INEFFICIENCY AND SOCIAL EXCLUSION IN A COALITION FORMATION GAME: EXPERIMENTAL EVIDENCE*

Akira Okada‡ and Arno Riedl§

March, 2002

Revision: June, 2003

---

*We are indebted to an associate editor and two anonymous referees of this journal for their valuable comments and remarks on an earlier version of this paper. Former versions of this paper have been presented at the ESA 1998 meeting in Mannheim, the 1999 annual meeting of the EEA in Santiago de Compostella, the International Conference on Experimental Economics 1999 in Osaka, the Game Theory World meeting 2000 in Bilbao, and at seminars in Dortmund, Munich, and Tilburg. We thank the participants for their useful comments. We also thank Dirk Engelmann and Gary Charness for their useful remarks. This paper is part of a research project on Strategic Bargaining and Coalition Formation financed by the Oesterreichische Nationalbank (Project number: 6933) and the EU-TMR Research Network ENDEAR (FMRX-CT-98-0238). It was also partially financed by CREST (Core Research for Evolutional Science and Technology) of Japan Science and Technology Corporation (JST), Asahi Glass Foundation, and the Oesterreichische Postsparkasse. A large part of the research was undertaken while the second author was at the Institute for Advanced Studies, Vienna. The second author also thanks Jana Vyrastekova for invaluable research assistance. The usual disclaimer applies.

‡Institute of Economic Research, Kyoto University, Sakyo, Kyoto 606-8501, Japan

§Tinbergen Institute and CREED, Department of Economics and Econometrics, University of Amsterdam, Roetersstraat 11, NL-1018 WB Amsterdam, The Netherlands; tel.: +31-20-525-4250, fax: +31-20-525-5283, e-mail: a.m.riedl@uva.nl (corresponding author)

Abstract

This paper experimentally investigates the impact of reciprocal behavior in multilateral bargaining and coalition formation. Our results show that reciprocal fairness strongly affects the efficiency and equity of coalition formation. In a large majority of cases, inefficient and unfair coalitions are chosen when their coalition values are relatively high. Up to one third of the experimental population is excluded from bargaining and earns nothing. In monetary terms economically significant efficiency losses occur. We find that the interplay of selfish and reciprocal behavior unavoidably leads to this undesirable consequences. We also compare the predictions of recently developed models of social preferences with our experimental results. We find that some of these models capture the empirical regularities surprisingly well.

JEL Classification Number: A13, C91, D61, D63.

Keywords: Coalition formation, inefficiency, reciprocity, social exclusion.

1 Introduction

Bargaining is one of the central aspects in economic activity. Some bargaining is bilateral and negotiations take place only between two economic agents such as a buyer and a seller. There are, however, many multilateral bargaining situations in which agents are free to form coalitions. Examples are abundant: private-ownership firms, cartel of firms, labor unions, clubs, networks, international trading blocks, coalitions of political parties, etc. An important aspect in multilateral bargaining is the possible conflict in coalition formation. In particular, it may play an important role for the efficiency and equity of agreements. Standard literature on bargaining theory analyses this conflict under the assumption of narrow selfishness of bargainers. The literature on bargaining experiments, however, leaves little doubt that behavior of people is also influenced by considerations of fairness and reciprocity.[1] The main purpose of this paper is to experimentally investigate the impact of reciprocal fairness on multilateral bargaining.

To experimentally study coalition-forming behavior in the most clear-cut way, we introduce a simple non-cooperative bargaining procedure of a three-person super-additive game in coalition form. In the game, a group benefit is assigned to each possible coalition while any single player produces zero benefit. In the experiment a 'proposer' has to choose between the efficient three- and an inefficient[2] and unfair two-person coalition. Thereafter, the proposer makes a proposal about the division of the coalition value. Only if all members of the chosen coalition accept the proposal the allocation is implemented. Otherwise the surplus is destroyed and nobody earns anything. Subjects who are not members of a coalition earn nothing for sure. This set-up not only allows us to investigate (ultimatum) bargaining behavior in two- *and* three-person coalitions within one setting but also whether people are ready to forego resources *and* increase inequality simultaneously.

To investigate the effect of different gains from cooperation on coalition choices and payoff distributions we systematically varied the value of an inefficient coalition, keeping the value of the efficient coalition unchanged. We implemented four different values of two-person

[1] When using the terms reciprocity, reciprocal behavior, or reciprocal action we do not only mean reciprocity in the narrow sense of responding (un)kindly to (un)kind behavior but also behavior that may be interpreted as reciprocal though it is not driven by intentions. The reason is that some purely outcome-based models of social preferences predict behavior that is not distinguishable from intentional rewarding or punishment. We discuss these models in Section 4.2.

[2] As an anonymous referee rightly remarked one has to be careful when talking about inefficiency in the presence of social or other regarding preferences. Whenever we are using this term it should be interpreted in the material or monetary sense, which we believe is in terms of resources an important measure.

coalitions within four experimental conditions. The conditions differed only with respect to the value of the two-person coalition. The grand coalition was always worth 3000 points and the two-person coalition values varied between 2800, 2500, 2100, and 1200 points.

Our results indicate a clear link between two-person coalition values and coalition formation. In the two conditions with efficiency losses of 7 and 17 percent, respectively, an overwhelming majority of up to 95 percent of 'proposers' take up the inefficient two-person coalition. They thereby exclude almost one third of the population from participation. In the condition where the two-person coalition induces an efficiency loss of 40 percent still about 40 percent of the proposers choose this small coalition. The actual behaviorally induced efficiency losses are economically significant and vary between 6 and 15 percent. We provide evidence that these economically and socially undesirable results are an unavoidable consequence of reciprocal behavior of responders and (seemingly) selfish behavior of proposers.

While reciprocity is often identified as a force which leads to more even income distributions (like in standard ultimatum games[3]) and/or increases efficiency (like in gift-exchange and trust games[4]) our study shows that the same behavioral predisposition can have the exact opposite consequences in other institutional environments.

For the first time, our study experimentally investigates how potential efficiency losses - due to inefficient subcoalitions - relate to coalition formation in a systematic way. Earlier experimental studies - shortly discussed below - on multilateral bargaining either focus on three-person ultimatum bargaining with an inactive player and no coalition formation or do not vary the efficiency loss of an inefficient coalition decision. None of these studies relates coalition formation to reciprocal behavior and its consequences for allocative efficiency and distributional concerns.

The experimental work closest to ours are the studies by Bolton and Chatterjee (1996), Bolton *et al.* (2003), Güth and Van Damme (1998), and Riedl and Vyrastekova (2003). The latter two studies are three-person ultimatum games without coalition decision. Güth and van Damme (1998) investigate proposer and responder behavior in an ultimatum game setting where the third player is inactive. In particular, they examine how different information about the proposal affects rejection behavior. In the experiment of Riedl and Vyrastekova (2002) all players are active. The authors are primarily interested in acceptance behavior when the rejection consequences are varied. Bolton and Chatterjee (1996) and Bolton *et al.*

[3]See the seminal paper of Güth *et al.* (1982), and for overviews Roth (1995) and Camerer (2002).

[4]See e.g. Berg *et al.* (1995) and Fehr *et al.* (1993, 1997, 1998). For one of the first accounts of efficiency increasing reciprocal behavior, see the seminal work of Axelrod, 1984.

(2003) conducted experiments with a three-person coalition-form game similar to our game. They investigated how different bargaining procedures and communication structures affect coalition formation.[5]

In the next section we describe the design of our experiment, including a portray of our three-person coalition-form game and a description of the experimental procedures. In section 3 our experimental results are presented. In section 4 we discuss shortly them in the light of some recently developed models of social preferences. Section 5 summarizes and concludes. Proofs are given in the Appendix.

2 Experimental Setup

2.1 A Non-Cooperative Coalition Formation Game

The game implemented in the laboratory is a non-cooperative three-person coalition formation game with an ultimatum bargaining stage. The three players involved are called *proposer*, *responder 1*, and *responder 2*. The sequence of the play is the following (see also Figure 1):

1. The *proposer* P chooses either a two-person (small) coalition or the three-person (grand) coalition. The grand coalition has a value of $V(P, R1, R2)$, where $R1$ and $R2$ stands for *responder 1* and *responder 2*, respectively. The value of the two-person coalition, denoted $V(P, Ri)$, is strictly smaller than the value of the grand coalition.

2. After P has chosen her coalition, she proposes how to divide the coalition value between her and the *chosen* bargaining partner(s).

    (a) If she has chosen the grand coalition, she proposes (x_P, x_{R1}, x_{R2}) with $x_P + x_{R1} + x_{R2} = V(P, R1, R2)$ to both responders.

    (b) If she has opted for a small coalition, she proposes (x_P, x_{Ri}) with $x_P + x_{Ri} = V(P, Ri)$ only to the chosen responder Ri.

[5] There exist some other ultimatum-like bargaining studies involving three players which are less close to our work. Güth *et al.* (1996) and Güth and Huck (1997) investigate proposer behavior in two-stage ultimatum games with uncertainty about the pie size; Knez and Camerer (1995) run experiments where a proposer plays two ultimatum games simultaneously and responders have asymmetric outside options; Kagel and Wolfe (2001) investigate how acceptance behavior changes if upon rejection an inactive third player receives different consolidation prizes.

3. If $R1$ has been chosen as a member of either the three- or two-person coalition he has to decide whether to accept or reject the proposal. If he has not been chosen he has nothing to decide on.

4. If the grand coalition was chosen and $R1$ has accepted the proposal, $R2$ decides whether to accept or reject the proposal. Otherwise, for $R2$ the same holds as for $R1$.

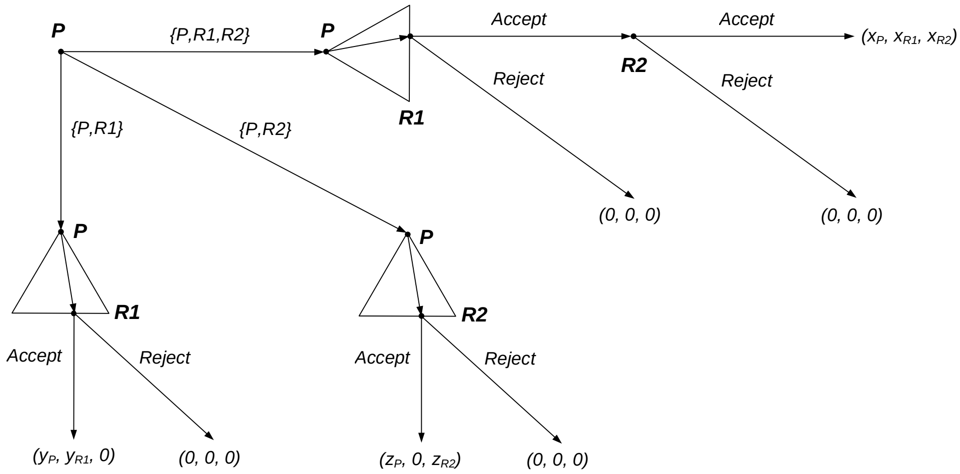


Figure 1 – A Non-Cooperative Three-Person Coalition Formation Game

The payoffs are allocated as follows: (i) If P has chosen the grand coalition and *both* responders *accept* the proposal then all players receive their shares according to the proposal. If any of the responders *rejects* nobody earns anything. (ii) If P has opted for a two-person coalition and *the chosen* responder *accepts* the proposal then these two players receive their shares according to the proposal. If he *rejects* both earn nothing. The responder who has not been chosen always has a payoff of zero.

