DO WEALTH DIFFERENCES AFFECT FAIRNESS CONSIDERATIONS?

Olivier Armantier

May 2003

Abstract

The influence of relative wealth on fairness considerations is analyzed in a series of ultimatum game experiments in which proposers and receivers are given large and widely unequal initial endowments. Subjects initially demonstrate a concern for fairness. With time however, the dynamics of behavior become at odds with both subgame perfection and fairness. Evidence of learning is detected for both proposers and receivers in the estimation of a structural reinforcement learning model. The estimation results suggest that, guided by foregone best responses and an acquired sense of deservingness, rich subjects become more selfish, while poor subjects, influenced only by their past plays and outcomes, learn to tolerate this behavior.

Keywords: Fairness, Learning, Ultimatum Game, Experimental Economics.
JEL Classifications : C91, C78, C13, C15.

1. Introduction

During the last two decades, experimental economics has clearly demonstrated that agents do not maximize only their monetary payoffs.[1] In particular, players in ultimatum games are willing to sacrifice their own payoffs and often refuse "free money".[2] Such behavior is widely believed to be motivated by equity or fairness concerns.[3] Experimental studies have also established that behavior in ultimatum games may be affected by learning, as well as cultural, social, psychological and/or environmental factors. The present paper is an attempt to analyze the influence of one such factor. Namely, a series of ultimatum games experiments is conducted to try to determine whether fairness considerations, and in particular their stability over time, are affected by observed wealth differences between subjects.

The analysis of observed wealth differences in ultimatum games, although challenging, is important as it may help us better understand the nature of fair behavior. In particular, it offers the possibility to test the robustness of recently proposed fairness models compared to the more traditional game theoretic approach constructed around self-interested agents. In addition, the influence of relative wealth on fairness considerations may have important, and not yet well determined, real life consequences. Indeed, traditional ultimatum games do not explicitly account for wealth differences that typically exist between participants in most real life bargaining situations such as wage negotiations, international trade between developed and less developed countries, or pre-trial settlements between (e.g.) a patient and a hospital in a malpractice case. The analysis of wealth differences as one of the determinants of bargaining outcomes and fairness perception, is therefore of academic and practical interests.

The ultimatum game corresponds to the last round of a two player bargaining process, in which a proposer offers an ultimatum in the form of a share of a given pie. This

---

[1]See also Andreoni et al. (2001) or Bewley (1998) for field evidence.

[2]See Camerer and Thaler (1995), Roth (1995) or Güth (1995) for surveys of the ultimatum game literature.

[3]The term fairness is used here broadly, and it encompasses several theories such as reciprocity, inequality aversion, or altruism.

offer can be either accepted or rejected by a receiver. If the offer is accepted, then an agreement is reached and the pie is shared according to the offer. If the offer is rejected, then the bargaining process stops and players receive no payoff. If players are assumed to maximize their own monetary payoff, then subgame perfection predicts that proposers offer the smallest possible positive portion of the pie and receivers always accept. In sharp contrast with this theoretic prediction, the following facts have been consistently observed in ultimatum experiments, independent of the experimental design: most offers lay between 40% and 50% of the pie; there are almost no offers above 50% and below 15%; the probability of rejection decreases with the offer; and offers below 20% are rarely accepted.

These experimental results were originally perceived as anomalies since it is traditionally assumed that the Homo economicus is self-interested (see Camerer and Thaler 1995). In the wake of mounting experimental evidence, some models incorporating static fairness considerations have been recently developed.[4] These models are able to organize a large part of experimental data. However, they do not account for social, psychological, and learning effects that have be shown to influence behavior in the ultimatum game. For instance, anonymity (Hoffman et al. 1994), chivalry (Eckel and Grossman 2001), and deservingness or morally justifiable entitlement (Hoffman et al. 1994, Eckel and Grossman 1996, Ruffle 1998) have be shown to affect behavior, either toward or away from the subgame perfect equilibrium. In addition, we shall see in section 2 that evidence of learning has been detected in ultimatum game experiments, suggesting that fairness considerations may not be stable over time.

A previously unexplored psychological factor that may affect the outcomes in ultimatum games is the observed wealth difference between players. Indeed, one may expect subjects to act more fairly when it is known that they are wealthier. Alternatively, wealth may also provide some bargaining power that will be used to extract more favorable and possibly more inequitable outcomes. The present paper proposes an exper-

---

[4]See e.g. Rabin (1993), Levine (1998), Fehr and Schimdt (1999), or Bolton and Ockenfels (2000).

iment in an attempt to test these hypotheses, and to verify whether fairness preferences are stable over time in the presence of wealth inequalities. As we shall see however, the analysis of relative wealth in experimental economics is non-trivial and presents a number of challenges. The experiment proposed to address these challenges consists in three treatments. Treatment 0 is similar in nature to the traditional ultimatum experiments. In treatment 1 (respectively 2) proposers (respectively receivers) are initially endowed with a large amount of money. Then, they play the same ultimatum game as in treatment 0. The allocation of the endowments by the experimenter is common knowledge and purely arbitrary. It is therefore difficult to justify it morally. In addition, the endowment is substantial enough that it can be assumed to create adequately wealth differences between participants.

In a nutshell, the results are as follows: Subjects initially attempt to reach an egalitarian division of their overall earnings (i.e. the initial endowment plus the revenue from the ultimatum game); with experience, however, rich (poor) proposers make smaller (larger) offers, while rich (poor) receivers are more likely to reject (accept) small offers.[5] In other words, the dynamics of behavior not only contrast with previous ultimatum game experiments, but they are also at odds with both subgame perfection, and the notion that fairness preferences at stable over time. To explain the dynamics of behavior, a structural reinforcement learning model accounting for experience and foregone best responses is estimated. The estimation results suggest that both proposers and receivers learn to adjust their strategies. This result is remarkable since evidence of learning on the part of receivers has rarely been collected in traditional ultimatum games. The estimation results also suggest that the evolution of strategies may be essentially explained by a combination of four factors: first, rich subjects acquire most bargaining power as their opponents lose their ability to punish greedy behavior; second, the rich players' learning process is mainly driven by foregone best responses rather than the actual payoff received; third, unlike their opponents, the monetary expectations of rich subjects

---

[5]To simplify, subjects (not) receiving the large initial endowment will be labelled as "rich" ("poor").

increase with time; fourth, poor players select their strategies based on their past payoffs and actions. In other words, the dynamics of behavior in the experiment was affected by wealth differences, as rich subjects appeared to become more selfish, while poor players learned to tolerate their opponents increasingly aggressive behavior.

The paper is organized as follows: the related literature is summarized in section 2; the design of the experiment is presented in section 3, and discussed in section 4; the experimental outcomes are analyzed in section 5; the general reinforcement model is briefly explained in section 6; the estimation results are commented in section 7; finally, section 8 concludes.

2. Related Literature

To the best of our knowledge, the effect of observed wealth differences on the behavior of subjects in bargaining experiments has never been explicitly analyzed. Equal initial endowments, however, are often implicitly provided to experimental subjects in the form of a participation fees.[6] Such fees usually vary between $3 and $10, which may represent up to 100% of the pie to be divided in the bargaining game. No significant treatment effect has been observed in bargaining experiments when participation fees are equally provided to both parties. In other words, these experiments suggest that behavior in ultimatum games is not affected by traditional absolute wealth effects. It may be argued, however, that the participation fees may have been too modest to represent adequately wealth. Although the object of the present paper is to concentrate on observed wealth differences, it has to be acknowledged that the question of absolute wealth effects has not been fully resolved.

Wealth differences between participants have rarely been introduced in bargaining experiments. Güth et al. (1992) tried to induce behavior conforming to subgame perfection by providing receivers in ultimatum experiments with modest initial endowments.

---

[6]See e.g. Ochs and Roth (1989), Roth et al. (1991), Ruffle (1998), Duffy and Feltovich (1999), or Cooper et al. (2003).

The data suggested that initial endowments appeared to lower slightly offers, but no significant support was provided in favor of either fairness or subgame perfection. However, the authors concluded that the experimental outcomes may be explained by the combination of two factors: first, the roles of proposers and receivers were auctioned prior to the ultimatum game, and second, subjects only played one round. In the present paper, positions are randomly attributed and the experiment consists of 60 periods which provides ample opportunity to learn. Goeree and Holt (2000) use small asymmetric endowments in a two stage bargaining model to differentiate fair from random behavior. They find a prevalence of fair proposals that tend to equalize the overall earnings. Subjects, however, were not given the opportunity to learn. Finally, Eckel and Grossman (1996) analyze a dictator game where proposers face "needy" and/or "deserving" responders such as the American Red Cross.[7] The authors find that altruism is more common than usual, but non-negligible greedy behaviors are still observed. Unlike Eckel and Grossman (1996), the differences in wealth, and the needs of participants are difficult to justify morally in the present study.

The first wave of bargaining experiments yielded conflicting results regarding the importance of learning in ultimatum games. Indeed, Binmore et al. (1985) found an effect of experience, while Güth and Tietz (1988), as well as Neelin et al. (1988), did not find any evidence of learning. More recent studies however, clearly indicate that proposers learn to lower slightly their offers (e.g. Slonim and Roth 1998). On the other hand, with the notable exception of Cooper et al. (2003), very little evidence of learning on the part of receivers has been collected. Erev and Roth (1998) argue that this stylized fact is consistent with the prediction of a choice reinforcement learning model, since accepting or rejecting small offers provide basically the same amount of reinforcement. As we shall see in section 5.1, this observation may also be partially explained by the fact that the receivers' strategies are typically imperfectly observed in traditional ultimatum experiments.

---

[7]The dictator "game" is similar to the ultimatum game except that the receiver has no alternative but to accept the proposal.

3. The Experimental Design

We present in this section the different experimental treatments. The main differences with traditional ultimatum experiments are then discussed in a subsequent section.

The experiment was conducted with volunteers at the State University of New York at Stony Brook. There were three experimental sessions, one for each treatment and each with 60 subjects and 60 rounds. No subject participated in more than one session. Prior to any session, the pool of 60 subjects was equally divided in two groups of proposers and receivers. To preserve anonymity, each group was asked by E-mail to meet in a different laboratory located in separated buildings. A small number of additional subjects were also invited in each group in order to anticipate any potential withdrawal. At the beginning of the session, players were assigned to an isolated computer and informed of their role as receivers or proposers. Subjects remained in the same role for the entire session. Subjects were told in advance how many rounds would be played, and they knew that the experiment would not exceed 2 hours. Instructions were then read aloud, followed by participants' questions, and two dry runs in which the outcome did not count toward the players' final earnings.[8] The analysis of the data begin with the first round involving cash payoffs. At the beginning of each round, players were randomly matched in pairs. To avoid reputation building, the subjects were informed that the assignment was such that it was not possible to identify the other member of the pair, and no pair was identical in two successive rounds.[9]

The design of treatment 0 is a hybrid between the ultimatum game in strategic and extensive forms. As further discussed below, this design enables the experimenter to observe fully the receiver's strategy, while preserving subgame perfection as a relevant equilibrium concept. In each round, each proposer and receiver participated in a two stage game. In stage 1, the proposer makes an offer in cents between 0 and $50 while

---

[8]The complete list of instructions are available at this web address: http://www.sunysb.edu/economics/research/workpap.html.

[9]During the course of the experiment, each proposer was matched twice with each receiver. However, the one-shot nature of the game was essentially preserved since the probability of playing against a given player in a given period was negligible.

the receiver simultaneously announces a minimum acceptable offer (hereafter MAO) corresponding to the smallest offer he is willing to accept. At the beginning of stage 2, the proposer's offer is revealed to the receiver. The receiver is then given the opportunity to revise his strategy by accepting (rejecting) an offer below (above) the MAO he announced in stage 1. To promote truthful announcements in stage 1, the revision is only randomly approved by the computer with probability 0.5. In other words, a receiver could reject (accept) in stage 2 an offer he originally declared he would accept (reject), but the revision had only a fifty percent chance of being implemented.

The money is allocated exactly according to the proposer's offer under two scenarios. In the first scenario, the offer exceeds the MAO and either no revision takes place in stage 2, or a revision to reject the offer is not approved. In the second scenario, a revision to accept the offer is approved. As we shall see, to ensure that the hybrid and traditional ultimatum games have the same (subgame) perfect equilibrium, we have to impose that, although considered accepted, an offer exactly equals to the MAO yields only half of the offer to the receiver, while the proposer receives his entire share of the proposed division (i.e. $50 minus his offer). In such cases, the receiver could still revise his strategy by rejecting the offer in stage 2, in order to try to deprive both players of their payoffs.[10] Finally, in all situations not previously mentioned, neither players receives any money for this round.

After choices are made, both members of the pair are informed of each others actions (offer, MAO and possible revision) and payoffs. In addition, the computer screen displayed the subject's own history of plays for the last five rounds. At the end of the session a round and six pairs of subjects who were matched during that round were drawn randomly. Each selected player was paid in cash the amount of money earned during the selected round. In addition, every subjects received $5 just to participate.

The experimental design was slightly modified in treatments 1 and 2. The object of these treatments is to analyze the effect of known wealth differences between pro-

---

[10]Given the wide range of offers and MAOs available to subjects, a situation in which the proposer's offer equal the receiver's MAO never arose in the subsequent experiments.

posers and receivers. This implies that, prior to the ultimatum game, both players should have positive but unequal levels of wealth. Therefore, the experimental design remained identical to treatment 0, except that widely unequal endowments were now allocated at the beginning of each round. Namely, proposers (receivers) received $50 in treatment 1 (2) while receivers (proposers) were given only $10.[11] The distribution of initial endowments, as well as the rules of the game were common knowledge. The payment method was identical to treatment 0, except that the selected players received their initial endowments for the randomly chosen round in addition to their earnings from the ultimatum game. In addition, every subject received $5 just to participate. Note also that rich players (if selected) received a more than adequate remuneration for their participation in the experiment, even when the ultimatum game resulted in a disagreement. The difference in initial endowments is therefore assumed to represent adequately wealth inequality between players.

