No. 30

U N I V E R S I T Y O F C O L O G N E

W O R K I N G P A P E R S E R I E S I N E C O N O M I C S

HOW DO COALITIONS GET BUILT - EVIDENCE FROM AN EXTENSIVE FORM COALITION GAME WITH RENEGOTIATION & EXTERNALITIES

GARY E BOLTON
JEANNETTE BROSIG

Department of Economics
University of Cologne
Albertus-Magnus-Platz
D-50923 Köln
Germany

http://www.wiso.uni-koeln.de

How do coalitions get built

Evidence from an extensive form coalition game with renegotiation & externalities

Gary E Bolton
Penn State University, USA

Jeannette Brosig*
University of Cologne, Germany

We investigate a three-person coalition game in which one bargainer, the builder, can propose and build a coalition over two stages. In equilibrium, coalition building ends with an efficient grand coalition, while the equilibrium path is contingent on the values of the two-person coalitions and associated externality payoffs. Considering relative payoffs need not change the equilibrium path. Nevertheless, outcomes in the experiment are often inefficient. One explanation is that bargainers have difficulties anticipating the future actions of other bargainers. This problem might be mitigated by allowing bargainers to communicate prior to each stage. A test finds that communication does in fact increase efficiency, although unevenly, and at the cost of the builder. The study implies that the nature and pattern of communication among bargainers is a critical factor in efficient coalition building.

Keywords: coalitional bargaining, communication, game theory, experiment

JEL classification: C7, C9, D7

* Bolton: Smeal College of Business, Penn State University, University Park, PA 16802, gbolton@psu.edu. Brosig: Department of Economics, University of Cologne, Albertus-Magnus-Platz, D-50923, Köln, Germany, jbrosig@uni-koeln.de.

1 Introduction

What is involved in building a coalition? In this paper, we study behavior in a three-person extensive form game in which coalitions can be endogenously renegotiated. The formation of a coalition imposes commonly known externalities on non-members. The game allows one bargainer, the coalition "builder", to shape the process of negotiation by choosing to negotiate with both of the other bargainers together or one after another. The game has a unique equilibrium that is sensitive to the size of the externalities, but always leads to the efficient grand coalition. The driving elements that shape coalition building in this game - renegotiation and externalities - are commonly observed features of many actual coalition building problems; Lax and Sebenius (1991) discuss a number of examples including intra-organizational negotiations, mergers and acquisitions, and international trade treaties.

There is a long, if fragmented, history of behavioral study of coalition building. Coalitions have been central to the theory of games since the subject's inception. In the first chapter of their famous book, Von Neumann and Morgenstern (1944, p. 15) state that, "*Our subsequent discussion of "games of strategy" will show that the role and size of "coalitions" is decisive throughout the entire subject.*" Early behavioral work was guided by cooperative game theory which at the time provided the most highly developed models of coalition formation. Selten (1987) reviews the early experimental literature and finds none of the (cooperative) theories "*completely satisfactory in light of the data*."

One of the limitations of cooperative models is that, by their nature, they provide little guidance for studying the process of coalition formation. It seems likely that this process has important explanatory power: On the one hand, making some commitments early on allows the builder to credibly tie his hands in a way that can be advantageous in later negotiations (i.e., Schelling, 1960); indeed the theory of the influence of externalities we benchmark with here picks up on this. On the other hand, the early commitments entail strategic risk in that the builder constrains his ability to satisfy interests in ways that might derail later agreements. Hence the process of coalition building may have implications both for the efficiency of the negotiation as well as for individual allocations.

There have been a number of recent advances in the non-cooperative modeling of coalition building. The game we study here borrows much of its architecture from a model by Gomes (2005). He develops an infinite horizon coalition bargaining model of sequential offers and counteroffers, incorporating both externalities and renegotiation. We simplify the process to a finite horizon game in much the same spirit that Rubinstein's (1982) bilateral sequential bargaining game was simplified to a finite horizon for the purpose of behavioral study.[1]

Non-cooperative models of coalition building have sparked a new, still nascent, round of behavioral studies. One line in this research compares non-cooperative and cooperative model forecasts. These studies focus on relatively free form bargaining games, to provide a fairer comparison of cooperative and non-cooperative solution concepts. Bolton, Chatterjee, and McGuinn (2003) investigate how various communication configurations affect a three-person coalition negotiation with (implicit) renegotiation but no externalities. Neither the core nor the Shapley value show any particular fit, but the solution to an infinite horizon non-cooperative game implies many of the qualitative characteristics of the data. Croson, Gomes, McGinn, and Nöth (2004) study three-person coalition building in the context of mergers and acquisitions, where there are externalities and renegotiation. They conclude that the predictions of Gomes' infinite horizon model are a better approximation to what they observe than are cooperative game solution concepts such as the nucleolus and the Shapley value.[2]

While these studies are indicative of the promise of non-cooperative models, the evidence is indirect in the sense that the free form game experiments introduce a number of factors that are extraneous to the non-cooperative models but potentially important to the actions actually observed. For example, they introduce relatively free communication between bargainers, the influence of which will be investigated here. Moreover, it is difficult to tease the influence of these factors apart without greater control over the action space. A second, still nascent, line of laboratory research looks directly at non-cooperative game forms. Frechette, Kagel, and Morelli (2005) study winning coalition formation under majority rule in Baron-Ferejohn as well as demand models of political coalition building. They find that neither model fits the data particularly well. Notably, they show the data is consistent with field data on parliamentary voting - data taken from essentially free form bargaining, and previously interpreted as being consistent with the non-cooperative models. Okada and Riedl (2005) investigate a three-person

[1] Other theoretical work on non-cooperative coalition models considers renegotiation or externalities, but not both; see Gomes (2005) for an overview of this literature.
[2] See also McGinn, Milkman, and Nöth (2007) who investigate the free form communication patterns obtained in Croson, Gomes, McGinn, and Nöth (2004) in more detail.

ultimatum game, and characterize the role that reciprocal fairness plays. The coalition games in these studies do not include renegotiation or externalities, the issues that are our focus here.

We can anticipate that one obstacle to building a grand coalition – a factor not considered by the benchmark model – will be the (privately known) preferences individual bargainers have for relative payoffs. In fact, demands for relative payoffs are a known cause of the inefficient outcomes observed in non-cooperative, bilateral sequential bargaining games (Güth, Schmittberger, and Schwarze, 1982; Ochs and Roth, 1989; and De Bruyn and Bolton, 2006). In anticipation of this, our experiment includes a treatment in which relative and absolute payoffs are aligned such that the renegotiation path proscribed by relative payoff theories, and that proscribed by absolute-payoff-only theory are the same, and have just slightly different payoff allocations. The builder in this game gains virtually nothing from knowing how the other bargainers he or she faces trade-off absolute and relative payoffs; the optimal action, and one that leads to the successful building of a grand coalition, is essentially independent of this information, and is essentially the same as what results from considering absolute payoffs only.

Nevertheless, the data shows a rather large amount of inefficiency, even in the treatment where relative payoffs are controlled as described. A close analysis of the data suggests that, while relative payoffs are clearly responsible for much of the disagreement we see, bargainers gauge their relative payoff more myopically than either observations from bilateral bargaining games of similar length, or standard theories of relative payoffs, would suggest.

