Nash demand game

In the Nash demand game, due to Nash (1953), each party i simultaneously makes a claim (or demand), $c_i \in [0, G]$. The payoffs are

$$\pi_i^N = \begin{cases} c_i & \text{if } c_B + c_S \leq G; \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

As is well known, the Nash demand game itself has many equilibria. Any combination (c_B, c_S) such that $c_B + c_S = G$ is an efficient equilibrium. The combination $(1, 1)$ is an example of an inefficient equilibrium; there are many other inefficient equilibria in mixed strategies. In other words, this bargaining game offers ample room for coordination failure.

The whole game, call it Γ^N, is now compactly described by the sets of pure strategies $P_S = \{0, 1\} \times [0, G]$ and $P_B = [0, G]$ and the payoff functions

$$u_S^N = \begin{cases} c_S - F & \text{if } I = 1 \text{ and } c_B + c_S \leq G; \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

and

$$u_B^N = \begin{cases} c_B & \text{if } I = 1 \text{ and } c_B + c_S \leq G; \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

What is the set of equilibria of the whole game? It is clear that we can rule out seller strategies of the form $(1, c_S)$ where $c_S < F$. The seller is better off by not investing than by investing and claiming less than the investment cost. The pure strategy equilibrium outcomes of the game can now be easily described.

Proposition 1 *(i) In the game Γ^N, the set of efficient equilibria is given by $I = 1$ in combination with any pair of claims (c_B, c_S) such that $c_B + c_S = G$ and $c_S \geq F$. (ii) The inefficient pure strategy equilibria are given by $I = 0$ in combination with any buyer claim $c_B \geq G - F$. (iii) All these equilibria are subgame perfect.*

A remarkable feature of this result is that there are equilibria, indeed subgame perfect equilibria, in which the seller invests. At the same time, there appears to be a distinct possibility of underinvestment. However, unlike in the classical hold–up literature there is no critical level of investment costs at which underinvestment becomes dramatically accentuated.

In order to make more specific predictions, we may consider, for example, iterated elimination of weakly dominated strategies. After elimination of the seller's weakly dominated strategies, the strategy set of the seller is $P_S^1 = \{0, 1\} \times (F, G]$. When the buyer observes investment, it is then weakly dominated to claim $G - F$ or more, since this will yield a payoff of zero. Eliminating these claims, it is optimal for the seller to invest, since the seller can secure a positive payoff by doing so.

If we think of the strategy space as being discrete, due for example to the fact that money has a smallest unit, the result can be strengthened even further. Then it is enough to apply one round of elimination of weakly dominated strategeies followed by iterated elimination of *strictly* dominated strategies. Since it is weakly dominated for the seller to invest and then claim F or less, it becomes strictly dominated for the buyer to claim more than $G - F + \epsilon$, and in the next round it is then strictly dominated for the buyer not to invest, since investment yields a profit of at least ϵ. This procedure of one round's elimination of weakly dominated strateges followed by iterated elimination of strictly dominated strategies is known as the *Dekel–Fudenberg procedure*, and has gained considerable credence among game theorists after Dekel and Fudenberg (1990) showed that the procedure satisfies a set of appealing requirements concerning player rationality.[7]

Proposition 2 *(i) In the game Γ^N, three rounds of elimination of weakly dominated strategies implies investment. (ii) If there is a smallest unit of account, the Dekel–Fudenberg procedure implies investment.*

With self–interested and sufficiently rational players, there should hence not be a hold–up problem when there is bargaining according to the Nash demand game.

The above analysis makes no prediction with respect to the actual bargaining split. However, Ellingsen and Robles (2000) have shown that there is a unique stochastically stable outcome in this game, which gives the seller slightly above F and the buyer slightly below $G - F$. Hence, in this particular situation, evolutionary game theory predicts that investment is efficient, but that the buyer reaps essentially all the net surplus.

2.2 The ultimatum game

The ultimatum bargaining game gives one party the power to propose a split and the other party the power to accept or reject. We let the buyer be the

[7] For subsequent contributions to the literature on rational behavior in extensive form games, see Ben–Porath (1997), Asheim (1999) and the references therein.

proposer. A proposal is a pair (x_B, x_S), with the restriction that $x_B + x_S = G$. Hence, without loss of generality a proposal is a number x_B. A response is denoted $a(x_B) \in \{Accept, Reject\}$.

Let $\mathcal{A}$ be the set of all functions $a : [0, G] \to \{Accept, Reject\}$. The sets of pure strategies in the bargaining game are then $P_B = [0, G]$ and $P_S = \mathcal{A}$. The payoff functions in the bargaining game are

$$\pi_i^U = \begin{cases} x_i & \text{if } a(x_B) = Accept; \\ 0 & \text{if } a(x_B) = Reject. \end{cases} \tag{4}$$

Observe that this bargaining game too has many Nash equilibria. As in the Nash demand game, there are several efficient equilibria. For example, consider any proposal $x_B^* \in [0, G]$ in combination with the seller's strategy

$$a(x_B) = \begin{cases} Accept & \text{if } x_B \leq x_B^*; \\ Reject & \text{otherwise.} \end{cases}$$

However, among all the efficient equilibria of the ultimatum bargaining game only one is subgame perfect, namely that in which $x_B^* = G$, where the buyer keeps all the surplus. There is also an inefficient equilibrium; the proposal $x_B = G$ in combination with the response function $a(x_B) = Reject$ for all x_B. This equilibrium is not subgame perfect.

Turning to the whole game, denoted Γ^U, the sets of pure strategies are $P_B^U = [0, G]$ and $P_S^U = \{0, 1\} \times \mathcal{A}$. The payoff functions are

$$u_B^U = \begin{cases} x_B & \text{if } I = 1 \text{ and } a(x_B) = Accept; \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

and

$$u_S^U = \begin{cases} G - x_B - F & \text{if } I = 1 \text{ and } a(x_B) = Accept; \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Let us first characterize the pure strategy equilibrium outcomes.

Proposition 3 *(i) In the game Γ^U, there exists a set of efficient pure strategy equilibrium outcomes, given by $I = 1$ in combination with any proposal $x_B^* \leq G - F$ and any response function with the property that $a(x_B^*) = Accept$ and $a(x_B) = Reject$ for all $x_B > x_B^*$. (ii) There also exists an inefficient pure strategy equilibrium outcome with $I = 0$, which is supported by any offer $x_B^* \geq G - F$. (iii) The unique subgame perfect equilibrium is $(I = 0, x_B = G, a = Accept)$.*

It is interesting to note how similar this set of pure strategy equilibrium outcomes is to the set of pure strategy equilibrium outcomes in the Nash demand game. In particular, they span exactly the same combinations of payoffs. The differences only become pronounced when we apply refinements of Nash equilibrium:

First, in the ultimatum game there is never investment in a subgame perfect equilibrium. The point is that the seller has an incentive ex post to accept any offer greater than zero. Hence, this is what the buyer will offer. As a result, the seller does not invest, regardless of how low the investment cost is. Second, the Dekel–Fudenberg procedure dictates no investment. To see this, observe that it is weakly dominated for the seller ever to play a strategy where he rejects a positive offer. Eliminating these responses makes it strictly dominated for the buyer to offer any positive amount.