Each session lasted approximately an hour and a half. Although no time limit was imposed on subjects, the length of a round (roughly a minute) was constant within and across sessions, suggesting that subjects exercised the same amount of introspection at the beginning and the end of each experiment. On average, the selected proposers (receivers) earned $20.53 ($17.21) in treatment 0, $78.33 ($21.4) in treatment 1 and $24.7 ($72.5) in treatment 2.[12]

4. Comments on the Experimental Design

The design of an ultimatum game experiment to analyze relative wealth effects and learning presents a number of challenges. We now discuss some of the solutions proposed to address these challenges.

The first challenge consists in artificially creating in a laboratory wealth differences between subjects. The solution adopted is to provide half of the subjects with an initial

---

[11] The initial endowment should not be confused with an outside option which is allocated only after a negotiation breakdown (e.g. Binmore et al. 1989 or Kahn and Murnigham 1993).

[12] These amounts include the initial endowments and the ultimatum game outcomes.

endowment sufficiently large to represent adequately wealth inequality. However, the potential profits of rich players in the ultimatum game must remain commensurate to their initial endowment, in order to promote rational-like behavior, and avoid the so called "satiety effect" (see Friedman and Sunder 1994). Moreover, a pilot experiment indicated that providing poor subjects with no initial endowment was perceived as to unfair, which led a substantial number of poor subjects to either refuse to participate in the experiment, or to act in an apparently random manner.[13] It was therefore decided to conduct the experiment with i) unequal but strictly positive endowments for both types of subjects, and ii) a large amount of money to be divided in the ultimatum game.

The second challenge is to design an experiment that may be financed with a reasonable research budget. Indeed, it would be extremely costly to provide every subjects with sufficiently large initial endowments to create artificially wealth differences. To address this problem, a random payment approach was adopted, in which only 1 out of 5 subjects received a payment associated with his performance.[14] This payment method, however, differs from most ultimatum games where subjects are either paid their cumulative payoffs during the entire session (e.g. Bolton and Zwick 1995 or Abbink et al. 2001b), or one round is drawn at random and every subject receives the payoff he earned for this specific round (e.g. Ochs and Roth 1989 or Slonim and Roth 1998). Although the payment mechanisms have not been found to generate a significant treatment effect in ultimatum games, one may still wonder whether the random payment method may affect how subjects behave in the experiment.[15] Indeed, a potential drawback of this method is that subjects may become unmotivated realizing that their decision in each

---

[13]This pilot experiment already provides some information on the effect of observed wealth inequality on behavior in the ultimatum game. However, one should be prudent not to extrapolate unreasonably from this result, as it may be argued that providing poor subjects with no endowment, while their opponents receive $50, is a too extreme treatment.

[14]An alternative approach, adopted for some costly experiments, consists in conducting the experiment in a country with lower standard of living (see e.g. Slonim and Roth 1998). This approach, however, raises the question of potential cultural effects.

[15]Starmer and Sugden (1991), as well as Cubitt et al. (1998) find no treatment effect when a randomly selected period is used to pay subjects. See also Camerer and Hogarth (1999) for an analysis of the effects of incentives on experimental outcomes.

period has in fact little bearing, in expectation, on their final earnings. Although this issue has not been fully resolved in experimental economics, Bolle (1990), Straub and Murnighan (1995), as well as Murnighan and Saxon (1998) found no treatment effect in ultimatum game experiments when adopting a random payment method similar to the one used here.[16] Likewise, we shall see in the next section, that the outcomes observed in treatment 0 are consistent with previous ultimatum game experiments. Note also that subjects were allowed to leave the experiment with the show-up fee of $5 immediately after they were informed of the payment method, or at any point during the experiment. Only two subjects out of 182 decided to do so before the experiment started. Finally, a post-experiment survey, and the statistical analysis of the data in the subsequent section, did not provide any evidence of boredom, lack of motivation, or random behavior on the part of subjects.

The third challenge consists in finding a proper balance between conducting a sufficient number of periods per subject to analyze adequately learning, and preserving the one shot nature of the ultimatum game. To avoid reputation building stemming from rematching the same subjects together, it has been decided to conduct a single session per treatment, consisting of a large number of periods, and including all the subjects recruited for that treatment. The number of rounds played in the present experiment may appear large by traditional standards, but it is not unprecedented (see e.g. Winter and Zamir 1997, Duffy and Feltovich 1999, Cooper et al. 2003). Likewise, it is infrequent but not uncommon to group all subjects in a single session (see e.g. Ochs and Roth 1989 or Bolton and Zwick 1995 for examples in the ultimatum game literature). However, multiple sessions per treatment are preferred in general, as it may be argued that the behavior of subjects interacting within the same session cannot be assumed to be fully independent. Indeed, with a small pool of symmetric players frequently matched together, the actions of a participant may have a lasting effect on the behavior of the

---

[16]Bolle (1990) actually argues that paying experimental subjects large amounts of money with low probability is preferable to paying them small amounts with certainty. Moreover, other recent applications of a similar random payment method include Franck and Schulze (2000), and Fershtman et al. (2000).

other subjects, after they are paired with him, and then with each other. However, the risk of propagation is lower in the present context, as the number of subjects is large, and the same participants are rarely matched with each other. Note also that the players facing each other in the ultimatum game (i.e. the proposers and receivers) are not symmetric. Therefore, the contamination of the proposer's (respectively receiver's) population is more difficult in this context, as it can only be carried out indirectly (e.g. from a proposer, to a receiver, back to a proposer). As we shall see, this intuition is confirmed by an econometric analysis which provides no evidence of a systematic dependency in the data collected in each treatment.

The fourth challenge consists in designing an experiment producing data that enables to analyze adequately learning. The overwhelming majority of ultimatum experiments have been conducted under the extensive form of the game, in which the receivers observe the offer before making a decision. A drawback of this approach is that it creates a data imbalance since, unlike the proposer strategy, the receiver's actions are not observed for all possible offers but only for the actual offer made by the proposer. This data restriction makes it difficult to analyze properly the dynamics of the strategies for each type of players, and it prevents in particular the estimation of structural models such as adjustments or learning models. The implementation of the strategic form of the game enables one to correct the data imbalance. Indeed, both players' strategies are fully observed by the experimenter, as offers and MAOs are stated simultaneously. However, the timing of the ultimatum game is modified from a two stage game to a one-shot simultaneous move game. As a consequence, subgame perfection is lost as a solution concept. In addition, Roth (1995) argues that the strategy method may generate different behaviors, as receivers may regret their choice when they observe the actual proposer's offer. In fact, a recent comprehensive study of thirty two published papers on ultimatum games, indicates that the strategic implementation does not affect the actions of proposers, but it influences how responders behave (see Oosterbeck et al. 2001). This study therefore reinforces the need to use the hybrid design, as it appears to offer the advantages of both the strategic and extensive approaches, without some of the obvious drawbacks. Indeed,

12

both players strategies should be fully observable, since subjects have a strict incentive to make truthful announcements in stage 1. In particular, the receivers must reveal his strategy in stage 1 in order to avoid the potential cost associated with the revision lottery in stage 2. However, receivers also have an incentive to correct any mistake or regret in stage 2.[17] In addition, subgame perfection remains a relevant equilibrium concept in this two stage game. Indeed, if subjects are self interested, then receivers have a strictly dominant strategy consisting in announcing a MAO strictly smaller than the offer. In addition, an offer equal to the MAO is strictly dominant for proposers. Therefore, the (subgame) perfect equilibrium consists for proposers to offer the smallest possible share of the pie in stage 1, and for receivers to always accept this offer by setting a MAO of zero in stage 1, and by accepting any strictly positive offer in stage 2.[18] In other words, the hybrid model has the same (subgame) perfect equilibrium, and the same equilibrium payoffs as the ultimatum game.

5. Experimental Results

The experimental results are summarized in Graphs 0.1 to 2.6 and Tables 1 to 4.[19] For the analysis to be consistent with previous studies (in which the pie is often $10) offers and MAOs have been divided by 5. Therefore, the players' strategies may also be interpreted as deciles of the total pie.

[Table 1 Here]

Before describing the experimental outcomes, we conduct a statistical analysis of the data collected in each treatment to verify whether i) conducting a unique session per

---

[17] As we shall see in section 5.1, however, receivers appear to reveal their true preferences in stage 1, since they rarely exercised their right to revise their strategy in stage 2.

[18] A proposer may not want to take the risk of making a $0 offer, since it may be rejected in stage 2 by the receiver who is indifferent in this situation between accepting and rejecting the offer.

[19] Graphs 0.1 to 0.6 are associated with treatment 0. Graphs 1.1 to 1.6 (2.1 to 2.6) are associated with treatment 1 (2).

13

treatment generated a significant correlation between subjects' actions; ii) the random payment method, and/or the large initial endowments led some subjects to behave in a random manner. To capture potential systemic effects across individuals, the regressions conducted in this section will rely on a random-effect specification. This reduced-form approach, enables to exploit the panel structure of the data to account for possible correlations between the strategies selected by a subject across time.

The strategy s_i^t selected by subject i in period t is modelled as a function of the subject's past plays, as well as the past plays of his previous opponents:

$$s_i^t = \delta_1 s_i^{t-1} + \delta_2 s_i^{t-2} + \delta_3 s_j^{t-1} + \delta_4 s_j^{t-2} + \delta_5 s_{j'}^{t-2} + \delta_6 s_{i'}^{t-2} + \nu_i + \varepsilon_{it} \qquad , \qquad (5.1)$$

where $\left(s_i^{t-1}, s_i^{t-2}\right)$ are the strategies of subject i in the two previous periods; $\left(s_j^{t-1}, s_j^{t-2}\right)$ are the strategies in the two previous periods of subject j, where subject j was the opponent of subject i in period $t-1$; $s_{j'}^{t-2}$ is the strategy in period $t-2$ of subject j', where subject j' was the opponent of subject i in period $t-2$; $s_{i'}^{t-2}$ is the strategy in period $t-2$ of subject i', where subject i' was the opponent of subject j in period $t-2$; ν_i is the individual random-effect parameter; and finally, ε_{it} is an error term. In other words, the econometric model accounts for all the strategies in the history tree of player i, backing up two periods. Note in particular that subject i and i' are of the same type (i.e. both proposers, or both receivers). In other words, δ_6 significantly different from zero will indicate that a subject's actions contaminate his entire population. Moreover, $\{\delta_k = 0, \forall k = 1, ..., 6\}$ will indicate that subjects acted randomly, and did not adjust their strategy over time.

To gain efficiency, the model is estimated with the general method of moment approach developed by Blundel and Bond (1998). Table 1 reports the regressions results for the sample of proposers and receivers in each treatment.[20] For both types of players in each treatment, the parameters associated with the actions played beyond the previ-

---

[20]For the sake of brevity, the parameters associated with the random-effect are not reported as they have no economic interpretation. Note, however, that the random-effect term is always significantly different from zero at a 5% level, which provides support for the random effect-specification adopted.

ous period, and in particular δ_6, are not significantly different from zero. In other words, it appears that a subject's current choice depends only on the previous period actions, and in particular, it is not directly affected by decisions taken by other subjects in his own population. To confirm this result, the regression in (5.1) has been estimated after excluding the previous period actions (i.e. imposing $\delta_1 = \delta_3 = 0$). The results in Table 1 indicate that in this case, a player's strategy in period t can only be explained by his own action in period $t-2$. In other words, we cannot detect any dependency within the samples of proposers and receivers, and the correlation between proposers and receivers actions does not appear to extend beyond the previous period.

Moreover, to test for random-like behavior (i.e. $\delta_k = 0$, $\forall k = 1, ..., 6$) we apply the extension to the general method of moment framework of the Wald test to each sample of proposers and receivers in each treatment (see e.g. Newey and West 1987). The Wald statistics vary between 45.694 and 59.464, which correspond to P-values ranging from 3.406E-8 to 5.782E-11. In other words, there is no indication that subjects, and in particular rich proposers in treatment 1 and rich receivers in treatment 2, acted randomly during the experiment. To verify whether subjects became un-motivated by the end of the experiment, we test the same hypothesis on the samples collected during the last ten periods of each treatment. The Wald statistics now vary between 38.852 and 53.264, which correspond to P-values ranging from 6.751E-7 to 1.038E-9. In other words, subjects did not exhibit any apparent random-like behavior, even by the end of the experiment. To summarize, we find no evidence that the experimental design created a significant dependency in the data, or generated random behavior on the part of subjects.

5.1. Treatment 0

The sole objective when analyzing the experimental outcomes in treatment 0 is to verify whether the introduction of the design modifications outlined in the previous section created a significant treatment effect compared to traditional ultimatum games. Several approaches may be considered to address this issue. The most thorough approach would

consist in conducting a control treatment with the traditional ultimatum game design, and to compare it with alternative treatments in which the design modifications would be individually or jointly incorporated. The indispensability of a control treatment may be questioned, since the main behavioral characteristics of the traditional ultimatum game have been well established in the multiple experiments conducted over the years. Moreover, we are in fact only interested in knowing whether the combination of all design modifications produced a significant effect on behavior. Therefore, we adopted an alternative approach consisting in comparing the outcomes in a treatment including all the design modifications, with those observed traditionally in ultimatum game experiments.