While these games are no longer than the two-round bilateral bargaining games where standard relative payoff theory is known to work well, they are nevertheless more complex in that there are more second stage scenarios, as well as more bargainers, to be taken account of. A natural hypothesis is that this greater complexity makes it more difficult for bargainers to anticipate the second round actions of others. If so, it seems likely that allowing nonbinding communication between bargainers might reduce inefficiency and lead to a higher rate of successful coalition building. An extension of our design to allow nonbinding communication among bargainers prior to play of each round of the game supports the claim. It appears that communication is a critical factor in the amount of efficiency observed in these games.

2 The new experiment

2.1 The coalition game

The three bargainers in the game are *A*, *B*, and *C*, where *A* is "the builder". Each possible coalition results in a certain earning for the coalition as a whole, as well as a certain earning for any player who is not in the coalition (we will get to specific values in a moment). The game begins in the *No Coalition* condition. The negotiation then proceeds in two stages: In stage 1, *A* proposes a coalition and a split of the coalition's earnings among the coalition members. Proposed-to members are then simultaneously asked to either accept or reject the proposal. If all accept, then the coalition forms and the accepting members exit the negotiation with the earnings they accepted. If any rejects, then there is *No Coalition*. If the grand coalition, *ABC*, forms, then the game ends with the implied payoffs. Otherwise, the game proceeds to stage 2, where *A* either proposes to enlarge the coalition that formed in Stage 1 or keeps it as is. To enlarge the coalition, *A* proposes a split of the earnings between himself and the *new* members. The earnings of a member who already formed a coalition with *A* in stage 1 cannot be changed. Each *newly* proposed-to member is asked to either accept or reject the proposal. If all accept, then the coalition forms and all players exit the game with the earnings they accepted. If any rejects, then all players exit with the coalition and split determined by stage 1.

*Table 1. Coalition configurations and predicted outcomes by treatment**
(all payoffs in U.S. dollars)

All treatments
No coalition payoffs: **V(A) = V(B) = V(C) =** 14 Grand coalition: *V*(*ABC*) = 66

Coalition values					Threat coalition (if no 1st stage coalition)	SPE** coalitions 1st stage → 2nd stage	SPE** final payoffs (*A*, *B*, *C*)
Treat.	*V*(*AB*)	*V*(*AC*)	*V*(*B*)	*V*(*C*)			
T1	46	38	21	17	[*ABC*]	[*ABC*]	(36, 15, 15)
T2	46	38	11	7	[*ABC*]	[*AB*] → [*ABC*]	(42, 16, 8)
T3	56	38	11	7	[*AB*]	[*AC*] → [*ABC*]	(46, 12, 8)

* *V*(*XY*) is the value of coalition [*XY*] and *V*(*Z*) is the externality payoff that *Z* receives when coalition [*XY*] forms.
** Subgame perfect equilibrium.

Payoffs to two-person coalitions and associated externality payoffs vary by treatment as stated in Table 1. In all treatments, a *No Coalition* outcome pays $14 per player, while the grand coalition is worth a total of $66. A two-person coalition imposes an externality on the excluded player in the sense that, for example, in Treatment 1, if *A* and *B* form a coalition, *C*'s payoff increases to \$17 from \$14 when there is no coalition.

2.2 Equilibrium analysis (non-relative payoff version)

The subgame perfect equilibrium path of coalition formation, as well as the associated outcome, vary with the values of the two-person coalitions and associated externality payoffs. In this analysis, we will assume that each bargainer attempts to maximize his or her (absolute) pecuniary payoff from the game. Each treatment of the experiment features a game with a unique equilibrium coalition formation path.

Let's begin with Treatment 2: To derive the equilibrium path, we first consider the *threat coalition*; that is, the coalition that should form if no coalition forms in the first stage. In such circumstances, the second stage is essentially an ultimatum game in which the proposer *A* needs to get both responders *B* and *C* to accept. Inspection shows that the threat coalition is the grand coalition with associated allocation 36-15-15, since this offers the minimum amount that makes *B* and *C* better off from accepting than rejecting, thus maximizing *A*'s payoff.[3] We then roll back to the first stage. *A* can best improve on his threat coalition payoff by proposing, in stage 1, *AB* with associated allocation 30-16, which *B* should accept, and then proposing *ABC* and 42-16-8 to *C* in stage 2, which should also be accepted (recall that, by the rules of the game, *B* is out of the game after accepting in stage 1).

Observe that, in equilibrium of the Treatment 2 game, *A* improves on his threat coalition payoff by sequencing the coalition building to exploit the negative externality incurred by *C* when *AB* forms in the first stage. The Treatment 1 variation of the game has positive externalities, adding 10 to each external payoff in Treatment 2, which leaves the threat coalition and associated payoffs unchanged. But now the threat coalition is the best that *A* can hope to realize, and so the threat coalition is also the final coalition. The Treatment 3 variation of the game, relative to Treatment 2, leaves the externalities unchanged but increases the payoff value of *AB* to the point that it becomes the threat coalition. *A* can improve on the threat coalition payoff by first proposing *AC* with associated allocation 30-8 in stage 1, which *C* should accept

[3] We assume that bargainers accept only if they make more than if they reject.

given his expectation of receiving 7 under the threat coalition, and then proposing *ABC* and 46-12-8 to *B* in stage 2, which should also be accepted.

The above theory does not consider what role relative payoffs might play in bargainer behavior. We will return to discuss this analysis in section 3.2 below.

2.3 Laboratory procedures

The experiment involved 159 subjects. Each of the three treatments T1, T2, and T3 was played in two sessions with 18 subjects each (using a between subject design). The only difference in the treatments was game payoffs. All other aspects of the design, and all procedures, were the same. Subjects played a sequence of five games. A subject's role (*A*, *B*, or *C*) was assigned randomly at the beginning of the session and remained the same for all five games. No one ever negotiated with the same person more than once (publicly announced at the beginning of the session). Actual game matches were anonymous before, during, and after the experiment.

In addition, we ran two treatments T1_C (one session with 18 subjects) and T3_C (two sessions with 18 and 15 subjects, respectively) allowing bargainers to communicate prior to the issuing of proposals and responses.[4] Specifically, prior to each stage of the game, the three bargainers entered a chat-room where they could exchange messages for a period of three minutes. The only constraint imposed was that bargainers could not reveal their identity in their messages (sent messages were tagged by role of sender: *A*, *B*, or *C*). We discuss the communication treatments in section 4.

The experiment was conducted at the LEMA laboratory, Smeal College of Business, Penn State University. The subjects were all Penn State University students, recruited via a web site solicitation. Cash was the only incentive offered. At the beginning of the session, subjects read instructions and played some practice games with the computer as partner to illustrate how the computer interface worked from the perspective of all three players (see the reprinted instructions in the Appendix for details). Upon completion of the five rounds of games, subjects

[4] Due to technical difficulties, in one session of T3_C subjects played only a sequence of 2 games instead of the 5 done in all other sessions.

were paid a show-up fee of \$5 plus their earnings from one, randomly chosen game. The average total payout was \$26, ranging from \$5 to \$45. Sessions lasted about 90 minutes.

