

# Power and preferences: an experimental approach

Fuad Aleskerov\*, Alexis Belianin<sup>†</sup>, Kirill Pogorelskiy<sup>‡</sup>

March 7, 2009

Preliminary and incomplete

## Abstract

An experimental approach to study the distribution of power in a voting body is described. Laboratory experiments in collective decision-making provide a relatively new way of measuring voting power that 1) is based on empirical data, 2) is analytically tractable and 3) allows to control voters' preferences in a theoretically unambiguous environment. Taking the experiment by Montero, Sefton and Zhang (2008) (MSZ) as starting point, we confirm their basic findings, and explain some of their empirical paradoxes. In particular, we show that the asymmetry between the resulting shares of some otherwise identical players is a feature of the experimental design of MSZ, and develop an experimental way to correct for this. The main contribution of our research deals with the question of how voters' preferences to coalesce influence their behaviour and the resulting allocation of shares (as measured by the average payoffs of the players). To tackle the issue experimentally we extend the basic design to allow for asymmetric voters' preferences depending on the coalitions they take part in. The results of the experiments show that even small modifications of preferences lead to statistically significant differences in players' shares. This result supports preference-based power indices as proposed in Aleskerov (2006), rather than the classical Banzhaf and Shapley-Shubik indices.

Keywords: voting power, preferences, experiments

---

\*Department of Higher mathematics for economics, Higher School of Economics, and Institute of Control Science of Russian Academy of Sciences, email: alesk@hse.ru

<sup>†</sup>International College of Economics and Finance, Higher School of Economics, and Institute for World Economy and International Relations, Russian Academy of Sciences, email: icef-research@hse.ru

<sup>‡</sup>Department of Higher mathematics for economics, Higher School of Economics, email: kpogorelskiy@hse.ru

# 1 Introduction

Voting power studies have evolved considerably since the first classic papers were published in the 1950-70s (see Penrose (1946); Shapley&Shubik (1954); Banzhaf (1965); Coleman (1971) ). However, the central question - that of the power possessed by members of a voting body, remains unsettled. Despite several competing attempts made in political science (Dahl (1969); Lukes (1974); Parsons (1986)), there is even less agreement on how it should be properly measured. Some classes of power measures which possess desirable properties or satisfy certain natural axioms may not be universally applicable in other contexts.

In our viewpoint, the main reason for such situation is that the proposed measures could hardly be verified in real-life conditions. This is due to not only lack of the appropriate data, but also to theoretically convenient but practically unobservable properties of certain measures (e.g., how can one reliably estimate the probability distribution on all possible divisions of a voting body into those voting ‘aye’ and ‘nay’?). There exist approaches that deal with these issues (e.g., in Aleskerov et al (2007) a very thorough analysis was made based on the data on voting behavior of individual voters collected for a long period), but in general most properties of power measures (and thus their appropriateness for use for voting power analysis) are based on theoretical assumptions, examples or simulation and not on empirical data. Indeed, one cannot come and ask, for instance, a member of parliament, what her real power is, not only because she might be unwilling to answer, but also because, depending on the situation, her power may be different and even unknown to herself, and/or perceived in a biased manner.

In the present paper we explore experimentally the predictive power of voting power indices in different contexts and two generically different settings. The first of these is a classical coalitional game with or without veto power (see Montero, Sefton & Zhang (2008), henceforth referred to as MSZ ), together with the enlarged treatment in which the addition of an extra player changes the voting power indices of the original members notwithstanding their weights and the decision rule (Brams & Affuso (1976)). The second, and novel part of experiment, measures predictive power of the generalized power indices, which incorporates into the model a possibility that players have asymmetric preferences towards each other, and thus may be more inclined to form some coalitions rather than others. We confirm experimentally that these preferences, however small, may have large effects and change the likelihood of observing different coalitions and different gains accrued to the players.

The remainder of the paper is organized as follows. Section 2 describes the main theoretical notions and voting indices in both classical and preference-adjusted cases. Section 3 outlines hypotheses designed to be checked by the experiments, and the general experimental setup. Section 4 deals with the experimental design and procedures. Section 5 contains our main results (still under completion). Section 6 concludes.

## 2 Main notions

Consider the following example (Aleskerov (2006)), which we preclude by a few definitions. A *coalition*  $S$  is any subset of  $N$  players,  $|N| = n$ . The set of all possible coalitions is denoted by  $2^N$ . Each player  $i$  has a certain number of votes  $w_i$  he or she may use to support a bill. A *quota*  $q$  is the least number of votes required to pass a bill. A coalition  $S \subseteq 2^N$  is *winning* iff  $\sum_{i \in S} w_i \geq q$ . Similarly, a *losing* coalition is the one that lacks enough votes for a bill to pass. A player  $i \in S$  is said to be *pivotal* in a coalition  $S$  if  $S$  is winning while  $S \setminus \{i\}$  is losing. Let the payoff of a coalition  $S$  be denoted by  $v(S)$ ; we define  $v(S) = 1$  iff  $S$  is winning and  $v(S) = 0$  iff  $S$  is losing in that vote. A dual notion of *swing* is also useful: a coalition  $S : i \notin S$  is a swing for player  $i$  if  $S$  is losing, while  $S \cup \{i\}$  is winning. We denote the set of coalitions to which player  $i$  is pivotal (resp., swing) through  $\varsigma_i$  (resp.,  $\xi_i$ ).

Suppose now we have a parliament with three parties  $A$ ,  $B$  and  $C$  having each 50, 49 and 1 votes, respectively. Let the quota be set at 51 votes. Then the winning coalitions are  $\{A, B\}, \{A, C\}, \{A, B, C\}$  and  $A$  is pivotal in all coalitions,  $B$  is pivotal in the first, and  $C$  – in the second one. To quantify somehow the power the players possess, several power indices have been proposed, of which the Banzhaf index  $\beta$  (Banzhaf (1965)) is perhaps the most intuitive. This (normalized) index shows the relative proportion of winning coalitions in which player  $i$  is pivotal with regards to all other players, or

$$\beta_i = \frac{\sum_S (v(S) - v(S \setminus \{i\}))}{\sum_{i=1}^N \sum_S (v(S) - v(S \setminus \{i\}))} \quad (1)$$

The Banzhaf index  $\beta$  in our example yields the following power distribution:  $\beta(A) = 3/5, \beta(B) = \beta(C) = 1/5$ .