All this is known by all players, and all players are informed about the decisions of all other players in previous moves. Assuming for the moment that there is no smallest money unit, then it can be easily seen that this game has a unique subgame perfect equilibrium (payoff). In the subgame starting after the proposer has opted for the three-person coalition there exists a unique subgame perfect equilibrium where P demands the whole pie for herself and both responders accept. In the subgame after P has chosen the two-person coalition with Ri $(i = 1, 2)$ the unique subgame perfect equilibrium implies that the proposer demands the whole pie $V(P, Ri)$ for herself, leaving Ri a payoff of zero which he will accept. Since the value of the two-person coalition is strictly smaller than the value of the grand coalition, the unique best decision for the proposer is to opt for the grand coalition. Hence, standard game theory predicts that P chooses the efficient coalition and makes the proposal $(x_P^*, x_{R1}^*, x_{R2}^*) = (V(P, R1, R2), 0, 0)$ which is accepted by both responders.[6]

2.2 Experimental Procedures and Parameters

We conducted ten experimental sessions involving 240 subjects. In each session we implemented one of two treatments (called T1 and T2). Both treatments consisted of two phases. In a session the 24 subjects were divided into two separate groups of twelve. Within these groups in each phase eight rounds of the above described game were played with random matching in each round. These created 10 statistically independent observations per treatment on the group level. If not mentioned otherwise we will base our tests and estimates on these independent units of observation.

The two phases within a treatment differed only with respect to the value of the two-person coalition. All values were described in points. The value $V(P, R1, R2)$ of the grand coalition was always 3000 points. In T1 the value $V(P, Ri)$ of the two-person coalition was 2800 points in phase 1 and 1200 points in phase 2. In treatment T2 the value of the two-person coalition was only 2100 points in phase 1 and 2500 points in phase 2. Note, that our experimental procedures (see Box 1 for details) ensured the following: (i) during phase 1 subjects did not know that there would be a second phase, and (ii) they only heard about their actual earnings at the end of both phases. Furthermore, the use of a with-in subject design in each treatment enabled us to examine whether the same persons make different choices, depending only on the value of the two-person coalition.

In the following we shall refer to the different conditions by T1-2800, T1-1200, T2-2100, and T2-2500. All ten sessions were computerized and run in English language at the CREED laboratory at the Faculty of Economics and Econometrics of the University of Amsterdam in March 2003.[7] Proposals had to be made in steps of 10 points. The exchange rate from points to euro was 250 points = 1 euro. (At the time of the experiment 1 euro was worth approximately 1 U.S. dollar.) Hence, the grand coalition was worth approximately \$ 12,-.

[6] In our experiment there is a smallest money unit. This destroys the uniqueness of the equilibrium. It can be shown, however, that in any subgame perfect equilibrium proposers always choose the grand coalition if the difference $V(P, R1, R2) - V(P, Ri)$ is larger than twice the smallest money unit, and that any proposal which gives each responder at least the smallest money unit is accepted.

[7] About two-thirds of the participants were undergraduate students in economics, econometrics, business administration or management studies. The remainder came from various fields. Slightly above 60 percent of the students held the Dutch nationality. The rest came from different countries, mostly within Europe. None of the subjects had participated in a similar experiment before.

Box 1: Experimental procedures

Phase 1: After arriving in the laboratory subjects were randomly assigned letters "R's", "M's", and "L's", which assigned them the different roles in the experiment. The "R's" were the proposers, the "M's" the first responders and the "L's" the second responders. A bargaining group consisted of one "R"-, one "M"-, and one "L"-subject. Subjects had the same role throughout the whole experiment. Subjects were seated in cubicles with sight shields. Any form of communication other than via the computer net was made impossible. Neither during nor after the experiment was the identity of bargaining partners revealed. Subjects received the instructions on-screen and they were also read aloud. One practice round was conducted. Thereafter, eight real rounds were carried out with random re-matching in each round. In the instructions subjects were informed that - in addition to the showup fee - they would be paid in cash the sum of their earnings in two out of the eight rounds after the experiment. These two rounds were randomly selected at the end of the experiment and subjects were aware of this procedure.

After the last round of *Phase 1* subjects were informed that there will be another experiment. Subsequently *Phase 2* started. The instructions were again given on-screen and read aloud. Participants were informed that there will be another eight rounds and that, thereafter, the experiment would definitely be finished. Furthermore, they were informed that they would be paid in cash the sum of their earnings in two randomly chosen rounds. The earnings of the first phase were unaffected by those of the second phase. The matching of subjects was the same as in phase 1 and subjects were informed about that.

3 Experimental Results

Subjects' average earnings (net of a show-up fee of 5 euro) were 12,15 euro in T1 and 11,75 euro in T2. Sessions lasted between 75 and 100 minutes. In the following we present first the results concerning the coalition decisions. Thereafter, we analyze bargaining behavior of responders and proposers within the chosen two- and three-person coalitions.

3.1 Coalition Decisions

The following result reports the coalition decisions.

Result 1 *(i) If the value of the two-person coalition is* 2800 *or* 2500 *an overwhelming majority of proposers opts for the two-person coalition. (ii) If the value of the two-person coalition is* 2100 *still about 40 percent of the proposers chooses the two-person coalition. (iii) Only for a value of the two-person coalition of* 1200 *the grand coalition is almost always formed. (iv) The frequency of two-person coalitions does not decrease over time.*

Evidence for this result is provided by Figure 2, which shows the percentage of chosen two-person coalitions per condition and round. The figure clearly indicates that in T1-2800 and

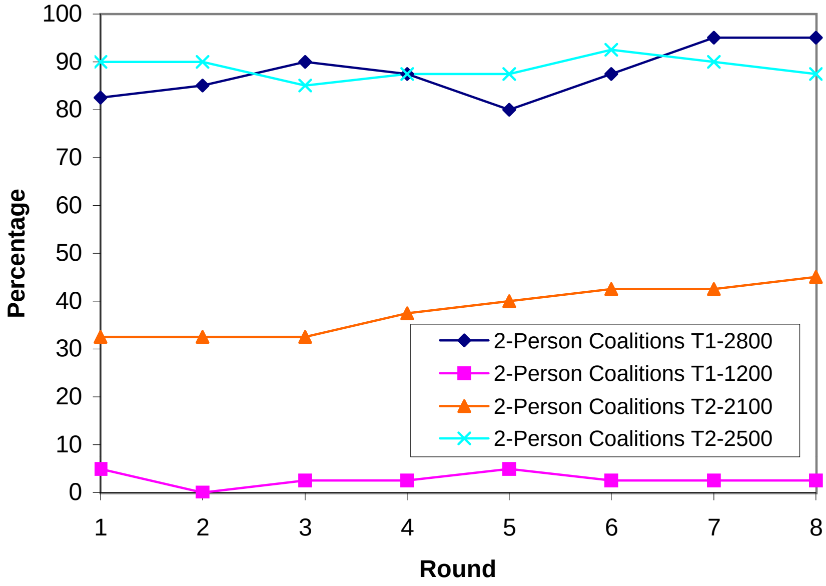


Figure 2 – Coalition Decisions

T2-2500 most proposers choose the two-person coalition right from the beginning. In the first round of T1-2800 82.5 percent (33 out of 40 proposers) choose the two-person coalition. In the first round of T2-2500 even 90.0 percent (36/40) opt for the small coalition. Hence, social exclusion takes place right from the beginning. The frequency of two-person coalitions does not decrease over rounds. For T1-2800 the Spearman rank-order correlation coefficient of the average number (across all proposers) of two-person coalitions on rounds yields a value of $r_s = 0.64$ ($p = 0.044$, one-sided test). For T2-2500 this correlation coefficient is zero.

If the value of the two-person coalition is only 2100, about one third of the proposers (13/40) choose the small coalition in the first round. This number increases to 45 percent (13/40) in round eight leading to 38.1 percent across all rounds. The Spearman rank order coefficient indicates a significantly increasing trend in the frequency of two-person coalitions ($r_s = 0.97$; $p < 0.0001$, one-sided test). Only for the very low coalition value of 1200 (almost) no small coalitions are observed. Over all rounds, in only 9 cases (out of 320 decisions) the two-person coalition is chosen.

Result 2 *There is no difference in the frequency of two-person coalitions between the values of* 2800 *and* 2500 *of the small coalition. If the value of the small coalition is* 2100 *the frequency of two-person coalitions is smaller than for the higher values but higher than for the small value of* 1200.

Statistical support for this result comes from round by round comparisons with the frequency of two-person coalitions per round and group as unit of observation. When comparing T1-

2800 with T2-2500 the Mann-Whitney rank-sum test does not reject the null hypothesis of no difference for any round ($p \geq 0.282$; 2-sided tests). When testing T2-2100 against T1-2800 and T2-2500, respectively, the difference in frequencies of two-person coalitions turns out to be statistically significant in each round ($p \leq 0.0032$, 2-sided Mann-Whitney test, and $p \leq 0.0056$, 2-sided Wilcoxon signed-ranks test, respectively). A comparison of T2-2100 with T1-1200 also reveals significant differences in each round ($p \leq 0.0052$, 2-sided Mann-Whitney test).

3.2 Bargaining Behavior in Dividing Coalition Values

Behavior in Two-Person Coalitions. Table I summarizes the behavior of responders and proposers in two-person coalitions. Although relative offers are a bid on the low side, on average, observed behavior is within keeping earlier results in stand-alone ultimatum games.[8] In the following we analyze responder and proposer behavior in more detail.

Table I – SUMMARY OF BEHAVIOR IN TWO-PERSON COALITIONS

Means and Medians of Offers to Chosen Responder and Disagreement Rates[a]

T1-2800				T2-2100				T2-2500			
# of 2-PC	Dis. in %	Mean (in %)	Med. (in %)	# of 2-PC	Dis. in %	Mean (in %)	Med. (in %)	# of 2-PC	Dis. in %	Mean (in %)	Med. (in %)
281	18.2	865 (30.9)	800 (28.6)	122	9.8	797 (38.0)	800 (38.1)	284	8.8	818 (32.7)	800 (32.0)

Note: [a]... the number of two-person coalitions is the total across all triads and rounds; means, medians, and disagreement rates are calculated across two-person coalitions and rounds. Due to the lack of two-person coalitions in T1-1200 no statistics for this condition are presented.

Responder behavior in two-person coalitions: Figure 3 depicts the rejection rates by offer ranges (in points) and condition. The figure shows that in all conditions responders behave reciprocally, in the sense that lower offers are rejected with a higher probability. The figure also suggests that in all conditions relatively low offers are accepted with a relatively high frequency. In T2-2100 *any* offer above 40.5 percent (850 points) is accepted (42 offers), in

[8]In this paper we do not investigate how responder and proposer behavior in two-person coalitions compares to behavior in standard stand-alone ultimatum game experiments. In another paper we examined this issue by comparing data from experiments we have run in Japan and Austria with those from Slonim and Roth (1998). The main result there is that for a comparable pie size responders in our study seem to be significantly 'softer'. For details we refer the interested reader to Okada and Riedl (2001).

In the figure empty squares indicate that no offers in the respective range have been made and bars with zero height indicate that all offers have been accepted.