If we believe players to be perfectly rational, in the precise sense that the Dekel–Fudenberg procedure is applicable, there is a strong prediction from our model: If agents bargain according to the Nash demand game, there should always be investment. If they bargain according to the ultimatum game, there should never be investment.[8]

2.3 Social preferences

Above, we assumed that agents are only concerned about their own monetary payoff. A considerable amount of experimental evidence suggests that many people are not so narrowly self–interested. In particular, many appear to care about how well they do relative to other agents that they interact with. Recently, several authors have proposed formulations which can account for "social" behavior while preserving much of the simplicity and predictive content of earlier models. In this subsection we investigate how social preferences alter the theoretical predictions. The two kinds of social attitudes we consider are *altruism* and *inequity aversion.*

Once we allow agents to hold social preferences, it seems quite unnatural to insist that everyone must share precisely the same preferences, and that preferences are perfectly observable. Hence, we shall also discuss the implications of incomplete information.

2.3.1 Altruism

Suppose that agent i has the utility function

$$w_i = u_i + \gamma_i u_j, \;\; i \neq j, \tag{7}$$

[8] On the other hand, as shown by Ellingsen and Robles (2000), the set of stochastically stable outcomes does include outcomes with efficient investment even in this case.

with $\gamma_i \in [0, 1)$. Recall that u_i is agent i's monetary payoff. If γ_i is positive, we shall say that agent i is altruistic. Since $\gamma_i < 1$, we do not allow agents to care as much about their opponent as about themselves.

To start with, let us assume that there is complete information, i.e., agents know each others' preferences. It is then easily shown that Proposition 2 carries over.

Proposition 4 *Suppose that both agents have preferences satisfying $\gamma_i \in [0, 1)$ and that these preferences are common knowledge. Let strategies and material payoffs be given by Γ^N. (i) Three rounds of elimination of weakly dominated strategies implies investment. (ii) If there is a smallest unit of account, the Dekel–Fudenberg procedure implies investment.*

The proof is essentially the same as before, and we therefore omit it. The only notable difference is that an altruistic seller might now be willing to invest even if he expects to get less than the investment cost F back in the negotiations.

If there is bargaining according to the ultimatum game, on the other hand, sufficient seller altruism might change the result.

Proposition 5 *Suppose that both agents have preferences satisfying $\gamma_i \in [0, 1)$. Let strategies and material payoffs be given by Γ^U. (i) If $\gamma_S < F/G$, the unique subgame perfect equilibrium is $(I = 0, x_B = G, a = Accept)$. (ii) If $\gamma_S > F/G$, the unique subgame perfect equilibrium is $(I = 1, x_B = G, a = Accept)$.*

To prove this result is easy. We start from the last stage and apply backward induction. Suppose the seller has invested and made some offer x_B. Regardless of the exact value of the parameter γ_S an altruistic seller's best response is to accept the offer. Since the buyer values own monetary payoff higher than the seller's monetary payoff, even an altruistic buyer will thus demand the whole pie, $x_B = G$. It remains to check whether the seller will invest. If the seller does not invest, $w_S = 0$. If he invests, $w_S = -F + \gamma_S G$. Comparing these payoffs yields the result.

The assumption that agents perfectly know each other's preferences, in particular each other's parameter γ_i, is unrealistic. We therefore also investigate how incomplete information about the γ's affects our results. In particular, let the buyer hold some prior on γ_S, with support $[0, \bar{\gamma}]$. Let p denote the probability that $\gamma_S = \bar{\gamma}$.

Proposition 6 *Suppose that there is incomplete information about γ_S, with $Prob(\gamma_S = \bar{\gamma}) = p$. Let strategies and material payoffs be given by Γ^N. (i) Three rounds of elimination of weakly dominated strategies implies that there is investment with probability of at least p. (ii) If there is a smallest unit of account, the Dekel–Fudenberg procedure implies investment with probability of at least p.*

The idea here is that the smallest seller claim which is consistent with rational behavior is a decreasing function of γ_S. Let $\underline{c}_S$ denote the lowest seller claim which may give a seller with parameter $\bar{\gamma}$ a positive profit. Moreover, for given seller beliefs about buyer behavior, more altruistic sellers get higher expected utility from investing. Hence, conditional on investment, the only seller strategies that can be eliminated are offers lower than $\underline{c}_S$.

Observe that Proposition 6 does not guarantee a positive probability of investment. Unless the density function over γ_S has positive mass at $\bar{\gamma}$, so that $p > 0$, our rationalizability criteria have no cutting power whatsoever. This again means that a tiny amount of incomplete information may destroy arguments based on deletion of dominated strategies.

On the other hand, when bargaining is carried out according to the ultimatum game, incomplete information does not have any influence on the result; Proposition 5 does not invoke the assumption that information is complete.

2.3.2 Inequity aversion

While altruism is a simple and plausible form of social attitude, there are many phenomena which it cannot explain. For example, as noted above, it cannot explain why responders reject meager offers in ultimatum games. To explain such behavior through social preferences, it appears that people must care about how their payoff compares to that of their opponent. One model with this feature is proposed by Fehr and Schmidt (1999), who assume that an agent i has a utility function

$$w_i = u_i - \alpha_i \max\{u_j - u_i, 0\} - \beta_i \max\{u_i - u_j, 0\},\ i \neq j, \tag{8}$$

with the parameter restriction $\alpha_i \geq 0$. As above, u_i denotes agent i's net monetary payoff. A narrowly self–interested agent is given by the special case $\alpha_i = \beta_i = 0$. If $\alpha_i > 0$, agent i dislikes having a lower monetary payoff than his opponent. If $\beta_i > 0$, agent i dislikes having a higher payoff than his opponent.

Like altruism, inequity aversion turns out to have little influence on our analysis when agents bargain according to the demand game and preferences are common knowledge.

Proposition 7 *Suppose that agent i has preferences satisfying $\alpha_i \geq 0$ and that these preferences are common knowledge. (i) In the game with material payoffs given by Γ^N, three rounds of elimination of weakly dominated strategies implies investment. (ii) If there is a smallest unit of account, the Dekel–Fudenberg procedure implies investment.*

To prove that Proposition 2 extends in this way, it suffices to observe that a seller will only invest if he believes that the net payoff is non–negative. If the seller's α is zero, he must demand F or more, as shown above. If the seller's α is positive, he must demand strictly more than F in order to get non–negative utility. Hence, if anything, social preferences that allow the seller to dislike being unfairly treated narrows the set of rationalizable outcomes in this case.

When bargaining proceeds according to the ultimatum game, buyer inequity aversion may expand the set of admissible outcomes.