Suppose now that for some reason parties  $A$  and  $B$  are 'implacable enemies' and thus together will never form a coalition on their own. Then the only winning coalition is  $A + C$ . In this case  $B$  has zero voting power despite its 49 votes, whereas the power of  $C$ , having only one vote equals that of  $A$  with 50 votes. This example shows that estimates of the voting power can change significantly if voters' preferences to coalesce are taken into account (Aleskerov (2006)).

Let the function  $f_i(S)$  be the *intensity of connections* between the player  $i$  and the coalition  $S \subseteq N, f_i(S) : N \times 2^N \rightarrow \mathbb{R}$ . For each player  $i$ , let  $\chi_i = \sum_{S \in \xi_i} f_i(S)$  be the sum of intensities of connections of player  $i$  over all those losing coalitions which are swings for  $i$  (alternatively, this definition may be stated in terms of coalitions in which  $i$  is pivotal). Then the voting power index of the agent  $i$  is defined in the manner of Banzhaf index as

$$\alpha_i = \frac{\chi_i}{\sum_{j \in N} \chi_j} = \frac{\sum_{S \in \xi_i} f_i(S)}{\sum_{j \in N} \sum_{S \in \xi_j} f_j(S)} \quad (2)$$

The very idea of  $\alpha_i$  is similar to that of the Banzhaf index, the difference being that in the definition of the latter we evaluate the number of coalitions which

are swings for  $i$ , and not the intensity of  $i$ 's connections with such coalitions. Analogous definition can be stated in terms of pivotal players, with obvious renormalization. The main question remaining is how to construct the intensity functions  $f_i(S)$ .

Assume that the desire of the agent  $i$  to coalesce with  $j$  is given by a real number  $p_{ij}, i, j = 1, \dots, n$ , which we refer to as *modifiers*. (In general, it is not required that  $p_{ij} = p_{ji}$ ). The following forms of intensity functions may be defined:

a) Mean intensity of  $i$ 's connection with other members of  $S$ :

$$f_i^+(S) = \frac{\sum_{j \in S} p_{ij}}{|S|} \quad (3)$$

b) Mean intensity of connection of other members of  $S$  with  $i$ :

$$f_i^-(S) = \frac{\sum_{j \in S} p_{ji}}{|S|} \quad (4)$$

Naturally, other forms of intensity functions are possible (see Aleskerov (2006)). As an illustration, recall again our example of three voters with the number of votes of 50, 49, and 1, and the quota of 51. The set of all winning coalitions is  $\{S_1 = \{1, 2\}, S_2 = \{1, 3\}, S_3 = \{1, 2, 3\}\}$ , of which player 1 is pivotal in all three ( $\varsigma_1 = \{\{1, 2\}, \{1, 3\}, \{1, 2, 3\}\}$ ), player 2 is pivotal in the first ( $\varsigma_2 = \{\{1, 2\}\}$ ), and player 3 in the second coalition ( $\varsigma_3 = \{\{1, 3\}\}$ ).

Assume the preferences are ordinal such that  $3 \succ_1 2, 3 \succ_2 1, 1 \succ_3 2$ . Since the scale of preferences does not matter<sup>1</sup>, we let the value of  $p_{ij}$  for the preferred partner be 2, for the inferior partner be 1, for oneself, 0:

$$\|p_{ij}\| = \begin{Bmatrix} \begin{array}{|c|c|c|c|} \hline & 1 & 2 & 3 \\ \hline 1 & 0 & 1 & 2 \\ \hline 2 & 1 & 0 & 2 \\ \hline 3 & 2 & 1 & 0 \\ \hline \end{array} \end{Bmatrix}$$

Table 1. Preferences of the players.

---

<sup>1</sup>Although, one should note the possible need to change the normalizing constant in (3) and (4) such that the equations are adjusted for the 'swing-based' definition as well as the 'pivotal-based' one.

Intensities of connections across coalitions with  $f^+$  intensity function defined in terms of pivotal players are

	$S_1$	$S_2$	$S_3$
player 1	$f_1^+(S_1) = \frac{1}{2}$	$f_1^+(S_2) = \frac{2}{2} = 1$	$f_1^+(S_3) = \frac{1+2}{3} = 1$
player 2	$f_2^+(S_1) = \frac{1}{2}$	–	$f_2^+(S_3) = \frac{1+2}{3} = 1$
player 3	–	$f_3^+(S_2) = \frac{2}{2} = 1$	$f_3^+(S_3) = \frac{2+1}{3} = 1$

Table 2. Intensities of connections.

Sums of intensities of connections over the winning coalitions are:

- Player 1 is pivotal in 3 coalitions, so  
 $\chi_1 = f_1^+(S_1) + f_1^+(S_2) + f_1^+(S_3) = 1/2 + 1 + 1 = 5/2$
- Player 2 is pivotal in 1 coalition, so  
 $\chi_2 = f_2^+(S_1) = 1/2$
- Player 3 is pivotal in 1 coalition, so  
 $\chi_3 = f_3^+(S_2) = 1$

The generalized power indices are then:

$$\alpha_1 = \frac{\chi_1}{\sum_{i=1}^3 \chi_i} = 5/8,$$

$$\alpha_2 = \frac{\chi_2}{\sum_{i=1}^3 \chi_i} = 1/8,$$

$$\alpha_3 = \frac{\chi_3}{\sum_{i=1}^3 \chi_i} = 1/4.$$

As we see,  $\alpha_i \neq \beta_i$ : in particular,  $\alpha_1 - \beta_1 = \frac{1}{40}$ ,  $\alpha_2 - \beta_2 = -\frac{3}{40}$ ,  $\alpha_3 - \beta_3 = \frac{2}{40}$ . Player 3, who has just one vote, becomes more powerful than player 2, who has almost as many votes as player 1. This is due to the fact that player 3 is the preferred partner to both players 1 and 2, while player 2 is not preferred to anyone. Also, player 1 is most preferred by one of the players (player 3), and hence receives even more voting power. These conclusions, however, are clearly conditional on our specification of intensities of preferences, and are *ad hoc* in that sense.

