The dynamics of coalition formation - a multilateral bargaining experiment with free timing of moves☆

James Tremewan[a], Christoph Vanberg[b,*]

[a] *University of Vienna and Vienna Center for Experimental Economics, Austria*
[b] *University of Heidelberg, Germany.*

Abstract

We experimentally investigate behavior in a finitely repeated coalition formation game played in continuous time. Subjects interact in groups of three, bargaining over the distribution of payments which occur at regular time intervals. During a given interval, payments occur if and only if a majority is in agreement about their allocation. Aside from these rules, we purposefully impose little structure on the bargaining process. We investigate the frequency and stability of different types of agreements, as well as transitions between them. Two-thirds of payments involve divisions where one player receives nothing, almost half of which are equal splits of the entire surplus between two players. The most stable division is the three-way equal split. Transitions between agreements are frequent and are generally consistent with myopic payoff maximization, in the sense that subjects do not accept short-term losses. We also find that transitions between coalitions are not Markovian. In particular, players more often forgo short-term gains in order to remain in a coalition if it has proven stable in the past.

Keywords: Multilateral bargaining, group choice, experiments, continuous time, cooperative game theory, coalition formation

JEL-Codes: C7, C9, D7

☆We thank the editor and two anonymous reviewers for constructive comments.

*Corresponding author at: Department of Economics, University of Heidelberg, Bergheimer Str. 58, 69115 Heidelberg, Germany; Tel: +49-6221-542912.

Email addresses: `james.tremewan@univie.ac.at` (James Tremewan), `vanberg@uni-hd.de` (Christoph Vanberg)

1. Introduction

A large literature in economics and political science studies coalition formation as a bargaining game in which a majority of players must agree on a division of an exogenously available surplus. Most of this literature investigates situations in which the interaction ends once a coalition is formed. In many real-world settings, coalition formation occurs in the context of repeated interaction over an extended period of time. Examples include the formation and maintenance of government coalitions, alliances between factions in international or civil conflict (Nolutshungu, 1996), and firms cooperating on supply chain management (Nagarajan and Sošić, 2008). In each of these examples, the members of a coalition reap benefits repeatedly or continuously for as long as agreement persists. Over time, coalitions may dissolve if agreement erodes and new coalitions may be formed. These dynamic aspects introduce new strategic considerations and raise additional questions concerning the stability of coalitions.

This paper reports on an experiment designed to study the formation and stability of coalitions in such a setting. Specifically, we investigate behavior in a finitely repeated three-player majoritarian bargaining game played in continuous time. In our game, players may propose and agree to divisions at any point in real time, and no structure is imposed on the order in which they do so. Payoffs flow at one-second intervals when agreements are in place, and the game ends after a fixed surplus is fully distributed. The goal of our research is to observe what types of divisions arise most frequently and are most stable in this environment.

We find that most payments are two-way splits that completely exclude one player. Among these, nearly half are equal divisions between the two members of the "minimum winning coalition". However, these agreements are typically highly unstable. The most stable agreement is the three-way equal split, which consistently accounts for 20% of realized payoffs. We find that transitions between agreements, *when they occur*, are consistent with myopic payoff maximization. That is, players agree only to transitions which lead to higher immediate payoffs. However, there is also evidence of far-sighted behavior, as not all such (proposed) transitions are agreed to. In particular, players appear to condition their response to tempting proposals on the past behavior of their (current) coalition partners. We find that players are less likely to agree to a new proposal that promises higher immediate payoffs, the more 'loyal' their coalition partner has proven in the past. This observation is important because such history dependent behavior is commonly excluded in theoretical analyses, which typically make the simplifying assumption that strategies are stationary and transitions are Markovian.

2. Related Literature

The literature on coalition formation comprises contributions from diverse fields, including sociology, social psychology, economics, and political science. The multitude of approaches within both the traditions of cooperative and non-cooperative game theory, testifies to the complexity of the problem under investigation. Bargaining behavior and outcomes are likely to be affected by subtle institutional, environmental, and personal factors. This complexity makes experimental investigation of an unstructured environment particularly relevant, as it can help to test and inspire theory in the face of so many reasonable approaches.

Early theoretical contributions to the problem of coalition formation used the axiomatic approach of cooperative game theory. The relative strength of this approach is that it avoids the imposition of a particular structure on the bargaining process, a feature shared by unstructured bargaining experiments. A variety of solution concepts were developed, including the Shapley value (Shapley, 1952), kernel (Davis and Maschler, 1965), nucleolus (Schmeidler, 1969), core (Aumann, 1961), and bargaining set (Aumann and Maschler, 1961). What these concepts imply in our bargaining environment will be detailed in section 4.1.

In keeping with the "institution-free" spirit of cooperative approaches, early experiments were typically unstructured. In Kalisch et al. (1996), subjects bargain face-to-face, and the only rule imposed is that they agree by majority vote. Fiorina and Plott (1978) follow a similar approach, arguing that this "allows (...) procedures to be essentially endogenous and as 'natural' as possible (...)." Thus, these authors felt that experiments on coalition formation should induce preferences and enforce majority rule, but otherwise leave subjects "free to do what they want" (ibid.). One of the conclusions from such face-to-face experiments was that personality plays an important role, with more talkative and aggressive subjects being more successful. There are numerous disadvantages associated with face-to-face experiments: bargaining partners can be identified, so factors such as gender and personal appearance need to be controlled for; face-to-face communication may be more likely to induce other-regarding concerns that interfere with monetarily induced preferences; completely free communication allows for the influence of personality as discussed above. These concerns suggest a role for computer-mediated experiments which exclude face-to-face interaction.

As far as we are aware, the earliest such experiments were performed using a set of programs called *Coalitions*, first described in Kahan and Helwig (1971). The program was designed to implement one-shot bargaining games in a characteristic function game framework. Communication was limited to a small vocabulary, and although players could send messages only in a predefined order, the options a player had during their turn, and the necessity of everybody first accepting then ratifying a coalition meant that little meaningful structure was imposed. A number of papers used this program to test and compare cooperative solution concepts with a variety of different games, for example Rapoport and Kahan (1976) which finds support for the individually rational bargaining set model and Horowitz (1977) for the core.[1]

In contrast to the axiomatic and institution-free approach of cooperative game theory, more recent theoretical contributions have followed a non-cooperative approach. The method is to explicitly specify a bargaining *procedure* as an extensive form game. The structure of a such a game imposes strict rules regarding, for example, who may make a proposal, when and in what order votes are taken, and so on. The most well-known theory in this category is the legislative bargaining model of Baron and Ferejohn (1989). Despite strong *procedural* assumptions, it admits multiple subgame perfect equilibria. Concrete predictions can therefore be derived only by imposing additional *behavioral* assumptions which restrict the kinds of strategies employed (e.g. symmetry and stationarity).[2] The general point is that all non-cooperative models impose rigid procedural rules on the timing of moves (offers, votes) as well as strong behavioral restrictions regarding the strategies players employ (typically symmetry and stationarity). Together, such restrictions yield results concerning equilibrium

[1] Anbarci and Feltovich (2013) implements computer-mediated unstructured bargaining, but in the much simpler bilateral case.

[2] Since its publication, a number of extensions and alternatives to the BF model have been developed. A detailed review of this literature is beyond the scope of the current paper.

play, the properties of which are interpreted as predictions concerning actual behavior in real-world situations to which the theories are meant to apply.