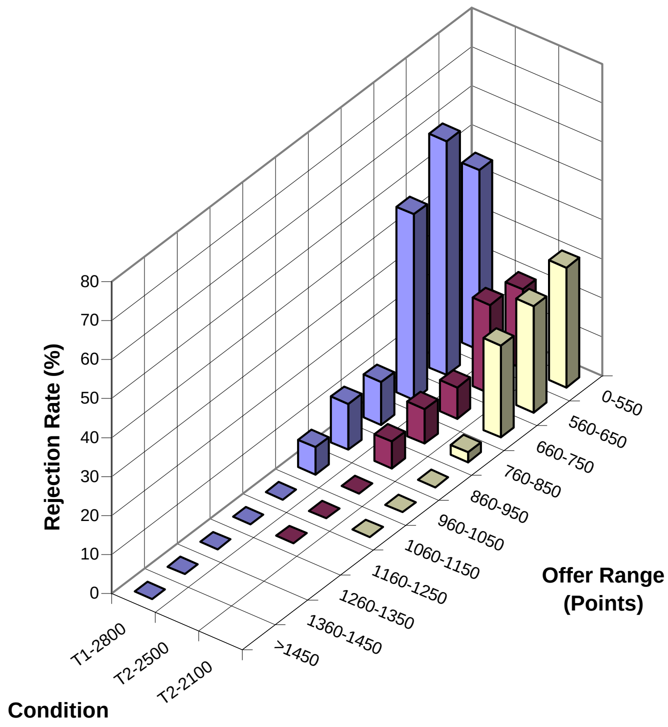


Figure 3 – Rejection Rates in Two-Person Coalitions

T2-2500 *any* offer above 38 percent (950 points) is accepted (153 offers), and in T1-2800 all offers above 37.5 percent (1050 points) are accepted for sure (56 offers).

With the help of probit estimates we also investigate whether there is a difference in responder behavior across the different conditions. The obtained estimation results clearly indicate that reciprocal behavior is prevalent in all conditions. We do, however, not find strong differences across the different coalition values.[9]

[9]Specifically, we estimate two probit models with robust standard errors and allowing for observations being not independent within groups. The specification is as follows: $Accept = f(\alpha + \beta_{relof} * relof + \beta_{v\star} * v\star + \beta_{int\star} * int\star + round)$, where '$\star$' stands for the coalition values 2100 and 2500, respectively. $Accept = 1$ if the offer was accepted and 0 otherwise, $f(x)$ denotes the probit function, and $relof$ is the offer measured relative to the value of the respective two-person coalition. If reciprocal fairness is at work higher offers should be accepted more often (i.e., $\beta_{relof} > 0$). The dummies $v2100 = 1$ ($v2500 = 1$) if the value of the coalition is 2100 (2500) (0 otherwise) measure the marginal change in acceptances in two-person coalitions from T1-2800 to T2-2100 and T1-2800 to T2-2500, respectively. The interaction variable $int2100$ ($int2500$) between relative offers and the value of the two-person coalition measures if the responsiveness to a change in the relative offer differs across the different values of the two-person coalition. *round* controls for a time trend. In both models β_{relof} turns out to be significantly positive ($p < 0.001$). In the model comparing T1-2800 with T2-2500, $v2500$ is significantly positive at the 5 percent confidence level. The interaction variables and the dummy $v2100$ are not significantly different from zero ($p \geq 0.106$) We also run random effects probit estimates. The results do not differ from those reported above.

Proposer behavior in two-person coalitions: Inspection of Table I reveals that the average relative offer is monotonically decreasing with the coalition value. Statistical tests only partly corroborate this visual impression. Using the average relative offer per group as unit of observation neither Mann-Whitney tests nor t-tests detect significant differences in offers between T1-2800 and T1-2500 and T1-2800 and T1-2100, respectively. However, when comparing relative offers in T2-2100 with those made in T2-2500 a Wilcoxon signed-rank test and a t-test reject the null hypothesis of equal offers at the 5 percent significance level (two-sided).[10] Hence, statistically proposers are only slightly more demanding in two-person coalitions with higher values.

Result 3 *Responder behavior in two-person coalitions does not substantially differ across different coalition values. Consistent with responder behavior proposers show only a slight tendency towards lower offers in coalitions with higher values.*

Behavior in Three-Person Coalitions. Table II summarizes responder and proposer behavior in three-person coalitions for T1-1200 and T2-2100.[11]

Table II – SUMMARY OF BEHAVIOR IN THREE-PERSON COALITIONS

Means and Medians of Offers (Percentages of the Grand Coalition Value) to Responder 1 and Responder 2 and Disagreement Rates[a]

T1-1200						T2-2100					
		Resp. 1		Resp. 2				Resp. 1		Resp. 2	
# of 3-PC	Dis. in %	Mean (in %)	Med. (in %)	Mean (in %)	Med. (in %)	# of 3-PC	Dis. in %	Mean (in %)	Med. (in %)	Mean (in %)	Med. (in %)
311	24.8	748 (24.9)	750 (25.0)	740 (24.7)	750 (25.0)	198	37.9	702 (23.4)	750 (25.0)	696 (23.2)	750 (25.0)

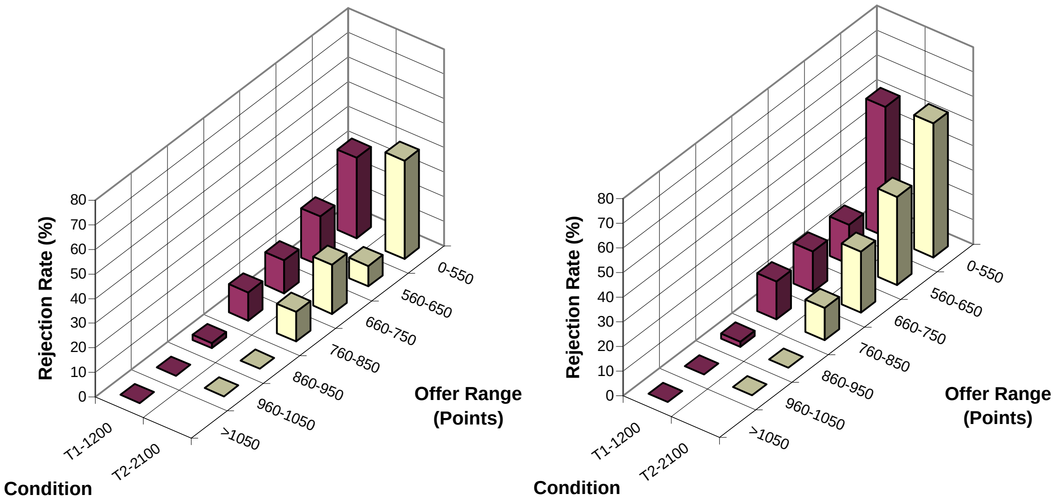
Note: [a]... the number of three-person coalitions is the total across all triads; means, medians, and disagreement rates are calculated across three-person coalitions and rounds.

Responder behavior in three-person coalitions: Table II shows the disagreement rates (that is, the frequency with which at least one responder rejected the proposal) across three-person coalitions and rounds. For both conditions disagreement rates are higher than in two-person

[10]Round-by-round comparisons lead to similar conclusions. When comparing relative offers in T1-2800 with relative offers in T2-2100 (T2-2500) only rounds 2 to 5 turn out to be significantly different (in no round significant differences are found) ($\alpha = 0.05$; 2-sided Mann-Whitney tests). For T2-2100 versus T2-2500 Wilcoxon signed rank tests find significant differences in round 4 to 7 ($\alpha = 0.05$; 2-sided tests).

[11]Since in T1-2800 and T2-2500 almost no three-person coalitions have been chosen no meaningful statistics can be presented for these subgames.

In the figure empty squares indicate that no offers in the respective range have been made and bars with zero height indicate that all offers have been accepted.



(a) Responder 1

(b) Responder 2

Figure 4 – Rejection Rates in Three-Person Coalitions

coalitions. In particular, in T2-2100 rejections are quite frequent. Interestingly, the higher disagreement rate in T2-2100 seems mainly due to the second responders' decisions. Given an acceptance by the first responder, in T2-2100 the proposal was declined by responder 2 in 22 percent of the cases whereas this was the case in only 13 percent of the cases in T1-1200. Similar numbers are obtained when we control for unequal offers.

Figures 4(a) and (b) show the (unconditional) rejection rates of first and second responders by offer range and condition. They indicate that, generally, both responders accept lower offers less often than higher offers. Furthermore, offers in the neighborhood and above one third of the grand coalition value are (almost) always accepted. Hence, reciprocal fairness considerations seem to be at work in three-person coalitions, too.

An interesting feature of the presence of two instead of only one active responder is that it allows us to examine whether acceptance behavior is influenced by the offer made to the other responder. We run probit regressions with acceptance behavior of responder i ($i = 1, 2$) as dependent variable and the relative offer, the difference in relative offers, and the round number as explanatory variables. The estimates for the second responder take only those observations into account where the first responder accepts the proposal. In all regressions robust standard errors are calculated also allowing for observations that are not independent within groups.[12]

[12]More formally, we run the following regressions: $AcceptR1 = f(\alpha + \beta_{relofR1} * relofR1 + \beta_{R1-R2} * (relofR1 -$

Table III – Probit Regressions:
Responder Behavior in Three-Person Coalitions

	T1-1200		T1-2100	
Coefficient	Responder 1	Responder 2	Responder 1	Responder 2
Constant	-1.147	-1.390^{*}	-1.526^{**}	-2.515^{**}
	($p = 0.163$)			
$\beta_{relofRi}$	9.485^{***}	10.94^{***}	9.355^{***}	14.48^{***}
β_{Ri-Rj}	9.687^{**}	7.828	3.703	-3.341
		($p = 0.065$)	($p = 0.105$)	($p = 0.420$)
round	0.006	-0.001	0.062	0.006
	($p = 0.912$)	($p = 0.978$)	($p = 0.288$)	($p = 0.889$)
Observations	311	268	198	157
Log Likelihood	-107.5	-81.34	-87.84	-62.97
Wald χ^2_3	36.90	29.00	71.15	56.47

Notes: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$; two-sided tests. All estimates are with robust standard errors also allowing for observations that are not independent within groups. In regression for responder 1 (responder 2), $i = 1$ and $j = 2$ ($i = 2$ and $j = 1$).

Table III depicts the regression results for three-person coalitions in T1-1200 and T2-2100. The coefficients $\beta_{relofR1}$ and $\beta_{relofR2}$ are all significantly greater than zero ($p < 0.001$), corroborating the visual impression of reciprocal behavior from Figure 4.

In T1-1200, for both responders the coefficients measuring the impact of the difference in offers are significantly positive (at least at the 10 percent level; 2-sided tests). In T2-2100 the coefficient is positive for responder 1 and negative for responder 2 (for both responders they are not significantly different from zero, however).[13] This indicates that, deviations from

$relofR2) + \beta_{round} * round$ for the first responder, and $CondAcceptR2 = f(\alpha + \beta_{relofR2} * relofR2 + \beta_{R2-R1} * (relofR2 - relofR1) + \beta_{round} * round$ for the second responder. $AcceptR1 = 1$ ($CondAcceptR2 = 1$) if the offer is accepted by the first (second) responder; 0 otherwise. $f(x)$ denotes the probit function, and $relofR1$ ($relofR2$) is the relative offer (as share of the value of the grand coalition) made to the first (second) responder. The coefficient β_{R2-R1} measures the influence of the relative standing with respect to the other responder. Since, in three-person coalitions an agreement is reached only if both responders accept, the analysis of the second responder's behavior is restricted to the cases where the first responder accepts the proposal ($CondAcceptR2$).

[13] We also run a different specification with dummy variables indicating whether one responder is strictly better or strictly worse off than the other responder as explanatory variables. It turns out that an offer

equal treatment - with respect to the other responder - increases a responder's acceptance likelihood when treated favorable and decreases it when treated unfavorable.