These observations suggests several questions. First, although the generalized power indices  $\alpha$  are intuitively quite compelling, it should be checked whether they can be validated experimentally; and if confirmed, how robust is their influence on the players' behaviour in different contexts where the players' number of votes and the quota are varied.

Then, it is interesting to see which of the numerous ways to define the intensity functions is justified in practice; and whether the scale of intensities matters: if yes, cardinal intensities should be deemed superior to ordinal ones. Finally, some quantitative estimates of the effects of modifiers on the outcomes of the votes and players' payoffs are worth investigating.

To address these questions, we use laboratory experiments in collective decision-making because this approach 1) is based on empirical data, 2) is analytically tractable and 3) allows for direct control over the relevant aspects of a decision-making process in the theoretically unambiguous circumstances.

### 3 Experimental setup

The starting point for this paper was the experimental setup by MSZ, although the main emphasis of their work was somewhat different<sup>2</sup> In the experiment carried out by MSZ, participants were able to propose and vote on how to distribute a fixed budget among them, using the average share of the budget accrued to each voter as an empirical measure of the voting power. They have considered three treatments, corresponding to the examples in Brams and Affuso (see Table 3).

VETO treatment				
Player#	1	2	3	
Votes#	3	2	2	
Quota	5			
SYMMETRIC treatment				
Player#	1	2	3	
Votes#	3	2	2	
Quota	4			
ENLARGED treatment				
Player#	1	2	3	4
Votes#	3	2	2	1
Quota	5			

Table 3. Three experimental treatments (MSZ).

In all treatments there is a 'strong' player (with three votes) and two 'weak' players (with 2 votes). In addition, in the enlarged game there is also a newcomer – player 4 with 1 vote.

Whether the game is VETO or SYMMETRIC is defined by the quota value: in the VETO game it is set at 5 votes meaning that player 1 has veto power, whereas in the SYMMETRIC game the power of each player is a priori equal. In the ENLARGED game a new member is added and this setting was used to test against the paradox occurrence with regards to the VETO and SYMMETRIC games. We will refer to these treatments for short as V-game, S-game and E-game, respectively.

MSZ found significant differences among treatments, which exhibit the paradox of new members: the empirical voting power of some existing members is increased with the addition of a new member.

In all three games the players divide the fixed budget of 120 points among them and at the end of experiment were paid in accordance with their earnings during the game. The average payoffs of the players were taken as empirical

<sup>2</sup>MSZ studied the so-called *paradox of the new members*: when a new member is added to a voting body, the power indices of some original members may increase even if their weights and the decision rule remain the same. Brams&Affuso (1976) argued that this is a paradox related to voting power rather than to the mathematical properties of the power index chosen. This conclusion, however, was questioned in the literature (Barry (1980)).

measures of their power and compared to the theoretical estimates, given by the Banzhaf and Shapley-Shubik (Shapley&Shubik (1954)) power indices as well as by the nucleolus (Schmeidler (1969)).

Each game consisted of 10 rounds. In each round the players were able to make proposals about how to divide 120 points among them and cast votes in favour of proposals already on the table (whether their own, or made by other players). The first proposal to receive enough votes to satisfy the quota (that is, 5 in V- and E-games, or 4 in S-game) was enforced, i.e. the round ended and each person earned the number of points specified in that proposal. The total earnings of each player in 10 rounds were then converted to money. Our experiments followed this process. The results are given in Section 5.

Let us compare this case to the situation when players have different preferences to coalesce, that is, introduce the modifiers  $p_{ij}$ . Assuming cardinal preference model, if the modifiers are set to 1 for all players, everyone is having equal preferences to coalesce with all other players, and the generalized indices  $\alpha$  will clearly coincide with the Banzhaf power index  $\beta$ . This is our control setting for the three games studied in MSZ, which we use as benchmark case.

Now assume that the modifiers can be different for some players, depending on other players she is in coalition with. In particular, the payoff a player receives is multiplied by her average preference towards other players of her coalition. For instance, if the proposal was passed by the winning coalition  $\{1, 2\}$ , and player 1 has modifier 0.5 with regards to player 2, while the latter player's respective modifier equals 1 (i.e.,  $p_{12} = 0.5, p_{21} = 1$ ), then player 1 will receive only half of the share specified in that proposal, while player 2 will receive her share in full. Such restrictions are introduced to bind players to vote in accordance with the modifiers ascribed to them. Moreover, we can model the game with imperfect information if the modifiers are not commonly observable and known by the players, or a player knows her own preferences only.

Below we state the a priori conjectures to be verified by our experiments.

### 3.1 Symmetric game

In the symmetric case (see Table 3), the following coalitions are swings for the players:

For player 1:  $\{2\}, \{3\}$

For player 2:  $\{1\}, \{3\}$

For player 3:  $\{1\}, \{2\}$

The Banzhaf index, which treats all coalitions as equiprobable, predicts that payoffs will be equal to 40 (out of 120 points) for each player.

Now consider the S-game with the modifiers as defined in Table 4. These modifiers define our treatment condition, which we refer to as the 1-game.

Preferences of $i \setminus$ towards $j$	player 1	player 2	player 3
player 1	-	1	1
player 2	1	-	1.01
player 3	1	1	-

Table 4. Modifiers for the S-game.