Most of the recent experimental research on majoritarian bargaining is motivated as a test of theoretical predictions of non-cooperative models. The dominant approach is to faithfully implement all procedural aspects of these models. For example, a number of authors have tested the predictions of the BF model by implementing, within a laboratory environment, rigid procedural rules closely resembling those assumed in that model (see McKelvey, 1991; Diermeier and Morton, 2005; Frechette et al., 2005; Miller and Vanberg, 2013, and many others). Within this approach, the study most closely related to our own is Battaglini and Palfrey (2012), who implement a repeated version of the BF game, first analyzed in Kalandrakis (2004). In all such experiments, the interaction is computer mediated in order to maintain anonymity and to make sure that subjects cannot take actions that are not part of the structure of the model, for example by negotiating verbal agreements.

When viewed as a test of non-cooperative theories, an advantage of rigidly structured experiments is that the failure of a model's predictions may be attributed to the failure of *behavioral* rather than procedural assumptions. Another is that structured experiments are well controlled in the sense that the range of possible behaviors is limited to a small set of easily quantifiable action choices and the interaction is simple. A disadvantage is that the rigidly structured experimental environment differs substantially from the real world settings of ultimate interest. In other words, such experiments have very little ecological validity, casting doubt on the external validity of the corresponding findings.[3]

Not all structured bargaining experiments are intended as a test of non-cooperative theories. Thus, Nash et al. (2012) implement a multilateral bargaining game which allows players to cede bargaining rights to another player. Despite the clearly defined structure of the bargaining process, they also derive hypotheses from cooperative game theory, arguing that in their repeated game any division can be supported as an equilibrium, so non-cooperative game theory can provide no predictions.

An intermediate approach combines the imposition of certain procedural rules with a wide scope for individual behavior. For example, Diermeier et al. (2008) conduct

[3]For a discussion on the relationship between ecological and external validity, see Frechette (2011).

relatively free-form bargaining experiments involving communication. In contrast to the earlier unstructured experiments, these authors impose a set of procedural rules, allowing subjects to form ‘protocoalitions’ and negotiate freely in subgroups. This ‘in between’ method nicely combines the openness and realism of unstructured experiments with experimental control over substantive causal factors such as communication and some formal rules.

To our knowledge, the only prior experimental paper investigating a *repeated* bargaining game is that of Battaglini and Palfrey (2012). The focus of that study is on the effects of an endogenously evolving status quo. The experiment is structured in accordance with a non-cooperative model based on the Baron-Ferejohn game. Like that game, the interaction stretches over a potentially infinite number of discreet rounds. In each round, one player is randomly chosen to suggest a division of a surplus. Unlike in the BF game, however, payoffs are realized in each round. If a proposal fails, payoffs are determined by a status quo option (initially exogenous).[4] If it passes, payoffs are determined accordingly, and the division becomes the status quo at time $t + 1$. As a consequence, an agreement reached at time t determines not only the in-period payoffs, but also the power of the proposer at time $t + 1$. In an earlier paper, Kalandrakis (2004) constructed a subgame perfect equilibrium in which players vote myopically and proposers have complete power, resulting in a ‘rotating dictatorship’ where one player secures the entire in-period payoff. In contrast, Battaglini and Palfrey (2012) find that most agreements are either ‘universal’ (equal or nearly equal three-way splits) or ‘majoritarian’ (equal or nearly equal splits in a minimum winning coalition). The universal outcome is by far the most stable. Transitions away from it usually lead to a two-way split, which is the most frequent category of outcome. Two-way splits, in turn, are less stable, and are most often replaced by other two-way splits (ibid. Table 8). Thus, most committees seem to rotate through minimum winning coalitions, and the predicted ‘rotating dictatorship’ does not emerge.

Our approach is also related to early experiments on unstructured (usually face-to face) *bilateral* bargaining. As noted by Roth and Malouf (1979), these experiments

[4] This differs from our game, where payoffs flow only for as long as agreement on a division is maintained (see below).

typically resulted in agreement on equal distributions. This stands in sharp contrast to more structured (alternating-offers or ultimatum) bargaining experiments, which often result in unequal distributions - albeit less unequal than predicted by theory. A possible explanation is that unstructured experiments are inherently more symmetric, as no player enjoys a procedural advantage. The same is true in our setting. Thus, one may expect that our experiments will result in (even) more equal distributions than those observed in comparable structured experiments, e.g. Battaglini and Palfrey (2012). We will come back to this conjecture in the conclusion.

Finally, our work is related to other recent experiments involving games played in continuous time. For example, Berninghaus et al. (2006) and Teteryatnikova and Tremewan (2015) investigate the predictive power of various theoretical stability concepts in continuous time network formation games, with unilateral and bilateral link formation, respectively. Friedman and Oprea (2012) investigate two-person prisoner's dilemma games in which players can change their strategies at any time during a 60 second interval, and payoffs flow continuously. They find that the ability to rapidly react to defection leads to very high rates of mutual cooperation and thus efficiency.[5] Since our game involves the distribution of a fixed aggregate surplus, efficiency is not an issue in our context.

3. Experimental Design

Our experiment implements a finitely repeated multilateral bargaining game played in continuous time. Subjects bargain in groups of three over payments that occur once per second when agreements are in place, i.e. payments flow almost continuously. All subjects can propose allocations and agree to existing proposals at any time. Payments occur each second, if and only if at least two players agree on an allocation at that time. The game ends when a predetermined surplus is exhausted or after a total of five minutes has passed. This feature excludes efficiency costs of delay, reflecting our interest in the frequency and stability of coalitions rather than efficiency.

The design of an unstructured experiment mediated by computers presents two

[5] This result is replicated by Bigoni et al. (2015), who additionally investigate shorter (20 second) games as well as games with a stochastic end point. They find that cooperation rates are lower in shorter games, especially when the end point is stochastic.

important challenges. First, the program interface must be simple enough to be easily operated, yet versatile enough to allow for a natural and procedurally unrestricted exchange of offers, counteroffers, and votes. Kahan and Rapoport (1974) report that the *Coalitions* program interface required several hours of training, a fact that is probably due to some extent to the technology of that time. A second challenge is that the recording and transmission of subjects' decisions must not take up too much time. This is especially important if one wishes to study bargaining over multiple rounds or an extended period of time, and also to allow for many repetitions of each game in case learning effects play a significant role (which is likely in a strategic environment as complex as multilateral bargaining).

In order to address both of these challenges, we designed a simple mouse-operated graphical interface. The bargaining set was represented in the form of a two-dimensional simplex consisting of a finite number of circles (see Figure 1). Each circle represents a feasible allocation of the available (per-second) surplus. The allocation associated with a given circle was displayed if a subject hovered over that circle with the mouse. At any time, subjects could individually select a circle by clicking on it. This was made visible on all subjects' screens by highlighting the chosen circle with a color associated with that player. Once selected, a circle remained so unless and until the subject clicked elsewhere (either selecting another circle, or outside of the simplex, in which case no circle would be selected by that subject). Payments occurred if and only if at least two subjects were selecting the same circle at the end of a one second time interval. In what follows, we refer to the first click on a circle as a "proposal" and the second click on an already occupied circle as an "acceptance" of that proposal.