3 Results

We begin the analysis by reviewing the end results of the games. We then back up to examine each stage of play, with an aim of understanding how the results came about.

Table 2. Frequencies of final coalitions

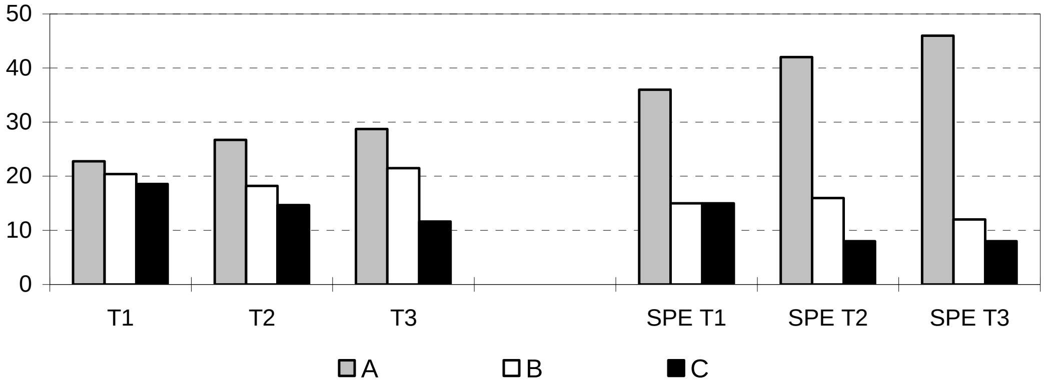
Treatment	*ABC*	*AB*	*AC*	*No Coalition*
T1	31 (52%)	20 (33%)	0 (0%)	9 (15%)
T2	35 (58%)	17 (28%)	4 (7%)	4 (7%)
T3	16 (27%)	37 (62%)	4 (7%)	3 (5%)
Total	82 (46%)	74 (41%)	8 (4%)	16 (9%)

3.1 The final coalitions and associated player payoffs

Table 2 displays the distribution of final coalitions. There are three things to observe. First, many of the final coalitions are inefficient; only 46% of all negotiations end with the grand coalition *ABC*. Second, most of the deviations from the grand coalition are to *AB* which, overall, is observed almost as frequently as *ABC*. In all treatments, *AC* and *No Coalitions* are observed at only very low frequencies. Third, the frequency distribution of final coalitions is about the same in T1 and T2, but different in T3. In T1 and T2, *ABC* is roughly twice as likely as *AB*. In T3, just the reverse: *AB* is roughly twice as likely as *ABC*.

Statistical tests on observed frequencies of final coalitions are consistent with these observations (χ^2-tests, n = 12 per round and treatment). Comparing T3 to the other treatments, pair-wise, indicates significant differences for the frequency of *ABC* formation ($p < 0.050$ for 2 comparisons, $p < 0.100$ for 2 comparisons, $p > 0.100$ for 6 comparisons) as well as *AB* formation ($p < 0.050$ for 2 comparisons, $p < 0.100$ for 3 comparisons, $p > 0.100$ for 5 comparisons). Analogous tests between T1 and T2 show almost no significant difference ($p = 0.098$ for 1 comparison, $p > 0.100$ for 19 comparisons). There is also almost no significant difference across the three treatments with regard to the frequency of *AC* or *No Coalition* outcomes ($p = 0.064$ for 1 comparison, $p > 0.100$ for 29 comparisons).

Figure 1. Actual mean payoffs versus subgame perfect equilibrium (SPE), by treatment



Moving to final player payoffs, Figure 1 displays the average payoffs across roles. There are two things to observe. First, the roles' payoffs are more equal than theory anticipates. Nevertheless, and second, the differences that there are between roles largely conform to the ordinal differences anticipated by the theory: In all three treatments, there is a persistent first mover advantage. The ordinal differences across treatments for both *A* and *C* role payoffs are also consistent with theory. Ordinal differences in *B* role payoffs across treatment, however, deviate somewhat from theory.

Statistical tests on observed frequencies of final payoffs are consistent with these observations (n = 12 per round and treatment). Players *A* payoffs were lower than those predicted by theory, the payoffs for players *B* and *C* were higher than the predicted ones ($p < 0.050$ for 39 comparisons, $p < 0.100$ for 4 comparisons, two-tailed one-sample *t* test).[5] Players *A* do better than the other roles ($p < 0.050$ for all 5 rounds in all three treatments, two-tailed exact Mann Whitney *U* test). Players *A* receive higher payoffs in T2 and T3 than in T1 (T1 vs. T3: $p < 0.050$ for 3 rounds, $p < 0.100$ for 2 rounds, T1 vs. T2: $p < 0.100$ for 3 rounds). Players *C* receive lower payoffs in T3 than in T2 and lower payoffs in T2 than in T1 (T1 vs. T3: $p < 0.050$ for 4 rounds, T1 vs. T2: $p < 0.050$ for 2 rounds, T2 vs. T3: $p < 0.100$ for 1 round). Players *B*, however, receive

[5] The total number of comparisons is 45 and results from multiplying the number of players (3) with the number of rounds (5) with the number of treatments (3).

the highest payoff in T3 (T1 vs. T3: $p < 0.100$ for 1 round, T1 vs. T2: $p < 0.050$ for 1 round, T2 vs. T3: $p < 0.050$ for 3 rounds).

In sum, there are clear treatment effects within the set of observed outcomes, although not all of these effects are anticipated by theory. In what follows, we will focus on the major deviations from theory: The bias towards the inefficient *AB* coalition, the tendency for this bias to be more pronounced in Treatment 3, and the compression of payoffs towards equality.

3.2 A brief digression on the application of relative payoff theory

We will see in a moment that relative payoffs are an important factor in the behavior we observe. However, it turns out that standard analyses of the role of relative payoffs in (bilateral) bargaining fall short in particular ways. Specifically, the following analysis shows that while preferences for relative payoffs may well compress differences in bargainer payoffs relative to the purely monetary equilibrium analysis, the coalition formation process and its final outcome need not be affected. That is, relative payoffs may affect the final allocation, but nevertheless be *process neutral*:

We can see this most clearly by analyzing the Treatment 3 game.[6] Suppose, for example, bargainers have ERC social preferences (Bolton and Ockenfels, 2000).[7] This means that bargainer *i*'s utility is of the form $u_i(x_i, \sigma_i)$, where x_i is the monetary payoff *i* receives and σ_i is the proportion of the sum payout to all bargainers that *i* receives. As x_i increases, holding σ_i fixed, u_i increases. The social reference point for the 3-person game is 1/3 of the pie (we will see in a moment that this fits well with our data). Holding x_i fixed, u_i increases as σ_i increases to 1/3 and then decreases as σ_i increases beyond 1/3.