The experimental outcomes in treatment 0 are summarized in Table 2. Pooling data over all periods, the average offer is 4.38 while the average MAO is slightly below with a mean of 3.42 (see Table 2). These actions resulted in an overall rejection rate of 23%.[21] There are few offers above 6 (3.94%) and below 2 (2.06%). Graph 0.4 shows that offers and MAOs slightly decrease over time, but they mostly remain within a narrow interval (see also the comparison of the first and last 10 periods in Table 2). To confirm this observation, let us run the following random-effect regression separately for each sample of proposers and responders:[22]

$$s_i^t = \delta_1 t + \delta_2 s_i^{t-1} + \delta_3 s_j^{t-1} + \nu_i + \varepsilon_{it} \qquad . \qquad (5.2)$$

The regressions results in Table 3 indicate that the time trend parameters in treatment 0 are small, but significantly smaller than 0 in both samples. This suggests that both proposers and receivers slowly adjust their strategies over time by announcing smaller offers and MAOs. As noted previously, the finding of adjustments on the part of receivers is remarkable since it has rarely been detected in ultimatum experiments, even when subjects participated for a large number of rounds. This result should be essen-

---

[21] The overall rejection is rather high, but it is consistent with previous studies (see e.g. Roth et al. 1991, Abbink et al. 2001b, Slembeck 1999).

[22] This specification has been preferred over alternatives on the basis of a test of the over-identifying restrictions.

tially credited to the strategic implementation of the ultimatum game, as it allows the experimenter to observe fully the strategy of receivers.

[Tables 2 and 3 Here]

Graph 0.1 indicates that the MAOs are symmetrically distributed around their mode located in the range 3-4. The offers have a mode of comparable magnitude in the interval 4-5, but they are essentially concentrated between 2 and 6. The analysis of the first and last 10 periods (Graphs 0.2 and 0.3) confirm the slight decrease of strategies over time. However, the general shapes of the strategy distributions are essentially preserved.

The evolution of the rejection rate is volatile, and a series of regressions confirms that it does not exhibit any obvious trend (see Graph 0.5). Graph 0.1 also indicates that the rate of rejection decreases rapidly with the offer made.[23] For instance, offers above 4 have a 91.2% chance of being accepted, while offers below 2.5 are rejected 85.2% of the time. Following Slembeck (1999), let us define a measure of the proposers' bargaining power as the income difference (in percentage of the pie) between proposers and receivers. Graph 0.5 indicates that this measure of the proposers bargaining power is slightly positive, and remains roughly stable over time.

Finally, receivers revise their strategy in stage 2 of the game only 5.61% of the time. The curve in Graph 0.6 indicates that most of these revisions (54%) occur within the first 10 periods of plays. The number of revisions as a function of the difference (in dollars) between the offer and the MAO announced in stage 1 is also plotted in Graph 0.6.[24] This bar graph is concentrated around zero, indicating that most receivers revise their strategy when the proposer's offer is close to the MAO originally announced. Notice,

---

[23]The rejection rate for low and high offers should be interpreted with caution through the paper because they are often based on few observations.

[24]This graph should be read as follows: the highest bar indicates that 25 receivers decided to revise their strategy in order to accept an offer to divide the $50 they initially declined, that was between $0 and $1 smaller than the MAO they originally announced. Note also that unlike the other graphs and tables, Graphs 0.6, 1.6, and 2.6 have not been constructed with the offers and MAOs divided by 5, but with the actual dollar values.

however, that the revisions are not symmetrically distributed around zero. Indeed, the five bars to the right, representing the rejections of offers originally accepted, account only for 1/4 of the total number of revisions. In contrast, 58.4% of the revisions consist of receivers deciding to accept an offer that was between $0 and $5 smaller than the MAO announced in stage 1. In other words, after a few periods of practice receivers learn to make few and only small mistakes in stage 1, and they appear to have very few regrets in stage 2.[25]

To conclude, the outcomes appear to be consistent, both qualitatively and quantitatively, with previous ultimatum game experiments. More specifically, the distribution of offers, the probability of rejection, the distribution of payoffs, and the evolution of behavior is analogous to what is typically observed in traditional ultimatum games. In addition, the low rate and the distribution of strategy revisions suggest that, with experience, the extensive and strategic implementations of the game generate similar behavior from experimental subjects. In other words, not only the experimental design modifications did not change the theoretic predictions, but they also did not appear to create any significant treatment effect compared to the traditional ultimatum game experiment. Nevertheless, it cannot be excluded that in a different context the two designs may produce different behaviors. To test whether the modified design is in general behaviorally equivalent to the traditional ultimatum game is beyond the scope of the present paper. What is truly important for the purpose of our study is that we have established that, like in an ultimatum game, subjects in treatment 0 appear to express a concern for fairness that is sustained over time.

5.2. Treatment 1: Rich Proposer

The overall average offer (4.29) and MAO (3.11) are slightly smaller than in treatment 0. These figures, however, only tell part of the story. Indeed, Table 2 indicates that,

---

[25]The tables, graphs, and econometric estimations in this section are based on the MAO announced, and they do not account for possible strategy revisions by receivers in stage 2. However, given the small number and the distribution of the revisions, the tables, graphs and econometric estimations differ only slightly, and the nature of the results presented is preserved when the revisions are taken into account.

compared to treatment 0, both players strategies are higher (lower) in the first (last) 10 periods. As illustrated in Graph 1.4, and confirmed by regressions of the form (5.2) (see Table 3), strategies now sharply decrease during the course of the experiment. MAOs still have a mode in the usual range 3-4 but they are now much more uniformly distributed (Graph 1.1). Note that a substantial number of receivers initially make large demands (i.e. between 4 and 6 in Graph 1.2). By the end of the session however, most receivers are willing to accept very small offers (i.e. lower than 2 in Graph 1.3). The mode of the offers is still in the interval 4-5 but larger (smaller) offers are now much more frequent in the early (late) periods of the game (see Graphs 1.1 to 1.3).

The rejection rate is slightly larger than in treatment 0, but it still does not reveal any specific trend (Graph 1.5). Graph 1.1 indicates that small offers are much more likely to be accepted than in treatment 0. For instance, offers below 2.5 are now accepted almost one third of the time. Graph 1.5 also indicates that the bargaining power is initially on the receiver's side, but it is increasingly captured by proposers over time. Finally, the number of revisions made by receivers increases slightly compared to treatment 0, but its distribution over time and offers remains similar (Graph 1.6).

To confirm the presence of a treatment effect the following regression was conducted after merging the two samples of proposers (respectively receivers) in treatments 0 and 1:

$$s_i^t = \delta_1 s_i^{t-1} + \delta_2 s_j^{t-1} + \delta_3 Early\_T0_i^t + \delta_4 Late\_T0_i^t + \nu_i + \varepsilon_{it} \qquad , \qquad (5.3)$$

where $Early\_T0_i^t$ (respectively $Late\_T0_i^t$) is a dummy variable equal to 1 for data collected in treatment 0 within the first (respectively last) 10 periods. The results presented in Table 4 indicate that the parameters associated with $Early\_T0_i^t$ (respectively $Late\_T0_i^t$) are significantly smaller (respectively larger) than zero for both proposers and receivers. In other words, compared to treatment 0, rich proposers (poor receivers) are initially asking less (more) for themselves, but with time, they end up making larger (smaller) demands.

5.3. Treatment 2: Rich Receiver

The results in treatment 2 are almost the opposite image of those in treatment 1. Compared to treatment 0, offers and MAOs are initially smaller but they rapidly increase over time (Table 3 and Graph 2.4). The distribution of offers is now bi-modal (Graph 2.1) with a first mode in the range of 3-4, and a second mode in the range of 6-7. The first (second) mode is the result of low (high) offers in early (late) periods (see Graphs 2.2 and 2.3). Proposers are now willing to make offers above 6 more than one third of the time overall, and a surprising 54% of the time during the last 10 periods (Table 2). This is a striking difference with treatment 0, and behavior observed in traditional ultimatum game experiments. Note that the distribution of offers in the last 10 periods is also bi-modal with a first mode in the interval 6-7, and a smaller mode in the range 3-4. An inspection of individual data reveals that subjects cannot be divided in two sub-samples of "greedy" and "generous" proposers. Instead, it appears that the same players alternate between small and larger offers with a clear predominance of large offers. The distribution of MAOs still has its usual mode in the range 3-4, but receivers become much more demanding with experience (see Graphs 2.1 to 2.3). As an illustration, the MAOs of 6 and above increase to 26.1% in the last 10 periods.

The evolution of the rejection rate remains as volatile as in the previous two treatments (Graph 2.5). Graph 2.1, however, indicates that small offers are rejected more often, while large offers are not always accepted. For instance, offers above 4 are now only rejected 15.1% of the time. The initial bargaining power of proposers shifts rapidly in the receivers' hands, in sharp contrast with traditional ultimatum games experiments (Graph 2.5). Moreover, the number and the distribution of the revisions made by receivers are once again comparable to treatment 0. Finally, the results of the regression (5.3) presented in Table 4 confirm the presence of a treatment effect. Indeed, compared to treatment 0, the offers and MAOs are significantly smaller in the early periods, but they become substantially larger by the end of the experiment.

5.4. Interpretation of the results

It is usually accepted that subjects behavior when they begin an experiment is predominantly influenced with real life experiences with comparable games. In real life bargaining situations, agents are typically not anonymous, the game is often repeated, the outcomes are frequently observed by others, and traditional norms as well as conventions apply. It has been demonstrated in several experiments that the combination of these factors is likely to generate fair or egalitarian outcomes in ultimatum games. This may therefore explain the wealth egalitarian selection of strategies observed in the first periods of the different treatments.

On the other hand, the evolution of behavior is at odds with subgame perfection, and existing fairness models. Indeed, according to subgame perfection, wealth differences should not affect the players' decision process. In other words, the behavior of participants should remain the same across the three treatments. This prediction is strongly rejected by the data, since we have just established that the introduction of large and unequal initial endowments creates a significant treatment effect on the dynamics of plays. It has to be noted however, that behavior in treatment 1 is leaning toward the game theoretic prediction. This result is remarkable, as it has been notoriously difficult to induce subjects in ultimatum game experiments to conform to subgame perfection (see e.g. Weg and Smith 1993).

The predictions of inequality aversion and altruistic models (see e.g. Fehr and Schimdt 1999 and Bolton and Ockenfels 2000) depend on how wealth differences are perceived by agents. If the arbitrary distribution of initial endowments is not seen as unfair by subjects, then behavior should remain constant across treatments. Otherwise, these models predict that the differences in initial endowments will result in more generous behaviors on the part of the rich players. More specifically, rich (poor) proposers should make larger (smaller) offers, and rich (poor) receivers should (not) be willing to accept small proposals. Such a wealth egalitarian approach appears to be a good predictor of initial behavior. Indeed, proposers tend to offer some amounts that would

equalize, or at least reduce the difference between both players total earnings. The distribution of MAOs in the first periods of plays suggests that the wealth egalitarian outcome is also initially expected by the receivers. With time however, the experimental outcomes in treatments 1 and 2 become at odds with the fairness predictions. Indeed, rich players ask more for themselves, while poor subjects are willing to accept widely unequal divisions.

The reciprocity approach (see e.g. Rabin 1993) roughly assumes that agents are willing to punish hostile behavior, but they are inclined to reward other agents who intend to reward them. Such predictions are not consistent with the dynamics of behavior observed in the present experiment. Indeed, poor receivers are unfairly treated after a limited number of periods, but nevertheless, they accept to lower their demands over time. In addition, reciprocity suggests that revisions to reject an offer in stage 2 should be more frequent after the receiver observes an unequal proposal. Such behavior however, is rarely observed since most revisions consist in accepting an offer initially declined. In addition, revisions essentially occur at the beginning of the session, when offers are arguably more equitable.

It has to be noted however that the inequality aversion, altruistic, and reciprocity models have not been developed to predict how behavior may evolve, since they implicitly assumed that fairness preferences are stable over time. Therefore, failure to explain the dynamics of behavior in the experiment should not be seen as a definitive rejection of fairness models. Indeed, fairness still appears to play a significant role, since offers and MAOs remain non-negligible, even at the end of each treatment. The present experiment may therefore only suggest that in certain environments, fairness preferences may be quite unstable over time, and that existing fairness models may need to be generalized to account for learning and/or other dynamic aspects. To do so however, one would first need to identify the determinants affecting the stability of fairness considerations. In the next section, we will estimate a learning model in an attempt to address this issue.

Before closing this section, let us mention a number elements revealed through post-experiment interviews, that provide an insight into the subjects' motivations when se-

lecting their strategy. Several authors have suggested that the outcomes in ultimatum experiments may be explained by the combination of three factors: a fear of disagreement on the proposers' side, an ability for receivers to punish greedy behavior, and a desire for receivers to prove their "toughness" by rejecting free money. Interviews conducted after the experiments suggest that these factors may have some explanatory power in the present context. Indeed, the majority of poor receivers declared that they could not punish proposers since rejections felt relatively more costly to themselves. Rich receivers, on the other hand, said that their initial endowments led them to punish proposers by rejecting free money if their initial demands were not met. Some even added that, although they understood that the next subject they would be matched with would be different, they rejected positive and sometime large amount of money to demonstrate their determination to the proposer. A large number of rich proposers reported that they did not fear rejection, since they felt satisfied with only the initial endowment even when the ultimatum game resulted in a disagreement. As a consequence, they were willing to take some risks by making small offers. On the other hand, poor proposers felt like they lost their first mover advantage and that they could only acquiesce to the receivers requests.

It is important to note that these interviews do not explain the evolution of strategies in the experiment. For instance, were poor receivers willing to accept less money, and proposers reacted accordingly? Or were rich proposers making smaller offers, which constrained receivers to lower their demands? These questions are addressed in the next section by estimating a structural learning model.

6. General Reinforcement Learning Model

6.1. Preliminaries

The object of this section is to develop a learning model to estimate how behavior evolved in the experiment. A potential drawback of this structural approach however, is that subjects are assumed to behave according to the learning model considered.

As we shall see, several competing models of learning have been proposed in recent years. The strategy adopted therefore consists in estimating a comprehensive structural model, capturing most relevant features of these learning models, in order to identify the determinants explaining the dynamic of plays in each treatment.