Proposer behavior: From Table II we see that, on average, proposers treat both responders more or less equally. In all cases the median offer is 25 percent of the grand coalition value. Based on groups as unit of observations, neither a non-parametric Wilcoxon signed-rank test nor a t-test detects a significant difference in offers made to the responders. This holds across all eight rounds as well as for each round separately. We also do not detect any significant difference in offers across the two reported conditions. We have also no indications that proposers change behavior across rounds within a condition.

Result 4 *(i) Responders in three-person coalitions behave reciprocally with respect to their own received offer. There is also evidence that the acceptance likelihood decreases with disadvantageous treatment compared to the other responder.*
(ii) In three-person coalitions proposers treat the two responders equally. They offer them on average about 25 percent of the grand coalition value. This holds independently of the value of the two-person coalition and the experience level.

3.3 Inefficiency

Contrary to stand-alone ultimatum game experiments observed material efficiency losses in our experiment are not (only) the direct result of rejections of unfair offers. Rather, they are the consequence of proposers' inefficient choices because of anticipated reciprocal responses by responders. Recall that the frequency of two-person coalitions in T2-2500 is as high as in T1-2800. Thus, the increase of the material efficiency loss from 6.67 to 16.67 percent does not retain proposers from choosing the inefficient and unfair allocation. Furthermore, even if the value of the two-person coalition is only 2100 points still about two-fifth of the proposers choose the inefficient allocation thereby inducing an efficiency loss of 30 percent.

In Figure 5 these inefficient decisions are reflected by the actually induced material efficiency losses of 5.9 percent in T1-2800, 11.4 percent in T2-2100, and 14.8 percent in T2-2500. These material efficiency losses are economically not negligible. As we will argue in the following, in coalitional bargaining situations as in our experiment, inefficient small coalitions are unavoidable as long as responders behave reciprocally and proposers act as (as if) income

where one responder is strictly better off than the other responder is always accepted (with one exception for responder 1 in T2-2100). The impact of being worse off turns out to be significantly negative in T1-1200 for both responders. For T2-2100 we do not find a significant effect.

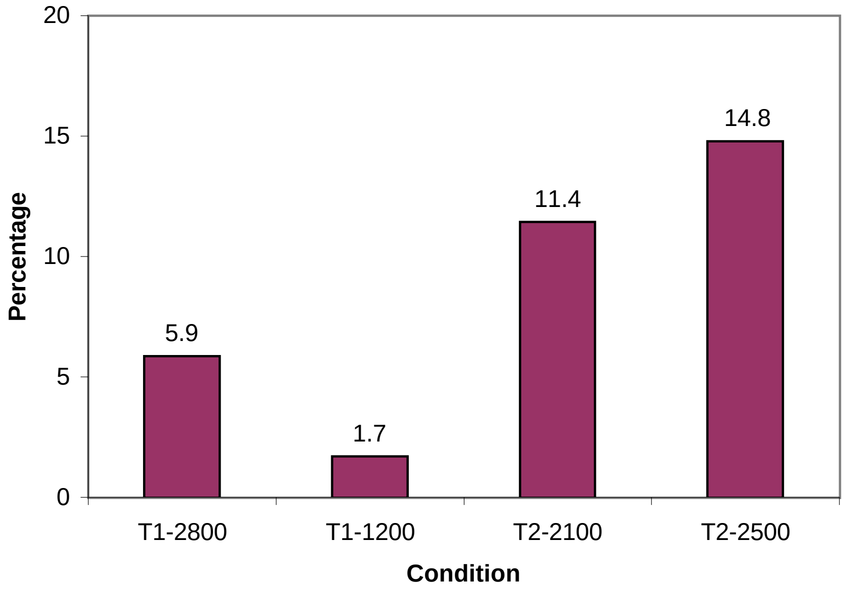


Figure 5 – Induced Material Efficiency Losses across Rounds

maximizers. Additionally, we show that, for a wide range of social preferences, this result is also predicted by outcome oriented models of reciprocal fairness.

4 Discussion

In this section we first provide arguments that reciprocal fairness considerations on the responders side and income maximization on the side of proposers are consistent with the observed inefficient coalitions and its adverse distributional consequences. Thereafter, we relate our empirical outcomes to the predictions of recently developed models of social preferences.

4.1 Reciprocal Behavior and Income Maximization

Result 5 *Income maximization of proposers dictates the choice of the two-person coalition whenever its coalition value is sufficiently high.*

First support for this result is provided by Figure 6. It shows the average earnings (in points) of proposers by condition and coalition across rounds. The three leftmost bars depict the average earnings in two-person coalitions in T1-2800, T2-2500, and T2-2100, respectively.[14] The four rightmost bars show the average earnings in three-person coalitions for all four values of the small coalition. The figure clearly indicates that in T1-2800 and T2-2500 a proposer

[14]T1-1200 is not shown since we have observed too little two-person coalitions in this condition.

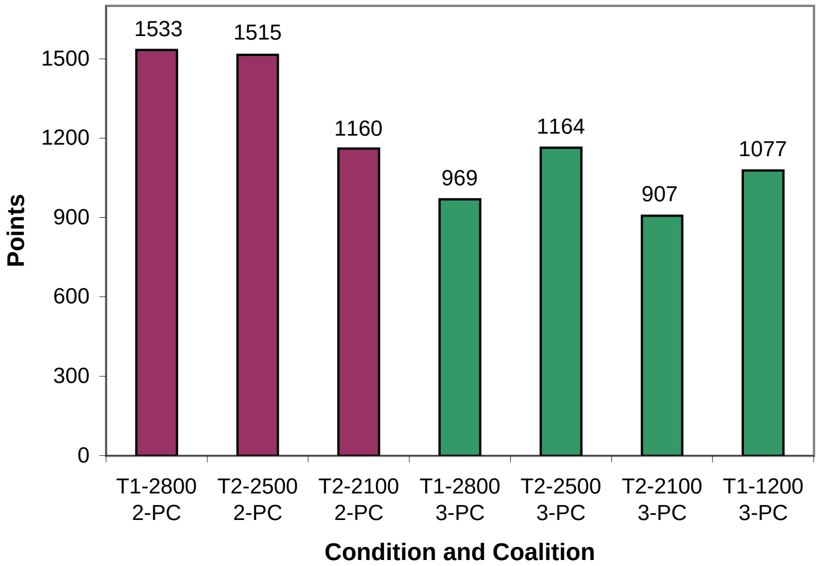


Figure 6 – Average Actual Earnings of Proposers across Rounds

earns considerably more when choosing a two-person coalition than when choosing the grand coalition; and even for the low value of 2100 proposers serve better in two- than in three person coalitions. Based on this information it seems obvious that an income maximizing proposer should choose the two-person coalition if its coalition value is high. One might argue, however, that only fair-minded proposers choose the three-person coalition leading to a downward bias on proposer earnings in these coalitions. We have therefore calculated the income maximizing offer ranges using the empirically observed acceptance rates. The results of this exercise are given in Table IV. The table depicts the income maximizing offer ranges x^*, and the maximal expected income (in points) belonging to the maximizing offers, for two- and three-person coalitions in the different conditions.[15] For the sake of comparison the table also contains the average offers across the last two periods in the respective condition and coalition.

The values of $maxE\pi$ clearly show that given the responders' behavior expected money income for proposers is maximized in two-person coalitions with high coalition value. Even for the intermediate coalition value the maximal proposer income in a three-person coalition is smaller than in a two-person coalition. Note that, on average, proposers make offers that come surprisingly close to the optimal offers. In our view, these observations together with

[15]For the calculation of the maximizing offers in three-person coalitions we assumed for simplicity equal offers to both responders. In T1-2800 and T2-2500 we observe to few three-person coalitions to obtain any statistically reliable results.

Table IV – Proposer's Actual and Income Maximizing Offers

	Offers		
	Mean	x^*	$maxE\pi$
2-Pers.Coal.:			
T1-2800	0.303	$]0.25, 0.30]$	$[1715, 1835[$
T1-2500	0.310	$]0.30, 0.35]$	$[1543, 1660[$
T2-2100	0.345	$]0.35, 0.40]$	$[1213, 1311[$
3-Pers.Coal.:			
T2-2100	0.232	$]0.25, 0.30]$	$[1030, 1282[$
T1-1200	0.252	$]0.25, 0.30]$	$[961, 1196[$

Note: Actual mean offers are based on the last two rounds. For all calculations the independent groups are used as units of observation. For the calculation of x^* in three-person coalitions equal treatment of both responders is assumed. x^* denotes therefore an offer made to both responders.

the other results provide ample evidence that allows us to conclude that proposers act as if they are expected money maximizers under the constraint of negatively reciprocally behaving responders.[16]

Result 6 *Together, reciprocal behavior of responders in two- and three-person coalitions and (seemingly) selfish behavior of proposers necessarily lead to social exclusion and (materially) inefficient coalition formation if the value of the small coalition is sufficiently high.*

4.2 Models of Social Preferences

Recently developed models of social preferences (Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999; Charness and Rabin, 2002) assume that people are not only motivated by their own money income but also by a taste for equity.[17] Can these models explain the results

[16] This result is in line with the theoretical model of Bolton (1991) and also matches experimental evidence from standard ultimatum game experiments (Roth *et al.*, 1991; Prasnikar and Roth, 1992). Charness and Rabin (2002), however, find in a different context no evidence that supports the view that first mover behavior is an optimizing response to the fear of rejection.

[17] We are focusing on these outcome oriented models, which as will turn out predict the behavior in our game quite well. For models taking intentions into account see Rabin, 1993; Dufwenberg and Kirchsteiger, 1998; Falk and Fischbacher, 1998.

obtained in our experiment? In the following we present theoretical predictions of the model by Fehr and Schmidt (1999) (hereafter FS) and Bolton and Ockenfels (2000) (hereafter BO). Thereby we confine ourself to those predictions we can relate to the results obtained in our experiment. Thereafter we shortly discuss the model of Charness and Rabin, 2002. The formal proofs of the statements are relegated to Appendix A.

The Fehr and Schmidt model. In the FS model, every player $i(=1,\cdots,n)$ is assumed to be endowed with a utility function of the following form: for a monetary payoff distribution $x=(x_1,\cdots,x_n)$ among n players the utility function of player i is given by

$$U_i(x) = x_i - \alpha_i \frac{1}{n-1}\sum_{j\neq i}|x_j - x_i|^+ - \beta_i \frac{1}{n-1}\sum_{j\neq i}|x_i - x_j|^+ \tag{4.1}$$

where $z^+ = max(z,0)$. The two parameters α_i and β_i ($\beta_i \leq \alpha_i$ and $0 \leq \beta_i < 1$) are considered to measure player i's utility loss from disadvantageous and advantageous inequality, respectively. In the context of our game $n=3$ and $i=P,1,2$ for the proposer and the two responders.