For the moment, suppose that the builder, bargainer *A*, incurs a negligible loss of utility when $\sigma_A > 1/3$; that is, *A* does not like to be treated unfairly but does not mind (much) to treat others unfairly. This sort of asymmetry assumption has proven to be a good approximation for bilateral bargaining behavior (Bolton, 1991; De Bruyn and Bolton, 2006). The subgame perfect equilibrium for the Treatment 3 game is derived by applying the same procedure as for the

[6] For Treatments 1 and 2, the relative payoff equilibrium path is somewhat less precise. The game in treatment 3 suffices to make our point.

narrowly self-interested preferences. We first look for the threat coalition: In the second round, the best safe offer *A* can make is to propose *AB* with split 35-21. This offer should not be turned down by *B* since 21 is 1/3 of the pie of 63 (recall *C*'s externality payment is 7), and $u(21, 1/3) > u(14, 1/3)$ for all ERC players. The best safe offers of *ABC* or *AC* give *A* a smaller payoff, and so are dominated by the *AB* proposal, making the latter the threat coalition. The reader can verify that, generally, *AB* is the threat coalition independent of *A*'s risk posture since any feasible offer of another coalition is dominated by a feasible *AB* coalition in terms of both *A*'s payoff and the probability that the offer is rejected. (Of course, less risk averse builders may choose a more aggressive offer than the safe one, a point we will return to below.)

Then the equilibrium offer in the first round has *A* proposing *AC* with a split of 30-8. *C* should accept, since rejection leads to the threat coalition and $u_C(8,8/66) > u_C(7,7/63)$ for all ERC players. Observe that this is the exact same first stage proposal implied by the absolute payoff version of the equilibrium.

Then, in the second round, *A* proposes *ABC* with a split of 43-15-8. *B* should accept in the second round since $u_B(15,15/66) > u_B(11,11/63)$ for all ERC players. So the equilibrium outcome is 43-15-8. Relative payoffs are process neutral in that the sequence of negotiations and the final coalition are the same as described in Table 1; but the payoffs are a bit compressed.

By this equilibrium, the game ends efficiently. Yet, this equilibrium does not compare well with the data. From Table 3, only 27% of the games in T3 end with the grand coalition, and among these, none form via *AC* in the first step.

There are other ERC-equilibrium paths for Treatment 3, but none square well with the data. Some involve *A* making offers that risk rejection. But all of these involve *A* proposing *AC* in the first stage. In the next section (3.3), we will see that *AC* is virtually never proposed. All the other alternative equilibria involve relaxing the assumption that bargainer *A* cares little about fairness towards others. But we will see in the next section that *A*'s actual proposals are well explained as optimal, self interested response to the other bargainers' actual accept/reject behavior. So, while the compression of the payoff data is indicative of a role for relative payoff, there is something going beyond standard relative payoff theory explanations.

[7] Other social preference formulations probably lead to a similar analysis.

3.3 *The coalition formation process and the role of relative payoffs*

Turning back to the data, we will see that once their role is properly understood, relative payoffs do explain a great deal of what we see here. To begin, however, observe from Table 3 that not only the outcome, but also the process of coalition formation typically differs from that anticipated by the theory (both pure absolute and relative payoff theory); most coalitions form in one step even where the theory anticipates two.

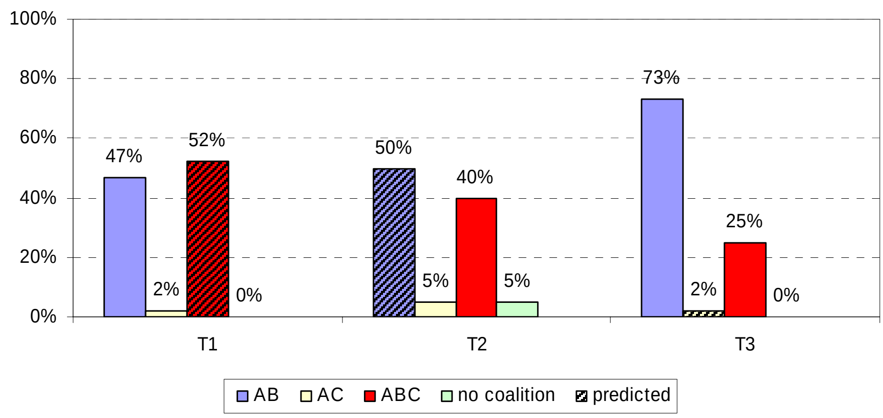
*Table 3. Frequencies of process of grand coalition formation **

Treatment	*ABC (1 step)*	*stage 1/stage 2*	*AB* $\rightarrow$ *ABC*	*AC* $\rightarrow$ *ABC*
T1	**27 (45%)**	**5/12**	4 (7%)	0 (0%)
T2	23 (38%)	19/4	**11 (18%)**	1 (2%)
T3	15 (25%)	10/5	1 (2%)	**0 (0%)**
Total	65 (36%)	44/21	16 (9%)	1 (1%)

* Final coalitions consistent with equilibrium are in boldface.

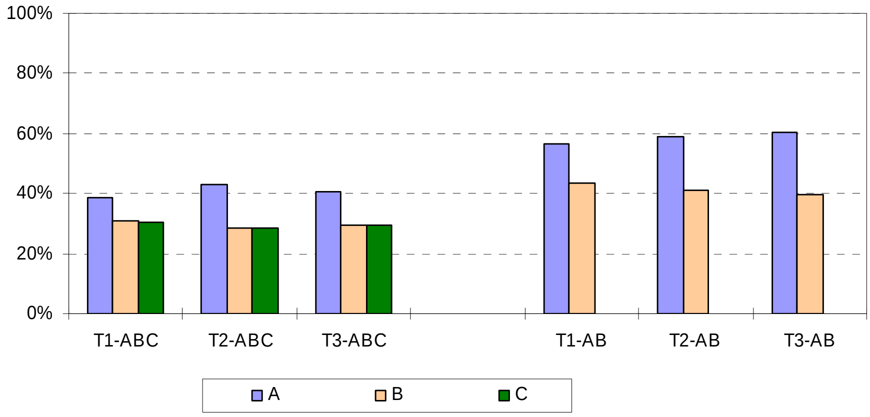
Figure 2 provides a comparison of observed stage 1 proposals with theory. Observe that in all treatments the vast majority of players *A* propose to form either a coalition with player *B* alone or a coalition with both players *B* and *C*. While in T1 and T2 the number of proposals of *AB* and the number of proposals of *ABC* are about the same, in T3 proposals of *AB* are more prevalent than proposals of *ABC* by about a 3-to-1 ratio.

Figure 2: Coalitions proposed in stage 1



Recall from Table 1 that the manipulation that distinguishes T1 from T2 is the externality payoff that goes to a player left out of a coalition. This manipulation appears to have little impact on coalitions proposed in stage 1; in particular, the manipulation does not meaningfully move stage 1 proposals towards more *AB*. In contrast, the manipulation that distinguishes T3, the higher payoff that goes to coalition *AB* has a clear impact on the coalition proposed; but the observed movement is towards more *AB* not towards *AC* as the theory predicts.