If we take into account the voter's preference to coalesce, and measure the intensities by intensity function  $f^+$  (3), we receive the following estimates of the voters' desire to form coalitions:

$$\begin{aligned}
f_1^+ (\{2\}) &= \frac{1}{|\{2\}|} = 1; \\
f_1^+ (\{3\}) &= \frac{1}{|\{3\}|} = 1; \\
f_2^+ (\{1\}) &= \frac{1}{|\{1\}|} = 1; \\
f_2^+ (\{3\}) &= \frac{1.01}{|\{3\}|} = 1.01; \\
f_3^+ (\{1\}) &= \frac{1}{|\{1\}|} = 1; \\
f_3^+ (\{2\}) &= \frac{1}{|\{2\}|} = 1.
\end{aligned}$$

The predicted shares as defined by  $\alpha$  power index are then 39.9334 for players 1 and 3, and 40.1331 for player 2.

On the other hand, if we measure the intensities by intensity function  $f^-$  (4), we receive the following estimates:

$$\begin{aligned}
f_1^- (\{2\}) &= \frac{1}{|\{2\}|} = 1; \\
f_1^- (\{3\}) &= \frac{1}{|\{3\}|} = 1; \\
f_2^- (\{1\}) &= \frac{1}{|\{1\}|} = 1; \\
f_2^- (\{3\}) &= \frac{1}{|\{3\}|} = 1; \\
f_3^- (\{1\}) &= \frac{1}{|\{1\}|} = 1; \\
f_3^- (\{2\}) &= \frac{1.01}{|\{2\}|} = 1.01
\end{aligned}$$

Now  $\alpha$  power index predicts that the pie shares are 39.9334 for players 1 and 2, and 40.1331 for player 3.

Arguably, both cases can be viewed as minor changes of payoffs. Yet we expect the coalition  $\{2, 3\}$  to occur significantly more frequently than other coalitions in the 1-game (H1). In addition, earnings of players 2 and/or 3 will be higher in the 1-games (with modifiers from Table 4) than in the control S-game (H2).

If either of these hypotheses is correct, we have evidence that preference do matter: small changes in material payoffs make big difference. Moreover, H2 is supported by some recent works in biased preferences — e.g., Ariely (2008); Warber et al (2008). These works imply that small (indeed immaterial) changes in stimuli may have their 'symbolic value', leading to visible and significant differences in the behaviour observed. Drawing on these conjectures, we expect player 1 (the strong player) to receive less than his predicted share in the 1-game treatment. Hence, it is possible that, *upon getting experienced, player 1 will try to 'buy' one of the other players' courtesy, offering them more than in the control S-game.* This is the third hypothesis (H3)<sup>3</sup>. One would not bet on whether this will be player 2 or 3, but the size of this payment off player 1's fair share (and the share we observe in the standard game without modifiers) would give us an idea of what the price of the preferences is.

### 3.2 Veto game

The veto game differs from its symmetric counterpart in that now five votes are needed to pass the proposal (see Table 3). In this case player 1 is the veto player in the sense that no coalition can be winning without him. The following coalitions are then swings for the players:

For player 1:  $\{2\}, \{3\}, \{2, 3\}$

For player 2:  $\{1\}$

For player 3:  $\{1\}$

The Banzhaf index is  $\beta_1 = 3/5, \beta_2 = \beta_3 = 1/5$ , predicting that player 1 gets 60 and 2 and 3 – 20 points each. Now consider the V-game with the modifiers as defined in Table 5 (referred to as the 2-game):

Preferences of $i \setminus$ towards $j$	player 1	player 2	player 3
Player 1	-	1	1
Player 2	0.99	-	1
Player 3	0.99	1	-

Table 5. Modifiers for the V-game

Our first hypothesis (H4) is that *earnings of player 1 will be significantly higher than in V-game.* Furthermore, (H5) *as players 2 and 3 dislike player 1 just a bit, reached agreements (which necessarily involve player 1) will involve significantly higher shares for player 2 and 3 in the treatment 2-games than in the*

<sup>3</sup>Data collected allows us to track the preliminary proposals of the players as well, even if they are not accepted and implemented, although this analysis requires more time and effort.

control V-games with neutral modifiers of 1. Finally, we may also hypothesize that (H6) bargaining time needed to reach an agreement in the 2-games will be significantly higher than in the control V-games.

### 3.3 Enlarged games

As more players come into play, more alternative configurations are possible. Consider an enlarged treatment of MSZ as given in Table 3. The following coalitions are swings for the players:

For player 1: {2}, {3}, {2, 3}, {2, 4}, {3, 4}

For player 2: {1}, {1, 4}, {3, 4}

For player 3: {1}, {1, 4}, {2, 4}

For player 4: {2, 3}

The Banzhaf index yields  $\beta_1 = 5/12, \beta_2 = \beta_3 = 3/12, \beta_4 = 1/12$ .

Let the multiplicative modifiers be given by Table 6, which define the 3-game:

Preferences of $i \setminus$ towards $j$	player 1	player 2	player 3	player 4
player 1	-	1	1	1
player 2	0.99	-	1	1
player 3	1	1	-	1
player 4	1	1	1	-

Table 6. Modifiers for the E-game.

In this case, player 2 (or, if we switch  $p_{21}$  and  $p_{31}$ , player 3) just a bit dislikes player 1, hence hypothesis H7 is that *in the 3-game, the winning coalition {2, 3, 4} will be significantly more frequent than the winning coalitions involving player 1 in the E-game*, or (H8) *the winning coalitions involving both players 1 and 2 will yield significantly higher gains to player 2 in the 3-game than in the E-game*.

Of course, this setup leaves ample room for more treatment options with the E-game, as well as with other games.