The game ended when a total of 360 points had been allocated, or after a total of five minutes had passed. In our experiment, the available surplus was always exhausted well within five minutes, so this deadline was never binding.

If no allocation was supported by a majority at the end of the one second time interval, no payment occurred. Since the aggregate payment was fixed, this rule is equivalent to 'pausing' the game whenever there is no agreement, until a new allocation (or the previous status quo) is agreed upon. Since the 5 minute time limit was never even close to binding, in practice there was no exogenously imposed pressure on subjects to quickly arrive at an agreement. This reflects our interest in

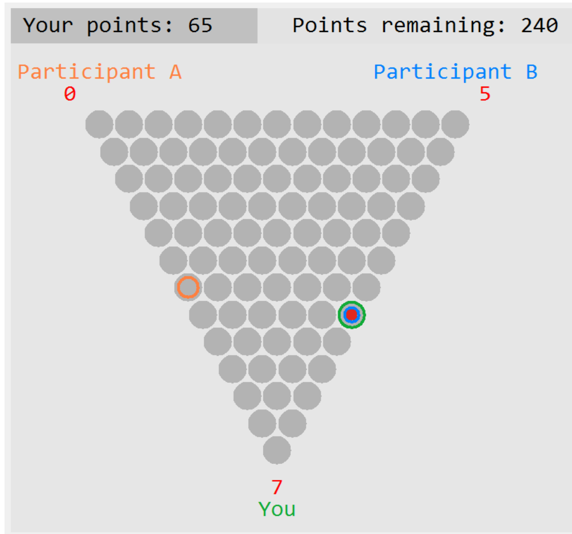

Figure 1: Screen Shot

a setting where the process of coalition formation is quick as compared to the time scale at which benefits accrue. The total number of points available was such that the game ended after 30 seconds of 'agreement time'.[6]

The experiments were conducted in the Vienna Center for Experimental Economics. 72 subjects took part, divided into 8 matching groups of 9 subjects each. Prior to the paid games, subjects completed a tutorial, including control questions, which lasted approximately 15 minutes. During the tutorial, subjects first familiarised themselves with the interface, learning how each circle defined the points earned by each player. They then interacted with simulated computer players which were programmed to select random circles to help understand how agreements could be formed. Onscreen and printed instructions can be found in Appendix A.

The game was repeated 20 times with stranger matching. (We will refer to these repetitions as rounds). One randomly chosen round was paid, and subjects received a 3 EUR show-up fee. The total surplus being divided was worth 36 Euros (360 'points'). Not including the show-up fee, subjects earned between 0 EUR and 24.50 EUR, with a standard deviation of 4 EUR. By definition, average earnings were 12 EUR. Games lasted for 30-60 seconds and sessions lasted approximately one hour.

[6] To better understand the rules of the game, the reader may wish to view a video replay of one of the actual games which can be found in the supplementary material.

The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007). Subjects were recruited using ORSEE (Greiner, 2004).

4. Frequency and Stability

Our empirical analysis will focus on three categories of divisions, each of which is predicted by one or more prominent cooperative solution concepts.[7] We will then test two types of hypotheses. First, that each category of division is more common and more stable than other divisions (i.e. its complement). Second, that there are systematic differences in the frequency and stability between the three categories of divisions, when compared to one another.

Our analysis will be structured around sets of divisions of the per-second surplus that are suggested by concepts from cooperative game theory. The three sets of divisions we consider are:

(i.) *Two-way splits*: These divisions are "coalitionally rational"[8], i.e. no subcoalition of an existing coalition can enforce a division that improves its collective payoff.[9]

(ii.) *Even two-way splits*: The three even two-way splits comprise the bargaining set. Roughly speaking, this is the set of divisions where any profitable deviation by a subset of a coalition can be followed by another profitable deviation where the original deviators are worse off and those they abandoned are no worse off. It is also the kernel and nucleolus, and focal in that it is an equal division among members of a minimum winning coalition.

(iii.) *The three-way even split*: This is the Shapley value: each player earns the average of what they could contribute to all possible coalitions. It is also focal as an equal division among all players.

[7] Some of these categories are consistent with several solution concepts. However we are not interested in distinguishing between theories but rather in identifying divisions of interest around which to structure our analysis.

[8] Aumann and Maschler (1961)

[9] Note that this requirement is less stringent than for the core, which allows for deviations by *any* coalition. The core here is empty.

4.1. Hypotheses

We use the divisions identified by cooperative solution concepts to generate two types of hypotheses. First, in order to test whether the solution concepts have any predictive power in our experimental environment, we treat each type of division in isolation by comparing its properties to the average over *all* other divisions, i.e. including the other theoretically identified divisions (hypotheses 1 a-c). We then make comparisons *between* the types of divisions which have theoretical reasons to be of interest (hypotheses 2 a-c). Because we have no a priori reason to favor any of the solution concepts, we assume that they all perform equally well in predicting outcomes.

We study two properties of divisions: their frequency and their stability. Frequency is defined as the proportion of time spent in a type of division. Clearly, more time may be spent in a given category of divisions, simply because it contains a larger number of individual divisions. To account for this, we divide the time spent in a category by the number of divisions in that category (e.g. 33 for the 2-way splits, 3 for the even two way splits) and compare the average frequency of individual divisions within each category. Stability is defined simply as the average length of a type of division in seconds.

Hypothesis 1. *Divisions identified by cooperative solution concepts are more common and/or more stable than other types of divisions. More precisely:*

a) A two-way split is more common on average and/or more stable than other divisions (coalitional rationality).
b) A two-way even split is more common and/or more stable than other divisions (bargaining set, kernel, nucleolus).
c) A three-way even split is more common and/or more stable than other divisions (Shapely value).

Hypothesis 2. *There is no difference in the stability of divisions identified by different solution concepts.*

a) A two-way even split is equally common and/or stable as the average two-way split (coalitional rationality).
b) A three-way even split is equally common and/or stable as a two-way even split (bargaining set, kernel, nucleolus).
c) A three-way even split is equally common and/or stable as the average two-way split (Shapely value).

4.2. Results

Table 1 summarizes the proportion of time spent in the different categories of divisions that relate to our hypotheses, and the average time spent in individual divisions within those categories. As explained above, our statisitical tests will be based upon the latter. Table 2 reports the the average duration of divisions contained in each category. These statistics constitute simple measures of the frequency and stability of agreements. The data are split into early (1-10) and late (11-20) rounds of the experiment.

	Proportion of time (Category)		Proportion of time (Average division)	
Division Type	Rounds 1-10	Rounds 11-20	Rounds 1-10	Rounds 11-20
4 - 4 - 4	0.21	0.20	0.21	0.20
6 - 6 - 0	0.21	0.31	0.07	0.10
Uneven Two-Way	0.25	0.35	0.01	0.01
Uneven Three-Way	0.33	0.14	0.01	0.00
Two-Way	0.46	0.66	0.01	0.02
Three-Way	0.54	0.34	0.01	0.01

Table 1: Frequency of divisions

	Average Length	
Division Type	Rounds 1-10	Rounds 11-20
4 - 4 - 4	3.74	6.34
6 - 6 - 0	1.98	2.01
Uneven Two-Way	1.63	1.55
Uneven Three-Way	1.61	1.40
Two-Way	1.77	1.73
Three-Way	2.07	2.56

Table 2: Stability of divisions

As Tables 1 and 2 reveal, both the frequency and stability of different divisions change over time. In particular, we see that three-way splits become more rare and two-way splits more common. This pattern is driven by a decline in *uneven* three-way splits. It is worth emphasizing that the majority of the points in the simplex are uneven three-way splits (see Figure 1). In the final rounds of the experiment, the vast majority of agreements occur somewhere along the edge of the simplex (two-way splits), and most of the remaining agreements are at the very center of the simplex $(4-4-4)$. This suggests that subjects are learning either to split equally in a 'grand' coalition, or to form minimum winning coalitions and completely exclude one player. Another observable pattern is that the *stability* of three-way even splits (as measured by the length of time that they last) increases substantially in later rounds.