Theorem 1 RESPONDER BEHAVIOR
The subgame perfect equilibrium point of the three-person ultimatum coalition formation game prescribes the following behavior of responder $i=1,2$ in two- and three-person coalitions:

1. *Suppose that the proposer P chooses the grand coalition and proposes a monetary distribution $x=(V-x_1-x_2,x_1,x_2)$ of the coalition value V, where x_i, $0\leq x_i\leq V$, is the offer to responder Ri.*
   (a) *If $\frac{V-x_j}{2}\leq x_i$ and $x_j\leq x_i$ then accept the proposal.*
   (b) *If $\frac{V-x_j}{2}\leq x_i$ and $x_i<x_j$ then accept the proposal if and only if $x_i\geq\frac{(\alpha_i+\beta_i)x_j-\beta_iV}{\alpha_i+2(1-\beta_i)}$.*
   (c) *If $x_i<\frac{V-x_j}{2}$ and $x_j\leq x_i$ then accept the proposal if and only if $x_i\geq-\frac{(\alpha_i+\beta_i)x_j-\alpha_iV}{2(1+\alpha_i)-\beta_i}$.*
   (d) *If $x_i<\frac{V-x_j}{2}$ and $x_i<x_j$ then accept the proposal if and only if $x_i\geq\frac{\alpha_i}{2+3\alpha_i}V$.*
2. *Suppose that the proposer P chooses a two-person coalition and proposes a monetary distribution $x=(v-x_i,x_i,0)$ of the coalition value v, where x_i, $0\leq x_i\leq v$ is the offer to the chosen responder Ri, then Ri accepts the offer if and only if $x_i\geq\frac{\alpha_i}{2(1+\alpha_i)-\beta_i}v$.*

The first part of Theorem 1 is illustrated in Figure 7. The shaded area shows the acceptance region for responder $R2$. It nicely shows that generally the responder's behavior depends not only on his own offer but also on the offer to the other responder (and thus on the proposer's demand, too). If we assume that the inequity parameters are randomly distributed then it is easy to deduce from the figure that, given the offer to the other responder, the likelihood that an offer is rejected decreases with the offer. Note that a responder's rejection region is much larger when he is treated worse compared to his fellow responder than when he is better off. Hence, we can also say that the rejection likelihood is larger when a responder receives a lower offer than the other responder. All this is in accordance with our Result 4.

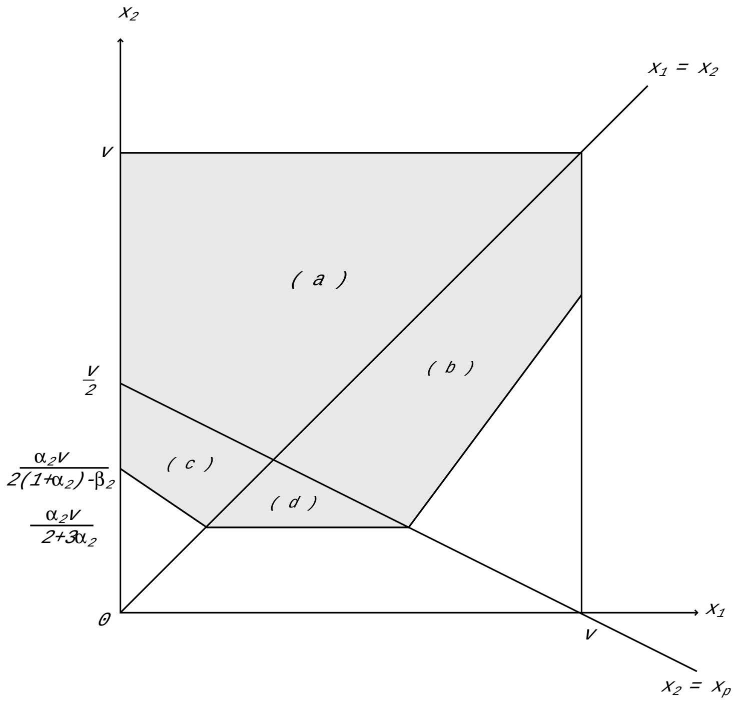


Figure 7 – Acceptance Region of $R2$ in the Three-Person Coalition

Part 2 of Theorem 1 shows that acceptance behavior in two-person coalitions depends only on the own relative share. That is, with respect to the relative offer responder behavior should be constant across two-person coalitions with different values. We have only very weak indications that responders' behavior may not be constant across treatments. We therefore conclude that the prediction of the FS model is consistent with our Result 3.

Theorem 2 Proposer behavior

Suppose that $\alpha_1 \geq \alpha_2$ *and* $v < V$. *Let* $K_i := K_i(\alpha_1, \alpha_2, \beta_P, \beta_1, \beta_i)$ $(i = 1, 2)$ *be given. The subgame perfect equilibrium point of the three-person ultimatum coalition formation game prescribes the following behavior of the proposer:*

1. *If* $2/3 \leq \beta_P < 1$ *then the proposer* P *chooses the three-person coalition and proposes the equity distribution* $(V/3, V/3, V/3)$, *regardless of the two-person coalition values* $v < V$.
2. *If* $0 \leq \beta_P < 2/3$ *then the proposer's behavior depends on the two-person coalitional value* v *as follows:*

(a) If $VK_i \geq v$ $(i = 1, 2)$ then the proposer chooses the three-person coalition and proposes the payoff distribution $x = (x_P, x_1, x_2)$ satisfying

$$1/3 < x_P = V - x_1 - x_2, \ x_2 = \frac{\alpha_2}{2 + 3\alpha_2} V \leq x_1 = -\frac{(\alpha_1 + \beta_1)x_2 - \alpha_1 V}{2(1 + \alpha_1) - \beta_1} < 1/3.$$

(b) If $v_i > VK_i$ (for at least one i) then the proposer chooses the two-person coalition with the responder Ri $(i = 1, 2)$ satisfying $v_i > VK_i$. If both $v_1 > VK_1$ and $v_2 > VK_2$ hold then the proposer chooses the responder Ri $(i = 1, 2)$ satisfying $\alpha_i/(2 - \beta_i) < \alpha_j/(2 - \beta_j)$ $(i \neq j)$. The proposer offers the payoff distribution $x = (x_P, x_i, x_j)$ satisfying

$$x_P = v_i - x_i, \ x_i = \frac{\alpha_i}{2(1 + \alpha_i) - \beta_i} v_i < v_i/2, \ x_j = 0.$$

The first part of Theorem 2 implies that only proposers who strongly dislike it to be better off always choose the grand coalition. In view of our Result 1 we may conclude that the advantageous inequality measure is smaller than 2/3 for basically all our proposers.

The second part of Theorem 2 shows that a two-person coalition is chosen if and only if its value v exceeds a particular threshold value that depends on the players' inequality parameters. Hence, the FS model allows for two-person coalitions in equilibrium. In this sense the model's prediction is consistent with our Result 1. Additionally, if we assume that the utility loss measures α_i and β_i are randomly distributed, the likelihood that condition $VK_i < v_i$ is satisfied for some $i = 1, 2$ increases as the two-person coalitional values v_i becomes larger. In this sense the model's prediction is also consistent with our Result 2.

When $\alpha_1 = \alpha_2 = \alpha$, the proposer treats the responders equally in the three-person coalition and offers them a relative share that is independent of the small coalition value v. Furthermore, a proposer demands more than one-third of the pie. All these are consistent with our Result 4(ii). Notice also that the relative share offered by the proposer in a two-person coalition is independent of the value of this coalition (Theorem 2 2.(b)). We only find weak evidence that proposers are more greedy in coalitions with higher values. Hence, the model's prediction can also be regarded as consistent with our Result 3.

Observe that, due to the 'self-centered' fairness notion implicit in the FS model (when $\beta < 2/3$), a proposer choosing the grand coalition offers the less (disadvantageous) inequity avers responder, $R2$ $(\alpha_2 \leq \alpha_1)$, the smaller part of the pie. Similarly, a proposer choosing a two-person coalition chooses the 'more selfish' responder as bargaining partner. The reason for this is simply that such responders accept lower material offers and, hence, give the proposer the opportunity of higher monetary earnings. One might say that such proposers behave as if they were selfish money maximizers and they will choose the two-person coalition with high coalition value for this reason. This squares nicely with our Results 5 and 6.

The Bolton and Ockenfels model. The BO model also presumes that subjects are motivated by both their pecuniary payoff and their relative material standing. Specifically, every player $i(= 1, \cdots, n)$ is assumed to maximize the expected value of her motivation function $\nu_i = \nu_i(y_i, \sigma_i)$ where $y_i \geq 0$ is player i's monetary payoff, and σ_i is player i's relative share of the payoff which is defined by

$$\sigma_i = \sigma_i(y_i, c, n) = \begin{cases} y_i/c & \text{if} \quad c > 0 \\ l/n & \text{if} \quad c = 0 \end{cases} \qquad \text{(where } c = \textstyle\sum_{j=1}^n y_i\text{)}$$

BO make the following assumptions (ν_{ij} denotes the partial derivative of ν_i with respect to the j-th variable):

Assumption 1 *The function* ν_i *is continuous and twice differentiable on the domain of* (y_i, σ_i).

Assumption 2 $\nu_{i1}(y_i, \sigma_i) \geq 0$ *and* $\nu_{i11}(y_i, \sigma_i) \leq 0$. *Also, fixing* σ *and given two choices where* $\nu_i(y_i^1, \sigma) = \nu_i(y_i^2, \sigma)$ *and* $y_i^1 > y_i^2$, *player* i *chooses* (y_i^1, σ).

Assumption 3 $\nu_{i2}(y_i, \sigma_i) = 0$ *for* $\sigma_i(c, y_i, n) = 1/n$, *and* $\nu_{i22}(y_i, \sigma_i) < 0$.

Assumption 4 *Fixing* c, *let* $\nu_i^c(\sigma_i)$ *denote* $\nu_i(c\sigma_i, \sigma_i)$. $\nu_i^c(\sigma_i)$ *is strictly concave in* σ_i *for all* $c > 0$, *and* $\nu_i^c(1) > \nu_i(0, 1/n)$.[18]

Together these assumptions guarantee that, for each $c > 0$, there exists a unique $r_i = r_i(c) \in [1/n, 1]$ such that

$$r_i = arg \max_{0 \leq \sigma_i \leq 1} \nu_i(c\sigma_i, \sigma_i). \tag{4.2}$$

and a unique $s_i = s_i(c) \in]0, 1/n]$ such that

$$\nu_i(cs_i, s_i) = \nu_i(0, 1/n), \tag{4.3}$$

These threshold functions $r_i(c)$ and $s_i(c)$ characterize players' types and are assumed to be random variables with density functions f^r and f^s. It is assumed that for all $c > 0$, $f^r(r|c) > 0$ for $r \in [1/n, 1]$ and $f^s(s|c) > 0$ for $s \in]0, 1/n]$.

Following BO we analyze the properties of a perfect Bayesian equilibrium of the coalition formation game where each player's thresholds r and s are private information but the densities f^r and f^s are common knowledge.

[18]We need this inequality to guarantee that $\nu_i(c\sigma_i, \sigma_i) > \nu_i(0, 1/n)$ for any $\sigma_i > s_i(c)$.

Theorem 3 Responder behavior

In the perfect Bayesian equilibrium of the three-person ultimatum coalition formation game behavior of responder Ri ($i = 1, 2$) in two- and three-person coalitions is characterized as follows:

1. *Suppose that the proposer P chooses the grand coalition and proposes a monetary distribution $x = (V - x_1 - x_2, x_1, x_2)$ of the coalition value V, where x_i, $0 \leq x_i \leq V$ ($i = 1, 2$), is the offer to responder Ri. Let $p_i = p_i(V, \sigma_i)$ denote the probability that a randomly selected responder $Ri(i = 1, 2)$ rejects such a proposal, then*
   (a) *$p_i(V, 0) = 1$ and $p_i(V, \sigma_i) = 0$ for all $\sigma_i \in [1/3, 1]$; furthermore, $p_i(V, \sigma_i)$ is strictly decreasing in σ_i on the interval $]0, 1/3[$.*
   (b) *Given the offer to Ri, the rejection probability p_i is independent of the offer to the other responder.*
2. *Suppose that the proposer P chooses the small coalition and proposes a monetary distribution $x = (v - x_i, x_i, 0)$ of the coalition value v, where x_i, $0 \leq x_i \leq v$ ($i = 1, 2$), is the offer to the chosen responder Ri. Let $p_i = p_i(v, \sigma_i)$ denote the probability that a randomly selected responder $Ri(i = 1, 2)$ rejects such a proposal, then*
   (a) *$p_i(v, 0) = 1$ and $p_i(v, \sigma_i) = 0$ for all $\sigma_i \in [1/3, 1]$; furthermore, $p_i(v, \sigma_i)$ is strictly decreasing in σ_i on the interval $]0, 1/3[$.*
   (b) *Fixing $\sigma_i \in]0, 1/3]$, $p_i(v, \sigma_i)$ is non-increasing in v.*

Part 1 of this theorem states that responders behave reciprocally in three-person coalitions and, specifically, predicts that offers above one-third of the grand coalition value are always accepted. Both these predictions are in accordance with the behavior observed in our experiment (Result 4 and Figure 4). The theorem also tells us that responder's acceptance behavior should be independent of the offer to the other responder, given his own offer. Theoretically this is the case, because the BO model assumes that people only take their relative standing with respect to the whole group into account. This implies that, given an offer, redistributing money between the proposer and the other responder does not affect the value of the motivation function of the responder and, thus, also not his behavior. Since, we have evidence that the relative standing influences rejection behavior in three-person coalitions this prediction is not consistent with this part of our Result 4.