## 4 Experimental design

The experiments were conducted in the autumn semester of 2008 in the computer laboratory of Higher School of Economics (HSE), using the originally developed experimental software. In each of the S-1 and V-2 session there were 12 participants playing in 4 groups of 3 players. Each game lasts 10 rounds, and all players were mixed in roles and across groups in each round. Six experimental sessions of that kind were conducted, 2 games are played in each experimental session in randomized block order as represented in Table 7 (sessions labeled SC and 1C will be explained below). In each of the two E-3 sessions there were 16 participants in 4 groups of 4 players; these games lasted for 20 rounds. In one of these sessions game E was played first, followed by game 3; in another the

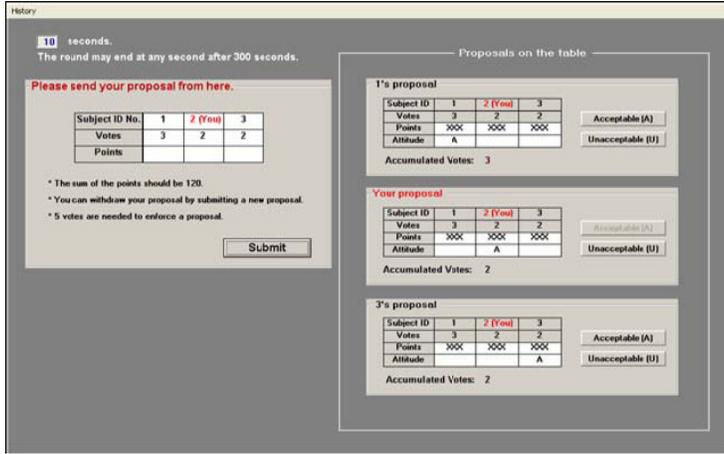


Figure 1: MSZ session screenshot.

order was the opposite. In each round of each game all groups have to agree on how to divide a pie of the size 120 points, with a flat exchange rate 1 point = 0.4 Russian rubles (RuR).

Session 1	Session 2
S	1
1	V
2	S
SC	1C
1C	2
V	SC
E	3
3	E

Table 7. Experimental sessions.

Our experimental software and procedures were explicitly aligned with those of MSZ, including the appearance of the game in the screen (compare Figures 1 and Figure 2), with obvious additions in the case of modifiers (Figure 3).

The game proceeded as follows. Participants entered the room, signed the registration forms and logged into the game. Experimental instructions were read aloud and available on the screen should the participant need it at any time. After all questions are answered, the game began. Each group has to vote for any proposal in no more than 300 seconds<sup>4</sup>; if no agreement was reached, all players in that group receive zero.

<sup>4</sup>This feature of our design is the only substantial difference from the experiment by MSZ. The reason for this difference is purely technical; however, as we shall see shortly, it did not cause any significant divergence of our results.

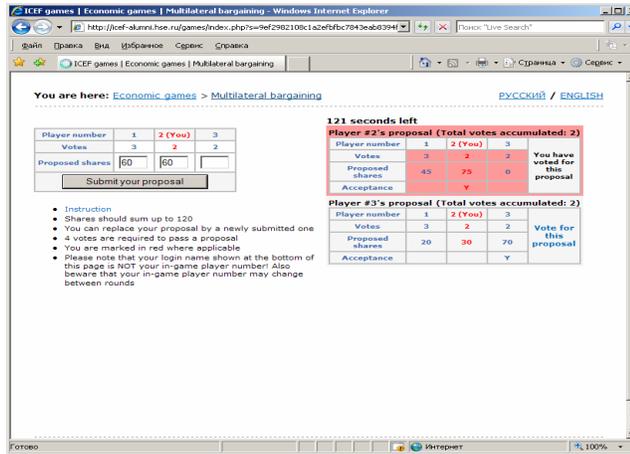


Figure 2: ABP session (without modifiers) screenshot.

In games S-1, we used the  $2 \times 2 \times 2$  design, controlling for order of the game in the session, modifiers and position of players on the screen (C), as will be described shortly. In games V-2, we used the  $2 \times 2$  design, controlling for order in the game and modifiers. In games E-3, there was just one control for the order of the game, which difference did not matter for players 2 to 4, and was only marginally significant for player 1 according to Kolmogorov-Smirnov test in all sessions, which allow us to combine the datasets in one. Participants were 104 students of the various departments of HSE recruited through posters and the announcement on the web, volunteers being requested to register online on <http://games.hse.ru>. Subjects were invited by email to a particular game. 51 of the participants were females, 53 – males, the average age of participants was 19.11 years. The average gain of participants in the three-players games was 340 RuR<sup>5</sup>, minimum – 170 RuR, maximum – 610 RuR per 1- to 1.5-hour session. Corresponding figures for the longer 4-players sessions were 485 RuR<sup>6</sup>, with minimum of 240 RuR and maximum of 750 RuR. The money were paid in cash at the end of each session.

## 5 Results

### 5.1 S-1 games

Figure 4 and Table 8 represents the outcomes of the S-1 sessions.

<sup>5</sup>Roughly equivalent to US\$12, although the rouble was rather heavily depreciating at the time of experiments.

<sup>6</sup>Roughly equivalent to US\$13.5, at the time of experiments

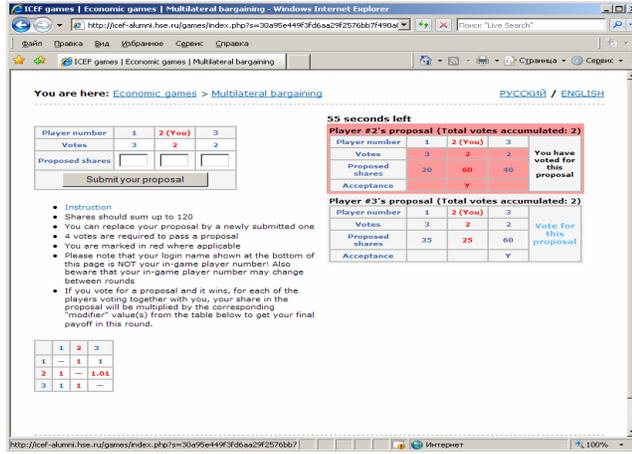


Figure 3: ABP session (with modifiers) screenshot.

All ( $N = 320$ )	mean	s.d.	min	max
player 1	35.36	29.04	0	80
player 2	44.53	24.42	0	100
player 3	40.1	27.56	0	111
Game S				
player 1	37.40	29.44	0	80
player 2	46.25	23.89	0	100
player 3	36.34	28.05	0	110
Game 1				
player 1	33.32	28.57	0	80
player 2	42.81	24.91	0	99
player 3	43.85	26.62	0	111

Table 8. S-1 games, summary statistics.