Figure 3: Relative payoffs proposed at the first stage



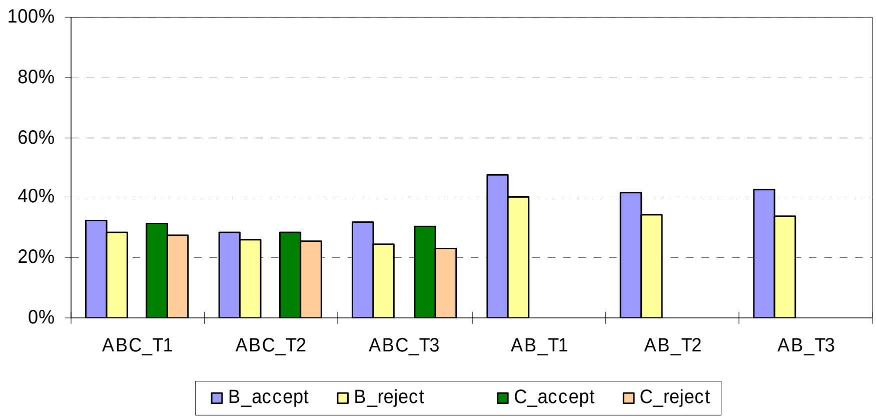
Why is there a persistent attraction to proposing *AB* and *ABC* in T1 and T2? Second, why does the attraction to proposing *AB* grow for T3? Towards answering this, first observe that the relative payoffs players *A* propose in stage 1 are revealing (see Figure 3; given the very small number of proposals for *AC* and for *No Coalition*, we focus only on proposals for *AB* and for *ABC*). Observe the consistent first mover advantage for both *ABC* and *AB* coalitions.[8] Also observe, for proposals of *ABC*, that the average relative payoff offered to player *B* is not particularly different from the average relative payoff offered to player *C* ($p > 0.100$ for all rounds in all three treatments). Most importantly, there is a good deal of consistency in relative payoffs across treatments. Comparing each coalition members' relative payoff between all three

[8] That is, for all coalitions and treatments, the relative payoff for the first mover was higher than the equal split, while the payoff for the other coalition members was lower (coalition *AB*: player *A*/*B*: $p < 0.050$ for 10 comparisons, $p < 0.100$ for 5 comp.; coalition *ABC*: player *A*: $p < 0.050$ for 4 comp., $p < 0.100$ for 2 comp., player *B*: $p < 0.050$ for 2 comp., $p < 0.100$ for 4 comp., player *C*: $p < 0.050$ for 4 comp., $p < 0.100$ for 3 comp., two-tailed one-sample *t*-test). Note that we found no significant results for coalition *ABC* in treatment T3. This might be due to the fact that, in this case, the number of observations was very low.

treatments we find significant differences neither for coalition *AB* (p > 0.100 for all rounds and treatment comparisons) nor for coalition *ABC* (p > 0.100 for all players, rounds, and treatment comparisons, except 1 comparison where p < 0.100).

Observe from Figure 3 that the average relative payoffs for proposals of *ABC* approximate a 40-30-30 split; for *AB* they approximate a 60-40 split. The tendency towards these values is well explained by what the responders to the proposal are willing to accept and reject, as displayed in Figure 4. There are almost no significant differences between treatments with regard to the average payoffs that were accepted and rejected by the proposed-to coalition members (p > 0.100 for all treatment comparisons, coalition members, and coalitions in all 5 rounds, except 1 comparison, in which p < 0.05, two-tailed exact Mann-Whitney-*U* test). Moreover, we observe no significant differences regarding the accepted and rejected amounts between player *B* and *C* in coalition *ABC* (p > 0.100 for all 3 treatments in all 5 rounds, two-tailed exact Mann-Whitney-*U* test).

Figure 4: Stage1 average relative payoffs that were accepted and rejected by players B and C



It is important to note the critical difference between the responder behavior we observe and that anticipated by the standard theory of relative payoffs. Following on the analysis in section 3.2: Observe from Figure 4 that the amount of a grand coalition that *C* players accept on average in Treatment 3 is about 30% of the grand coalition payoff – even though this is more than twice, in both absolute and relative terms, what they would receive if the threat coalition *AB* forms in stage 2.[9] Hence responding bargainers appear to evaluate relative payoffs myopically. That is, they appear to be responding to offers as if this is the final coalition decision, *ignoring the consequences and (relative) payoffs that might result from stage 2 activity*. We will come back to the potential reasons for this in the summary section; for the moment, we simply accept the empirical fact and follow its consequences.

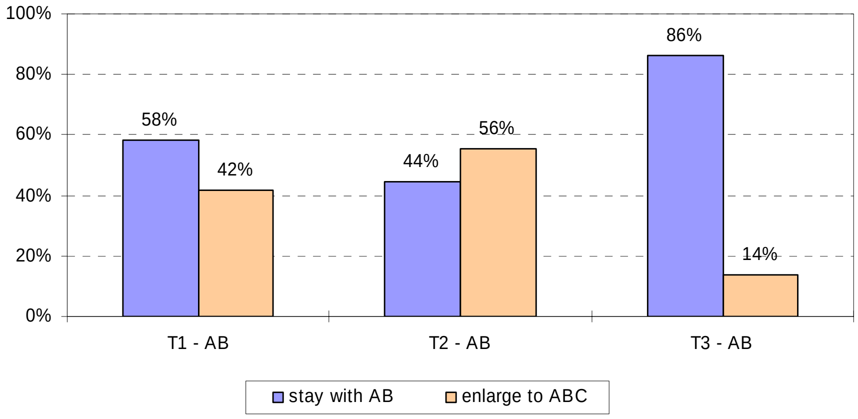
Given the facts represented in Figures 3 and 4, we can understand why players *A* favor proposals of *AB* and why *AB* proposals increase in T3. Specifically, a successful grand coalition proposal pays *A* an average of 0.4 x \$66 = \$26.4. For T1 and T2, a successful *AB* coalition pays 0.6 x \$46 = \$27.6, and in T3, it pays 0.6 x \$56 = \$33.6. In contrast (supposing an average 40% offer is necessary for success here as well), proposing *AC* would pay 0.6 x \$38 = \$22.8.

There is nevertheless a good number of builders in T1 and T2 who propose the grand coalition in stage 1, even though their own payoff goes down on average about \$1. These proposers are plausibly opting to sacrifice the \$1 to purchase greater equity. It is also possible that the difference in expected payoff is not salient to them. Either way, this observation also implies an explanation for why, in T3, the balance tips decisively away from proposing grand coalitions to proposals of *AB*: In T3, the average payoff that resulted from forming coalition *ABC* was \$5 lower than the average payoff that resulted from forming coalition *AB*; hence greater equity is more expensive to purchase, and the purchase price is more difficult to miss.

By the argument in the last paragraph, we might have expected more *ABC* coalitions in T2, in which the non-member *C*'s payoff was only \$7, than in T1, in which the non-member *C*'s payoff was \$17. But we do not observe such pattern in stage 1. However, as Figure 5 illustrates, at the second stage 56 percent of players *A* wanted to enlarge coalition *AB* in T2, while only 42 percent wanted to do so in T1. Unfortunately, the number of observations per round is quite low, making it difficult to support this interpretation of the data with significant results (i.e., for all comparisons in all rounds: $p > 0.100$, two-tailed Fisher test).

[9] In fact, there is *no* player *C* in Treatment 3, who accepts less than about 20% of the grand coalition payoff even though this is twice, in both absolute and relative terms, what this player would receive if the threat coalition *AB* forms in stage 2.