The second part of Theorem 3 states that also in two-person coalitions responders behave reciprocally, which is clearly consistent with our observations. Specifically, it also predicts that all offers above one-third of the small coalition value are accepted for sure. We actually observe relatively high acceptance rates for relatively low offers. For $v = 2100$, $v = 2500$, and

$v = 2800$ the rejection rates for offers above $v/3$ are 8.2, 1.2, 3.5 percent. In our view, the model's prediction approximates real behavior quite well. Hence, this prediction of the BO model are in accordance with our Result 3.

The last part of Theorem 3 states that the rejection rate in two-person coalitions are at least not increasing in the two-person coalition value v. Since we do not have strong indications that responder behavior changes with v we regard this prediction as in accordance with our observations (Result 3).

Theorem 4 Proposer behavior

In the perfect Bayesian equilibrium of the three-person ultimatum coalition formation game proposer behavior in two- and three-person coalitions with values v and V is characterized by the maximization of the expected motivation functions

$$E\nu_P(v, \sigma_P) = (1 - p_i(v, \sigma_i)) \cdot \nu_P(v\sigma_P, \sigma_P), \;\; \sigma_i = 1 - \sigma_P, i = 1, 2,$$

$$E\nu_P(V, \sigma_P) = (1 - p_1(V, \sigma_1)) \cdot (1 - p_2(V, \sigma_2)) \cdot \nu_P(z\sigma_P, \sigma_P), \;\; \sigma_i = 1 - \sigma_j - \sigma_P, \;\; i = 1, 2, i \neq j.$$

The maximum value $E\nu_P{}^*(v)$ of the proposer's expected motivation function in a two-person coalition is nondecreasing in the value v.

The proposer chooses a two-person coalition if and only if her optimal expected motivation value in the two-person coalition is greater than that in the three-person coalition. The above Theorem shows that the former is nondecreasing in the two-person coalition value v, while the latter is independent of v. This implies that the likelihood of two-person coalitions increases with the two-person coalition value v. In this sense the prediction of the BO model is consistent with our Results 1 and 2.

Note, that if two responders are drawn from the same pool of subjects, which means that they have the same rejection probabilities, then the proposer's optimal behavior is to treat them equally. Furthermore, since offers above one-third of the coalition value are always accepted the assumptions made by BO imply that offers in three-person coalitions will not exceed $V/3$. A similar argument holds for two-person coalitions. Hence, in this respect, proposer behavior predicted by the BO model is consistent with our observations reported in Results 3 and 4.

In summary we can conclude that in qualitative terms both, the FS model and the BO model, predict real behavior of proposers and responders in our three-person ultimatum coalition formation game surprisingly well.

In another recent model Charness and Rabin (2002) argue that experimental subjects are more concerned with increasing social welfare and helping worst-off people rather than by self-centered inequality aversion. They show that in a number of predominantly two-person games their model indeed performs better than other models of (non-)social preferences. They admit, however, that their (basic) model cannot explain rejections in ultimatum games. Similarly, the evidence observed in our experiment stays in stark contrast to the predictions of their model. The choice of a two-person coalition decreases social welfare and makes one person as worse-off as possible. We do, of course, not deny the existence of people who take the welfare of the worst off into account. What our results, however, suggest is that this disposition is not strong enough to overcome the anticipated increased rejection probability, and thus the decreased expected material welfare, in case of a three-person coalition.[19]

The inefficient outcomes observed in our experiment may also be compared with theoretical results in the literature on non-cooperative sequential bargaining models of coalition formation, initiated by Selten (1981). In this literature it has been shown that inefficient sub-coalitions may be formed in equilibrium in Rubinstein (1982) type sequential bargaining models of coalition formation even under complete information about coalition values (Chatterjee *et al.*, 1993; Okada, 1996).[20] In these models the reason of inefficienct outcomes is that the minimum acceptance levels of responders become larger than zero, being equal to the (discounted) value of their continuation payoffs in future negotiations. With rational expectations about responders' continuation payoffs it may be optimal for proposers to choose inefficient allocations. We point out, however, that inefficiency in these sequential models is very different from that induced by (anticipated) rejections of low offers as observed in our experiments. The anticipation of further negotiations can not play any role in our experiments by definition.

5 Conclusions

In this paper we provide experimental evidence that anticipated and actual reciprocal actions strongly affect coalition formation in multilateral bargaining. In particular, we observe that an overwhelming majority of subjects choose inefficient subcoalitions. They are ready to forego resources and to increase distributional inequality, by excluding other subjects from bargaining. We argue that the undesirable result of inefficiency and social exclusion is unavoidable when responders behave reciprocally and proposers act as income maximizers. We

[19]In line with our observations, Kagel and Wolfe (2001) also report evidence of a sort of third-party-neglect where people involved in strategic interactions seem not to take into account the well-being of a third inactive player.

[20]Uhlich (1989) reports data of an experiment based on the bargaining model of Selten (1981).

also compare the predictions of recently developed models of social preferences (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000) with our experimental results. We find that these models capture the empirical regularities surprisingly well, at least in a qualitative sense.

Other experimental studies have shown that under some important institutions reciprocal fairness can be a powerful force that leads to quite efficient and rather fair outcomes. Our study provides the complementary evidence that the same force may lead to precisely opposite consequences, under another important institutional environment. The evidence given in this paper may also shed some new light on the on-going debate about efficiency, inequality-aversion, and reciprocity, in particular, in the context of coalition formation. Needless to say that much more work is necessary for a better understanding of the interaction of reciprocal behavior and economic institutions and its likely consequences.

6 References

Axelrod, R. (1984). *The Evolution of Cooperation.* Basic Books, New York.

Berg, J., Dickhaut, J., and McCabe, K. (1995). "Trust, Reciprocity and Social History," *Games and Economic Behavior*, **10**, 122-142.

Bolton, G.E. (1991). "A Comparative Model of Bargaining: Theory and Evidence, *American Economic Review*, **81**, 1096-1136.

Bolton, G.E., and Chatterjee, K. (1996). "Coalition Formation, Communication, and Coordination: An Exploratory Experiment," in *Wise choices: decisions, games, and negotiations* (R.J. Zeckhauser, R.L. Keeney and J.K. Sebenius, Eds.), pp. 253-271. Boston: Harvard Business School Press.

Bolton, G.E., Chatterjee, K., McGinn K.L. (2003). "How Communication Links Influence Coalitional Bargaining: A Laboratory Investigation", *Management Sciences*, **49**, 583-598.

Bolton, G., and Ockenfels, A. (1998). "Strategy and Equity: An ERC-Analysis of the Güth-van Damme Game", *Journal of Mathematical Psychology*, **42**, 215-226.

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

Camerer, C. (2002). *Behavioral Game Theory: Experiments on Strategic Interaction*, forthcoming.

Charness, G., and Rabin, M. (2002). "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics*, **117**, forthcoming.

Chatterjee, K., Dutta, B., Ray, D., and Sengupta, K. (1993). "A Noncooperative Theory of Coalitional Bargaining," *Review of Economic Studies*, **60**, 463-477.

Dufwenberg, M., and Kirchsteiger, G. (1998). "A Theory of Sequential Reciprocity," Discussion Paper No. 37, CentER, Tilburg University.

Falk, A., and Fischbacher, U. (1998). "A Theory of Reciprocity," Working Paper No. 6, Institute for Empirical Research in Economics, University of Zürich.

Fehr, E., Gächter, S., and Kirchsteiger, G. (1997). "Reciprocity as a Contract Enforcement Device: Experimental Evidence," *Econometrica*, **65**, 833-860.

Fehr, E., Kirchsteiger, G., and Riedl, A. (1993). "Does Fairness Prevent Market Clearing? An Experimental Investigation," *Quarterly Journal of Economics*, **108**, 437-460.

Fehr, E., Kirchsteiger, G., and Riedl, A. (1998). "Gift Exchange and Reciprocity in Competitive Experimental Markets," *European Economic Review*, **42**, 1-34.

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

Güth, W., and Van Damme, E. (1998). "Information, Strategic Behavior and Fairness in Ultimatum Bargaining - An Experimental Study," *Journal of Mathematical Psychology*, **42**, 227-247.

Güth, W., and Huck, S. (1997). "From Ultimatum Bargaining to Dictatorship - An experimental Study of Four Games Varying in Veto Power," *Metroeconomica*, **48**, 262-279.

Güth, W., Huck, S., and Ockenfels, P. (1996). "Two-Level Ultimatum Bargaining with Incomplete Information," *The Economic Journal*, **106**, 593-604.

Güth, W., Schmittberger, R., and Schwarze, B. (1982). "An Experimental Analysis of Ultimatum Bargaining," *Journal of Economic Behavior and Organization*, **3**, 367-388.

Kagel, J.H., and Wolfe, K. (2001). "Tests of Fairness Models Based on Equity Considerations in a Three-Person Ultimatum Game," *Experimental Economics*, **4**, 203-219.

Knez, M., and Camerer, C. (1995). "Social Comparison and Outside Options in Three-Person Ultimatum Games," *Games and Economic Behavior*, **10**, 65-94.

Okada, A. (1996). "A Noncooperative Coalitional Bargaining Game with Random Proposers," *Games and Economic Behavior*, **16**, 97-108.

Okada, A., and Riedl, A. (2001). "Reciprocity, Inefficiency, and Social Exclusion: Experimental evidence," working paper, CREED, Department of Economics, University of Amsterdam.

Prasnikar, V., and Roth, A.E. (1992). "Considerations of Fairness and Strategy: Experimental Data from Sequential Games," *Quarterly Journal of Economics*, **107**, 865-888.

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

Riedl, A., and Vyrastekova, J. (2003). "Responder Behavior in Three-Players Ultimatum Game Experiments," CentER for Economic Research discussion paper 2002-5, Tilburg University.

Roth, A.E. (1995). "Bargaining Experiments," in *The Handbook of Experimental Economics* (J. Kagel, and A.E. Roth, Eds.), pp. 253-348. Princeton: Princeton University Press.

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

Rubinstein, A. (1982). "Perfect Equilibrium in a Bargaining Model," *Econometrica*, **50**, 97-109.

Selten, R. (1981). "A Noncooperative Model of Characteristic-Function Bargaining," in *Essays in Game Theory and Mathematical Economics in Honor of Oscar Morgenstern* (V. Böhm, and H. Nachthanp, Eds.), pp. 139-151. Mannheim: Bibliographisches Institut Mannheim.