As Figure 4 reveals, all players receive about 40, in line with the Banzhaf index prediction. There are no treatment effects for players 1 and 2. By contrast, player 3 who on average receives systematically more in the 1 treatment (42.08) than in S treatment (32.25), which difference is significant (Student  $t = 2.24$ ,  $Prob < 0.0264$ ; Kruskal-Wallis  $\chi^2 = 5.89$ ,  $Prob < 0.0122$ ). Hence our modifiers do indeed work for player 3: colloquially speaking, 'being loved is better than love'. This finding is even more striking given that all games in this treatment were played quite quickly: on average, it took about 30 seconds to complete each session. In their comments, many participants were also saying that they did not pay attention to the modifiers at all; yet summary statistics clearly suggests this feeling is wrong: the psychological 'invisible hand' tends to lead their actions in a regular and predicted way, confirming our hypothesis H2; other hypotheses raised in connection to this game await further research.

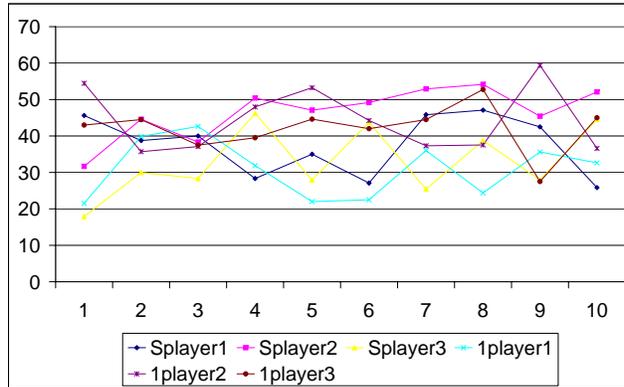


Figure 4: S-1 games, average shares of the players.

One more striking feature follows from Figure 4: player 2 receives systematically more than player 3 in both S and 1 treatments combined (49.89 vs. 37.14), which difference is also significant. This same fact has been observed by MSZ, who attribute it to the ‘framing effect’, without going into further explanations. In our view, this fact has a more explicit and clear explanation: it is the position of player 2 in the middle of the table. In S-1 games, players quickly realize they all have equal effective bargaining power, and find it more profitable to share the pie with just one of the other players than with two (120 divided by 2 is greater than 120 divided by 3).

Given the quick pace for this game, each player seeks one of the neighbours to whom she can offer a coalition. Note that by a neighbour here we mean the player who is closest in the sense of both place and natural order. Thus, player 2 has two neighbours (1 and 3), and hence may be offered a coalition by two players, whereas the other two players have just one neighbour (player 2), so their supply of offers is two times less, as is the expected number of coalitions involving the noncentral players. We call this positioning effect an ‘implicit modifier’ applied to player 2’s payoff; in contrast of ‘explicit modifier’  $p_{ij}$  that we introduce in the experiment in an explicit way. Having said that, we concur that we are mostly interested in the latter ones; implicit modifiers are experimental artifacts that are hardly relevant outside of the experimental lab. To get rid of them, we introduced the following modification of the original setup: instead of player 2, *each* player was shown in the middle of the table, with clockwise displacements of the other two players. That is, if the player is in role 1 for a particular round, it is shown in the middle, player 2 is shown to her right, and player 3 to her left, and so on (see Figure 5 for a screenshot).

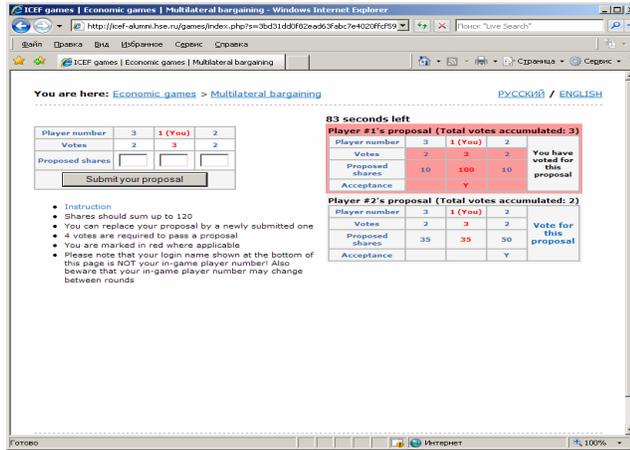


Figure 5: Standard Centered Game screenshot

We refer to this treatment as the Standard Centered (SC in Table 7) and 1-Centered (1C) games; their outcomes are summarised in Figure 6).

In the SC-1C games, the difference in the earnings of players 2 and 3 is mitigated (in particular, player 2's share is 43.03 vs. 39.41 of player 3), and becomes insignificant, although it did not disappear altogether. Explicit modifiers persist in these sessions for player 3 as well, although to a somewhat smaller extent and over the last rounds. We attribute the residual effect of implicit modifiers to another experimental artifact: for the European peoples it is more common to 'read' texts from left to right, and thus the rightmost player is more likely to receive an offer than the leftmost one. The effect of an implicit modifier is most likely to fully disappear in a completely symmetric treatment, i.e., had we used a random clockwise and counterclockwise rotation of the players. However, upon consideration, we did not find this role of implicit modifiers especially interesting outside of the particular experimental treatment.

To summarise, our S-1 experiments confirm our main hypothesis: even a very small modifier of 1.01 for the preferences of player 2 to coalesce with player 3 would cause substantial change of agreements in favour of players 2 and 3 (the weak players) at the expense of player 1 (the strong player) as compared with the standard S-game. Average payoffs of player 3 went up from 36.34 in the S-game to 43.85 in the 1-game, which means that 1% of preferential change brings to that player about 7.5 points, or 3 RuR per round, or about 1US\$ per session. In percentage points, the 'implicit multiplier' of the 1% increase of preferences yields a fair 20% increase in revenues, which we view as quite substantial. Altogether, this means that we observe the treatment effect of even very small modifiers on players' shares of the final payoff, which supports the use of the preference-based power indices rather than the standard Banzhaf.