To analyze learning, we must also decide whether to use a model for continuous or discrete strategies. In the experiment, players had access to 5,000 different strategies, since they could make choices in cents between 0 and $50. Focal point strategies, such as integers, half dollars and offers ending in $.99, were used more frequently but they accounted only for 16.8% of subjects overall choices. Within the range of strategies selected most often (roughly between $10 and $35 for proposers, and between $5 and $30 for receivers), proposers (receivers) used 1,456 (1,724) different strategies representing approximately 58% (69%) of the strategies available to them. A discrete model would require to regroup arbitrarily the 5,000 possible strategies into a limited number of clusters. However, Monte Carlo simulations suggest that such arbitrary discretizations may erase some of the subtle nuances associated with the smooth evolution of behavior observed during the experiment.[26] A continuous model therefore appears more appropriate to analyze precisely the dynamic of plays.

We adopt the general reinforcement learning approach developed by Armantier (2003). This model not only explicitly accounts for continuous strategies, but it has also been shown to capture most relevant aspects of existing learning models. In addition, the general reinforcement learning model is sufficiently flexible, yet parsimonious, to be considered for a structural estimation. We now summarize the main features of the model.

6.2. The model

The general reinforcement learning model combines L different observational and experiential reinforcement rules, such as for instance choice reinforcement learning (Bush

---

[26]This result is consistent with simulations conducted for different games, in which continuous learning models were found to outperform discrete models to explain the evolution of strategies (see e.g. Armantier 2003 for an auction example).

and Mosteller 1955, Erev and Roth 1998), imitation (Vega-Redondo 1997, Schlag 1999), learning direction theory (Selten and Buchta 1999), or reinforcement of unchosen strategies (Camerer and Ho 1999). At round t, player i draws her strategy s_i^t from a continuous probability density function $g_i^t(.)$ defined over the interval $[\underline{s}, \overline{s}] = [0, 50]$. The strategy distribution of player i evolves according to the following law of motion,

$$g_i^{t+1}(s) = \frac{g_i^1(s) + \sum\limits_{k=1}^{t} \sum\limits_{l=1}^{L} R(s \mid m_{i,k}^l, r_{i,k}^l, t, \beta^l)}{1 + \sum\limits_{k=1}^{t} \sum\limits_{l=1}^{L} \int\limits_{\underline{s}}^{\overline{s}} R(u \mid m_{i,k}^l, r_{i,k}^l, t, \beta^l) du} \quad . \tag{6.1}$$

where t is the last period played; $R(. \mid .)$ denotes the reinforcement rule; $s \in [\underline{s}, \overline{s}]$; $m_{i,k}^l$ is the strategy reinforced in rule l by player i at any period $k \leq t$ based on a reinforcement factor $r_{i,k}^l$; finally, β^l is the vector of parameters of the reinforcement rule l ($l = 1, ..., L$). The function $g_i^1(.)$, commonly known as the initial propensity, may reflect players introspection or experience from previous games. How subjects select their initial strategy distribution is a question that has been briefly discussed in section 5.4. In the remainder $g_i^1(.)$ is a given distribution, possibly a function of parameters to be estimated.

The reinforcement rule $R(. \mid .)$ adopted here is based on a normal probability density function $f(. \mid \mu_i^k, \sigma_i^k(t))$ with mean μ_i^k and variance $\sigma_i^k(t)$,

$$R(s \mid m_{i,k}^l, r_{i,k}^l, t, \beta^l) = \left\{ \begin{array}{ll} f(\mu_i^k \mid \mu_i^k, \sigma_i^k(t)) - f(s \mid \mu_i^k, \sigma_i^k(t)) & \text{when } r_{i,k}^l < 0 \\ f(s \mid \mu_i^k, \sigma_i^k(t)) & \text{when } r_{i,k}^l > 0 \\ 1 & \text{when } r_{i,k}^l = 0 \end{array} \right\} \tag{6.2}$$

$$\text{where} \quad \mu_i^k = m_{i,k}^l \quad \text{and} \quad \sigma_i^k(t) = \left(\alpha^l\right)^{(k-t-1)} \left(r_{i,k}^l\right)^{-2} \quad . \tag{6.3}$$

In this simple model $\beta^l = \alpha^l$, where $0 < \underline{\alpha} \leq \alpha^l \leq \overline{\alpha} < 1$. Note that the negative reinforcement rule in expression (6.2) is always positive, which guarantees that $g_i^{t+1}(.)$ is a density function.

Heuristically, the purpose of the reinforcement rule $R(.)$ is to add some mass to the strategy density $g_i^{t+1}(.)$ around or away from the strategy reinforced, $m_{i,k}^l$, depending on the reinforcement factor $r_{i,k}^l$.[27] For instance, if the reinforcement factor is positive (negative), then the reinforcement rule has the shape of a (an inverse) normal density function centered on $m_{i,k}^l$; as a result, player i is more (less) likely to play strategies around $m_{i,k}^l$ in the future. As the reinforcement factor $r_{i,k}^l$ increases (decreases), the positive (negative) reinforcement rule becomes more concentrated, and more mass is added (subtracted) around $m_{i,k}^l$. There is no experiential learning when the reinforcement factor is zero, as the strategy distribution stays unchanged. Finally, $\sigma_i^k(t)$ becomes larger as k moves further away from the current period t, which implies that the reinforcement rule $R(.)$ becomes flatter and less influential with time. In other words, the exponent of α^l acts as a discount or forgetting parameter that reduces the influence of past reinforcements.

To illustrate how this general reinforcement learning model captures different forms of learning, consider the traditional choice reinforcement learning model as presented by Erev and Roth (1998). Players reinforce every strategy s_i^k ($\forall k \leq t$) they played in the past based on x_i^k the payoff received. In other words, the choice reinforcement learning rule reinforces at period t each strategy $m_{i,k}^l = m_{i,k}^{CRL} = s_i^k$ based on the reinforcement factors $r_{i,k}^l = r_{i,k}^{CRL} = x_i^k$ for any $k \leq t$. Similarly, the general reinforcement learning model may also nest the reinforcement of foregone best responses. In this case, player i partially reinforces the best response in period k to her opponents actual strategy profile s_{-i}^k (i.e. $BR(s_{-i}^k)$), based on the profit Π_{BR}^k this best response would have generated. The parameters of the rule reinforcing the foregone best strategies from periods $k \leq t$ are $m_{i,k}^l = m_{i,k}^{RBR} = BR(s_{-i}^k)$ and $r_{i,k}^l = r_{i,k}^{RBR} = \theta^{RBR}\Pi_{BR}^k$, where $\theta^{RBR} \geq 0$ represents the relative effect of foregone payoffs compared to actual payoffs. Other possible reinforcement rules include (e.g.) imitation of popular or successful behavior, payoff dependent imitation, direction learning, as well as other forms of exogenous adjustments (see Armantier 2003).

---

[27]See Armantier (2003) for a more detailed explanation of the model's properties.

Erev and Roth (1998) introduced a benchmark, typically representing a "subjective expected payoff", to classify payoffs as positive or negative reinforcements. Following this approach, let us respectively redefine the reinforcement factors in the previous two examples as $r_{i,k}^{CRL} = x_i^k - \rho_i^t$ and $r_{i,k}^{RBR} = \theta \Pi_{BR}^k - \rho_i^t$, where ρ_i^t is the reference point in period t. This internal reference point is assumed to follow a very general law of motion of the form

$$\rho_i^{t+1} = \gamma_0 + \gamma_1 x_i^t + (1 - \gamma_2) \rho_i^t \quad , \tag{6.4}$$

where $(\rho_i^1, \gamma_0) \in \mathbb{R}^2$, $0 \leq \gamma_j \leq 1$ ($\forall j = 1, 2$), and x_i^t is the payoff of player i at period t.

7. Estimation of the General Reinforcement Learning Model

7.1. Estimation Procedure

From an econometric perspective learning models have two interesting features. First, the actions observed during the experiment are neither identically nor independently distributed, since the strategy distributions are updated individually based on previous periods plays and outcomes. Second, learning models may converge toward a pure strategy equilibrium, in which case the asymptotic strategy distribution $g_i^\infty(.)$ is degenerate. These two characteristics are such that the analysis of learning model estimators is nontrivial and rarely addressed in the literature, with the notable exception of Cabrales and Garcia-Fontes (1999) and Armantier (2000). The Maximum Likelihood is a popular method to estimate learning models with a finite number of strategies (e.g. Camerer and Ho 1999, or McKelvey and Palfrey 1995). However, as noted by Stahl (1996) and Armantier (2000), the application of the Maximum Likelihood to continuous strategies may be hazardous. Indeed, as behavior converges toward an equilibrium, the strategy distribution becomes more concentrated, and later observations are much more influential on the likelihood function. This imbalance may result in a disproportionate contribution of later observations on the parameter estimate.

To circumvent this convergence problem Armantier (2000) proposes a moment esti-

mator $\widehat{\beta}$ corresponding to an objective function of the form:

$$SOM_2(\beta) = \sum_{t=1}^{T} \sum_{i=1}^{N} \left\{ \left([s_i^t] - \eta_{i,t}^1(\beta) \right)^2 + \left([s_i^t]^2 - \eta_{i,t}^2(\beta) \right)^2 \right\} \quad , \qquad (7.1)$$

where β is the parameter vector to be estimated, and $\eta_{i,t}^p(\beta)$ is the theoretic moment of order $p = 1, 2$ of player i's strategy distribution at period t.[28] Heuristically, the objective is to reconcile observations with their theoretical moments conditionally on the history of plays. Unlike the Maximum Likelihood, this method offers the key advantage of allocating the same weight in the objective function to any observation even when strategies converge.

The expression (7.1) requires the derivation of the theoretical moments at each period and for all bidders. The strategy distributions do not have tractable analytical forms, and the theoretical moments $\eta_{i,t}^p(\beta)$ are replaced with arbitrary precision by Monte Carlo simulation estimates, $\widehat{\eta}_{i,t}^p(\beta)$. In other words, the estimation method may be interpreted as a traditional method of simulated moments.

Nested learning models are typically compared on the basis of likelihood ratio tests (e.g. Stahl 2000 or Camerer and Ho 1999). As previously mentioned however, this approach is not well suited here due to the inadequate properties of the likelihood function. Instead, we will consider traditional Wald tests based on the unconstrained optimization of the objective function (7.1). The Wald test possesses the appropriate asymptotic properties, since the moment estimator is consistent and asymptotically normally distributed. The covariance matrix involved in the determination of the test statistic is evaluated with a Bootstrap technique based on the estimated parameter $\widehat{\beta}$.[29]

To conclude, it is important to note that the structural estimation procedure i) relies on individual observations, and ii) fully accounts for the interdependency in the data

---

[28]Higher moments may be included in the objective function. Monte Carlo simulations, however, suggest that the first two moments are sufficient in general to obtain precise estimates.

[29]The Bootstrap is a statistical technique consisting in repeatedly resampling the original data from the estimated distribution in order to make inferences from the resamples on parameters such as the standard deviation of the estimated parameters. For details on the Bootstrap technique see Shao and Tu (1995).

under the learning models assumptions.

7.2. Estimation Results

We first estimate a benchmark model including the traditional choice reinforcement learning, the reinforcement of foregone best responses, as well as a dynamic reference point. Alternative learning and adjustment models will be compared and tested against this benchmark in the next section.

Before we proceed, let us discuss how foregone best responses are approximated in the ultimatum game. At the end of each period, the foregone best response of a proposer consists in the MAO announced by the receiver with whom he was matched. In the rare cases in which a receiver decides to revise his strategy, the proposer's foregone best response is approximated by his own offer. This approximation has no significant consequence on the estimated parameters, since revisions are rare and typically occur when offers and MAOs are close. If the responders only maximize their own profits, then they always have a (weakly) dominant strategy consisting in announcing a MAO of zero. As we shall see later on, the experimental data clearly reject any attraction toward the subgame perfect equilibrium strategy. Instead, we assume that the foregone best response of a receiver may be represented by the proposer's offer, since it is the highest MAO that would have resulted in an agreement. This assumption appears reasonable for a group of boundedly rational subjects repeatedly playing the ultimatum game, since it is the best compromise for receivers between reaching an agreement and showing their determination.[30]

The positions of the proposers and receivers are not symmetric in the ultimatum game, and their learning process may therefore differ. Likewise, it has been shown that the environment in which a game is played may affect the way subjects learn. Consequently, we first estimate a general learning model with a specific set of parameters for proposers and receivers in each of the three treatments. The vector of parameters to

---

[30]Abbink et al. (2001a) show that receivers often reject positive offers in anonymous ultimatum games in order to establish a group reputation for being "tough".

estimate for each treatment is $\beta_j = \left(\beta_j^{CRL}, \beta_j^{RBR}, \beta_j^{RP}, \mu_j, \sigma_j^2\right)$ where $j \in \{p, r\}$ represents the player's position as proposer or receiver; $\beta_j^{CRL} = \alpha_j^{CRL}$ is the parameter of the choice reinforcement model; $\beta_j^{RBR} = (\alpha_j^{RBR}, \theta_j^{RBR})$ is the parameter associated with the reinforcement of best responses; $\beta_j^{RP} = (\rho_j^1, \gamma_{0,j}, \gamma_{1,j}, \gamma_{2,j})$ is the vector of coefficients of the reference point law of motion; and $\left(\mu_j, \sigma_j^2\right)$ are the parameters of the initial strategy distribution function. Proposers (receivers) are initially considered symmetric, and strategies in period 1 are assumed to be generated from a normal distribution truncated on $[0, 50]$ with parameters $\left(\mu_j, \sigma_j^2\right)$ $j \in \{p, r\}$. The histograms of both proposers and receivers strategies in the first period of play tend to support this assumption.