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

Uhlich, G.R. (1989). *Descriptive Theories of Bargaining*, Lecture Notes in Economics and Mathematical Systems, 341. Berlin: Springer Berlin.

A Formal proofs of the predictions of the models of social preferences (Theorems 1-4)

A.1 The Fehr and Schmidt Model (Theorems 1-2)

We now prove Theorems 1 and 2.

(1) Responders' behavior in the three-person coalition. Suppose that P proposes a payoff distribution $x = (V - x_1 - x_2, x_1, x_2)$ of V where $x_i, 0 \le x_i \le V,\ (i = 1, 2)$ is the material payoff for responder Ri. $R1$ and $R2$ either accept or reject the sequentially. We apply backward induction and first analyze the optimal response of $R2$. $R2$'s utility when accepting $x = (V - x_1 - x_2, x_1, x_2)$ is

$$u_2(x) = x_2 - \frac{\alpha_2}{2}\{|V - x_1 - 2x_2|^+ + |x_1 - x_2|^+\} - \frac{\beta_2}{2}\{|x_1 + 2x_2 - V|^+ + |x_2 - x_1|^+\}$$

and zero when rejecting.
We have to consider four cases with which we deal in turn:
(a) $(V - x_1)/2 \le x_2$ and $x_1 \le x_2$, (b) $(V - x_1)/2 \le x_2$ and $x_2 < x_1$, (c) $x_2 < (V - x_1)/2$ and $x_1 \le x_2$, (d) $x_2 < (V - x_1)/2$ and $x_2 < x_1$.

Case (a): $R2$'s utility reduces to $u_2(x) = x_2 - \frac{\beta_2}{2}(3x_2 - V)$, which greater than zero for all $\beta_2 \in [0, 1[$ $(0 \le x_2 \le V)$. Hence, the optimal response is to accept the proposal x.

Case (b): $R2$'s utility reduces to $u_2(x) = x_2 - \frac{\alpha_2}{2}(x_1 - x_2) - \frac{\beta_2}{2}(x_1 + 2x_2 - V)$ and the proposal is accepted if and only if $u_2(x) \ge 0$,[21] which is equivalent to

$$x_2 \ge \frac{(\alpha_2 + \beta_2)x_1 - \beta_2 V}{\alpha_2 + 2(1 - \beta_2)}.$$

Case (c): $R2$'s utility reduces to $u_2(x) = x_2 - \frac{\alpha_2}{2}(V - x_1 - 2x_2) - \frac{\beta_2}{2}(x_2 - x_1)$, and analogously to case (b), the proposal is accepted if and only if

$$x_2 \ge -\frac{(\alpha_2 + \beta_2)x_1 - \alpha_2 V}{2(1 + \alpha_2) - \beta_2}.$$

Case (d): $R2$'s utility $u_2(x)$ reduces to $u_2(x) = x_2 - \frac{\alpha_2}{2}(V - 3x_2)$, and $R2$ accepts the proposal if and only if $u_2(x) \ge 0$ or equivalently

$$x_2 \ge \frac{\alpha_2}{2 + 3\alpha_2}V.$$

Since the optimal response of $R1$ can be proved in an equivalent way this proves Part 1 of Theorem 1.

(2) Responder's behavior in a two-person coalition. Consider a two-person coalition, with coalition value v_i, of P and an Ri $(i = 1, 2)$. Ri's behavior can be analyzed by setting $V = v_i$ and $x_j = 0$ $(j \ne i)$ in his optimal response in cases (a) and (c) above. It follows then immediately

[21] $R2$ is indifferent between accepting and rejecting the proposal when the equality holds, and the proposer's equilibrium condition induces the acceptance of $R2$ in the case of equality.

that the optimal response to a proposal $x = (x_P, x_i, x_j)$ $(x_P + x_i = v_i,\ x_j = 0,\ 0 \le x_i \le v_i)$ in the two-person coalition is to accept it if and only if

$$x_i \ge \frac{\alpha_i}{2(1+\alpha_i) - \beta_i} v_i.$$

This proves Part 2 of Theorem 1.

(3) Proposer's behavior in the three-person coalition. Let the optimal responses $R1$ and $R2$ be given and assume without loss of generality $\alpha_1 \ge \alpha_2$. Note that the region of all proposals $x = (V - x_1 - x_2, x_1, x_2)$ accepted by both responders can be divided into six subregions (for an illustration see Figure8:

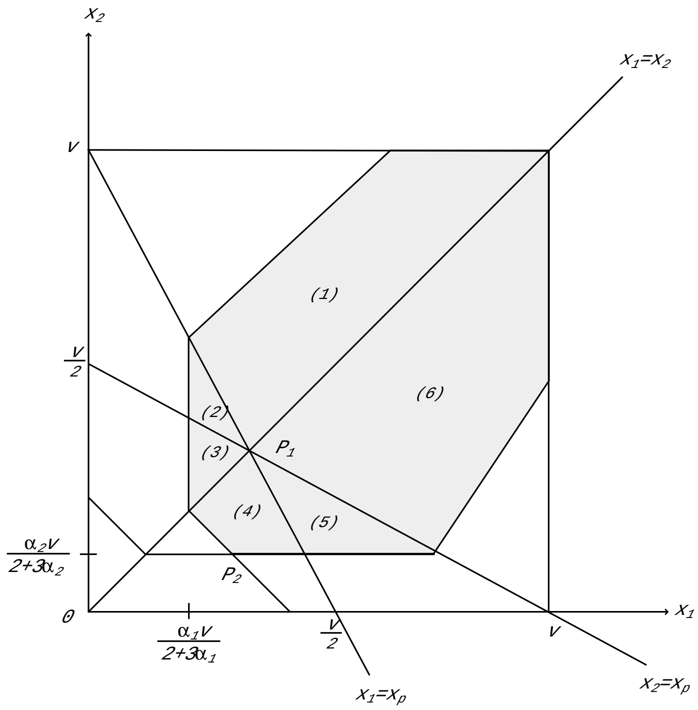


FIGURE 8 – REGIONS OF ACCEPTANCE IN THREE-PERSON COALITION $(\alpha_1 \ge \alpha_2)$

(1) $x_P \le x_1 \le x_2, \quad \dfrac{(\alpha_1+\beta_1)x_2 - \beta_1 V}{\alpha_1 + 2(1-\beta_1)} \le x_1,$

(2) $x_1 \le x_P \le x_2, \quad \dfrac{\alpha_1}{2+3\alpha_1} \le x_1,$

(3) $x_1 \le x_2 \le x_P, \quad \dfrac{\alpha_1}{2+3\alpha_1} \le x_1,$

(4) $x_2 \le x_1 \le x_P, \quad \dfrac{\alpha_2}{2+3\alpha_2} \le x_2, \ -\dfrac{(\alpha_1+\beta_1)x_2 - \alpha_1 V}{2(1+\alpha_1) - \beta_1} \le x_1,$

(5) $x_2 \le x_P \le x_1, \quad \dfrac{\alpha_2}{2+3\alpha_2} \le x_2,$

(6) $x_P \le x_2 \le x_1, \quad \dfrac{(\alpha_2+\beta_2)x_1 - \beta_2 V}{\alpha_2 + 2(1-\beta_2)} \le x_2.$

The proposer P's utility for $x = (x_P, x_1, x_2),\ x_P = V - x_1 - x_2,\ 0 \le x_1,\ x_2 \le V$, is given by

$$u_P(x) = x_P - \frac{\alpha_P}{2}\left\{|x_1 - x_P|^+ + |x_2 - x_P|^+\right\} - \frac{\beta_P}{2}\left\{|x_P - x_1|^+ + |x_P - x_2|^+\right\}.$$

We next characterize the optimal proposal in each subregion.

Regions (1) and (6): In these cases, the proposer's utility $u_P(x)$ from an accepted proposal $x = (x_P, x_1, x_2)$ reduces to

$$u_P(x) = x_P - \frac{\alpha_P}{2}(x_1 - x_P + x_2 - x_P) = (1+\alpha_P)V - (a + \frac{3}{2}\alpha_P)(x_1 + x_2).$$

Since $u_P(x)$ is decreasing in both x_1 and x_2 it follows that the optimal proposal in these regions is given by $x_P = x_1 = x_2 = \frac{1}{3}V$ leading to the utility $u_P = \frac{1}{3}V$.

Region (2): In this case the proposer's utility is given by

$$u_P(x) = x_P - \frac{\alpha_P}{2}(x_2 - x_P) - \frac{\beta_P}{2}(x_P - x_1) = (1 + \frac{\alpha_P}{2} - \frac{\beta_P}{2})V - A_1 x_1 - A - 2x_2,$$

with $A_1 := 1 + \alpha_P/2 - \beta_P > 0$ and $A_2 := 1 + \alpha_P - \beta_P/2 > 0$. Since $A_1, A_2 > 0$ the optimal choice lies on the line $x_2 = x_p$ and depends the values of A_1, A_2. It follows that the optimal proposal is given by

$$\begin{array}{ll} x_P = x_1 = x_2 = \frac{V}{3} & \text{if} \quad 2A_1 \leq A_2, \text{ i.e. } 2/3 \leq \beta_P < 1 \\ x_P = \frac{1+\alpha_1}{2+3\alpha_1}V = x_2,\ x_1 = \frac{\alpha_1 V}{2+3\alpha_1}, & \text{if} \quad 2A_1 > A_2, \text{ i.e., } 0 \leq \beta_P < 2/3, \end{array}$$

leading to a utility of

$$\begin{array}{ll} u_P(x) = \frac{V}{3} & \text{if} \quad 2/3 \leq \beta_P < 1 \\ u_P(x) = x_P - \frac{\beta_P}{2}(x_P - x_1) = \frac{2+2\alpha_1-\beta_P}{2(2+3\alpha_1)}V \quad (> \frac{V}{3}) & \text{if} \quad 0 < \beta_P < 2/3. \end{array}$$

Region (5): Similar to the above case the proposer's utility is given by

$$u_P(x) = x_P - \frac{\alpha_P}{2}(x_1 - x_P) - \frac{\beta_P}{2}(x_P - x_2) = (1 + \frac{\alpha_P}{2} - \frac{\beta_P}{2})x_P - A_2 x_1 - A_1 x_2,$$

with A_1 and A_2 as above. Now the optimal proposal lies on the line where $x_1 = x_P$ and is given by

$$\begin{array}{ll} x_P = x_1 = x_2 = V/3 & \text{if} \quad 2A_1 \leq A_2, \text{ i.e., } 2/3 \leq \beta_P < 1, \\ x_P = x_1 = \frac{1+\alpha_2}{2+3\alpha_2}V,\ x_2 = \frac{\alpha_2}{2+3\alpha_2}V & \text{if} \quad 2A_1 > A_2, \text{ i.e., } 0 \leq \beta_P < 2/3, \end{array}$$

giving a utility

$$\begin{array}{ll} u_P(x) = V/3 & \text{if} \quad 2/3 \leq \beta_P < 1, \\ u_P(x) = x_P - \frac{\beta_P}{2}(x_P - x_2) = \frac{2+2\alpha_2-\beta_P}{2(2+3\alpha_2)}V & \text{if} \quad 0 < \beta_P < 2/3. \end{array}$$

Regions (3) and (4): In these cases, the proposer's utility reduces to

$$u_P(x) = x_P - \frac{\beta_P}{2}(2x_P - x_1 - x_2) = (1 - \beta_P)V + (\frac{3}{2}\beta_P - 1)(x_1 + x_2).$$

If $\frac{3}{2}\beta_P - 1 > 0$, it follows that the equity proposal is optimal, and if $\frac{3}{2}\beta_P - 1 < 0$ the optimal proposal lies on the south-west boundary of region (4). Using this it follows that the optimal proposal is

$$\begin{array}{ll} x_P = x_1 = x_2 = V/3 & \text{if} \quad 2/3 \leq \beta_P < 1, \\ x_P = V - x_1 - x_2,\ x_1 = -\frac{(\alpha_1+\beta_1)x_2 - \alpha_1 V}{2(1+\alpha_1)-\beta_1},\ x_2 = \frac{\alpha_2 V}{2+3\alpha_2} & \text{if} \quad 0 \leq \beta_P < 2/3, \end{array}$$

with corresponding utilities

$$u_P(x) = V/3 \qquad \text{if} \quad 2/3 \leq \beta_P < 1, \tag{A.1}$$

$$u_P(x) = \left\{1 - \beta_P - \tfrac{2-3\beta_P}{2+3\alpha_2}(\alpha_2 + \tfrac{\alpha_1-\alpha_2}{2+2\alpha_1-\beta_1})\right\} V \qquad \text{if} \quad 0 \leq \beta_P < 2/3. \tag{A.2}$$

Comparing utility from the optimal proposals across the six acceptance regions implies that the equilibrium proposal (x_P, x_1, x_2) in the three-person coalition is characterized as follows (see also points P_1 and P_2 in Figure 8):

(1) If $2/3 \leq \beta_P < 1$, then the equilibrium proposal satisfies $x_P = x_1 = x_2 = V/3$.