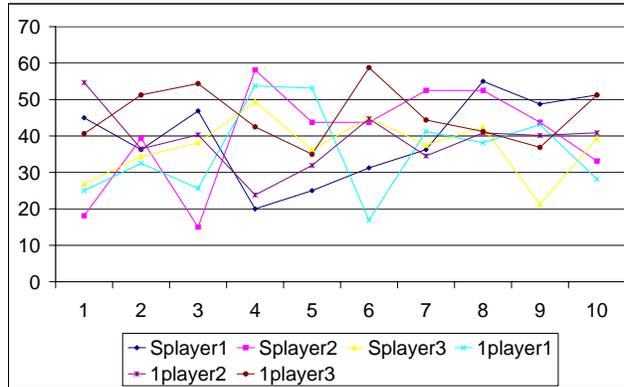


Figure 6: SC-1C games, average shares of the players.

## 5.2 V-2 games

Results of the veto sessions are presented in Figure 7 and Table 9.

All ( $N = 160$ )	mean	s.d.	min	max
player 1	84.29	24.99	0	120
player 2	19.76	20.97	0	70
player 3	13.56	18.42	0	60
Game V				
player 1	81.90	24.76	0	119
player 2	22.56	23.40	0	70
player 3	15.53	19.99	0	60
Game 2				
player 1	86.68	25.14	0	120
player 2	16.96	17.94	0	60
player 3	11.60	16.59	0	60

Table 9. V-2 games, summary statistics.

This figure is largely self-explanatory: the veto player 1 gets a vast majority of the pie, even more than the Banzhaf index predicts, and statistical tests confirm this at any reasonable degree of confidence. This time, although gains of players 2 and 3 fall in 2 treatment wrt the V treatment, none of the differences are statistically significant. This means that small distaste (of 0.99) of players 2 and 3 towards player 1 clearly cannot overturn the fact that this latter player

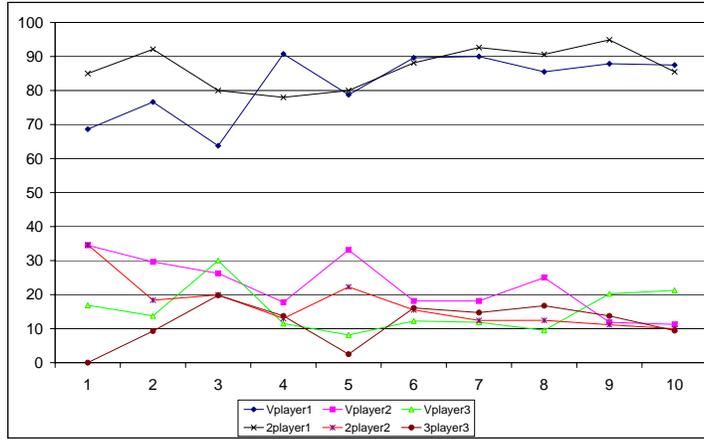


Figure 7: S-1 games, average shares of the players.

is crucial for any positive gain, and rejecting our hypotheses about the role of modifiers in the veto context.<sup>7</sup>

Participants again have quickly realized that: in many games, players 1 were simply asking a bulk of the pie, leaving some bit to one player and nothing to the other – and were waiting for the offered player to accept. Sometimes this strategy failed; yet the typical time to reach an agreement went up significantly, averaging to more than 3 minutes per session! This leads us to test another hypothesis: if the modifiers do not affect the outcome of the game, may they affect the timing of acceptance? On prior grounds, it seems plausible that players 2 and 3 who dislike the veto player 1 ought to be more hesitant to accept her offer (even though this fact goes unnoticed to the players themselves, as it was the case of S-1 games). This hypothesis can be checked by a fit of a duration model on the time of reaching an agreement — however, preliminary estimates of this model did not reveal any significant treatment effect.

### 5.3 E-3 games

Summaries of the E-3 games are presented in Figure 8 and Table 10 below.

<sup>7</sup>It is of course possible that larger modifiers would change the situation, making the difference statistically significant.

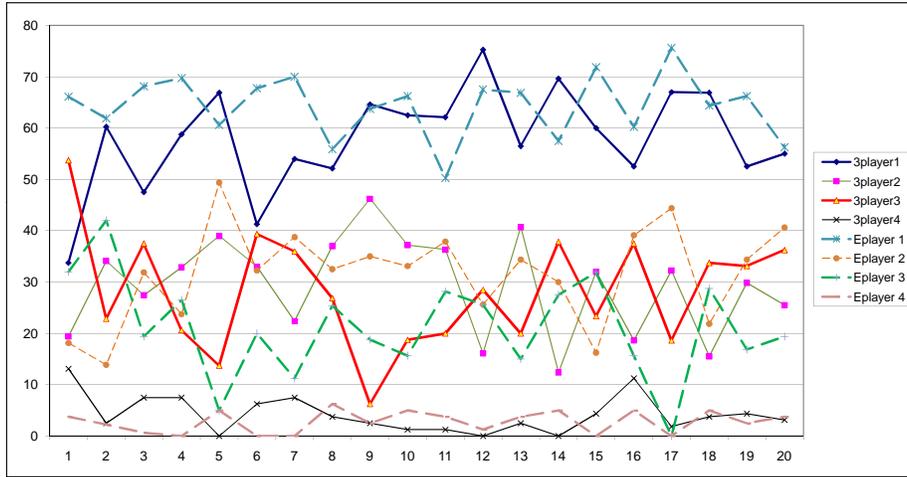


Figure 8: E-3 games, average shares of the players.