Results of the moment estimation method introduced in section 7.1 are presented in Table 5. As expected, the first moment of the proposers and receivers initial propensity distributions (which is slightly different from μ_j due to the truncation) are significantly larger (smaller) in treatment 1 (2) than in treatment 0.[31] This result confirms that subjects initially attempt to reach an egalitarian agreement that equalizes final wealth.

[Table 5 Here]

Let us examine first the parameters associated with the choice reinforcement learning model. The forgetting parameters $\widehat{\alpha}_j^{CRL}$, $j \in \{p, r\}$, are close to, but significantly smaller than one in all three treatments. This confirms that both types of players learn to adjust their strategies in all three treatments as the result of their own past plays and outcomes. The values of the forgetting parameters also imply that experiential learning occurs at a slow pace, as past and recent experience have almost the same influence on the current strategy selection. The forgetting parameter is, however, significantly larger for receivers than for proposers, and it is also significantly larger when subjects are rich. Therefore, the behavior of proposers and poor subjects can potentially change somewhat more abruptly as the result of new experiences. The fact that proposers and receivers

---

[31] The statistical comparisons in this section are conducted with Wald tests at a 5% significance level. Given the large numbers of comparisons, only the most relevant test statistics and p-values are reported.

both appear to learn from their own past plays and outcomes is consistent with previous studies (see e.g. Slonim and Roth 1998, as well as Cooper et al. 2003).

Let us now turn to the estimation of the parameters associated with the reinforcement of best strategies. The pair $\left(\widehat{\alpha}_j^{RBR}, \theta_j^{RBR}\right)$, $j \in \{p, r\}$, is only significantly different from zero for rich subjects, and for proposers in treatment 0. In other words, the foregone best responses act as an attractor when subjects possess some bargaining power provided by their wealth status and/or their first mover advantage. In addition, foregone best responses have a more immediate impact on the strategy selection of rich proposers, since the forgetting parameter $\widehat{\alpha}_p^{RBR}$ is significantly smaller in treatment 1. This result may be explained by the combination of wealth status and first mover advantage. Finally, the parameter θ_j^{RBR} is significantly larger than 1 for rich subjects. This indicates that rich players reinforce their foregone best response strategy with even more intensity than the strategy they actually played. Note that the magnitude of this effect is remarkably larger than in previous studies such as Camerer and Ho (1999) or Armantier (2003). This difference may be partially explained by the additional bargaining power provided by large initial endowments.

Finally, let us examine the parameters governing the evolution of the reference point. Except for the receivers in treatment 0, most parameters of the reference points laws of motion are significantly different from zero. As suggested by (e.g.) Erev and Roth (1998), this result reflects the key role played in the reinforcement learning mechanism by a reference point, transforming payoffs into reinforcement factors. Note that the sums of the proposers and receivers initial reference points $\left(\widehat{\rho}_p^1 + \widehat{\rho}_r^1\right)$ in all three treatments are roughly equal to the size of the pie to be divided (i.e. \$50). In other words, both types of players have realistic and matching initial expectations. This indicates also that rich players do not immediately feel that they deserve most of the pie. As expected, poor (rich) subjects have significantly higher (smaller) initial expectation in treatment 1 (2) than in treatment 0. The parameters $\widehat{\gamma}_{1,j}$ and $\widehat{\gamma}_{2,j}$, $j \in \{p, r\}$, in treatment 0 are small, and only significantly different from zero for the proposers. Therefore, the reference points, and the players expectations, barely evolve over time in treatment 0.

In the other two treatments, however, both $\widehat{\gamma}_{1,j}$ and $\widehat{\gamma}_{2,j}$, $j \in \{p, r\}$, are positive and highly significant. Given the profile of payoffs observed in these treatments, this implies that the reference point of a rich (poor) subject rises (declines) rapidly over time, as she sees her payoffs in the successive ultimatum games increase (decrease). In other words, a feeling of entitlement grows over time within rich subjects, while poor participants come to the realization that they should expect less.

These estimates provide an insight into the way known wealth differences affected the dynamics of behavior in treatments 1 and 2. Indeed, the estimated parameters suggest that guided by foregone best responses, and an acquired sense of deservingness, rich subjects increased their demands. In contrast, poor players only appear to react to their opponents increasingly aggressive behavior by scaling down their demands, since the strategy selection of poor players was essentially influenced by their own past plays and outcomes.

7.3. Alternative Specifications

The object of this section is to test whether alternative models may better explain the dynamics of behavior. The different hypotheses can be nested within the general reinforcement model, and comparisons will be conducted on the basis of Wald tests. Results are presented in Table 6.

We first verify whether subjects learn to behave homogeneously from their own experience and observation. To test for the presence of heterogeneity across proposers (receivers) within the same treatment we estimate the benchmark model with a different parameter $\beta_{i,j}$, $j \in \{p, r\}$, $i = 1, ..., N$, for each player. Then, we can test the restriction $H_0 : \left\{ \left(\beta_{i,j}^{CRL}, \beta_{i,j}^{RBR} \right) = \left(\beta_{1,j}^{CRL}, \beta_{1,j}^{RBR} \right), \forall i = 2, ..., N \right\}$ for $j \in \{p, r\}$. The p-values in Table 6 indicate that one cannot reject the null hypothesis at the usual significance levels. Therefore, we do not find conclusive evidence of heterogeneity across players. In other words, as mentioned in section 5, it is not obvious that we can differentiate subjects into two sub-samples of "fair" and "greedy" players. Rich players may alternate between generous and greedy offers, but they all appear to learn to behave in a more

selfish manner. This test also illustrates the ability of the general reinforcement learning model to describe the behavior of each proposer (receiver) in a treatment equally well.

We now test whether proposers (receivers) adjust their strategies according to the same learning model in all three treatments. Under this hypothesis, the difference in the dynamics of plays, would be essentially explained by the specific events experienced and observed by subjects in each treatment. The benchmark model is estimated jointly with the data observed in the three treatments, and we test the restriction $H_0 : \left\{ \left(\beta_j^{CRL}, \beta_j^{RBR}\right)_{T_0} = \left(\beta_j^{CRL}, \beta_j^{RBR}\right)_{T_1}, \left(\beta_j^{CRL}, \beta_j^{RBR}\right)_{T_0} = \left(\beta_j^{CRL}, \beta_j^{RBR}\right)_{T_2} \right\}$, where $\left(\beta_j^{CRL}, \beta_j^{RBR}\right)_{T_t}$ is the vector of all parameters common to subjects of type $j \in \{p, r\}$ in treatment $t \in \{0, 1, 2\}$.[32] The p-values in Table 6 indicate that one can reject the null hypothesis at the usual significance levels. Therefore, it appears that subjects in each treatment do not learn according to the same model. In other words, the environment in which the game was played seem to have shaped the way subjects learned. This result is consistent with previous studies analyzing the influence of environmental factors on learning (see e.g. Slembeck 1998 or Armantier 2003).

Next, we verify whether behavior in each treatment may be explained by a simpler learning model than the combination of learning rules estimated in the benchmark model. We successively test the following hypotheses for the proposers and receivers in each treatment: first, the data may be explained only by a simple choice reinforcement learning model (i.e. $H_0 : \left\{\theta_j^{RBR} = 0, \beta_j^{RP} = 0\right\}$); second, the data may be explained by a learning model reinforcing only the foregone best response strategies (i.e. $H_0 : \left\{\theta_j^{CRL} = 0, \beta_j^{RP} = 0\right\}$ in the modified benchmark model in which the reinforcement factor of the choice reinforcement learning rule is now written $r_{i,k}^{CRL} = \theta_j^{CRL} x_i^k$); third, the learning model does not include a reference point (i.e. $H_0 : \left\{\beta_j^{RP} = 0\right\}$). The p-values in Table 6 indicate that these three hypotheses are always rejected, except for the receivers in treatment 0 for which the choice reinforcement learning model cannot be rejected at a 5% significance level. In other words, the behavior of receivers in treatment

---

[32]To get the best chance to accept the test, we exclude from the null hypothesis the parameters associated with the initial propensity and reference point, as they differ significantly in each treatment.

0 may be explained solely by a simple choice reinforcement learning model. However, a proper modelization of the dynamics of plays in all other cases requires the combination of learning rules included in the benchmark model.

We now verify whether the experimental outcomes may be equally explained by an exogenous adjustment mechanism. Learning is typically assumed to occur when the strategy selection of a player at a given period is influenced by his own or any other players past actions and/or outcomes. In contrast, the general reinforcement rule can be modified to represent a completely exogenous adjustment of strategies. Indeed, if we impose $m_{i,k}^l = \lambda_1 m_{i,k-1}^l + s_\infty(1 - \lambda_1)$ and $r_{i,k}^l = \lambda_2 r_{i,k-1}^l$ (where $\lambda_1 \in [0, 1[$, $\lambda_2 > 1$, $s_\infty \in [\underline{s}, \overline{s}]$, $\alpha^l = r_{i,0}^l = 1$), then the strategy reinforced and the reinforcement factor evolve exogenously, and behavior systematically converges toward a pure strategy equilibrium s_∞. The equilibrium model with errors is a special case of exogenous adjustment in which s_∞ is equal in the present context to the subgame perfect equilibrium strategy $s_\infty = 0$. The model is tested twice under each of the following null hypotheses $H_0 : \left\{\theta_j^{CRL} = 0, \theta_j^{RBR} = 0, \beta_j^{RP} = 0\right\}$ and $H_0 : \left\{\theta_j^{CRL} = 0, \theta_j^{RBR} = 0, \beta_j^{RP} = 0, s_\infty = 0\right\}$. Table 6 indicates that both hypotheses are strongly rejected by the data. In other words, it appears that i) the adjustment process observed in each treatment is not exogenous; ii) the subgame perfect equilibrium does not act as an attractor; and iii) the evolution of strategies may be attributed to learning.

Learning direction theory (Selten and Buchta 1999) offers a potential alternative to explain the dynamics of behavior observed in each treatment. Indeed, according to this model, subjects may have simply adjusted their behavior in the direction of strategies that were, or would have been successful. To illustrate the concept of direction learning consider a proposer in the ultimatum game and the choice reinforcement learning model. When an agreement is not reached, the proposer should realize that his offer was too low; he should also understand that a smaller offer would not have improved his situation, while larger offer could have generated a better payoff. As a result, instead of symmetrically reinforcing strategies around the strategy played s_i^t, the proposer should reinforce negatively strategies below s_i^t, and reinforce positively strategies above s_i^t. By

doing so, the proposer is more likely to generate a strategy larger than s_i^t, and get a better payoff in the future. The notion of direction learning may also be applied to foregone best responses, in which case s_i^t is replaced by the best response $BR(s_{-i}^t)$ in the previous example.

Following Armantier (2003), we can test these hypotheses after incorporating direction learning within the foregone best response or choice reinforcement rule, by replacing μ_i^k in equation (6.3) by

$$\mu_i^k = m_{i,k}^l + \left(\omega_v^l\right)^2 \left(\underline{s}.I_{\alpha_\nu^l<0} + \overline{s}.I_{\alpha_\nu^l>0} - m_{i,k}^l\right) e^{-\left(r_{i,k}^l\right)^2} \quad , \qquad (7.2)$$

where $l \in \{CRL, RBR\}$, $-1 \leq \omega_\nu^l \leq 1$, $[\underline{s}, \overline{s}] = [0, 50]$, $\nu = 1$ ($\nu = 2$) when the reinforcement factor $r_{i,k}^l$ is positive (negative), and $I_{\omega_\nu^l<0}$ is the indicator function satisfying $I_{\omega_\nu^l<0} = 1$ when $\omega_\nu^l < 0$ (otherwise $I_{\omega_\nu^l<0} = 0$). The parameter ω_1^l $\left(\omega_2^l\right)$ accounts for any potential direction learning associated with a positive (negative) reinforcement factor. Indeed, when ω_ν^l is positive (respectively negative), μ_i^k lays in the interval $\left[m_{i,k}^l, \overline{s}\right]$ (respectively $\left[\underline{s}, m_{i,k}^l\right]$) and strategies slightly greater (respectively smaller) than $m_{i,k}^l$ are primarily reinforced. Note that μ_i^k gets closer to $m_{i,k}^l$ when $\left(r_{i,k}^l\right)^2$ increases. This implies that when the payoff (or the potential payoff) is large and positive (respectively negative), the center of the reinforcement rule gets closer to the strategy reinforced, so that the player is more (respectively less) likely to play again the same strategy in the future. Finally, $\omega_\nu^l = 0$ corresponds to no direction learning, while $\omega_\nu^l = \pm 1$ can be considered full direction learning, since players essentially reinforce either $\underline{s}$ or $\overline{s}$ for a reinforcement factor $r_{i,k}^l$ close to zero. Using this model, we can then test successively the null hypotheses $H_0 : \left\{\omega_1^{CRL} = 0, \omega_2^{CRL} = 0\right\}$ and $H_0 : \left\{\omega_1^{RBR} = 0, \omega_2^{RBR} = 0\right\}$. The p-values in Table 6 indicate that learning direction does not appear to play a significant role to explain the dynamics of plays observed in each treatment.

The benchmark learning model adopted therefore appears to outperform several alternative candidates. To evaluate further the empirical relevance of the estimates, and the fit of the model, the following regressions have been conducted for each type of

subjects in each treatment:

$$\overline{s}_j^t = \delta_0 + \delta_1 \eta_t^1 \left(\widehat{\beta}_j\right) + \varepsilon_{it} \qquad , \qquad (7.3)$$

where $\overline{s}_j^t$ is the average strategy played during a given treatment by subjects in group $j \in \{p, r\}$ at period t, and $\eta_t^1 \left(\widehat{\beta}_j\right)$ is the average strategy in group $j \in \{p, r\}$ at period t simulated with the general reinforcement model, and the estimated parameter $\widehat{\beta}_j$.[33] The F tests in Table 7 indicates that one cannot reject at a 5% significance level the null hypotheses $H_0 : \{\delta_0 = 0, \delta_1 = 1\}$. In addition, the R^2's reported in Table 7 illustrate that the learning model explains almost entirely the variation in the aggregated experimental data. In other words, the average path of strategies simulated with the general reinforcement learning model fits well the average behavior observed in the experiment.