(2) If $0 \leq \beta_P < 2/3$, then the equilibrium proposal satisfies

$$x_1 = -\frac{(\alpha_1+\beta_1)x_2 - \alpha_1 V}{2(1+\alpha_1)-\beta_1}, x_2 = \frac{\alpha_2}{2+3\alpha_2}V, \text{ with } x_2 < x_1, x_P.$$

It is easily shown that the corresponding utilities are larger than zero. Hence acceptance is better than rejection.

(4) Proposer's behavior in a two-person coalition. Without loss of generality, consider a two-person coalition of proposer P and responder $R2$. The proposer's utility for a payoff distribution $x = (v_2 - x_2, 0, x_2)$ is given by

$$u_P(x) = v_2 - x_2 - \frac{\alpha_P}{2}|2x_2 - v_2|^+ - \frac{\beta_P}{2}\left\{v_2 - x_2 + |v_2 - 2x_2|^+\right\}.$$

We have to consider the following two cases:

(i) $\frac{v_2}{2} \leq x_2 \leq v_2$ and (ii) $\frac{\alpha_2}{2(1+\alpha_2)-\beta_2}v_2 \leq x_2 \leq \frac{v_2}{2}$.

Case (i): In this case the proposer's utility is given by

$$u_P(x) = \left(1 + \frac{\alpha_P}{2} - \beta_P\right)v_2 - \left(1 + \alpha_P + \frac{\beta_P}{2}\right)x_2.$$

Since $1 + \alpha_P + \frac{\beta_P}{2} > 0$ this implies a optimal proposal $x_P = x_2 = \frac{v_2}{2}$ and the proposer can obtain the utility $u_P(x) = \frac{2-\beta_P}{4}v_2$.

Case (ii): Here the proposer's utility is

$$u_P(x) = (1-\beta_P)v_2 + \left(\frac{3}{2}\beta_P - 1\right)x_2.$$

In view of this, the proposer's optimal choice is

$$x_P = x_2 = v_2/2 \qquad \text{if} \quad 2/3 \leq \beta_P < 1,$$

$$x_P = \tfrac{2+\alpha_2-\beta_2}{2(1+\alpha_2)-\beta_2}v_2,\ x_2 = \tfrac{\alpha_2}{2(1+\alpha_2)-\beta_2}v_2 \qquad \text{if} \quad 0 \leq \beta_P < 2/3,$$

with corresponding utility

$$u_P(x) = \tfrac{2-\beta_P}{4}v_2 \qquad \text{if} \quad 2/3 \leq \beta_P < 1, \tag{A.3}$$

$$u_P(x) = \left\{1 - \beta_P - \tfrac{2-3\beta_P}{2}\tfrac{\alpha_2}{2(1+\alpha_2)-\beta_2}\right\} v_2 \qquad \text{if} \quad 0 \leq \beta_P < 2/3 \tag{A.4}$$

(5) The proposer's coalitional choice.

Consider first the case $2/3 \leq \beta_P < 1$. Comparing the proposer's maximal utility in the three-person coalition (see A.1) with the maximum utility in attainable in the two-person coalition (see A.3) implies that the optimal choice is the three-person coalition in this case, independent of the two-person coalitional values v_i $(i = 1, 2)$.

Now consider the case $0 \leq \beta_P < 2/3$. With the help of (A.4) (replacing the index 2 with i) it is easy to show that the utility u_P^2 a proposer can maximally attain when choosing $R2$ as bargaining partner in a two-person coalition is at least as large as her utility u_P^1 with $R1$ as partner, i.e. $u_P^1 \leq u_P^2$ if and only if

$$\frac{\alpha_1}{2-\beta_1} \geq \frac{\alpha_2}{2-\beta_2}.$$

In the subgame perfect equilibrium point, the proposer chooses the three-person coalition if and only if her maximal utility in the three-person coalition is at least as large as her maximal utility u_P^i $(i = 1, 2)$ attainable in the two-person coalition. By comparing A.2) with A.2) one can show that the three-person coalition is chosen if and only if

$$VK_i \geq v_i, \quad \text{where} \quad K_i = \frac{1 - \beta_P - \frac{2-3\beta_P}{2+3\alpha_2}(\alpha_2 + \frac{\alpha_1-\alpha_2}{2+2\alpha_1-\beta_1})}{1 - \beta_P - \frac{2-3\beta_P}{2}\frac{\alpha_i}{2+2\alpha_i-\beta_i}}, \quad i = 1, 2. \tag{A.5}$$

This proves Theorem 2.

A.2 The Bolton and Ockenfels Model (Theorems 3-4)

We now prove Theorems 3 and 4.

(1) Responder's behavior. Consider the three-person coalition and suppose that P proposes a payoff distribution $x = (V - x_1 - x_2, x_1, x_2)$, $0 \leq x_i \leq V$, $(i = 1, 2)$, where x_i is the offer to the responder Ri. By backward induction, we first investigate the optimal response of $R2$. Let $\sigma_2 = x_2/V$ be the relative share of $R2$. The responder accepts the proposal if and only if the motivation value from accepting is at least as high as the motivation value from a rejection, that is, if and only if

$$\nu_{R2}(V\sigma_2, \sigma_2) \geq \nu_{R2}(0, 1/3). \tag{A.6}$$

Similarly, responder $R1$ accepts the proposal if both (A.6) and

$$\nu_{R1}(V\sigma_1, \sigma_1) \geq \nu_{R1}(0, 1/3). \tag{A.7}$$

hold.[22]

[22]Remark that $R1$ is indifferent between accepting and rejecting when $R2$ is expected to reject the proposal. That is, when (A.6) does not hold. However, if (A.6) indeed does not hold, $R1$'s response is inessential the sense that the proposal is rejected by $R2$, even if $R1$ accepts it. Since we are primarily interested in the case when a proposal is agreed by both responders, we assume that $R1$ accepts the proposal if and only if (A.7) holds.

Let

$$p_i(V,\sigma_i) := Prob\{s_i | s_i(V) \geq \sigma_i\} \tag{A.8}$$

denote the probability that a randomly selected responder Ri ($i = 1,2$) rejects a proposal $x = (V - x_1, x_2, x_1, x_2)$.

We are now ready to prove statement 1.(a) of Theorem 3. From Assumptions 2 and 3 it follows that $\nu_i(V\sigma_i, \sigma_i)$ is strictly increasing in σ_i on the interval $]0,1/3[$ since $\frac{\partial \nu_i}{\partial \sigma_i} = V \cdot \nu_{i1}(V\sigma_i, \sigma_i) + \nu_{i2}(V\sigma_i, \sigma_i) > 0$.

For any $s_i > 0$, responder Ri rejects the proposal yielding $\sigma_i = 0$ since $\nu_i(0,0) < \nu_i(Vs_i, s_i) = \nu_i(0,1/3)$. Thus, $p_i(V,0) = 1$.

Assumption 2 implies $\nu_i(V/3, 1/3) \geq \nu_i(0,1/3)$. Therefore, responder Ri accepts such a proposal. Thus, $p_i(V,1/3) = 0$.[23] Assumptions 3 and 4 imply $\nu_i(V\sigma_i, \sigma_i) \geq \nu_i(0,1/3)$ for all $\sigma_i \in [1/3, 1]$.

That $p_i(V,\sigma_i)$ strictly increases in σ_i on $]0,1/3[$ follows directly from (A.8).

That the rejection probability (given an offer to the responder) does not depend on the offer to the other responder (3 1(b)) is clear from the definition of the motivation function.

Next consider two-person coalitions. Similarly to the three-person coalition, let $p(v,\sigma_i)$ denote the probability that a randomly selected responder Ri rejects a proposal $x = (v - x_i, x_i)$. Statement 2.(a) of Theorem 3 can be proved in the same way as the corresponding statement for the three-person coalition.

We now show that, when fixing $\sigma_i \in]0,1/3]$, $p_i(v,\sigma_i)$ is non-increasing in v. To see this, differentiate the threshold function (4.3) with respect to v to get

$$s_i'(c) = -\frac{s_i \nu_{i1}(cs_i, s_i)}{c\nu_{i1}(cs_i, s_i) + \nu_{i2}(cs_i, s_i)} \leq 0. \tag{A.9}$$

Together with the assumptions on the threshold value function and (A.8) this implies that $p_i(V,\sigma_i)$ is non-increasing in V. This proves Theorem 3.

(2) Proposer's behavior. The formulations

$$\begin{aligned} E\nu_P(v,\sigma_P) &= (1 - p_i(v,\sigma_i)) \cdot \nu_P(v\sigma_P, \sigma_P),\ \sigma_i = 1 - \sigma_P, i = 1,2, \\ E\nu_P(V,\sigma_P) &= (1 - p_1(V,\sigma_1)) \cdot (1 - p_2(V,\sigma_2)) \cdot \nu_P(z\sigma_P, \sigma_P),\ \sigma_i = 1 - \sigma_j - \sigma_P,\ i = 1,2, i \neq j \end{aligned}$$

of the expected motivation values follow directly from the definition of the motivation function and the normalization $\nu_P(0,1/3) = 0$, where σ_P is the proposer's relative share and $p_i(v,\sigma_i)$ is the probability that Ri rejects the proposal.

Assuming an interior solution and denoting the maximum value of the proposer's expected motivation function in a two-person coalition by $E\nu_P^*(v)$, gives

$$E\nu_P^*(v) = (1 - p_i(v,\sigma_i^*)) \cdot \nu_P(v\sigma_P^*, \sigma_P^*),\ \sigma_i^* = 1 - \sigma_P^*. \tag{A.10}$$

[23]Note that, even if the equality holds, Assumption 2 guarantees the acceptance.

By the envelope theorem we have

$$\frac{dE\nu_P^*(v)}{dv} = \frac{\partial E\nu_P}{\partial v}(v, \sigma_P) = -\frac{\partial p_i}{\partial v}\nu_P + (1 - p_i) \cdot \sigma_P \cdot \nu_{P1} \geq 0,$$

where the inequality follows from Assumption 1 and the monotonicity of p_i in v. This proves Theorem 4.