All ( $N = 320$ )	mean	s.d.	min	max
player 1	61.15	25.76	0	100
player 2	30.63	23.82	0	70
player 3	24.73	24.54	0	70
player 4	3.49	9.00	0	70
Game E				
player 1	64.34	22.36	0	95
player 2	31.65	23.17	0	70
player 3	21.23	23.72	0	70
player 4	2.76	7.40	0	40
Game 3				
player 1	57.95	28.47	0	100
player 2	29.59	24.48	0	70
player 3	28.23	24.90	0	65
player 4	4.21	10.33	0	70

Table 10. E-3 games, summary statistics.

Player 1 in this game receives systematically more than the Banzhaf index prediction of 50, and clearly does so at the expense of player 4, while gains to players 2 and 3 are basically in line with the index. Payoffs to players 2 and 3 also systematically increase in comparison to the V-2 treatment (cf. Tables 9 and 10). All these facts are in line with MSZ findings.

Two more tendencies peculiar for our design extend the above story. As is obvious from Figure 9, and confirmed confirmed statistically, share of player 1, who has most of the bargaining power, falls from 64.34 to 57.95, which change is

statistically significant (Student t-statistic 2.23,  $p < 0.0262$ ; Kruskal-Wallis Chi-square=3.07,  $p < 0.073$ ), meaning that player 1 loses in the 3 game as a result of modifier applied to player 2 payoffs. Player 3 gains most out of it, with average payoff increasing from 21.23 to 28.23, which effect is even stronger (Student t-statistic -2.57,  $p < 0.0104$ ; Kruskal-Wallis Chi-square=6.2,  $p < 0.0082$ ). Thus, a small negative modifier indirectly benefit player 3, whose gain per session increases by a quarter. One more interesting feature of the E game vs. 3 game: the frequency of coalitions involving three player (2, 3 and 4) are three times higher in the 3 treatment than in the E treatment, which effect is especially pronounced for the first of these sessions (E game followed by the 3 game). This means that players 2, realizing they do not like player 1, tend to switch to a larger coalition, even though building it is clearly more difficult technically, and probably involves lowering one's share of the pie, which now has to be divided among three players instead of two.

## 6 Conclusions

Voting power matters. We see everywhere: in the United Nations (should Iran be sanctioned?), in the boards of directors, in the shareholders' meetings, in the parliaments, even in overnight parties (what do we take – beer or wine?). Yet many important factors affecting the outcome of the vote remain unexplored. Our work builds on a formal treatment which captures important intuition about such factors: small preferences of the parties about whom to enter in coalition with have large implications in terms of the outcomes of the vote.

This paper uses an experimental approach to study the power distribution in a voting body. Results obtained so far may be summarised as follows. Explicit modifiers work in all treatments of the S-1 games, and increase the payoff of player 3 by about 20%, effects for the other players are not significant. This fact can be interpreted in two ways. Verbally, *'being loved is better than love'*, at least in terms of material payoffs. Formally, in terms of the model in Aleskerov (2006), this finding supports the use of  $f^-$  intensity function, at least over  $f^+$ . We have also found (and ruled out) the nuisance effect of implicit modifiers in the S game, and feel confident to attribute it to the peculiar visual representation of the game.

The case of the veto game is different: in this case, explicit modifiers (probably) do not matter at all, and the fact that one player has veto power overrules personal attitudes. This implies that the effect of modifiers (preferences) is of secondary importance to a player's number of votes and the quota. Of course, this effect may be not robust to the size of preferences, but for the small-scale preferences there is no effect.

Other effects await further data analysis and further games — e.g. the E-3 treatment is still under exploration.

## 7 Acknowledgments

The authors wish to thank Alexei Charkov for programming the experimental software, Irina Karpenko for excellent research assistance, Alexander Poddiakov for useful feedback, and the Scientific Fund of Higher School of Economics for supporting our research. The usual disclaimer applies.

## 8 References

- Aleskerov, F. (2006). Power indices taking into account agents' preferences. In: B. Simeone & F. Pukelsheim (eds), *Mathematics and Democracy*, Berlin: Springer, pp. 1-18.
- Aleskerov, F., Blagoveschenskiy, N., Satarov, G., Sokolova, A. and Yakuba, V. (2007), *Power distribution and stability in Russian Parliament (1905-1917, 1993-2004)*, Moscow: Fizmatlit (in Russian).
- Ariely D. (2008). *Predictably irrational: the power of illusion*. Harper Collins, 2008.
- Banzhaf, J. (1965). Weighted voting doesn't work: A Mathematical Analysis. *Rutgers Law Review* 19: 317–343.
- Barry, B. (1980). Is it better to be powerful or lucky? Part I and Part II. *Political Studies*, 28: 183–194, 338–352.
- Brams, S.J., Affuso, P.J. (1976). Power and size: a new paradox. *Theory and Decision*, 7:29–56.
- Coleman, J.S. (1971). Control of collectivities and the power of a collectivity to act. In: Lieberman B (ed) *Social choice*. Gordon and Breach, London.
- Dahl R. (1969) The concept of power. In: R.Bell, D.V.Edwards, R.H.Wagner, Eds. *Political power: a reader in theory and research*. N.Y.: Free press.
- Felsenthal D.S., Machover M. (1998). *The measurement of voting power: Theory and practice, problems and paradoxes*. Edward Elgar Publishers, London.
- Lukes S. (1974) *Power: a radical view*. London: Macmillan.
- Montero, M., Sefton, M., and Zhang, P. (2008) Enlargement and the balance of power: an experimental study. *Social Choice and Welfare*, 30:69–87.
- Parsons, T. (1986) Power and the social system. In: S.Lukes (ed.). *Power*. Oxford: Blackwell.
- Schmeidler, D. (1969). The nucleolus of a characteristic function game. *SIAM Journal of Applied Mathematics* 17:1163–1170.
- Shapley L.S., Shubik M. (1954) A method for evaluating the distribution of power in a committee system. *American Political Science Review*, 48: 787–792.
- Warber R., Shiv B., Carmon Z., Ariely D. (2008) Commercial features of placebo and therapeutic efficacy. *Journal of the American Medical Association*, v.299, March 2008, p.1016-1017.