Proposition 8 *Suppose agents are inequity averse and that preferences are common knowledge. Let strategies and material payoffs be given by Γ^U. If $\beta_B > 1/2$ the seller invests in any subgame perfect equilibrium.*

The point is quite simple. Consider $u_B > u_S$. Then the buyer's utility is

$$\begin{aligned} w_B &= u_B - \beta_B(u_B - u_S) \\ &= x_B - \beta_B(x_B - (G - x_B - F)). \end{aligned}$$

which is decreasing in x_B for all $\beta > 1/2$. Hence, the buyer will lower his offer until $u_B = u_S$. The intuition is that a large value of the buyer's β serves as a credible promise that the buyer will not make a greedy offer.

Could a large value of α_S also, by itself, motivate investment? The idea would be that a seller who is committed to rejecting low offers could extract a sufficiently large offer to cover the investment cost. However, this argument is faulty. While it is true that a large α_S would persuade the buyer that a high x_S is required, the lowest x_S that would be acceptable ex post is smaller than the lowest x_S that is acceptable ex ante. The former is the solution to the equation

$$x_S - \alpha_S(u_B - u_S) = 0,$$

whereas the latter solves

$$x_S - F - \alpha_S(u_B - u_S) = 0.$$

As long as β_B is sufficiently small, the buyer would of course make the lowest offer that is acceptable ex post. Thus, we conclude that the threat embodied in a seller's aversion to being badly treated is not *by itself* enough to sustain investment. However, a seller who is known by the buyer to have a high α_S may well be better off (in expected utility terms) by investing than a seller with a low α_S.

As with altruism, it is easily shown that incomplete information about parameters destroys the model's predictive power under Nash demand bargaining, whereas the result under ultimatum bargaining is less sensitive.

To conclude, the existence of social preferences lowers the predictive power of the model and also somewhat changes the predictions we can make. If the bargaining procedure is the Nash demand game, and players cannot explicitly coordinate, a small amount of asymmetric information about preferences means that we cannot confidently predict investment. If the bargaining game is the ultimatum game, we can no longer rule out investment. There may be investment if either (i) the seller is sufficiently altruistic, or (ii) the seller believes that, with high probability, the buyer is averse to demanding unfair outcomes.

This is not to say that any behavior can be rationalized through any model of social preferences. Even if people's preferences and beliefs differ, some behaviors are only consistent with either altruism or inequity aversion, and other behaviors may require very extreme beliefs to be rational.

2.4 Communication

In many realistic situations, parties can communicate before investments are made. Theoretically, communication in the form of cheap talk may improve agents' ability to coordinate on a particular equilibrium, but must not always do so.[9]

If bargaining is carried out through the Nash demand game, theory suggests that unilateral cheap talk should (almost) eliminate the risk of impasse and non–investment. To see this, suppose that one of the agents can send a message before play starts. Even if players have heterogeneous social preferences, and these are private information, a fool–proof strategy for the sender would be to suggest an equal split of the net surplus. Then, regardless of the opponent's preference parameter, as long as he is either altruistic or inequity averse in the manner described above, it will be in both agents' interest to stick to the suggested division: this split is always a Nash equilibrium, and it is always consistent with investment. Of course, depending on the own preferences and beliefs about the opponent, an agent might try to do even better than an equal split by suggesting a larger share of the surplus for himself, and with some probability such an attempt might backfire.

If bargaining is carried out through the ultimatum game, cheap talk is irrelevant. Here, there is generally a unique subgame perfect equilibrium outcome, so there is no role for a pure coordination device.

As mentioned in the Introduction, there is substantial empirical evidence that there is more to communication than the cheap talk paradigm admits. Psychological literature has considered a number of potential explanations,

[9]Farrell and Rabin (1996) survey the theoretical literature about "cheap talk." For experimental evidence on the coordinating power of communication, see Crawford (1998). Roth (1995) evaluates the effect of communication in bargaining games.

with recent work focussing primarily on two of them. The first is that communication creates *identification* among subjects, i.e., it can create altruism.[10] The second is that people can commit emotionally to act according to their statements, i.e., can make credible any threat or promise.[11] If emotional commitments work perfectly, the seller can send the threat "give me half the net surplus, otherwise I will reject your proposal" and safely invest. Likewise, the buyer can send the message, "I promise half the net surplus," and again the seller can safely invest.[12] Kerr and Kaufmann–Gilliland (1994) is a relatively recent attempt at distinguishing between the identification hypothesis and the commitment hypothesis within the framework of a five–player game of voluntary provision of a public good. They find that the commitment hypothesis is superior, and that commitment is internalized.

We shall distinguish between promises and threats. Following Schelling (1960), we imagine that threats and promises directly affect the utility of the threatener/promisor; with immutable preferences, threats and promises have almost no place within conventional game theory.[13] Like Schelling, we also say that a promise is a commitment which is costly (to the promisor) if it succeeds, whereas a threat is costly if it fails. The reason for the asymmetry is that a successful promise is carried out, whereas a successful threat is not; for a precise formalization, see Klein and O'Flaherty (1993). In our setting, this means that the buyer is in the position to give promises (not to abuse his bargaining power), while the seller is in a position to threaten (to reject meager offers).

3 Experiments

We have conducted six experiments, where bargaining is preceded by a decision whether to invest SEK 60 or not (SEK=Swedish crowns; Exchange rate

[10]Proponents of this view are, among others, Kramer and Brewer (1984), Dawes et al. (1988), and Orbell et al. (1991).

[11]Proponents are, among others, Ostrom et al. (1992), Braver (1995) and Kerr (1995). This idea has also been advocated by some economists, notably Frank (1987,1988) and Hirshleifer (1987).

[12]A third related possibility is that people's intentions (preference parameters) are given, and that they dislike to lie about them. Agents may then be able to truthfully communicate their predispositions. In this case, a buyer who has preference parameter $\beta_B \geq 1/2$ might credibly promise to split the net surplus equally. However, as discussed above, threats by the seller would not be sufficiently strong to sustain investment.

[13]The exception is if there are important decision nodes at which a player is indifferent; see Tranæs (1998).

March 1999 was USD 1=SEK 8.13). If the investment is made, the investor and a trading partner bargain over how to divide a SEK 100 revenue created by the investment. We carried out one set of experiments where the bargaining protocol was a Nash demand game and one experiment where the bargaining protocol was an ultimatum game (with the trading partner as proposer). Subjects were recruited among undergraduate business and economics students at the Stockholm School of Economics. They were paid a participation fee of SEK 100 in the experiments without communication and a fee of SEK 60 in the experiments with communication.[14] Below is a description of our procedures.

The subjects are recruited to two different rooms. In each room subjects are given a number between 1 and N, where N is the number of students in the room, and subjects with identical numbers form a pair. Anonymity is maintained throughout. The subjects are given the instructions and protocols and are asked to read them. When all subjects have read the instructions, the instructions are read aloud by the experimenter, who also answers clarifying questions, but does not answer questions regarding sensible strategies.