Given these visible learning effects, our empirical analysis will focus on the last ten rounds of the experiment. We will indicate in footnotes if results differ substantially when the earlier rounds are taken into account. We conduct binomial tests using the eight independent matching groups as units of observation. This means that a relationship holding in seven out of eight groups implies significance at $p = 0.035$ for one-tailed tests or $p = 0.070$ for two-tailed tests, and in all eight groups at $p < 0.01$ for both types of tests. We use one-tailed tests where we have a clear directional prediction (hypotheses 1 a-c), and two-tailed tests otherwise (hypotheses 2 a-c).

Coming back to Table 1, we see that two-thirds of the time is spent in two-way splits, with the average two-way split occurring significantly more frequently than the average three-way split($p < 0.01$)[10], but there is no difference in average length. Even two-way splits account for 31% of the time and are the second most stable type of division, lasting 2.01 seconds on average. These divisions are also significantly more frequent ($p < 0.01$) but not more stable ($p = 0.145$) than the average over all other types of divisions.[11] One fifth of the time is spent in three-way even splits, which are the most stable types of divisions, lasting 6.34 seconds on average. They are significantly more frequent ($p = 0.035$) and stable ($p = 0.035$) than the average over all other types of divisions.

Result 1. *a) Two-way splits are more common than the average over all other divisions. We find no statistical evidence that they are more stable.*

b) Two-way even splits are more common than the average over all other divisions. We find no statistical evidence that they are more stable.

c) Three-way even splits are more common and more stable than the average over all other divisions.

Next, we compare the different types of divisions of interest to one another. We find that two-way even splits are more frequent ($p < 0.01$) and more stable ($p =$

[10]There is no difference in the aggregate data.

[11]Looking at all 20 periods, even two-way splits are more stable than the average over all other types of divisions ($p = 0.035$).

0.07)[12] than unequal two-way splits, when each different uneven split is considered in isolation. (Note however, that the *set* of two-way uneven splits taken together are more common than the set of two-way even splits, although this difference is not statistically significant.) Comparing the two types of even splits, while there is no difference in frequency, even three-way splits are more stable than even two-way splits ($p = 0.07$). However, three-way even splits are both more frequent ($p < 0.01$) and more stable ($p = 0.07$)[13] than the average over all two-way splits.

Result 2. *a) The two-way even split is more common and stable than the average two-way split.*

b) The three-way even split is more stable than the two-way even split. We find no statistical difference in frequency.

c) The three-way even split is more common and stable than the average two-way split.

In summary, we see that most of the time (86%) is spent at one of the divisions identified by a cooperative solution concept. Overall, 2/3 of the time is spent in two-way splits, i.e. at circles somewhere along the edges of the simplex. This is significantly larger than the time spent at two-way splits in the first 10 rounds (46%), which suggests that subjects learned over time to build 'minimum winning coalitions' ($p < 0.01$). When looking at individual divisions, the three-way and two-way even splits are most common. However, it is interesting to note that the *set* of uneven two-way splits (edges minus two-way even splits) accounts for the largest amount of time of any category. Despite the complete symmetry of our experimental setup, we thus see a substantial amount of (temporary) agreement on uneven divisions within minimum winning coalitions. This pattern hints at the transition dynamics that occurred in the experiment, an issue to which we turn in the next section.

5. Transition Dynamics and Individual Behaviour

Having established patterns in the frequency and stability of different types of divisions, we now delve deeper into the underlying causes of these results by investigating individual behaviour and the transitions between divisions that are generated.

[12] $p < 0.01$ in the aggregate data.
[13] $p < 0.01$ in the aggregate data.

We begin by stating hypotheses regarding the rationality and motivations governing subject behaviour. We then provide an overview of the observed dynamics then formally test the hypotheses.

5.1. Hypotheses

First we examine to what degree transitions that occur are consistent with myopic payoff-maximisation. Specifically, we test the following hypothesis:

Hypothesis 3. *Transitions are such that two players receive at least as much as in the previous division.*[14]

In addition, we hypothesize that players are less likely to leave an existing coalition in favor of a (myopically) higher payoff if they they perceive the existing coalition to be particularly stable and thus valuable in the long term. An obvious measure for the expected stability of an existing coalition might be the length of time that it has existed. However, such a measure would be clearly endogenous. Therefore we instead measure the stability perceived by the player being 'tempted' by counting the number of times that the coalition *partner* has refused attractive offers in the past. This behavior may be motivated by a combination of reciprocity and the expectation that a stable partnership is more valuable in the long term.

Hypothesis 4. *The probability of a subject accepting an offer is decreasing in the number of offers an existing coalition member has not accepted.*

Note that these hypotheses are complementary rather than contradictory. Hypothesis 3 states that both coalition partners receiving higher payoffs is a *necessary* condition for the formation of a new coalition. Hypothesis 4 states that the probability of such a transition actually occuring is decreasing in the number of times that a coalition partner has refused such a transition in the past. Far-sighted players may exhibit behavior that is consistent with both of these hypotheses.

5.2. Results

We begin this section with a general description of the frequency of transitions between different types, then investigate what motivations might underlie the observed transitions.

[14] Note that this hypothesis may also be consistent with far-sighted payoff maximization (see below).

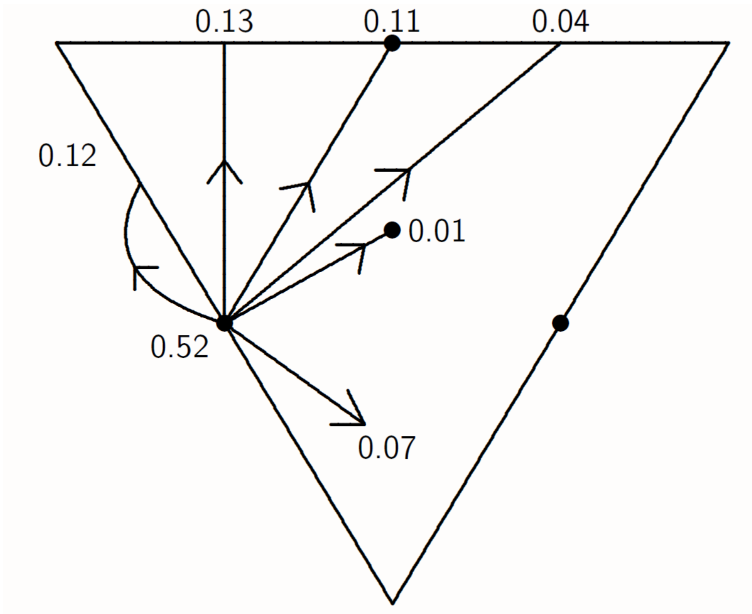

Figure 2: Transitions from 2-way equal splits

All of the figures reported in this subsection use only data from the last ten rounds.