Moreover, to evaluate individual fit, the following regressions have been estimated for each player in each treatment:

$$s_i^t = \delta_0 + \delta_1 \eta_{i,t}^1 \left(\widehat{\beta}_j\right) + \varepsilon_{it} \qquad , \qquad (7.4)$$

where s_i^t is the strategy of player i in period t, and $\eta_{i,t}^1 \left(\widehat{\beta}_j\right)$ is the simulated expected strategy in period t, conditional on the history of plays observed for player i up to period t. Table 8 indicates that the F tests for the null hypothesis $H_0 : \{\delta_0 = 0, \delta_1 = 1\}$ could not be rejected for the wide majority of subjects in each of the three different treatments. In addition, the average R^2's in Table 8 suggest that the benchmark model explains between 76% and 89% of the variance in the individual subjects decisions during the experiment. In other words, the general reinforcement learning model, along with the estimated parameters, replicate fairly well the individual behavior of the wide majority of subjects.

---

[33]In fact, the benchmark model is re-estimated for each type of subjects in each treatment, accounting only for the components of β_j significantly different from zero in Table 5. The Monte Carlo simulations $\eta_t^1 \left(\widehat{\beta}_j\right)$ in equation 7.3, and $\eta_{i,t}^1 \left(\widehat{\beta}_j\right)$ in equation 7.4, are based on these new sets of estimated parameters.

8. Discussion

The present paper was an attempt to analyze the important but challenging issue of the effect of known wealth differences on fairness considerations, and in particular on the stability of fairness preferences over time. To address this issue, an ultimatum game experiment was conducted by providing proposers and receivers with large and widely unequal initial endowments. The experimental outcomes indicate that subjects initially attempt to reach an egalitarian agreement that would equalize final wealth. With time, however, rich (poor) proposers make smaller (larger) offers, while rich (poor) receivers are willing to reject (accept) larger (smaller) offers. In other words, rich players become more greedy, and this behavior is tolerated by poor subjects.

The estimation of a general reinforcement learning model accounting for experience and foregone best responses indicates that i) independently of wealth differences, there is a clear evidence of learning by proposers and, which is even more remarkable, by receivers; ii) the monetary expectations of rich (poor) subjects increase (decrease) with time; iii) the strategy choices of rich players are mainly influenced by foregone best responses; and, iv) poor players essentially learn from their own past plays and outcomes. In other words, it appears that, guided by foregone best responses and an acquired sense of deservingness, rich subjects became more selfish, while poor subjects learned from their personal experience to accept this increasingly aggressive behavior.

We are now in a position to propose an answer to the question raised in the title of the paper. Indeed, the present study suggests that subjects start by adjusting their strategies in presence of known wealth inequalities, in a way that initially appears to leave fairness considerations unaffected. With time, however, relative wealth seems to modify the dynamics of behavior in a direction opposite to the fairness prediction.

These conclusions, however, should be interpreted with caution, and, as is often the case in experimental economics, the generality of the experimental results presented would need to be confirmed. Indeed, one cannot exclude that the experimental design adopted in order to enable the joint analysis of relative wealth and learning, partially

37

drove behavior. If it is premature to claim that our conclusions extend to other games, and in particular to the traditional ultimatum game, we have nevertheless constructed an experiment in which wealth differences had a significant impact on the stability of fairness preferences. Additional experiments may need to be conducted within the same, and/or slightly different environments, in order to verify whether our findings are robust, or to identify the factor(s) that drove behavior away from the fairness prediction in the present experiment.

If confirmed, the conclusions reached may be significant from a practical perspective. Indeed, there typically exist known wealth differences between the negotiating parties in real life bargaining such as wage negotiations or pre-trial settlements. The present paper suggests that a proper analysis of these situations may require one to take into consideration the bargaining power provided by wealth. The outcomes of the experiment may also partially explain and/or justify the success of class action lawsuits, as well as the intervention of a third party negotiator such as a mediator or a union representative. Indeed, this may, among other things, level the differences in bargaining power between the negotiating parties.

Finally, the present paper provides an insight into the notion of bargaining power. Economists typically define bargaining power by its effect (it allows agents to extract more favorable outcomes for themselves), and by the psychological or environmental factors from which it is derived (e.g. gender, entitlement, social status or wealth). However, the following questions have essentially been left unanswered: why do certain factors create bargaining power? How does bargaining power change a player's perspective of the game so that she modifies her behavior? The present paper sheds some light on the second question. Indeed, we have shown that the bargaining power provided by known wealth differences results in larger monetary expectations, and that the learning process of players with bargaining power is essentially driven by foregone best responses rather than personal experience.

Dept of Economics, S621 SBS Building, SUNY Stony Brook, Stony Brook NY, 11794-4384. E-mail: olivier.armantier@sunysb.edu.

REFERENCES

Abbink K., Buchta J., Sadrieh K. and R. Selten, 2001a, "How to Play 3x3-Games - A Strategy Method Experiment", *Games and Economic Behavior*, forthcoming.

Abbink K., Bolton G., Sadrieh K. and F-F. Tang, 2001b, "Adaptive Learning versus Punishment in Ultimatum Bargaining", *Games and Economic Behavior*, 37, 1-25.

Andreoni J., Smelser N.J. and P.B. Baltes, 2001, "The Economics of Philanthropy", *International Encyclopedia of the Social and Behavioral Sciences*, London: Elsevier.

Armantier O., 2000, "Estimation and Comparison of Learning Models", Mimeo SUNY Stony Brook.

Armantier O., 2003, "Does Observation Influence Learning?", *Games and Economic Behavior*, forthcoming, .

Bewley T., 1998, "Why Not Cut Pay?", *European Economic Review* 42, 459-490.

Binmore K., Shaked A. and J. Sutton, 1985, "Testing Noncooperative Bargaining Theory: A Preliminary Study", *American Economic Review*, 75, 1178-1180.

Binmore K., Avner S. and J. Sutton, 1989, "An Outside Option Experiment", *Quarterly Journal of Economics*, 104 (4), 753-770.

Blundell R. and S. Bond, 1998, "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics*, 87, 115-143.

Bolle F., 1990, "High Reward Experiments without High Expenditure for the Experimenter?", *Journal of Economic Psychology*, 11 (2), 157-167.

Bolton G. and R. Zwick, 1995, "Anonymity versus Punishment in Ultimatum Bargaining", *Games and Economic Behavior*, 10 (1), 95-121.

Bolton G. and A. Ockenfels, 2000, "ERC: A Theory of Equity, Reciprocity and Competition", *American Economic Review*, 90 (1), 166-193.

Bush R. and F. Mosteller, 1955, "Stochastic Models of Learning", New York: Wiley.

Cabrales A. and W. Garcia-Fontes, 1999, "Estimating Learning Models from Experimental Data: Quadratic Deviation and Maximum Likelihood", Mimeo, Universitat Pompeu Fabra.

Camerer C. and T-H. Ho, 1999, "Experience-Weighted Attraction Learning in Normal Form Games," *Econometrica* 67, 827-874.

Camerer C. and R. Hogarth, 1999, "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework", *Journal of Risk and Uncertainty*, 19, 7-42.

Camerer C. and R. Thaler, 1995, "Anomalies: Ultimatums, Dictators and Manners", *Journal of Economic Perspectives*, 9 (2), 209-219.

Cooper D., Feltovich N., Roth A. and R. Zwick, 2003, "Relative versus Absolute Speed of Adjustment in Strategic Environments: Responder Behavior in Ultimatum Games", *Experimental Economics*, forthcoming.

Cubitt R., C. Starmer and R. Sugden, 1998, "On the Validity of the Random Lottery Incentive System", *Experimental Economics*, 1 (2), 1998, 115-31.

Duffy J. and N. Feltovich. 1999, "Does Observation of Others Affect Learning in

Strategic Environments? An Experimental Study", *International Journal of Game Theory* v28, n1, 131-52.

Eckel C. and P. Grossman, 1996, "Altruism in Anonymous Dictator Games", *Games and Economic Behavior*, 16, 181-191.

Eckel C. and P. Grossman, 2001, "Chivalry and Solidarity in Ultimatum Games", *Economic Inquiry*, 39 (2), 171-188.

Erev I. and A. Roth, 1998, "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria", *American Economic Review*, 8, 848-881.

Fehr E. and Schimdt K., 1999, "A Theory of Fairness, Competition and Cooperation", *Quarterly Journal of Economics,* 114, 817-868.

Fershtman C., Gneezy U. and F. Verboven, 2002, "Discrimination and nepotism: The efficiency of the anonymity rule", Mimeo, University of Chicago Graduate School of Busisness.

Frank B. and G. Schulze, 2000, "Does economics make citizens corrupt?", *Journal of Economic Behavior and Organization*, Vol. 43, 101-113.

Friedman D. and S. Sunder, 1994, "Experimental Methods: A Primer for Economists", Cambridge University Press.

Goeree J. and C. Holt, 2000, "Asymmetric Inequality Aversion and Noisy Behavior in Alternating-Offer Bargaining Games", *European Economic Review*, 44, 1079-1089.

Güth W., 1995, "On Ultimatum Bargaining Experiments-A Personal Review", *Journal of Economic Behavior and Organization*, 27 (3), 329-344.

Güth W., Ockenfels P. and R. Tietz, 1992, "Distributive justice versus bargaining power", in New direction in Economic Psychology: Theory, Experiment and Application, S.E.G. Lea, P. Webley (eds.), Edward Edgar Ltd, 153-175.

Güth W. and R. Tietz, 1988, "Ultimatum bargaining for a shrinking cake–An experimental analysis," in Bounded Rational Behavior in Experimental Games and Markets, R. Tietz, W. Albers, R. Selten (editors), Springer, Berlin.

Hoffman E., McCabe K., Shachat K. and V. Smith, 1994, "Preferences, Property Rights, and Anonymity in Bargaining Games", *Games and Economic Behavior*, 7 (3), 346-380.

Kahn L. and J. Murnighan, 1993, "A General Experiment on Bargaining in Demand Games with Outside Options" *American Economic Review*, 83, 1260-1280.

Levine D., 1998, "Modeling Altruism and Spitefulness in Experiments", *Review of Economic Dynamics*, 1: 593-622.

McKelvey R. and T. Palfrey, 1995, "Quantal Response Equilibria for Normal Form Games", Games and Economic Behavior, 10, July, 6-38.

Murnighan J. and M. Saxon, 1998, "Ultimatum Bargaining by Children and Adults", *Journal of Economic Psychology*, 19, 415-445.

Neelin J., Sonnenschein H. and M. Speigel, 1988, "A Further Test of Noncooperative Bargaining Theory: Comment", *American Economic Review*, 78 (4), 824-836.

Newey W.K. and K. West, 1987, "Hypotheses Testing with Efficient Method of Moments Testing", *International Economic Review*, 28: 777-787.

Ochs J. and A. Roth, 1989, "An Experimental Study of Sequential Bargaining", *American Economic Review*, 79, 355-384.

Oosterbeck H., Sloof R. and G. Van de Kuilen, 2001 "Cultural Differences in Ultimatum Games Experiments: Evidence from a Meta-Analysis", Mimeo, University of Amsterdam.

Rabin M., 1993, "Incorporating Fairness into Game Theory and Economics", *American Economic Review*, 83, 1281-1302.

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

Roth A., 1995, "Bargaining Experiments", in The Handbook of Experimental Economics, edited by J. H. Kagel and A. E. Roth, Princeton: Princeton University Press, 253-348.

Ruffle B., 1998, "More is Better, but Fair is Fair: Tipping in Dictator and Ultimatum Games", *Games and Economic Behavior*, 23 (2), 247-265.

Schlag K., 1999, "Which One Should I Imitate?", *Journal of Mathematical Economics*, v31, n4 493-522.

Selten R. and J. Buchta, 1999, "Experimental Sealed Bid First Price Auctions with Directly Observed Bid Functions", In *Games and Human Behavior*, Mahwah NJ, L.Erlbaum Associates Inc., 79-102.

Shao J. and D. Tu, 1995, "The Jackknife and Bootstrap", New York: Springer-Verlag.

Slembeck T., 1998, "A Behavioral Approach to Learning in Economics. Toward an Economic Theory of Contingent Learning", Mimeo, University of St. Gallen.

Slembeck T., 1999, "Low Information Games: Experimental Evidence on Learning in Ultimatum Bargaining", Mimeo University of St.Gallen.

Slonim R. and A. Roth, 1998, "Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic," *Econometrica*, 66 (3), 569-596.

Stahl D., 1996, "Boundedly Rational Rules Learning in Guessing Games," *Games and Economic Behavior*, 16, 303-330.

Stahl D., 2000, "Rule Learning in Symmetric Normal Form Games: Theory and Evidence", *Games and Economic Behavior*, 32, 105-138.

Starmer C. and R. Sugden, 1991, "Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation", *American Economic Review*, vol. 81 (4), 971-78.

Straub P, and J. Murnighan, 1995, "An Experimental Investigation of Ultimatum Games: Information, Fairness, Expectations, and Lowest Acceptable Offers", *Journal of Economic Behavior and Organization*, 27 (3), 345-364.

Vega-Redondo F., 1997, "The Evolution of Walrasian Behavior", *Econometrica*, 65, 375-384.

Weg E. and V. Smith, 1993, "On the Failure to Induce Meager Offers in Ultimatum Games", *Journal of Economic Psychology*, 14 (1), 17-32.