The experiments without communication continue as follows.[15] At the first stage, the subjects in the "investor room" are asked to decide whether to invest SEK 60 to create a revenue of SEK 100 for himself/herself and a trading partner in the other room. The subjects record their investment decision on a form marked "Investment decision". When everybody has filled in the form, it is collected by the experimenter. The experimenter then hands out the forms to the respective trading partners in the other room. At the second stage, there is bargaining over the surplus, and this stage differs across experiments.

Observe that each subject plays only a single round. We did this for several reasons. First, our research is motivated by the desire to understand the willingness of people to make unique relationship–specific investments when contracts are inevitably incomplete. Hence, the isolation of the event was itself desirable. Second, the situation is relatively simple, so practice rounds would not necessarily improve subjects' understanding of the game much. (This is why the plain ultimatum game is often played without prior practice too.) Third, for a given level of incentive practice rounds introduce an element of boredom. Fourth, practice rounds must be carefully designed to avoid any strategic effects from repetition.

---

[14]The reason for the difference in participation fee is that in the latter case we were able to perform the experiments during lecture time (though not in our own course or with students we knew), whence the participation fee was not required to attract subjects. At the same time, we wanted to make sure that students did not abstain from investing due to a cash constraint.

[15]We describe experiments with communication below.

3.1 Nash demand game

In this treatment, bargaining is carried out the following way. Each subject writes down their claim for part of the revenue of SEK 100. This is done on a form marked "Bargaining". For pairs in which no investment was made, the subjects marks the box "No investment" on the bargaining form. The forms are then collected by the experimenters and handed out to the respective trading partners. The subjects in both rooms record the investment decision and the claims in their protocols. Both subjects also estimate their revenue from the bargaining and records their revenue on the protocol. The revenue is equal to 0 if the sum of the claims exceed SEK 100. If the sum of the claims are SEK 100 or less, the revenue of a subject equals the claim made by that subject. The subjects finally estimate their earnings from the experiment and records their earnings in the protocol. For subjects in the "investor room," the earnings equal the revenue from the bargaining less the investment cost of SEK 60. For subjects in the other room, the earnings are equal to the revenue from the bargaining. The subjects in the experiment are paid their earnings in the experiment plus the participation fee. The experiment is then over.

The exact instructions for this experiment are reproduced in Appendix 1 (translated from Swedish to English). We carried out two sessions of this version of the experiment (with 20 subject pairs in one session and 19 subject pairs in one session).

3.2 Ultimatum game

In this version of the experiment, the first stage is as above, whereas bargaining procedes as follows. The trading partner writes down a proposed division of the revenue of SEK 100. This is done on a form marked "Bargaining". For pairs where no investment was made, the trading partner marks the box "No investment" on the bargaining form. The forms are then collected by the experimenters and handed out to the respective investors. Investors accept or reject the proposed division by marking the "accept" or "reject" box on the bargaining forms. The forms are then collected again by the experimenter and handed out to the respective trading partners in the other room.

The subjects in both rooms record the proposed division and whether it was accepted or not in the protocol. Both subjects also estimate their revenue from the bargaining and record this in their protocols. The revenue is equal to 0 if the proposed division is rejected. If the proposed division is accepted, the revenue is equal to the proposed amount. The subjects finally estimate their earnings from the experiment and record their earnings on the protocol. For subjects in the "investor room" the earnings are equal to the revenue from the bargaining minus the investment cost of SEK 60. For subjects in the other room, the earnings are equal to the revenue from the bargaining. The subjects in the experiment are paid their earnings in the experiment plus their participation fee. The experiment is then over.

The exact instructions for this experiment are reproduced in Appendix 2. We carried out three sessions of this experiment (with 17 subject pairs in one session, 12 subject pairs in one session, and 11 subject pairs in one session).

3.3 Statistics

To investigate whether the proportion of investors differs between the experimental groups, we use a contingency table Pearson chi-square test.[16] The null hypothesis is no difference, and we report two–sided p–values.

We also want to test whether the investor's payoffs differ across experiments. Statistically, this is a tricky issue. Bargaining experiments usually lead to a highly skewed payoff distribution. For this reason, many authors have rejected standard parametric tests in favor of non–parametric tests, which do not invoke normality assumptions. For a careful discussion of this issue, see, for example, Roth et al. (1991). The drawback of non–parametric tests is that they do not utilize the rich cardinal information in the data. With recent advances in econometric theory and computer power, it has become possible to conduct parametric testing without imposing normality, i.e., by inferring the underlying distribution from which the data has emerged using bootstrap techniques.[17] The significance levels for the payoff comparisons that we report below have all been obtained by generating 1099 bootstrap replications. According to Davidson and MacKinnon (1998), this number of replications is high enough to guarantee a reasonable confidence in the estimated p–values, compared to the "ideal" bootstrap with infinitely many replications.[18] Although the bootstrap technique has been widely used recently, we are not aware of previous studies which have used it for experimental bargaining games.

---

[16]D'Agostino et al. (1988) have shown that the commonly used Yates correction, as well as Fisher's exact test, are overly conservative.

[17]For an introduction to bootstrap methods, see e.g. Efron and Tibshirani (1993). Some powerful recent characterizations are reported by Davidson and MacKinnon (1999).

[18]Davidson and MacKinnon (1998) argue that at least 399 replications is needed to avoid a power loss of more than 1 percent for computed p–values of about 0.05, whereas at least 1499 replications are required to avoid a similar power loss for p–values of about 0.01. (The reason for choosing numbers of replications ending with 99 is that Monte Carlo tests are exact only if $p(B+1)$ is an integer, where p is the significance level and B is the number of replications. The exact choice of B thus matters if special significance is attached to particular p–values, such as 0.05.)

3.4 Results without communication

The proportion of investors in each experimental group is shown in Table 1. (Table 1 in here). The proportion of subjects investing SEK 60 is 26% in the Nash demand game group and 35% in the ultimatum game group. The difference between the groups is not statistically significant (p=0.366).

Table 2 depicts the bargaining behaviour in the Nash demand game for the ten investing pairs. (Table 2 in here). The claims of the investors are in the interval [60,100] in all cases except one. The most common claim is SEK 80 that is made by 40% of the investors. The mean claim of the investors is SEK 72.00. The claims of the trading partners lie in the interval [0,40] in all cases. The most common claim is SEK 40, that is made by 50% of the trading partners. Observe that these claims are inconsistent with narrowly self–interested behavior, which only allows claims smaller than 40. (On the other hand, it would not have changed anything if all these claims had been 35 instead.) The second most common claim is SEK 20, that is made by 40% of the trading partners. The mean claim of the trading partners is SEK 30.50. On average the investors lose money. The average loss is SEK 28.50, which is significantly different from 0 (the p-value is 0.008).