5.2.1. Description of Transitions

To give an overview of the dynamics without introducing an unworkable level of detail, we will consider transitions between 11 different sets of circles: the three-way equal split, the three 2-way equal splits, the six segments of the edge of the simplex corresponding to uneven two-way splits, and all uneven interior circles. The main detail we lose is in transitions involving different types of non-focal interior circles and movements among uneven 2-way splits on the same edge. However these account for a relatively small number of movements overall. Since we are interested in the frequency of different *types* of transitions, we group 'symmetric' transitions in the obvious way. For example, when considering transitions from the three-way even-split to two-way even-splits, we group all such transitions rather than distinguishing between the three possible destinations.

Figure 2 illustrates the relative frequency of transitions from a two-way even-split to the other sets. In this and the following figures, each second that subjects remain in an agreement is treated as a transition to itself. The 52% probability of remaining at the equal split indicates the stability of such divisions. Conditional on moving, the majority of transitions ($\approx$ 58%) are to a circle on another edge of the simplex,

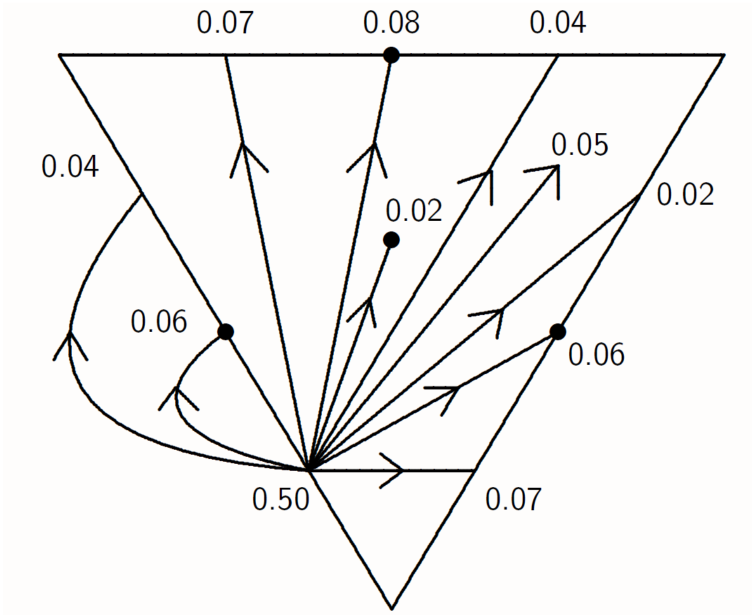

Figure 3: Transitions from uneven 2-way splits

i.e. a two-way split between the previously excluded player and one of the previously included players. Of these transitions, roughly half are to uneven splits which favor the previously included player. This indicates that such transitions are likely initiated by the previous outsider offering a 'better deal' to one of the insiders. However, roughly 1/4 of transitions originating at a two-way equal split are redistributions within the active coalition, implying a loss for one of the included players. This is much more common than a transition to a two-way split with the previously excluded player where the subject who is in both the old and new coalition loses out (8%).

Figure 3 summarizes transitions from uneven 2-way splits. Again, the 50% probability of transitioning back to the same set of circles in the next second indicates the stability of these divisions. Conditional on moving, the most likely destination is to a 2-way split between the excluded player and the disadvantaged included player (38% of transitions to other sets). This is consistent with the hypothesis that these transitions are initiated by the excluded player offering a 'better deal' to the 'cheaper' included player. However we also see a substantial amount of movement to other circles, including redistributions among the included players, mostly towards the center of the edge in question.

Finally, Figure 4 shows movements originating at the 3-way equal split. The 86% probability of remaining at this circle for another second indicates that this is the

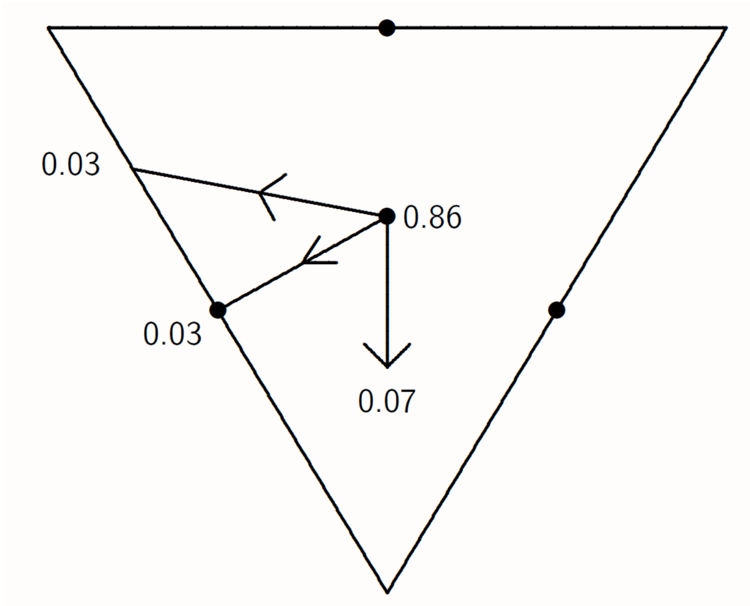

Figure 4: Transitions from the 3-way equal split

most stable division. Conditional on moving, roughly half of the transitions are to uneven internal circles, and half to either even or uneven 2-way splits.

5.2.2. Individual Behaviour

The majority of actions taken by subjects are consistent with myopic payoff-maximization, that is where subjects click on a division that would give them a per-second payoff no less than the previous agreement: 85% of offers and 75% of acceptances in the first ten rounds, rising to 87% and 78% respectively in the final ten rounds. The smallest figure for any of the eight matching groups is 72% (in both early and late rounds), so the proportion is statistically greater than 50% (p¡0.01).

Result 3. *Consistent with hypothesis 3, The majority of actions taken by subjects are consistent with myopic payoff-maximization.*

An interesting question is whether subjects can improve their long-term prospects by taking a short-term loss. For example a subject receiving seven points while their coalition partner receives five may wish to redistribute one of their points to their partner because they recognise the two-way split as being more stable. In order to approach this question in a tractable way we make the assumption that the transitions between outcomes are a Markov process: the probability of moving to a particular