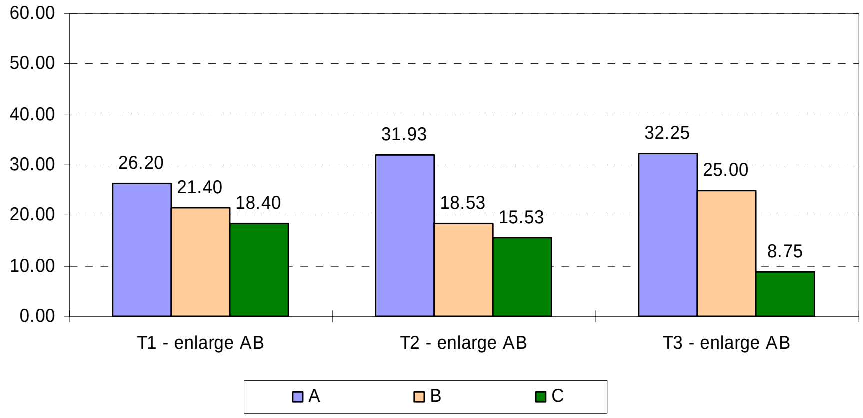
Figure 5: Stage 2 proposals of players A, who already formed coalition AB in stage 1



In T3 the number of players *A*, who attempted to enlarge coalition *AB*, was extremely low, though in this treatment player *C*'s payoff as a non-member was the same as in T2, i.e. only $7. This observation might be explained by the fact that the total payoff for coalition *AB* and player *C* was already $63 ($56 + $7) and, consequently, the maximum additional payoff player *A* could offer player *C* in order to enlarge the coalition was only $3. As a consequence, player *C*'s final payoff as a coalition member would be significantly lower than the final payoff for coalition member *B*.[10] Our observations regarding the average payoffs that were offered to player *C* in T1 and T2 suggests that, in fact, players *A* geared to the payoff for player *B* when making an offer to player *C*.

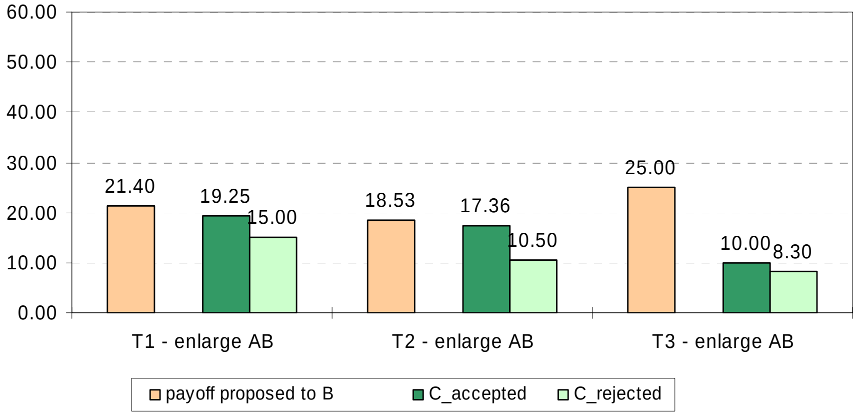
[10] Note that the maximum additional payoff player *A* could offer player *C* in order to enlarge the coalition was $3 also in treatment T1. However, proposing this amount to player *C* in treatment T1 would result in a final payoff of $20 – an amount that is very similar to the final payoff for player *B*.

Figure 6: Average payoffs proposed at the second stage after the formation of coalition AB



Analyzing player *C*s' responses seems to justify this strategy. In T1 and T2 only 25 percent of players *C* rejected player *A*s' proposal, while in T3 the frequency of rejections was 75 percent. Investigating the average payoffs that were accepted and rejected by players *C* seems to further support this interpretation (see Figure 7). In particular, players *C* always accepted a payoff that was similar to the payoff proposed to player *B*. The findings indicate that, in treatment T3, players *A* anticipated the high number of rejections and, therefore, abstained from proposing an enlargement of the coalition.[11]

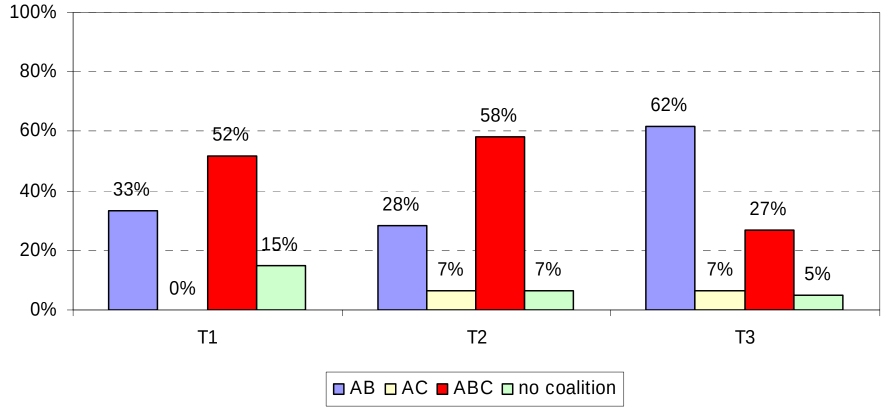
Figure 7: Player Cs' response to proposals made after the formation of coalition AB.



[11] Possibly, players *A* did not want to be faced with an almost sure rejection by players *C* and, therefore, did not want to give players *C* an opportunity to show their displeasure. Or perhaps they felt that what they had to offer would be treated as insulting and so refrained for this reason.

As a result, at the end of the coalition games the majority of players *A* formed a coalition with player *B* and player *C* in treatments T1 and T2, while in treatment T3 the majority of players *A* formed a coalition only with player *B*. The final outcomes are illustrated in Figure 8.

Figure 8: Coalition outcome per treatment



4 The game with communication

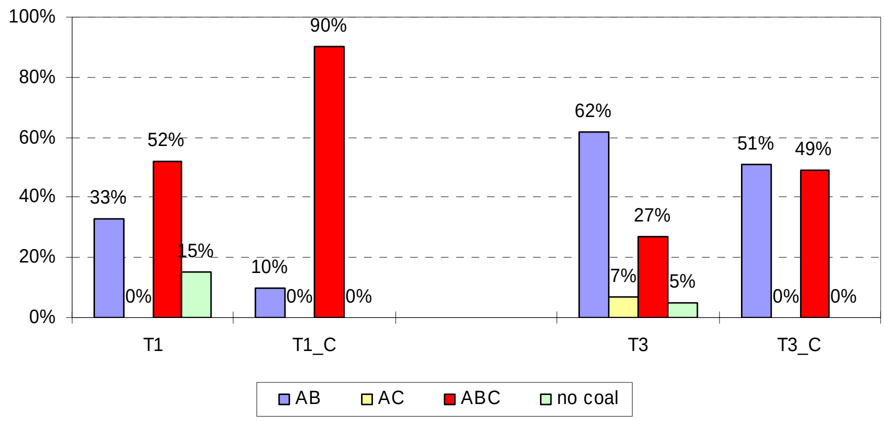
Section 3 demonstrates that once we understand the myopic nature of relative payoff considerations, most of the deviations from theory (whether the theory assumes purely absolute or admits relative payoff motives) fall into line. Perhaps most importantly, we can understand why the final outcome is often inefficient. Then what accounts for the myopia? The obvious explanation is that bargainers considering stage 1 proposals have trouble looking forward to stage 2 and reasoning back, which violates the backward induction assumption. There is evidence for a failure to backward induct in bilateral sequential bargaining games (Johnson, Camerer, Sen, and Rymon, 2002), although bargaining theory that assumes relative payoffs and backward induction nevertheless approximate much of what we observe in these games (Bolton, 1991; De Bruyn and Bolton, 2006). Still, two stage coalition bargaining is more complex, with more bargainers and larger strategy spaces (the builder must choose who to propose to as well as what to propose). Taking this greater complexity into account while backward inducting might well heighten the disconnect between theory and data. Indeed, incorporating relative payoffs into preferences makes the game one of incomplete information, and the greater complexity makes it all the harder to anticipate the future actions of other bargainers. Even if bargainers were quite good at backward induction, the greater strategic uncertainty would lead them to weight the stage