Winter W. and S. Zamir, 1997, "An Experiment with Ultimatum Bargaining in a Changing Environment", Mimeo, The Hebrew University.

TABLE 1
RANDOM-EFFECT REGRESSION
ANALYSIS OF POSSIBLE CORRELATIONS BETWEEN SUBJECTS' ACTIONS

	TREATMENT 0				TREATMENT 1				TREATMENT 2			
	Proposers		Receivers		Proposers		Receivers		Proposers		Receivers	
δ_1	0.461* (0.142)	.	0.450* (0.117)	.	0.447* (0.100)	.	0.428* (0.143)	.	0.482* (0.169)	.	0.479* (0.126)	.
δ_2	0.064 (0.174)	0.482 (0.310)	0.123 (0.256)	0.603* (0.229)	0.012 (0.111)	0.538* (0.219)	0.062 (0.182)	0.490 (0.262)	-0.026 (0.130)	0.571 (0.297)	0.017 (0.155)	0.589* (0.203)
δ_3	0.322 (0.188)	.	0.348* (0.113)	.	0.309* (0.112)	.	0.362 (0.196)	.	0.347* (0.121)	.	0.372* (0.130)	.
δ_4	-0.102 (0.355)	-0.062 (0.413)	0.055 (0.123)	-0.020 (0.219)	0.025 (0.228)	-0.089 (0.278)	0.082 (0.202)	-0.043 (0.231)	-0.080 (0.316)	0.087 (0.260)	0.021 (0.128)	-0.017 (0.202)
δ_5	-0.023 (0.482)	-0.051 (0.376)	-0.017 (0.502)	-0.013 (0.534)	-0.012 (0.422)	0.066 (0.475)	0.010 (0.379)	-0.008 (0.544)	0.008 (0.366)	-0.121 (0.627)	-0.013 (0.396)	0.034 (0.555)
δ_6	0.022 (0.266)	-0.063 (0.408)	-0.041 (0.533)	0.121 (0.318)	0.009 (0.469)	0.027 (0.243)	-0.089 (0.551)	-0.062 (0.251)	-0.056 (0.620)	-0.010 (0.393)	-0.119 (0.562)	0.077 (0.416)

Numbers in parenthesis refer to asymptotic standard deviations, asymptotically robust to heteroskedasticity.
* Indicates parameters significant at a 5% level.

TABLE 2
EXPERIMENTAL OUTCOMES

		TREATMENT 0			TREATMENT 1			TREATMENT 2		
		Offers	MAO	Rejection	Offers	MAO	Rejection	Offers	MAO	Rejection
Overall	Mean	4.38	3.42	0.23	4.29	3.11	0.28	5.11	3.69	0.72
	Std	1.22	1.14	0.42	1.57	1.78	0.45	1.92	2.09	0.45
	Median	4.40	3.50	0.00	4.40	3.30	0.00	4.60	3.20	1.00
	Above 6	3.94%	1.44%	.	9.17%	4.67%	.	36.00%	17.78%	.
	Below 2	2.06%	7.06%	.	7.39%	32.06%	.	2.72%	12.83%	.
First 10 Periods	Mean	4.56	3.64	0.28	4.87	3.97	0.30	4.29	3.28	0.76
	Std	1.17	1.31	0.45	1.40	1.56	0.46	1.42	1.60	0.43
	Median	4.60	3.70	0.00	4.80	3.90	0.00	4.10	3.10	1.00
	Above 6	3.87%	3.23%	.	14.84%	7.10%	.	9.68%	6.45%	.
	Below 2	0.97%	6.77%	.	2.26%	10.32%	.	2.90%	11.61%	.
Last 10 Periods	Mean	4.12	3.20	0.24	3.63	2.64	0.29	5.85	4.25	0.67
	Std	1.18	1.10	0.43	1.76	1.64	0.46	1.87	2.40	0.47
	Median	4.10	3.30	0.00	3.80	2.20	0.00	6.30	3.60	1.00
	Above 6	2.90%	0.32%	.	5.48%	2.90%	.	54.19%	26.13%	.
	Below 2	4.19%	8.39%	.	17.10%	40.00%	.	1.94%	10.97%	.

TABLE 3
RANDOM-EFFECT REGRESSION
INCLUDING A TIME TREND

	TREATMENT 0		TREATMENT 1		TREATMENT 2	
	Proposers	Receivers	Proposers	Receivers	Proposers	Receivers
δ_1	-0.013* (0.001)	-0.011* (0.002)	-0.025* (0.004)	-0.031* (0.005)	0.033* (0.006)	0.024* (0.004)
δ_2	0.054* (0.024)	0.032 (0.029)	0.093* (0.034)	0.023 (0.027)	0.095* (0.036)	0.066* (0.021)
δ_3	0.026 (0.014)	0.008 (0.020)	0.039* (0.018)	0.027 (0.022)	0.032 (0.018)	0.017 (0.031)

Numbers in parenthesis refer to asymptotic standard deviations, asymptotically robust to heteroskedasticity.
* Indicates parameters significant at a 5% level.

TABLE 4 RANDOM-EFFECT REGRESSION TREATMENT EFFECT				
	TREATMENTS 0 AND 1		TREATMENTS 0 AND 2	
	Proposers	Receivers	Proposers	Receivers
δ_1	0.452* (0.190)	0.409* (0.152)	0.476* (0.182)	0.462 (0.253)
δ_2	0.307 (0.271)	0.263 (0.187)	0.371* (0.171)	0.310 (0.228)
δ_3	-0.392* (0.045)	-0.264* (0.068)	0.414* (0.054)	0.389* (0.077)
δ_4	0.530* (0.022)	0.332* (0.031)	-1.316* (0.082)	-0.893* (0.065)

Numbers in parenthesis refer to asymptotic standard deviations, asymptotically robust to heteroskedasticity.
* Indicates parameters significant at a 5% level.

| TABLE 5 ESTIMATION OF THE GENERAL REINFORCEMENT LEARNING MODEL						
	TREATMENT 0		TREATMENT 1		TREATMENT 2	
	Proposers	Receivers	Proposers	Receivers	Proposers	Receivers
α_j^{CRL}	0.806* (0.033)	0.943* (0.008)	0.845* (0.022)	0.911* (0.013)	0.759* (0.027)	0.981* (0.013)
α_j^{RBR}	0.578* (0.091)	0.344 (0.212)	0.410* (0.126)	0.426 (0.333)	0.527 (0.314)	0.630* (0.151)
θ_j^{RBR}	0.943* (0.090)	0.387 (0.226)	1.478* (0.037)	0.634 (0.340)	0.531 (0.375)	1.135* (0.058)
ρ_j^1	29.104* (0.778)	25.831 (1.055)	21.287* (0.651)	28.920* (0.704)	35.234* (1.038)	20.34* (0.991)
$\gamma_{0,j}$	0.154 (0.098)	0.455 (0.298)	0.106* (0.043)	0.253* (0.112)	0.167 (0.105)	0.212 (0.148)
$\gamma_{1,j}$	0.187* (0.073)	0.201 (0.176)	0.254* (0.066)	0.196* (0.060)	0.207* (0.082)	0.314* (0.070)
$\gamma_{2,j}$	0.078* (0.030)	0.098 (0.067)	0.162* (0.040)	0.150* (0.059)	0.126* (0.029)	0.183* (0.047)
μ_j	23.732* (1.341)	20.012* (1.776)	28.711* (0.605)	24.394* (1.261)	19.490* (0.743)	15.366* (1.323)
σ_j	9.862* (2.034)	11.308* (1.729)	7.229* (2.404)	9.497* (2.007)	8.554* (1.550)	10.989* (1.908)

Numbers in parenthesis refer to standard deviations. * Indicates parameters significant at a 5% level.

TABLE 6
P-VALUES FOR NESTED MODEL COMPARISON

H_0	Degrees of Freedom	χ^2 Statistics for $\alpha = 0.05$	TREATMENT 0		TREATMENT 1		TREATMENT 2	
			Proposers	Receivers	Proposers	Receivers	Proposers	Receivers
No Heterogeneity	87	109.770	0.287 (93.925)	0.336 (92.012)	0.142 (101.172)	0.246 (95.670)	0.210 (97.356)	0.228 (96.492)
Same Learning Parameters Across Treatments	6	12.592	1.234E-16 (87.720)	4.675E-13 (69.716)	.	.	.	.
Choice Reinforcement Learning	5	11.070	2.462E-3 (18.422)	0.102 (9.183)	6.378E-5 (26.751)	1.400E-2 (14.267)	8.502E-6 (31.213)	2.489E-3 (18.396)
Reinforcement of Best Strategies	5	11.070	5.738E-7 (37.093)	6.039E-10 (51.761)	9.023E-5 (25.975)	2.812E-12 (63.072)	8.661E-8 (41.172)	4.284E-6 (37.717)
No Reference Point	4	9.488	2.741E-3 (16.217)	0.032 (10.558)	4.888E-5 (25.062)	5.642E-3 (14.586)	6.932E-5 (24.307)	1.273E-2 (12.720)
Exogenous Adjustments	6	12.592	3.803E-10 (55.426)	8.020E-08 (43.821)	2.565E-13 (70.987)	3.462E-10 (56.360)	9.470E-12 (63.326)	7.352E-11 (58.950)
Nash Equilibrium with Errors	7	14.067	8.372E-11 (61.281)	5.502E-12 (67.170)	7.275E-09 (51.515)	9.231E-10 (56.050)	8.371E-12 (66.265)	7.909E-12 (66.387)
CRL Direction Learning	2	5.991	0.474 (1.493)	0.267 (2.641)	0.408 (1.793)	0.180 (3.430)	0.245 (2.813)	0.386 (1.904)
RBR Direction Learning	2	5.991	0.551 (1.192)	0.293 (2.455)	0.519 (1.312)	0.276 (2.575)	0.342 (2.146)	0.405 (1.807)

Numbers in parenthesis refer to the Wald statistics.

TABLE 7
REGRESSIONS TO EVALUATE THE AVERAGE FIT OF THE MODEL

	TREATMENT 0		TREATMENT 1		TREATMENT 2	
	Proposers	Receivers	Proposers	Receivers	Proposers	Receivers
δ_0	0.034 (0.027)	0.012 (0.019)	-0.007 (0.022)	-0.011 (0.014)	0.016 (0.018)	-0.010 (0.020)
δ_1	0.987* (0.045)	1.014* (0.032)	0.973* (0.032)	0.995* (0.037)	1.021* (0.041)	0.982* (0.030)
$\Re^2$	0.973	0.986	0.966	0.994	0.981	0.976
F Statistic for $H_0:\{\delta_0=0,\delta_1=1\}$	2.924	1.917	3.506	1.305	2.417	2.670
P-Value for F Test	0.093	0.171	0.066	0.258	0.125	0.108

Numbers in parenthesis refer to standard deviations. * Indicates parameters significant at a 5% level.

TABLE 8
## REGRESSIONS TO EVALUATE THE INDIVIDUAL FIT OF THE MODEL

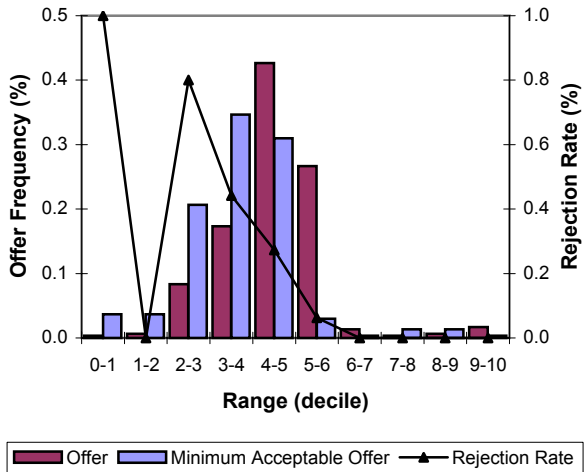|		TREATMENT 0		TREATMENT 1		TREATMENT 2	
		Proposers	Receivers	Proposers	Receivers	Proposers	Receivers
% of Subjects for Whom $H_0:\{\delta_0=0,\delta_1=1\}$ is not Rejected at a Significance Level	5%	70.00%	86.67%	63.33%	90.00%	76.67%	80.00%
	1%	86.67%	93.33%	83.33%	96.67%	90.00%	90.00%
$\Re^2$ in Individual Regressions	Average	0.766	0.837	0.753	0.879	0.815	0.822
	Min	0.448	0.613	0.410	0.642	0.545	0.577
	Max	0.926	0.938	0.893	0.924	0.907	0.913

