



FACULTY OF BUSINESS AND LAW

School of  
**ACCOUNTING, ECONOMICS AND FINANCE**

**School Working Papers – Series 2004**

**SWP 2004/22**

**Standard Voting Power Indices Work:  
An Experimental Investigation of Pure Voting Power**

Chris GELLER<sup>1</sup>,

Jamie MUSTARD  
and  
Ranya SHAHWAN

Faculty of Business and Law  
Deakin University  
Geelong VIC 3217 Australia

The working papers are a series of manuscripts in their draft form. Please do not quote without obtaining the author's consent as these works are in their draft form. The views expressed in this paper are those of the author and not necessarily endorsed by the School.

**URL: [http://www.deakin.edu.au/fac\\_buslaw/sch\\_aef/publications/default.htm](http://www.deakin.edu.au/fac_buslaw/sch_aef/publications/default.htm)**

---

<sup>1</sup> Please address comments to Chris Geller, [cgeller@deakin.edu.au](mailto:cgeller@deakin.edu.au). Title used by kind permission of Jonathan Katz.

**Standard Voting Power Indices Work:  
An Experimental Investigation of Pure Voting Power**

Chris GELLER<sup>2</sup>,

Jamie MUSTARD  
and  
Ranya SHAHWAN

Faculty of Business and Law  
Deakin University  
Geelong VIC 3217 Australia

18 June 2004

Earlier versions presented to:  
PISTA'04  
Orlando Florida  
21-25 July 2004

and submitted to  
The 33<sup>rd</sup> Conference of Economists 2004  
The University of Sydney  
27-30 September 2004

---

<sup>2</sup> Please address comments to Chris Geller, [cgeller@deakin.edu.au](mailto:cgeller@deakin.edu.au). Title used by kind permission of Jonathan Katz.

## ABSTRACT

We evaluate the accuracy of power indices by experimentally measuring the political power embodied in blocks of votes *per se*. The experiment incorporates several subjects interacting in online chat rooms under supervision. Chat rooms and processes for selecting subjects reduce or eliminate extraneous political forces leaving logrolling as the primary political force. Results show that two standard power indices reflect voting power while other power indices and proportionality do not.

**Keywords:** Voting, Power Index, P Power, Shapley-Shubik, Banzhaf, Johnston, Burgin, Experiment.

## 1. INTRODUCTION

For over half a century scholars, practitioners and courts have struggled to understand the relationship between sizes of voting blocks and influence over democratic decisions (see Felsenthal and Machover 1998 for an overview). This relationship has come to be formalized as ‘power indices’. Power indices are abstract mathematical representations of the ability of voters to affect group decisions focusing on votes *per se*, in isolation from other political considerations such as resources, traditional alliances, charismatic leadership, or shared agendas. The effort to refine power indices continues as theoretical efforts have not satisfactorily matched empirical results (Felsenthal and Machover 1998, 1995; Gelman, Katz and Bafumi, forthcoming). Indeed, a founder of the theory remains active developing new theoretical approaches (Burgin and Shapley 2001). Previous literature has been unable to resolve whether the unsatisfactory practical performance of indices is due to the complexity of real world political environments, or due to defects in the mathematical logic of the indices themselves (Felsenthal and Machover 1995). This paper presents the first empirical evidence that the mathematical logic of some power indices substantially captures voting power as manifested in a human institution.

Voting power is not a proportionate or simple translation of the size of voting blocks. Consider the following examples. Suppose there are four players with 55, 40, 3, and 2 votes each and a majority of at least 51 votes decides the outcome. We call them ‘players’ for simplicity since the real world versions can be political parties, shareholders, countries and so forth. In compact notation we write the vote information as {51; 52, 38, 6, 4}. The first player can determine all decisions and so has 100% of the electoral power while the others have none. With the profile {51; 45, 45, 6, 4} any two of the first three players can form a majority of votes, and the fourth cannot help any other players form a majority. Thus, the power in this case divides between the four players as follows: 33 1/3%, 33 1/3%, 33 1/3% and 0%. The first three players have equal power even though they are different in their votes, while the latter two with similar numbers of votes are very different in power. The fourth party is called a ‘dummy player’ because it cannot turn a losing coalition into a winning coalition. The power indices<sup>3</sup> – Shapley-Shubik (1954 and Shapley 1953), Banzhaf (1965) and Penrose (1946), Johnston (1977, 1978),

---

<sup>3</sup> Defined in Section 2.

Burgin-Shapley (2001), Deegan-Packel (1978), Intervals (Taylor and Zwicker 1997), Holler-Packel (1983), etc – agree with these interpretations. When the number of players increases, results become more complex and different indices can yield different outcomes, even opposite rankings of power (Saari and Sierberg 1999).

The aim of this paper is to measure experimentally the level of power held by blocks of voters. Specifically, we:

- Establish a practical, ethically responsible experimental procedure.
- Create an artificial environment in which power indices should work, that is an environment isolated from traditional alliances, adversaries, charismatic leaders, etc.
- Empirically measure the voting power of the various players in several weighted vectors of votes.
- Establish which, if any, power indices accurately reflect voting power in an environment isolated from other political considerations.