1 outcomes more heavily than potential stage 2 formations (since the uncertainty opens up the possibility that there will be no further coalition formation in stage 2).

One way to mitigate any backward induction problem, and perhaps any difficulties with strategic uncertainty, would be to allow (cheap talk) communication between bargainers prior to each stage of play. This would allow the builder to communicate his plans ahead of time and get feedback from the other bargainers before committing to that plan.

To test the influence of communication, we ran treatments T1_C and T3_C allowing bargainers to communicate prior to the issuing of proposals and responses. Figures 9 and 10 summarize the data.

Figure 9. Communication treatments: End coalitions

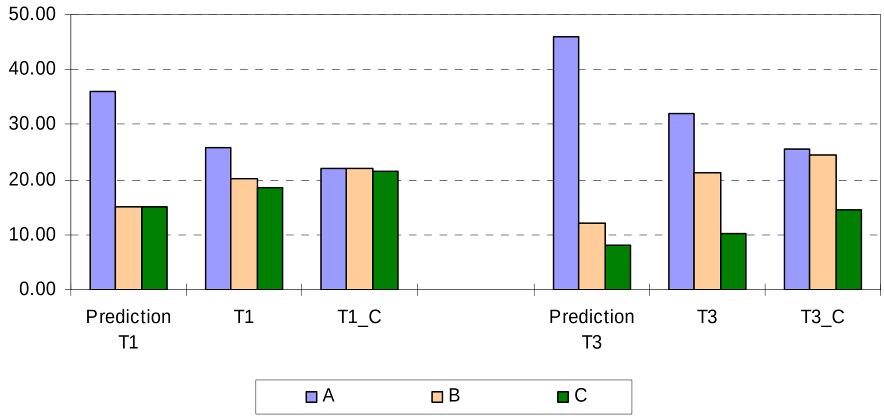


There are three principle observations to make. First, communication improves efficiency, although unevenly. Efficiency increases in both treatments because there are no *AC* or *No Coalition* outcomes. There is also growth in the frequency of grand coalitions at the expense of *AB* coalitions. On this last point, however, the two treatments differ with respect to the magnitude of change: Grand coalition outcomes grow greatly from T1 to T1_C (52% to 90%, $p < 0.050$ for 2 rounds, $p < 0.100$ for 1 round, $p > 0.100$ for 2 rounds), but less so from T3 to T3_C (27% to 49%, $p > 0.100$ for all rounds, two-tailed χ^2 test). The *AB* coalition is still the majority end coalition in T3_C (51%).

The second observation, not shown in the figures, is that communication changes the process of coalition building barely at all. For example, in T1_C, 23 of the 27 grand coalitions form in one, while in T1, 28 of 31 grand coalitions form in one step. (Numbers for T3 and T3_C are similar.)

Third, in Figure 10 we see that communication compresses differences in player payoffs. In particular, in treatment 1 players *C* and in treatment 3 players *B* significantly profited from communication (*C* in T1: $p < 0.050$ for 4 rounds, *B* in T3: $p < 0.05$ for 1 round, $p < 0.100$ for 1 round; $p > 0.100$ for all other comparisons, two-tailed exact Mann-Whitney-*U* test). The ordering of the average payoffs across roles however, remains the same. Though, only in treatment 3 and only with regard to players *C* we observe a significant first mover advantage after communication (*A* vs. *C* in T3_C: $p < 0.050$ for 5 rounds; $p > 0.100$ for all other comparisons).

Figure 10. Communication treatments: Average of final payoffs by role



To summarize, communication helps foster efficiency some but does not fully eliminate the inefficiencies we see, particularly in T3_C. It has virtually no influence on the pattern of coalition formation.

5 Summary

The most striking finding from our study is how challenging it turns out to be to build a grand coalition. Returning to Table 2, the overall rate of efficient, grand coalition formation is less than 50%. Analyzing the data for the underlying reasons reveals that the coalition formation process is not particularly consistent with the predicted equilibrium paths – regardless of whether we do the analysis with relative payoffs included or with absolute payoffs only. Most coalitions form in one step where the theory anticipates two and, irrespective of the payoff variation, the vast majority of coalition builders form a coalition either with player *B* alone or with both players *B* and *C*. The main difference between treatments is regarding the frequency of *AB* and

ABC coalitions: In treatments T1 and T2 (where the grand coalition yields the highest per capita payoff) *ABC* is the most frequent coalition outcome, in T3 (where *AB* yields the highest per capita payoff) it is *AB*.

Unsurprisingly, relative payoffs clearly matter in these games. This is most clear from Figure 4, which shows a strong correlation between relative payoffs and rejection behavior across treatments, even though the absolute payoff consequences of this behavior vary a good deal. Nevertheless, we observe the highest inefficiency in Treatment 3 (Table 3), where relative and absolute payoffs are aligned such that the equilibrium path and payoffs are essentially the same regardless of whether relative payoffs are taken into account. The inflexibility of rejection rates observed in Figure 4, along with other evidence, suggests that relative payoff calculations are being applied with greater myopia than theory suggests.

The communication we allowed in a second set of treatments, though it was cheap talk, increased efficiency: For Treatment 1 games efficiency rose from 52% to 90% and for Treatment 3 from 27% to 49% (Figure 9). It appears that communication is a crucial factor in successful coalition building.

An important limitation of the present study is that we examined only one communication network, one in which bargainers must all communicate publicly with one another. This may well explain why communication compresses payoffs in our study, and, in particular, why coalition builders do more poorly even as overall surplus rises. It would be useful to study other communication networks to see what influence this might have on coalition building behavior. Other studies, conducted on less controlled conditions than those studied here find a large communication effect (see references in section I). It would be beneficial to reproduce these effects in the more controlled non-cooperative game environment to better understand precisely what effect communication has, and how different communication systems might influence the ultimatum success or failure of coalition building.

References

Bolton, G. E., 1991. A comparative model of bargaining: theory and evidence. Amer. Econ. Rev. 81, 1096-1136.