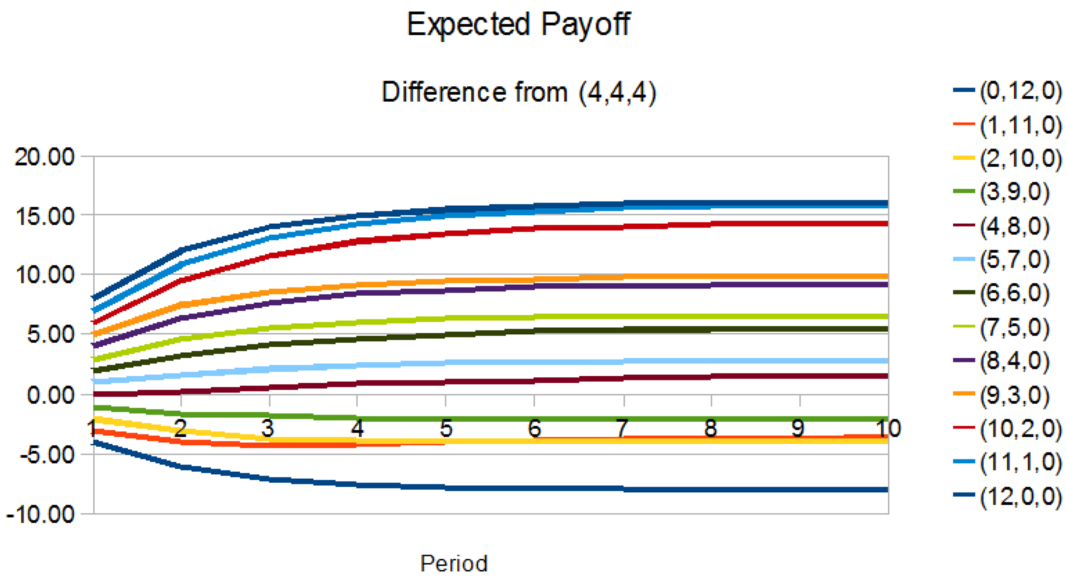

Figure 5: Expected long-term payoffs associated with divisions

division is assumed to depend only the current division.[15]

We estimate the Markov process using the empirical frequency of transitions, making the further assumption that symmetric transitions occur with the same probability (e.g. the probability of moving from any one of the two-way even-splits to either of the others is estimated using the average number of all such transitions). Using the estimated probability matrix we can calculate the expected payoff after any number of seconds given the current division.

Figure 5 shows the difference between the cumulative expected payoffs starting in all types of two-way splits and receiving the average of four points per second. Short term pain for long term gain would be seen if the lines cross. This happens marginally for the second and third lines from the bottom, meaning taking 10 rather than 11 is rational in the long term. In all other cases behaving myopically also maximises long run payoffs.

We now turn to Hypothesis 4. We do so by looking at the relationship between the probability an attractive offer by one player is accepted by a second player, and the number of times the third player has not accepted earlier beneficial offers (the

[15] We will test the Markov assumption later in this section.

variable "*prevoffs*"). Because in our setup offers are not made to a specific individual, we include only an offer where a player would get more than their current per-second payoff and consider this to be an offer to that player.[16] As shown in the first of the linear probability model regressions reported in Table 3, each additional beneficial offer not accepted by a subject's coalition partner is associated with a 7% reduction in the probability an offer is accepted.

In the second column we control for the number of points offered, the type of division that was previously in place, as well as the type of division that is suggested. The coefficient is marginally diminished, but still large (5%) and highly significant.[17] The regression in the third column interacts *prevoffs* with a dummy for the final ten rounds to see if the effect survives learning. The estimates imply a strong effect in the early rounds (10%) which is later much diminished (to 5%).

These regressions may overestimate the effect of *prevoffs* if some subjects have a fixed tendency making them less likely to accept offers. Since such a tendency will prolong the coalitions of which they are a part, their partners may tend to receive, and possibly reject, more offers. To address this possibility, we include individual fixed effects in columns 4-6. This reduces the estimated effect of *prevoffs.* However the last regression shows a significant impact in the early rounds (6%), which disappears in the later rounds.

Result 4. *Partially consistent with hypothesis 7, the probability of a subject accepting an offer is decreasing in the number of offers an existing coalition member has not accepted for inexperienced, but not experienced, subjects.*

5.3. Final payoff distributions

A natural question to ask is whether the frequent transitions between two-way splits ultimately result in an equal distribution of total payoffs, or are some subjects successful in securing larger shares than others? To investigate this question, we look at the distribution of total payoffs achieved by the three players in each of the 480 games. Figure 6 displays the result on a simplex. Fewer than half of final allocations (45%) are approximately equal, as represented by the central triangle,

[16] It may be the case that both of the two other players would be made better off, in which case it is included twice, in relationship to the number of offers rejected by each of the players.

[17] The regression also shows further evidence of the attractiveness of two-way splits and the stability of even three-way splits.

Table 3: Probability of accepting an offer (Linear Probability Models)

prevoffs	-0.0677***	-0.0538***	-0.0931***	-0.0199**	-0.0188**	-0.0631***
	(0.0146)	(0.0119)	(0.0205)	(0.00583)	(0.00760)	(0.0111)
prevoffs x last10			0.0476*			0.0544***
			(0.0243)			(0.0151)
pointsoffered		0.0442***	0.0445***		0.0389***	0.0392***
		(0.00622)	(0.00615)		(0.00511)	(0.00499)
rational		0.0543**	0.0536**		0.0143	0.0130
		(0.0214)	(0.0206)		(0.0233)	(0.0235)
even2way		0.0872**	0.0868**		0.0547**	0.0546**
		(0.0311)	(0.0312)		(0.0224)	(0.0223)
even3way		0.0141	0.0138		-0.0691	-0.0707
		(0.0615)	(0.0629)		(0.0598)	(0.0604)
prevrational		-0.00885	-0.0100		-0.0140	-0.0152
		(0.0301)	(0.0300)		(0.0213)	(0.0211)
preveven2way		0.0265	0.0257		0.0203	0.0197
		(0.0421)	(0.0417)		(0.0360)	(0.0358)
preveven3way		-0.152***	-0.153***		-0.123**	-0.122**
		(0.0406)	(0.0395)		(0.0511)	(0.0504)
Constant	0.482***	0.175***	0.176***	0.473***	0.235***	0.235***
	(0.0431)	(0.0287)	(0.0292)	(0.00110)	(0.0432)	(0.0435)
Observations	3,931	3,931	3,931	3,931	3,931	3,931
R-squared	0.011	0.075	0.076	0.001	0.040	0.041
Subject FE	NO	NO	NO	YES	YES	YES

Notes: prevoffs = number of earlier offers *not* accepted by partner in the current coalition. last10 = 1 for offers in last 10 seconds, otherwise 0. pointsoffered, rational, even2way, even3way describe the current offer. prevrational, preveven2way, preveven3way describe the current division. Standard errors (clustered by matching group) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

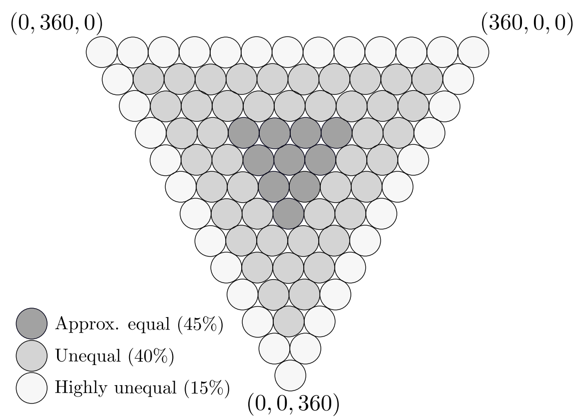

Figure 6: Distribution of final (total) payoffs

where all players receive at least $\frac{1}{4}$ of the available points. The remaining games end with unequal payoffs. In fact, 15% of games end with one player receiving less than $\frac{1}{12}$ of the available surplus. Thus, the frequent transitions between two-way splits do not (always) 'wash out', and indeed some coalitions manage to secure higher final payoffs.