Could bargaining impasse in the Nash demand game be caused by incomplete information about (social) preferences? We think so. The four trading partners who claim 40 might have believed that most of those who chose to invest would be slightly altruistic, meaning that they might rationally claim 60. Indeed, one of the investors did exactly this. The majority, who did not invest, might not have been altruistic, meaning that abstention was indeed a rational action under the circumstances. On the other hand, the observed outcome cannot be explained by a model which admits inequity aversion alone, because a trading partner's claim of 40 is then not rationalizable.

Table 3 displays the bargaining behaviour in the Ultimatum game for the 14 investing pairs (Table 3 in here). On average the proposers offer SEK 48.57 to the investors. The most common proposal is to offer SEK 80 to the investor. This offer is made by four of the proposers. The proposed division is rejected by three out of 14 investors (21%), and these rejections are for the three lowest offers (SEK 0, SEK 10, and SEK 10). On average the investors lose SEK 12.86 from the investment, but this loss is not significantly different from 0 (p=0.126). The investor's average loss is lower than in the Nash demand game group, but the payoff difference between groups is not significant (p=0.268). In the ultimatum game, the "anomaly" is that there is any investment at all. Investment is not obviously irrational, because trading partners sometimes do make quite generous offers; half of the investors make a profit. Why offers are generous is not quite clear from the experiment itself. Whereas very low offers are rejected, we observed no rejections of offers above 10. This suggests that fear of rejection is not the main reason for high offers. The behavior can not be explained by a model which admits altruism only, because as we have demonstrated above, such a model does not explain generous offers by the trading partner. A model of inequity aversion does better. If some investors believe that there is a high probability of being matched with a trading partner with $\beta \geq 1/2$, then these investors may rationally invest. For example, if the probability of being matched with such a trading partner is 3/4 or more, investment yields a positive expected profit. Clearly, there is also a distribution of β's (and of beliefs about α) which can explain the observed trading partner behavior. As it turned out, the investors were overly optimistic. However, we do not think the discrepancy between the required investor beliefs and actual trading partner behavior is so large that we can reject investor rationality.

Finally, and perhaps more subtly, the model with inequity aversion predicts that investment is an indication that the investor has relatively low α, i.e., he is unlikely to punish meager offers. This might explain why two out of the five meagrest offers were accepted by investors (one offer of 10 and one of 20).

Overall, we conclude from this first set of experiments that the propensity to invest is modest, and seemingly quite independent of the mode of bargaining. Both these results contradict the prediction based on rational self–interested behavior, which was that there should always be investment in the Nash demand game and never investment in the ultimatum game. In order to explain the observed behavior, we need to assume both that some agents are believed to be slightly altruistic and that about thirty per cent of the agents believe that there is a large probability of being matched with a partner who is strongly inequity averse, in the sense of having a large β. While inequity aversion alone is not enough to explain behavior, we only need to add in a slight amount of altruism to fit the data from the two experiments without communication.

3.5 Results with communication

We introduced communication in the simplest possible way, namely by having one of the parties sending a written message. In order to investigate whether it matters who sends the message, we conducted one set of experiments in which the investor had the opportunity to communicate and one set of experiments in which the trading partner had the opportunity to communicate. In the former case, the message was sent simultaneously with the investment decision; in the latter case, the trading partner sent the message (and the investor received it) before the investment decision had been made. No restrictions were put on the content of the message, nor did we propose what it may contain. Otherwise, the experiments were carried out exactly as in the case of no communication.

Table 4 summarizes the results. (Table 4 about here.) Generally, investment increases quite substantially when communication is allowed. In the Nash demand treatments on average 79% now invests, whereas on average 59% invests in the ultimatum game. Both rates are significantly higher than the rate under no communication (as shown in Table 7 below). Thus, the investment rate is roughly in line with the cheap talk prediction for the Nash demand game. However, the increase in investment rates for the ultimatum game is not predicted by a model based on cheap talk.

The allocation of authority (who sends the message) matters only slightly for the probability of investment, but perhaps surprisingly it appears that investors are better off when the trading partner has authority, in particular in the ultimatum game. Our interpretation of this finding is that promises are rarely broken, and that trading partners can induce investment by promising a high enough rate of return. (Possibly, trading partners stick to their promises out of fear that investors would not tolerate any reneging of a promise, but our data does not allow us to identify the motive directly.) On the other hand, investors threatening to reject low offers are less determined to carry out their threats.

An apparent incongruence in the data is that the probability of investment is at least as high when the investor sends a message as when the trading partner sends a message, although the investor's profit is higher in the latter case. One possible reason might be that investors have overconfidence in their ability to affect the behavior of trading partners, and too little confidence in the trading partners' assurances. Our current sample size is however to small to draw any sharp conclusion about this issue.[19]

Below is a more detailed description of the findings.

3.5.1 Behavior in the demand game

Observe that introducing communication into the demand game almost transforms it into an ultimatum game. If the sender indicates the planned claim in the message, and the message is believed, then the other party's best reply is to ask for the remainder, thereby in practice accepting the expected proposal. However, there is a very important distinction. When the trading partner sends the message, the investor has not yet made the investment decision. Hence, the trading partner has no possibility in this experiment to stage a hold–up.

In Table 5 we report the claims made in the demand game for each of

[19]For experimental evidence showing that overconfidence in own abilities may lead to excess entry, see Camerer and Lovallo (1999).

the two communication procedures. The most striking finding is the incidence of 80:20 splits, which are attained by more than half of all pairs. The fraction of bargaining breakdowns is lower than without communication. Five disagreements out of a possible 28, as found in the case where the investor communicates, is still high enough to almost wipe out the investors' profits. Interestingly, there are fewer disagreements when the trading partner communicates, but the sample is too small for this difference to be statistically significant. A possible explanation for the difference in bargaining impasses is that when the trading partner communicates, the investor can object to what he considers an unfair proposal by not investing; when the investor communicates, the trading partner can only object by creating an impasse.

In order to find out more about this, we read all of the messages. In all successful cases, there were explicit proposals, and agents played according to these proposals, as theory suggests. When the investor was sender, two of the five impasses arose when the investor did not propose a specific split (these were the only two messages where an explicit proposal was not made). In both cases, the investor ended up demanding 80, whereas the trading partner demanded 30 and 60 respectively. In one case, the trading partner appears to have misunderstood the rules of the game, something the investor cannot insure himself against under this treatment, and the outcome was an incompatible 80:50. In one case, the investor made an advanced proposal which entailed meeting up afterwards and redistribute the gains. Possibly the trading partner understood that the investor had every possibility and incentive to remain anonymous (the message contained a name etc, but this might of course have been the details of another participant), and the outcome was incompatible, 100:40. Finally, in one case the investor proposed 80:20, but the trading partner created an impasse by demanding 30. This behavior by the trading partner is not congruent with any preference patterns suggested in the literature; we conclude that this subject is either irrational (did not understand the game) or has very peculiar preferences. The two impasses when the trading partners sent the message were equally odd. In one case, the trading partner's message proposed 90:10. Then, the person went on to claim 30 instead of 10, which appears to be an act of either stupidity or spite. The second impasse was created by a message which said "Do not invest," but had a proposal of 80:20 crossed over. The outcome was 80:50, and perhaps this investor should have known better than to invest.