Bolton, G. E., Chatterjee, K., McGinn, K. L., 2003. How communication links influence coalitional bargaining: a laboratory investigation. Management Science 49, 583-598.

Bolton, G. E., DeBruyn, A., 2006. Predicting bargaining behavior: out-of-sample estimates from a social utility model with quantal response. Working Paper, Penn State University.

Bolton G. E., Ockenfels, A., 2000. ERC: A theory of equity, reciprocity and competition. Amer. Econ. Rev. 90, 166-193.

Croson, R. T. A., Gomes, A., McGinn, K. L., Nöth, M., 2004. Mergers and acquisitions: an experimental analysis of synergies, externalities and dynamics. Review of Finance 8, 481-514.

Fischbacher, U., 1999. Z-tree - zurich toolbox for readymade economic experiments. Experimenter’s manual. Working Paper, University of Zürich.

Frechette, G., Kagel, J., Morelli, M., 2005. Behavioral identification in coalitional bargaining: an experimental analysis of demand bargaining and alternating offers. Econometrica 73, 1893-1937.

Gomes, A., 2005. Multilateral contracting with externalities. Econometrica 73, 1329-1350.

Güth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimate bargaining. J. Econ. Behav. Organ. 3, 367-388.

Johnson, E. J., Camerer, C., Sen, S., Rymon, T., 2002. Detecting failures of backward induction: monitoring information search in sequential bargaining. J. Econ. Theory 104, 16-47.

Lax, D. A., Sebenius, J. K., 1991. Thinking coalitionally: party arithmetic, process opportunism, and strategic sequencing. In: Young, H. P. (Ed.), Negotiation Analysis. University of Michigan Press, Ann Arbor, pp. 153-193.

McGinn, K. L., Milkman, K. L., Nöth, M., 2007. Paths to equality: walking the talk in multi-party negotiations. Working Paper.

Ochs, J., Roth, A. E., 1989. An experimental study of sequential bargaining. Amer. Econ. Rev. 79, 355-387.

Okada, A., Riedl, A., 2005. Inefficiency and social exclusion in a coalition formation game: experimental evidence. Games Econ. Behav. 50, 278-311.

Rubinstein, A., 1982. Perfect equilibrium in a bargaining model. Econometrica 50, 97-110.

Schelling, T. C., 1960. The Strategy of Conflict. Harvard University Press, Cambridge, MA.

Selten, R., 1987. Equity and coalition bargaining in experimental three-person games. In: Roth, A. E. (Ed.), Laboratory Experimentation in Economics – 6 Six Points of View. Cambridge University Press, Cambridge, pp. 42-98.

Von Neumann, J., Morgenstern, O., 1944. Theory of Games and Economic Behavior. Princeton University Press, Princeton, N J.

Appendix. Below is a sample of the written instructions given to subjects for Treatment 1. Treatments 2 and 3 differed only in the stated payoffs. Additional instructions for the communication treatments appear in brackets.

General. Please read the instructions carefully. If at any time you have questions or problems, raise your hand and the monitor will be happy to assist you. From now until the end of the session, unauthorized communication of any nature with other participants is prohibited.

During the session, you will engage in a series of negotiations, carried out over the computer with other participants. Each negotiation gives you an opportunity to earn cash.

Description of the Negotiation. Each negotiation involves three participants labeled "Player *A*", "Player *B*", and "Player *C*". Player *A* can propose either to form a coalition with Player *B*, or to form a coalition with Player *C*, or to form a coalition with both *B* and *C*, or to form no coalition. Each possible coalition results in a certain earning for the coalition as a whole, as well as a certain earning for any player who is not in the coalition. The following schedule shows the earnings for all possible coalition formations.

[Coalition members]	Members *split* total of	Each non-member *earns*
[*A* and *B*]	$45.00	$17.00
[*A* and *C*]	$37.00	$21.00
[*A*, *B*, and *C*]	$66.00	–
No Coalition	–	$14.00

The game begins with No Coalition. In order to form a coalition, the members must negotiate and agree upon a split of the coalition earnings. The negotiation proceeds in two stages:

> Stage 1. Player *A* proposes a coalition and a split of the coalition's earnings among the coalition members. Then, each proposed member is asked to either accept or reject the proposal. If all accept, then the coalition forms and the accepting members exit the negotiation with the earnings they accepted. If any rejects, then there is No Coalition.
>
> If the coalition between all three players, [*A*, *B*, and *C*], forms, then the negotiation ends. Otherwise, Stage 2 of the negotiation begins.
>
> Stage 2. Player *A* either proposes to enlarge the coalition that formed in Stage 1, or keep it as it is. If he proposes to enlarge the coalition, he also proposes a split of the earnings between himself and the *new* members. The earnings of a member who already formed a coalition with Player *A* in Stage 1 cannot change. Each *new* proposed member is asked to either accept or reject the proposal. If all accept, then the coalition forms and all players exit the negotiation with the earnings they accepted. If any rejects, then all players exit with the coalition and split determined by Stage 1.
>
> [*In the communication version, the following paragraph began both the stage 1 and stage 2 description:* Players A, B, and C enter a shared chat room through the "In Touch" tab on their Angel screen. After entry, the players have 3 minutes to negotiate with each other in the chat room regarding forming coalitions and striking deals. Upon conclusion of the 3 minutes, stage two follows on the other active monitor screen.]

All negotiations are conducted by computer. Instructions on how to operate the computer program will be provided after everyone has finished reading the instructions.

Player Roles. Your role, *A*, *B*, or *C*, will be randomly determined prior to negotiating. You will keep the same role for all games; ex., if you are a Player *A* in the first game, you will be a player *A* for all games.

Information: In each stage of the negotiations the three players are informed about Player *A*'s proposal, about the acceptance or rejection of the proposal by each proposed member, and about any coalition that formed. There is also a box on the computer screen that provides you with a history of the past negotiations you participated in. At all times, an onscreen calculator is available to assist with your decisions. To use it, click the calculator icon. Scratch paper and a pen have also been provided for you.

Grouping Procedure. In each negotiation you are matched with different participants. You will never negotiate with the same person more than once.

During the negotiation, you will be referred to by your role: either Player *A*, *B* or *C*. Participants' identities are strictly confidential and will not be revealed either before, during or after the session.

Your earnings: You will negotiate more than once. You will actually be paid for just one negotiation. The paid negotiation will be selected by a lottery after all the negotiations have been completed. Each negotiation has an equal chance of being selected, so it is in your interest to make the most money you can in each and every negotiation. Immediately upon conclusion of the session, you will be paid your earnings plus a $5 show-up fee, in cash.

Consent Forms. If you wish to participate in this study, please read and sign the accompanying consent form.

Practice Games. After consent forms have been collected, and prior to playing the real games, we will play some practice games. Use the practice games to become accustomed to how the game is played and to how actions are entered into the computer. The practice games differ from the actual games in three ways. First, no money will be paid for the practice games. The purpose of the practice games is to give you experience with the game rules and the computer interface. Second, you will play each practice game from a different point of view, permitting you to experience the game from each player role: *A*, *B*, and *C*. Third, the computer will be your partner for the practice games and its actions will be generated at random (so do not worry if the actions do not make sense to you).