6. Conclusion

We experimentally investigate behavior in a three-player majoritarian bargaining game played in continuous time. Our experiment is largely unstructured, with players being free to make proposals and agree to existing proposals at any time. The interaction is mediated by a simple graphical interface, allowing subjects to make and respond to proposals very quickly. This permits us to simulate a reasonably long time horizon within each game, while keeping games short enough to allow for repetition and learning. Our primary goal was to investigate what types of divisions arise most frequently, and which are most stable. Our categorization of divisions was based on cooperative solution concepts.

1. Even three-way and even two-way splits are more frequent than other divisions, taken individually.

2. Unequal payoffs are common, both from the formation of minimum winning coalitions, and as a result of unequal splits between coalition partners. This holds true both for per-second payoffs and the aggregates over entire games.

3. The even three-way split is more stable than the two-way splits.

4. Transitions occur frequently, and are consistent with myopic payoff maximization.

5. Behavior is history dependent: Subjects are less likely to accept an attractive offer if their current coalition partner has proved 'loyal' in the past.

As mentioned above, the structured bargaining experiment most closely related to our own is the one conducted by Battaglini and Palfrey (2012). Although our experiment was not designed to be comparable, it is interesting to note that our results are quite similar in several respects. In particular, 'universal outcomes' (equal or nearly equal splits between all three players) are the most stable, followed by 'majoritarian outcomes'. As noted earlier, one might have expected that our unstructured experiment would result in even more egalitarian outcomes than observed by Battaglini and Palfrey (2012). However, this is not the case. As in their setting, and as in other structured majoritarian bargaining experiments, two-way splits within a minimum winning coalition are the most common.[18] And indeed, even the final payoff distributions are often highly unequal. Thus, these patterns from the structured multilateral bargaining literature appear to be relatively robust. This is in contrast to unstructured *bilateral* bargaining experiments, which typically result in equal distributions. Although we have no unanimity rule treatments to compare to, a possible interpretation is that *majority rule*, by itself, favors unequal outcomes, even in the absence of asymmetries in terms of procedural advantage.

Our analysis is also suggestive of the kinds of assumptions that might generate these patterns. First, despite the prevalence of minimum winning coalitions, the frequency of even splits within such coalitions, as well as of even three-way splits, suggest that fairness concerns are important. Further, loyalty in the rejection of some tempting offers suggests that reciprocity also plays an important role. Finally, increasing loyalty over the duration of a given coalition suggest that a Markov assumption on transitions, as is commonly used in theoretical work (e.g. Konishi and Ray, 2003),

[18]Remarkably, the observed proportion of minimal winning coalitions, two thirds, is very close to what is found in Frechette et al. (2003) in the context of a standard (one shot, closed rule) Baron-Ferejohn game.

may be inappropriate.

One of the contributions of this paper was to develop a simple interface for computer-mediated multilateral bargaining experiments. An important feature of our program is that it is time-efficient, requiring only a brief tutorial for subjects to understand its use, and allowing for a rapid exchange of offers and acceptances. The numerous learning effects we identify highlight the importance of designing experiments that allow for a reasonable number of repetitions to enable subjects to come to grips with a complex strategic environment. In this paper we have used the program to investigate a repeated environment with symmetric players, but it can easily be adjusted to look at other questions. In another experiment we have modified the program to implement a one-off division in an infinite horizon framework to test the robustness of experimental findings obtained using the Baron-Ferejohn framework. Future work will consider situations involving heterogeneous preferences and asymmetric information.

Appendix A. Instructions

Subjects received both printed and onscreen instructions. Onscreen instructions involved a tutorial in which subjects could try out the software used in the experiment.

Onscreen Instructions (1)

The first screen contained the following text.

In this experiment, participants will interact in groups of three. The interaction will be repeated a number of times. In what follows, we will call each repetition of the interaction a “round”.

During each round, the three members of a group will distribute 3600 points between themselves.

The points you receive will determine how much money you earn at the end of the experiment.

In addition to the money you earn during the interactions, you will also receive 3 Euros for filling in a questionnaire.

In order to help you understand exactly how the interactions work, we will now demonstrate the functioning of the program on your screen.

There will be a short tutorial followed by two practice rounds.

In the tutorial you will be shown the screen that will be used during the real rounds. How to understand this screen and use it to interact with the other participants in your group is explained to you in the printed instructions you have been given.

During this tutorial and the two practice rounds which follow no money will be awarded. The purpose of this tutorial is only to help you understand how the program works. You will be informed before the real interactions begin.

Please follow the printed instructions carefully. It is important to understand how the program works!

PLEASE CLICK “CONTINUE” IF YOU ARE READY TO BEGIN THE TUTORIAL

Printed Instructions

The following instructions were provided in hardcopy. The tutorial exercises were conducted onscreen while reading the instructions. See Figure 1 for a screen shot.

Tutorial: DO NOT CLICK CONTINUE UNTIL YOU HAVE COMPLETED AND UNDERSTOOD THE EXERCISES BELOW.

IF YOU HAVE ANY QUESTIONS, PLEASE RAISE YOUR HAND AND SOMEONE WILL COME TO ASSIST YOU AS SOON AS POSSIBLE.

- The participants in your group must decide how many points each will receive at any moment in time. A total of 12 points can be allocated every second.
- On your screen, you will see small circles arranged in the form of a large triangle. Each circle represents a different way of allocating the 12 points among you and the other two participants you are interacting with in a given round.

- The corners of the triangle are labeled "You", "Participant A", and "Participant B".
- The points in the corners correspond to allocations in which the indicated participant receives all 12 points, while the others receive no points.
- The closer a point is to a given corner, the more the corresponding allocation assigns to that participant.
- If you move your mouse over a circle, the corresponding points to be allocated to each subject are displayed in the appropriate corners of the triangle.
- Exercise:
  - Move your mouse around the triangle until you understand how the circles are arranged.
  - Click on a point to select it. Notice that the selected point is circled in green (for now, do not click on the circles that are already marked).
  - Click on a point outside the triangle. The green circle should disappear.
  - During the real interactions, when you click on a circle this will be shown on the screens of the two other participants you are interacting with, and it will disappear from their screens if you click outside the triangle. If two participants have their circles showing on the same circle, points will start to be distributed according to this division. This is explained in more detail below.
- During the real interactions, any of the three participants (including you) may click on any circle at any time.
- When you click on a circle, the other two participants in your group will see this circle marked on their screen.
- When another participant in your group clicks on a circle on their screen, it will be marked on your screen in the color corresponding to that participant: orange for"Participant A" blue for "Participant B". You can see how this looks now on your screen.
- Clicking on an unmarked circle is like suggesting that division.
- Clicking on a circle marked by another participant is like (temporarily) accepting the division they have suggested.
- When another participant clicks on a circle you have marked, they have (temporarily) accepted your suggestion.
- Points are distributed for as long as at least two members of a group select the same circle in the triangle.
- These agreements are not permanent and can be changed, as any participant can click on another circle at any time.
- If at least two participants are simultaneously choosing the same circle, the corresponding point allocation is implemented: each second, each participant receives the number of points assigned to them. All subjects will temporarily stop receiving points if subjects click on different circles such that all members of the group are now selecting different circles.
- The points a participant receives do not depend on whether she herself has selected the circle

being implemented. In particular, it is possible that a subject who has clicked on the circle receives no points, and that a subject who has not clicked on the circle receives points. The allocation of points depends only on the location of the point in the triangle, not on the subjects that are selecting it (as long as there are at least two).