Another somewhat surprising finding is that the 50:50 split of the gross return is chosen by about ten percent of the pairs, despite giving a very unequal division (-10,50) of the net surplus. This finding suggests that some of the investors are quite altruistic, with a γ of at least 0.2.

3.5.2 Behavior in the ultimatum game

Table 6 displays the proposals and responses in the ultimatum game for each of the two communication procedures.

When the investor has authority (sends the message), trading partners' proposals vary considerably. A quarter of the offers give the investor a net loss. The two most popular offers is 60:40, which just covers the investment cost and 80:20. Investors sometimes reject low offers, but not always. One investor rejects a 50:50 offer, one rejects a 60:40 offer, and one even rejects a 75:25 offer (leaving us slightly unsure whether this subject had actually understood the offer); at the same time, two loss–making 40:60 offers and one 30:70 offer are accepted. The rejections are most easily explained by inequity aversion. Since the investment rate is higher than without communication, there is reason to believe that some more inequity averse agents now have chosen to invest.

The crucial question is of course why the investment rate is significantly higher with investor communication than without communication (p=0.015). Since cheap talk has no explanatory power in the ultimatum setting, we must ask whether communication creates identification, conveys preferences, or carries an emotionally sustained threat that meager offers will be rejected. It is not trivial to discriminate between these explanations. Compared with the case without communication, the data shows both that the trading partners' offers are more generous and that the investors are more prone to reject meager offers. Nonetheless, we are tempted to dismiss the explanation that communication creates more altruistic or inequity averse trading partner behavior.[20] The fraction of very high offers (70:30 or better), does not increase much. Instead, there is a reduction of very meager offers and an increase in intermediate offers, notably 60:40. This is close to being the optimal selfish offer, given the empirical rejection behavior.

What then about the argument that the investors' messages contain credible threats? To investigate this issue more closely, we again need to examine the content of the messages. The first interesting observation is that nine out of the invesors' 21 messages contained explicit threats that meagrer offers than the investor's suggested split would be rejected. However, only in four of these cases did the trading partner's proposal respect the threat. In the remaining five cases, the trading partner made a lower proposal, which the investor accepted in all but one case. The most glaring neglect of a threat was perpetrated by the trading partner who successfully proposed a 30:70 split when

[20] However, we sympathise with a trading partner who proposed 100:0, giving away all his surplus, when the investor's message suggested 80:20. Another trading partner rewarded the modest message 60:40 by a proposal of 65:35.

the investor had threatened not to accept lower offers than 80:20. Hence, the experiment indicates that threats are not very credible. We observe that a vast majority, 16 out of 21, of the investors messages suggest 80:20, but in nine of these cases the trading partner's offer is lower. On the other hand, it is possible that a threat of not accepting anything less than 80:20 really means that the investor in many cases is not prepared to accept anything less than 60:40. This would be consistent with the observation that the one threat which was carried through rejected a 50:50 offer. The partial credibility of threats is an issue which we believe deserves further attention.

Turning to the case in which trading partners could send a message, the investment rate also goes up compared to the no communication treatment, but the effect is not statistically significant. However, the distributions of offers is radically different, as shown in the bottom panel of Table 6. Now, all proposals except one allows the investor to recoup his investment cost, and half of the offers are for an even split of the net surplus, i.e., 80:20. The explanation for the change in trading partner behavior is quite clear. Ten out of the sixteen investments followed explicit promises by the trading partner; in seven cases the promise was 80:20, in the remaining three cases 70:30. None of these promises were broken! (The other six investments followed messages which did not make explicit promises.)

The credibility of promises makes it slightly puzzling that not all promises of 80:20 or 70:30 lead to investment. Indeed, the investment rate given an 80:20 promise was not much higher than the investment rate following more obscure messages, although the latter tended to give a smaller profit for the investor. Of course, the fraction of promises kept would be lower than 1.0 in a larger sample. Nonetheless, the fact that investors do better on average when the trading partner communicates squares badly with the lower investment ratio under this scenario. Our results seem to suggest that the subjects are overly pessimistic about the credibility of promises.

4 Related work

Our main theoretical points appear to be novel; until now, the hold–up problem has been analysed on the presumption that the bargaining game has a unique equilibrium and that agents are narrowly self–interested.

Hackett (1994) appears to be the only earlier contribution to investigate experimentally the impact of ex ante investments on ex post bargaining outcomes. Hackett applies a version of alternating offer bargaining with a positive probability of breakdown. Based on the conventional analysis of this game, he argues that the unique subgame perfect equilibrium implies that bargaining is insensitive to sunk costs. His major experimental finding is that agreements are sensitive to transaction–specific investments whenever these investments are observable. Hackett interprets his result as evidence that agents have social preferences. Our theoretical analysis allows a different explanation, namely that his bargaining game has multiple subgame perfect equilibria. (To repeat, the alternating offer game is only guaranteed to have a unique subgame perfect equilibrium under the questionable assumption that money is infinitely divisible.) Our experiments suggest that there is something to both stories. Indeed, our experiments complements Hackett's in at least two respects. First, the details of the investment decisions are rather different. Hackett studies a much richer setting, with bilateral investments, uncertain investment returns, and a continuous choice of investment level. While our investment decisions and bargaining procedures are more extreme and perhaps less realistic, the deviations from conventional predictions are easier to identify and categorize. Second, and perhaps more importantly, we have systematically studied the impact of communication.

Binmore et al. (1998), experimentally investigate bargaining behavior in a modified version of Nash's demand game when one of the players has an outside option. Apart from a slight modification of the payoff functions, this game has the same normal form as our demand game experiments.[21] Despite the close relationship, the authors do not relate their work explicitly to the hold–up problem. Compared to our analysis of the demand game, Binmore et al. perform a larger number of experiments, with different outside options, but no communication. They also conduct a large number of practice rounds. Considering the experiment closest to ours, with an outside option of USD 6.40 and a gain of USD 10, the opt out frequency was 0.64, yielding a probability of investment of 0.36, which should be contrasted with a frequency of 0.26 in our demand game without communication. The slighly higher propensity to invest in their experiment could to some extent be explained by their modification of the demand game, which ensures that no money is ever left on the table. It is noteworthy that Binmore et al. interpret the bargaining impasses to be a sign that some people are boundedly rational, while we have pursued instead the explanation that people may have social preferences, and that these preferences are not common knowledge.