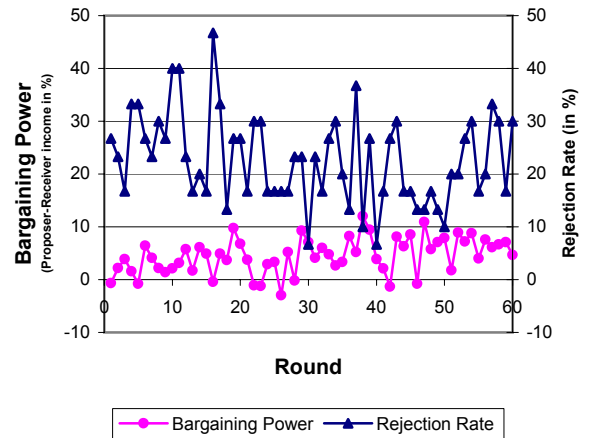# Treatment 0

Graph 0.1: 60 Periods
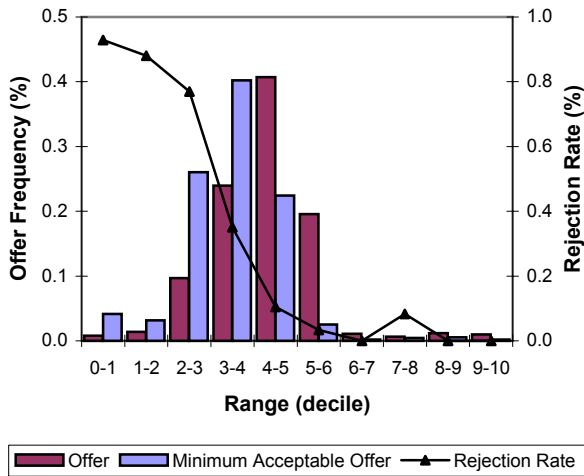
Graph 0.4: Evolution of Strategy
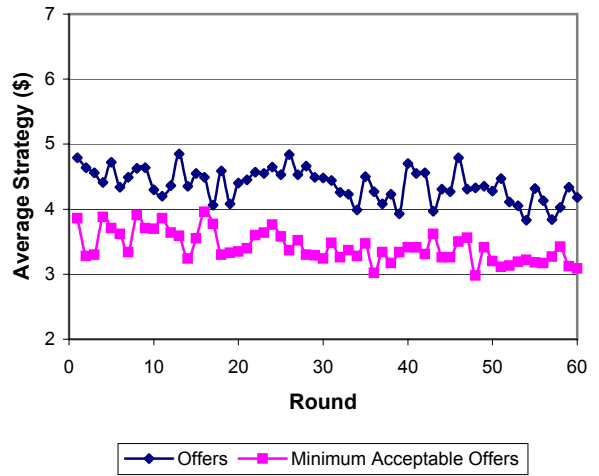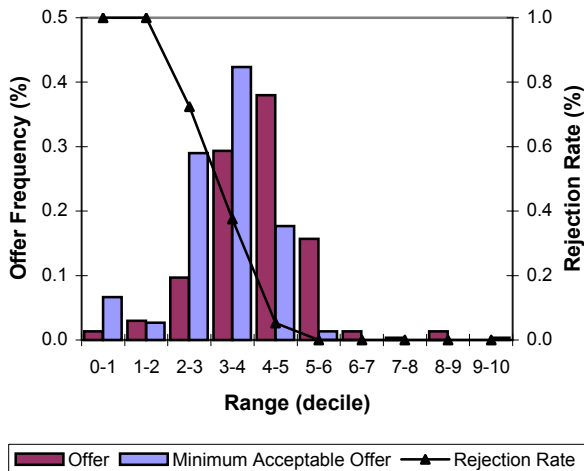

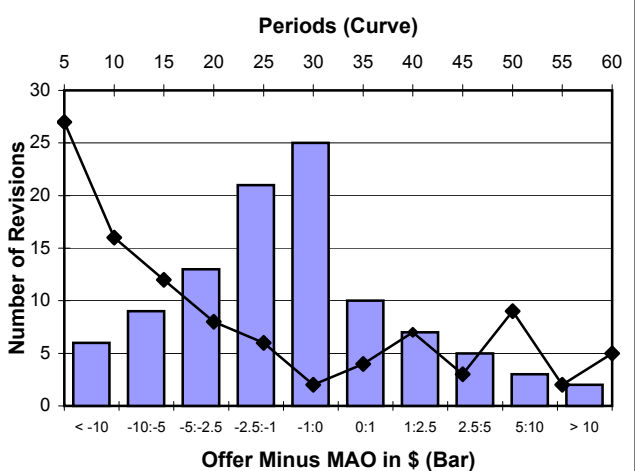## Graph 0.2: First 10 Periods

Graph 0.5: Evolution of Bargaining Power and Rejection Rate
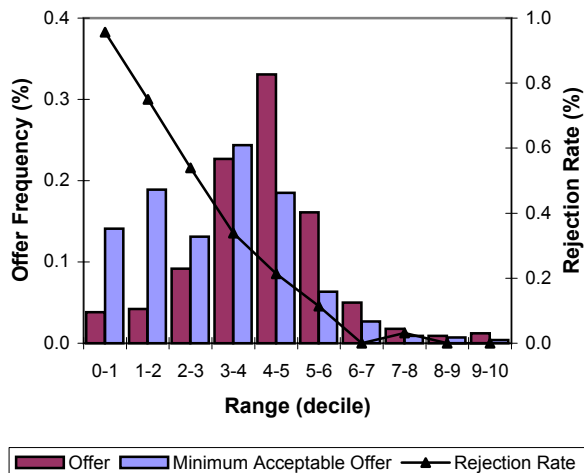

Graph 0.3 : Last 10 Periods
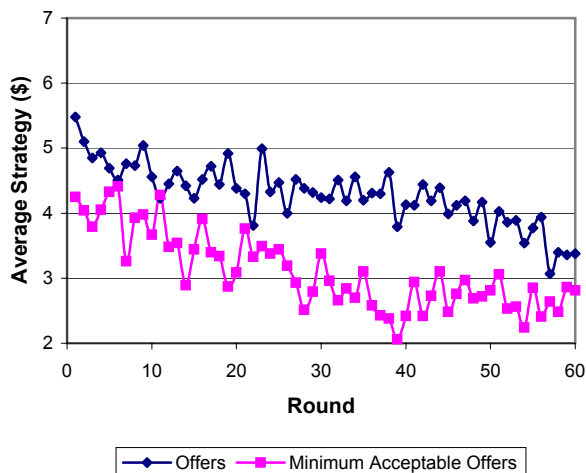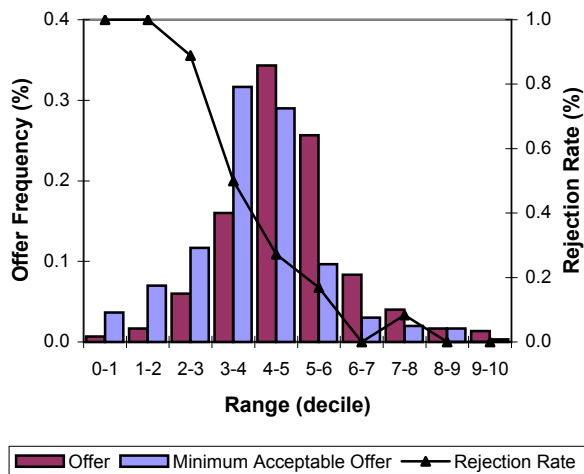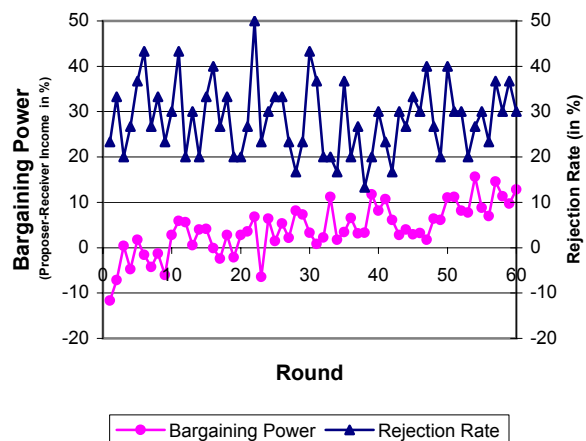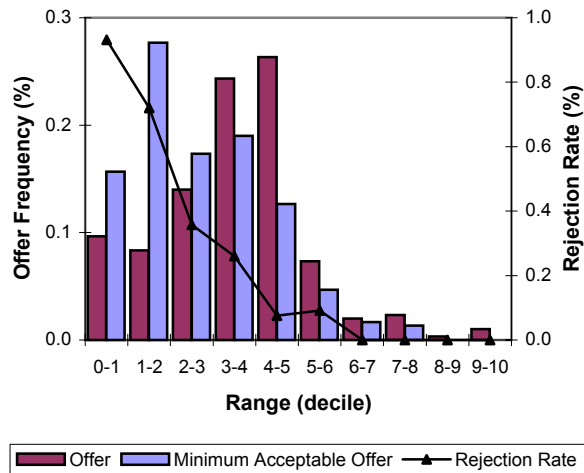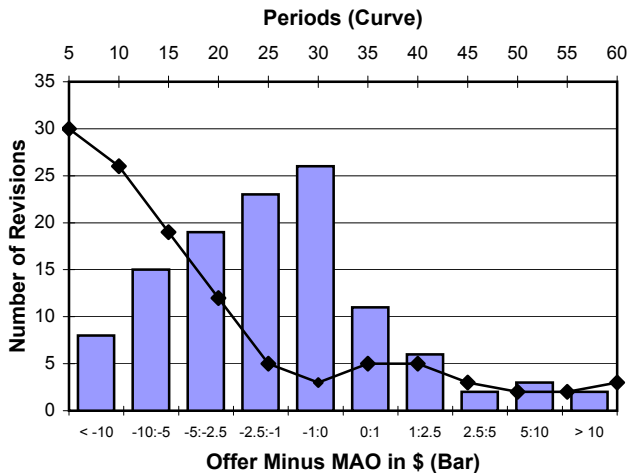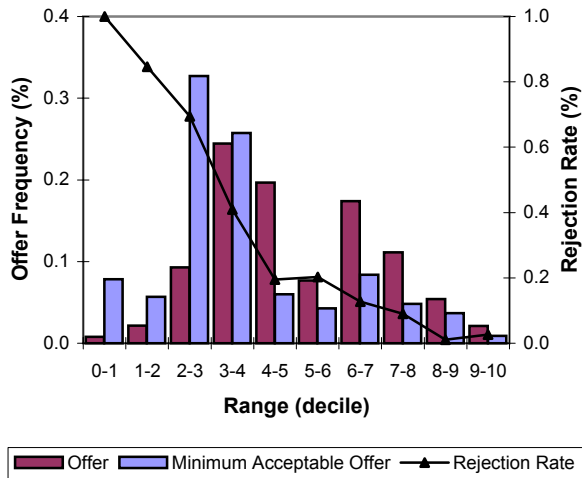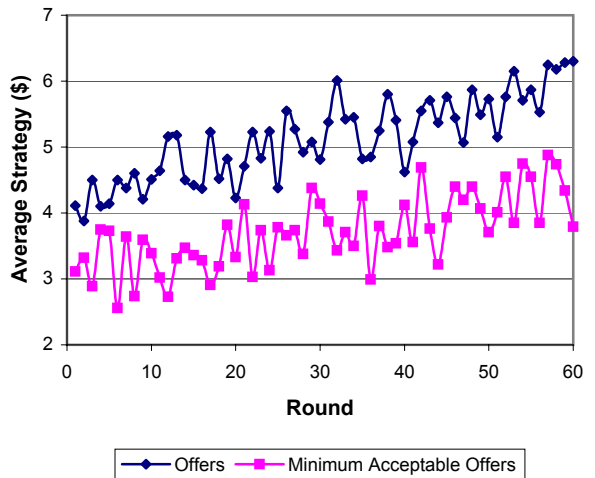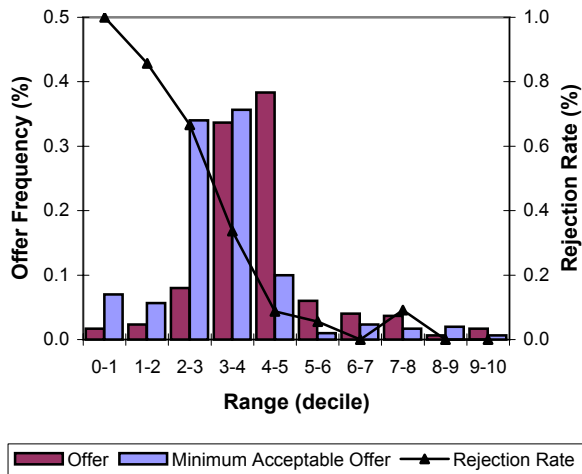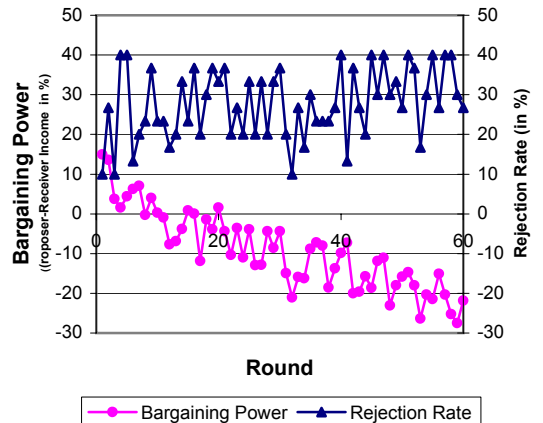

Graph 0.6: Revisions of Strategy

Treatment 1

Graph 1.1: 60 Periods

Graph 1.2: First 10 Periods

Graph 1.3: Last 10 Periods

Graph 1.4: Evolution of Strategy

Graph 1.5: Evolution of Bargaining Power and Rejection Rate

Graph 1.6: Revisions of Strategy

Treatment 2

Graph 2.1: 60 Periods

Graph 2.4: Evolution of Strategy

Graph 2.2: First 10 Periods

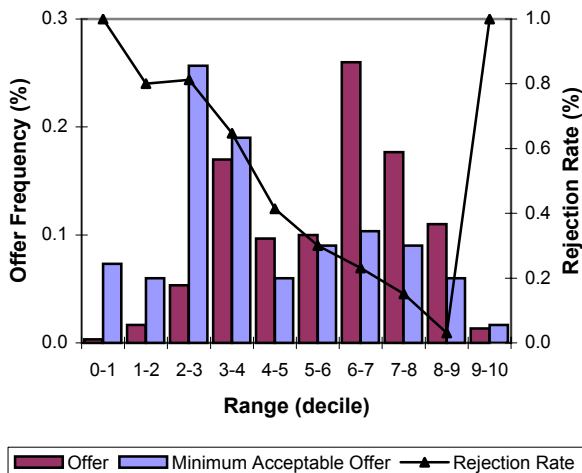## Graph 2.5: Evolution of Bargaining Power and Rejection Rate

Graph 2.3: Last 10 Periods

Graph 2.6: Revisions of Strategy