- In this tutorial there is no limit to the number of the points you can distribute. This is so you have time to understand how the program works. In the real interactions the round will end when 360 points have been distributed among you and the other two participants in your group.
- Exercise:
  - On your screen, the computer has randomly selected two circles on behalf of your hypothetical group partners.
  - Experiment by clicking on each of these circles, as well as other circles.
  - Watch how your points accumulate when you select the same circle as another group member, and how points stop accumulating when you subsequently click on a circle different from those selected by either of your partners. (Naturally, points are also allocated whenever the other two members of your group select the same circle. In that case, the points you receive will not depend on which circle you are selecting.)
  - Notice that an agreed / active circle is highlighted in red, and that the shares of points that correspond to the agreed circle appear in large red font at the appropriate corners when your mouse is outside the triangle.

WHEN YOU HAVE READ AND UNDERSTOOD THE INSTRUCTIONS AND COMPLETED THE EXERCISES, PLEASE CLICK CONTINUE.

THIS TUTORIAL WILL BE FOLLOWED BY TWO PRACTICE ROUNDS, AND YOU WILL HAVE THE OPPORTUNITY TO ASK FURTHER QUESTIONS BEFORE THE REAL ROUNDS BEGIN.

Onscreen Instructions (2)

After completing the turorial using the printed instructions, the following screen was displayed.

There will now be two practice rounds.

During these practice rounds you will be interacting with the computer. The divisions the computer chooses will be random. Everything else will be the same as in the real rounds which will follow.

There will be three differences between the previous screen and the practice (and real) rounds:

- The divisions chosen by "Participant A" and "Participant B" will move as they click on different circles. Sometimes you will not see their circles on the triangle: this is when they have not yet clicked on a circle, or have clicked off the triangle.

- There are only 360 points to be awarded. When these points have run out the round will end. The number of points remaining to be distributed will be shown next to the number of points you have currently received (check this on the picture on the last page of printed instructions).
- If not all points have been distributed after 300 seconds, the round will end automatically and any remaining points will be lost. The number of seconds remaining before the round ends automatically will be shown below the triangle (check this on the picture on the last page of printed instructions).

You will have an opportunity to ask questions after each of the two practice rounds.

During these practice rounds, no money will be awarded. The purpose of the practice rounds is only to help you understand how the program works. You will be informed before the real interactions begin.

It is more important during the practice rounds to learn to understand how the program works than to get the division of points that you want.

You may now ask questions about the way the program works, however please do not ask questions about strategies and divisions you or the computer may use. These types of questions will not be answered.

If you have a question, please raise your hand and ask it quietly when an experimenter comes.

If you have no questions, please click "Continue" and wait for the first practice round to begin.

References

Anbarci, N. and N. Feltovich (2013). How sensitive are bargaining outcomes to changes in disagreement payoffs? *Experimental economics 16*(4), 560–596.

Aumann, R. J. (1961). The core of a cooperative game without side payments. *Transactions of the American Mathematical Society*, 539–552.

Aumann, R. J. and M. Maschler (1961). *The bargaining set for cooperative games*. Defense Technical Information Center.

Baron, D. P. and J. A. Ferejohn (1989). Bargaining in legislatures. *The American Political Science Review*, 1181–1206.

Battaglini, M. and T. R. Palfrey (2012). The dynamics of distributive politics. *Economic Theory 49*(3), 739–777.

Berninghaus, S. K., K.-M. Ehrhart, and M. Ott (2006). A network experiment in continuous time: The influence of link costs. *Experimental Economics 9*(3), 237–251.

Bigoni, M., M. Casari, A. Skrzypacz, and G. Spagnolo (2015). Time horizon and cooperation in continuous time. *Econometrica 83*(2), 587–616.

Davis, M. and M. Maschler (1965). The kernel of a cooperative game. *Naval Research Logistics Quarterly 12*(3), 223–259.

Diermeier, D. and R. Morton (2005). Experiments in majoritarian bargaining. In *Social Choice and Strategic Decisions*, pp. 201–226. Springer.

Diermeier, D., R. I. Swaab, V. H. Medvec, and M. C. Kern (2008). The micro-dynamics of coalition formation. *Political Research Quarterly 61*(3), 484–501.

Fiorina, M. P. and C. R. Plott (1978). Committee decisions under majority rule: An experimental study. *The American Political Science Review*, 575–598.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics 10*(2), 171–178.

Frechette, G., J. H. Kagel, and M. Morelli (2005). Nominal bargaining power, selection protocol, and discounting in legislative bargaining. *Journal of Public Economics 89*(8), 1497–1517.

Frechette, G. R., J. H. Kagel, and S. F. Lehrer (2003). Bargaining in legislatures: an experimental investigation of open versus closed amendment rules. *American Political science review 97*(02), 221–232.

Friedman, D. and R. Oprea (2012). A continuous dilemma. *The American Economic Review*, 337–363.

Greiner, B. (2004). An online recruitment system for economic experiments.

Horowitz, A. D. (1977). A test of the core, bargaining set, kernel and shapley models in n-person quota games with one weak player. *Theory and decision 8*(1), 49–65.

Kahan, J. P. and R. A. Helwig (1971). Coalitions: A system of programs for computer-controlled bargaining games. *General Systems 16*, 31–41.

Kahan, J. P. and A. Rapoport (1974). *Test of the bargaining set and kernel models in three-person games.* Springer.

Kalandrakis, A. (2004). A three-player dynamic majoritarian bargaining game. *Journal of Economic Theory 116*(2), 294–322.

Kalisch, G., J. Milnor, J. Nash, and E. Nering (1996). Some experimental n-person games. *Essays on Game Theory*, 61.

Konishi, H. and D. Ray (2003). Coalition formation as a dynamic process. *Journal of Economic Theory 110*(1), 1–41.

McKelvey, R. D. (1991). An experimental test of a stochastic game model of committee bargaining. *Laboratory Research in Political Economy, T. Palfrey ed. University of Michigan Press: Ann Arbor*, 139–69.

Miller, L. and C. Vanberg (2013). Decision costs in legislative bargaining: an experimental analysis. *Public Choice 155*(3-4), 373–394.

Nagarajan, M. and G. Sošić (2008). Game-theoretic analysis of cooperation among supply chain agents: Review and extensions. *European Journal of Operational Research 187*(3), 719–745.

Nash, J. F., R. Nagel, A. Ockenfels, and R. Selten (2012). The agencies method for coalition formation in experimental games. *Proceedings of the National Academy of Sciences 109*(50), 20358–20363.

Nolutshungu, S. C. (1996). *Limits of anarchy: Intervention and state formation in Chad*. University of Virginia Press.

Rapoport, A. and J. P. Kahan (1976). When three is not always two against one: Coalitions in experimental three-person cooperative games. *Journal of Experimental Social Psychology 12*(3), 253–273.

Roth, A. E. and M. W. Malouf (1979). Game-theoretic models and the role of information in bargaining. *Psychological review 86*(6), 574.

Schmeidler, D. (1969). The nucleolus of a characteristic function game. *SIAM Journal on applied mathematics 17*(6), 1163–1170.

Shapley, L. S. (1952). A value for n-person games. Technical report, DTIC Document.

Teteryatnikova, M. and J. Tremewan (2015). Stability in network formation games with streams of payoffs: An experimental study.