Trust games, sometimes called gift–exchange games, concern a problem which is related to the hold–up problem. In these games, one party (the

[21] Oddly, Binmore et al. considers the "deal–me–out–outcome," in which the investor only recoups the investment (the outside option), to be the prediction from a fully rational model (what they call a perfect world). As argued in Section 2, fully rational and completely informed players should not ever settle on this outcome in the Nash demand game.

trustor) first decides whether or not to give a gift to the second party (the trustee), who then decides whether to reward the trustor. The first gift is typically worth more to the trustee than to the trustor, whence it is socially desirable that this transaction is made. However, the trustee is under no obligation to reward the trustor, and so the unique prediction if people are selfish is that there will be no reward and hence no gift in the first place. Trust games differ from the games that we have considered above in that the trustor has no way of punishing an ungrateful trustee; the second stage is a dictatorship game rather than an ultimatum game. As shown experimentally by Berg, Dickhaut and McCabe (1995) among others, some trustors give sizeable gifts and are often rewarded by trustees, although we should mention that these experiments tend to have much greater returns to the investment (gift) than we have assumed here. Our experiments raise the question whether communication could have an effect in trust games too. By investigating this issue, we might learn whether promises work because they affect the promisor's preferences directly (which would be necessary for promises to have an effect in trust games), or whether they work because broken promises are punished.

As mentioned above, social psychologists have documented that communication is of great help in mitigating social dilemmas. The meta–analysis of Sally (1995) considers 130 distinct experiments from 37 different studies. Of all the factors which according to the "economic" model should not matter, discussion is by far the most important, vastly raising cooperation rates. Almost all these studies concern dilemmas in which the relevant commitments are promises rather than threats. There is no scope for statements of the form: "if you do not cooperate, I will punish, even if punishment is costly to myself." Because an agent benefits from witholding the own private contribution to the public good, he can only give a promise. In contrast, we have studied an asymmetric situation, in which the seller can not promise anything, only threaten to reject low offers. We find that such threats are less credible than promises. This asymmetry is not predicted by standard psychological theories of why people keep promises. For example, the common view that promises are kept because people dislike to be inconsistent, appears to imply that threats should be carried out as well.[22] If our finding stands up in future experiments, this is a puzzle for psychologists to solve.

[22]Theories of behavior which emphasise the desire for consistency include Heider (1946), Newcomb (1953) and Festinger (1957). See also Cialdini (1993, Chapter 3).

5 Final Remarks

Our view is that bargaining is usually a game with many equilibria, and that if agents can communicate at the investment stage, there need not be a hold–up problem. The investment rates of about 0.8 in our demand game experiments, would in all probability be even higher with communication possibilities closer to real life conditions. And even when the investor is at a clear disadvantage at the bargaining stage, we find that non–binding promises by the trading partner are quite effective, although the investment rate is distinctly smaller than one.

One caveat, shared with most experimental work, is that we cannot know for sure what would happen if the stakes were larger. This objection is perhaps most valid for the asymmetric bargaining procedure. Would responders continue to reject low offers, and would proposers continue to be generous? Would promises still be credible? Recent experimental evidence suggest that larger stakes may not alter behavior too much. For example, Cameron (1999) let people play the ultimatum game with stakes of about three times the average participant's monthly income. Compared to experiments with low stakes, responders were somewhat more accommodating, while proposers' behavior was largely unaffected.

Does our paper prove that there isn't a hold–up problem? We think this conclusion is premature. Our analysis indicates that the hold–up problem is modest when agents have complete information about costs and benefits, including costs that have been sunk.[23] But complete information is a simplifying assumption, not a description of reality. It is easy to imagine how the hold–up problem resurfaces once this assumption is dropped. In reality, there is often asymmetric information about costs and benefits—not least about sunk costs. And even if information is symmetric, it may not be complete in the sense assumed above, because it is difficult to compute costs and benefits accurately and impartially. Investment costs are frequently opportunity costs that are hard to measure objectively. Once costs have been sunk, the investor's bias is to exaggerate them, while the trading partner's bias is to underestimate them. Babcock and Loewenstein (1997) survey a range of studies which show that such self–serving biases are pervasive, and that these biases lead to bargaining impasse both inside and outside the laboratory. With more appropriate foundations, the hold–up problem may keep us occupied for quite a while yet.

[23] When Oliver Hart suggested a literary illustration of the hold–up problem for the back cover of *Journal of Political Economy,* February 1999, the tale is of a seller which clearly does not know the buyer's valuation.

6 References

D'Agostino, Ralph B., Warren Chase and Albert Belanger (1988): The Appropriateness of Some Common Procedures for Testing the Equality of Two Independent Binomial Populations, *American Statestician* 42, 198–202.

Asheim, Geir B. (1999): On the Epistemic Foundation for Backward Induction, Department of Economics, Oslo University, Memorandum No 30/99.

Babcock, Linda and George Loewenstein (1997): Explaining Bargaining Impasse: The Role of Self–Serving Biases, *Journal of Economic Perspectives* 11 (1), 109–126.

Ben–Porath, Elchanan (1997): Rationality, Nash Equilibrium and Backwards Induction in Perfect–Information Games, *Review of Economic Studies* 64, 23–46.

Berg, Joyce, John Dickhaut and Kevin McCabe (1995): Trust, Reciprocity, and Social History, *Games and Economic Behavior* 10, 122–142.

Binmore, Ken, Chris Proulx, Larry Samuelson and Joe Swierzbinski (1998): Hard Bargains and Lost Opportunities, *Economic Journal* 108, 1279–1298.

Binmore, K, J. McCarthy, G. Ponti, L. Samuelson, and A. Shaked (1999): A Backward Induction Experiment, mimeo., Department of Economics, University of Bonn.

Braver, S.L. (1995): Social Contracts and the Provision of Public Goods, in D. Schroeder (ed.): *Social Dilemmas: Perspectives on Individuals and Groups,* New York:Praeger.

Caballero Ricardo J. and Mohamad L. Hammour (1998): The Macroeconomics of Specificity, *Journal of Political Economy* 106, 724–767.

Camerer, Colin F. (1998): Bargaining and Social Preferences: Dictator, Ultimatum, and Trust Games, Draft chapter of a forthcoming book *Experiments in Strategic Interaction.*

Camerer, Colin F. and Dan Lovallo (1998): Overconfidence and Excess Entry: An Experimental Approach, *American Economic Review* 89, 306–318.

Cameron, L.A. (1999): Raising the Stakes in the Ultimatum Game: Experimental Evidence from Indonesia, *Economic Inquiry* 37, 47–59.

Cialdini, Robert B. (1993) *Influence: Science and Practice, 3rd edition* New York: Harper Collins.

Crawford, Vincent (1997): Theory and Experiment in the Analysis of Strategic Interaction, in David Kreps and Kenneth Wallis (eds.), *Advances in Economics and Econometrics: Theory and Evidence, vol I* Cambridge: Cambridge University Press.

Crawford, Vincent (1998): A Survey of Experiments on Communication via Cheap Talk, *Journal of Economic Theory* 78, 286–298.

van Damme, Erik, Reinhard Selten, and Eyal Winter (1990): Alternating Bid Bargaining with a Smallest Money Unit, *Games and Economic Behavior* 2, 188–201.

Dawes, R.M., J. Orbell and A.J. van de Kragt (1988): Not Me or Thee but We: The Importance of Group Identity in Eliciting Cooperation in Dilemma Situations, *Acta Psychologica* 68, 83–97.

Davidson, Russell and James G. MacKinnon (1998): Bootstrap Tests: How Many Bootstraps?, mimeo., Department of Economics, Queens University, Kingston, Ontario. *Econometric Reviews* forthcoming.

Davidson, Russell and James G. MacKinnon (1999): The Size Distortion of Bootstrap Tests, *Econometric Theory* 15, 361–376.

Dekel, Eddie and Drew Fudenberg (1990): Rational Behavior with Payoff Uncertainty, *Journal of Economic Theory* 52, 243–267.

Efron, B. and R.J. Tibshirani (1993): *An Introduction to the Bootstrap,* Monographs on Statistics and Applied Probability, No 57, New York: Chapman and Hall.

Ellingsen, Tore (1997): The Evolution of Bargaining Behavior, *Quarterly Journal of Economics* 112, 581–602.

Ellingsen, Tore and Jack Robles (2000): Does Evolution Solve the Hold–up Problem? mimeo., Department of Economics, Stockholm School of Economics.

Farrell, Joseph and Matthew Rabin (1996): Cheap Talk, *Journal of Economic Perspectives* 10 (3), 103–118.

Fehr, Ernst and Klaus M. Schmidt (1999): A Theory of Fairness, Competition, and Cooperation, *Quarterly Journal of Economics* 114, 817–868.

Fernandez, Raquel and Jacob Glazer (1991): Striking for a Bargain Between Two Completely Informed Agents, *American Economic Review* 81, 240–252.

Festinger, L. (1957): *A Theory of Cognitive Dissonance,* Stanford: Stanford University Press.

Frank, Robert H., (1987): If Homo Economicus Could Choose His Own Utility Function, Would He Want One with a Conscience? *American Economic Review,* 77, 593–604.

Frank, Robert H., (1988): *Passions within Reason,* New York: W.W. Norton.

Grossman, Sanford J. and Oliver D. Hart (1986): The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration, *Journal of Political Economy* 94, 691–719.

Hackett, Steven C. (1994): Is Relational Exchange Possible in the Absence of Reputations and Repeated Contact? *Journal of Law, Economics and Organization* 10, 360–389.

Haller, Hans and Steinar Holden (1990): A Letter to the Editor on Wage Bargaining, *Journal of Economic Theory* 52, 232–236.

Hart, Oliver D. and John Moore (1990): Property Rights and the Nature of the Firm, *Journal of Political Economy* 98, 1119–1158.

Hart, Oliver D. and John Moore (1999): Foundations of Incomplete Contracts, *Review of Economic Studies* 66, 115–138.

Heider, F. (1946): Attitudes and Cognitive Organization, *Journal of Psychology* 21, 107–112.

Hirshleifer, Jack (1987): On the Emotions as Guarantors of Threats and Promises, in John Dupré (ed.): *The Latest on the Best: Essays in Evolution and Optimality,* Cambridge MA: MIT Press.

Kerr, Norbert L. (1995): Norms in Social Dilemmas, in D. Schroeder (ed.): *Social Dilemmas: Perspectives on Individuals and Groups,* New York:Praeger.

KERR, NORBERT L. AND CYNTHIA M. KAUFMANN-GILLILAND (1994): Communication, Commitment, and Cooperation in Social Dilemmas, *Journal of Personality and Social Psychology* 66, 513–529.

KLEIN, BENJAMIN, ROBERT G. CRAWFORD, AND ARMEN A. ALCHIAN (1978): Vertical Integration, Appropriable Rents, and the Competitive Contracting Process, *Journal of Law and Economics* 21, 297–326.

KLEIN, DANIEL B. AND BRENDAN O'FLAHERTY (1993): A Game–theoretic Rendering of Promises and Threats, *Journal of Economic Behavior and Organization* 21, 295–314.

KRAMER, R.M. AND M.B. BREWER (1984): Effects of Group Identity on Resource Use in a Simulated Commons Dilemma, *Journal of Personality and Social Psychology* 46, 1044–1057.

LOOMIS, JAMES L. (1959): Communication, the Development of Trust, and Cooperative Behavior, *Human Relations* 12, 305–315.

NASH, JOHN Two–Person Cooperative Games, *Econometrica* 21, 128–140.

NEWCOMB, T. (1953): An Approach to the Study of Communicative Acts, *Psychological Review* 60, 393–404.

NORTH, DOUGLASS AND BARRY WEINGAST (1989): Constitutions and Commitment: the Evolution of Institutions Governing Public Choice in Seventeenth–Century England, *Journal of Economic History* 49, 803–832.

ORBELL, J.M., A.J. VAN DE KRAGT AND R.M. DAWES (1991): Covenants without the Sword: The Role of Promises in social Dilemma Situations, in K. Koford and J. Miller (eds.): *Social Norms and Economic Institutions,* Ann Arbor: University of Michigan Press.

OSTROM, E., J.M. WALKER, AND R. GARDNER (1992): Covenants with and without the Sword: Self–governance is Possible, *Americal Political Science Review* 86, 404–417.

PITCHFORD, ROHAN AND CHRISTOPHER M. SNYDER (1999): Property Rights, Incomplete Contracts and Social Harm, mimeo. Australian National University.

ROTH, ALVIN E. (1995): Bargaining Experiments, Chapter 4 in John H. Kagel and Alvin E. Roth (eds.), *Handbook of Experimental Economics,* Princeton: Princeton University Press.

Roth, A.E., V. Prasnikar,, M. Okuno–Fujiwara, and S. Zamir (1991): Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study, *American Economic Review,* 81, 1068–1095.

Rubinstein, Ariel (1982): Perfect Equilibrium in a Bargaining Model, *Econometrica* 50, 97–109.

Sally, David F. (1995): Conversation and Cooperation in Social Dilemmas: Experimental Evidence from 1958 to 1992, *Rationality and Society* 7, 58–92.

Schelling, Thomas C. (1960): *The Strategy of Conflict,* Cambridge MA: Harvard University Press.

Segal, Ilya (1999): Complexity and Renegotiation: A Foundation for Incomplete Contracts, *Review of Economic Studies* 66, 57–82.

Shleifer, Andrei and Robert W. Vishny (1997): A Survey of Corporate Governance, *Journal of Finance* 52, 737–783.

Ståhl, Ingolf (1972): *Bargaining Theory,* Stockholm: Economic Research Institute, Stockholm School of Economics.

Tranæs, Torben (1998): Tie–Breaking in Games of Perfect Information, *Games and Economic Behavior* 22, 148–161.

Williamson, Oliver (1975): *Markets and Hierarchies: Analysis and Antitrust Implications,* New York: The Free Press.

Williamson, Oliver (1985): *The Economic Institutions of Capitalism,* New York: The Free Press.