The structure of the paper is as follows. Section 2 defines and explains the power indices. Section 3 presents the preliminary design of the experiment. Section 3 describes controls for variations among individual participants. Section 4 details the final form of the experiment. Section 5 presents and analyses the results. Section 6 concludes.

## 2. POWER INDICES

The central idea of power indices is that votes drive decision making in an environment effectively isolated from other political forces (Felsenthal and Machover 1998). In the world of power indices, players face a series of independent decisions, the players share no history, they have no overlap in their interests, they differ in their votes, but the players are all identical in other respects.

Power indices address two aspects of voter power called I and P power by Felsenthal and Machover (1998). The experiments we present address P power. P power is the capacity to capture a portion of a fixed sum of political rewards or purse, ‘measured by that [player’s] expected or estimated share in the fixed purse’ (pg 36). Examples of P power include the ability to control a portion of ministerial seats within a coalition government, receiving for one’s constituency ‘pork barrel’ government funding, and in our specific experiments, the power to capture money from a fixed pool.

For clarification by contrast, I power addresses the capacity to influence the outcome of a binary decision (Felsenthal and Machover 1998, 36), such as selecting one of two candidates for CEO or passage of a bill legalizing cannabis. Many real world applications combine P and I power, e.g. a university board creating a strategic plan mixing various levels of funding for many alternative priorities.

The Shapley-Shubik Index is the oldest P power index. Shapley (1953) developed an abstract measure of the value of playing a game based upon three assumptions<sup>4</sup>. 1) The game is abstract, that means that it is the number of votes controlled by a player that matters, not the player's personality, name, or other characteristic. 2) The game is efficient in that all possible gains are captured. 3) If two independent games are merged, then the value for each player in the merged game equals that sum of that player's value in the two games played separately. That is, value of two lottery tickets from different games, when purchased as a package, equals the sum of the values of the two tickets separately. Clearly, this assumption does not fit all games, for example two tickets from the same lottery when the purchase of the second ticket marginally reduces the probability of the first ticket winning.

"It is remarkable that no further conditions are required to determine the value uniquely" (Shapley 1953, 309). However, these are sufficient conditions, not necessary ones. As Shapley 1953 notes, other sets of sufficient conditions exist. By formula, the Shapley value  $\phi$  for a player  $i$  is:

$$\phi_i[v] = \sum_{S \subseteq N, i \in S} \left( \frac{(s-1)!(n-s)!}{n!} \right) v(S) - \sum_{S \subseteq N, i \notin S} \left( \frac{s!(n-s-1)!}{n!} \right) v(S), \quad 1$$

where  $v$  is a game expressed as a set function,  $S$  is a set of players,  $N$  is a carrier of  $S$  i.e.  $S$  together with any dummy players,  $s$  and  $n$  are the sizes of  $S$  and  $N$ .

This formula equates to an algorithm. Start all players with a value of zero. List all permutations of players. For each permutation, sum the players' votes starting from the left and note which player's votes change the sum from less than the majority to at least the majority. Each time a player changes a subset of players from losing to winning, increase that player's value by one. This algorithm generalizes to games with payoff values of other than all or nothing (1 or 0 above). In general, whenever a player's votes change the payoff for a subset of players within a permutation, add the change of payoff to the value for that player.

Shapley and Shubik (1954) apply the Shapley Value to politics. They normalize the Shapley value so that the sum of all players' values equals one, by dividing each player's value by the sum of all players' values. This normalized form of the Shapley value is the Shapley-Shubik power index.

Banzhaf (1963, pg 331 note) rejects the Shapley-Shubik index on the intuitive ground that using permutations implies that order of matters in the index, but is not essential in real political situations. He rejects the use of permutations and substitutes combinations in their place, thus replicating the method of Penrose (1946) and anticipating Coleman (1971) according to Felsenthal and Machover (1998). The Banzhaf index is defined by an algorithm. List all combinations of players. Identify the subset of combinations in which the total number of shares equals or surpasses the majority required in their corporation. Within this subset, count each time a player was pivotal, that is the player could, by leaving the combination, reduce the total number of

---

<sup>4</sup> The number of assumptions is somewhat contentious as Shapley specifies superadditivity in defining his value but does not use it in deriving the value, nor is the Shapley value superadditive generally. See Felsenthal and Machover 1998, 225 for example.

shares to less than the required majority. Each player's Banzhaf value, normalized as an index, is their count as a percentage of the total count for all players (Banzhaf 1965). The Banzhaf power index addresses I Power (Felsenthal and Machover 1998)

Johnston (1977, 1978), again on intuitive grounds, objected to ascribing full power to whichever players were pivotal. Instead, Johnston awards each player value in proportion to the reciprocal of the number of pivotal players in the coalition. Further, Johnston considers a player pivotal if it can turn a winning coalition into a losing one by withdrawing from the coalition. Instead Johnston's method is to list all combinations of players and sum, for each player and each coalition, the reciprocal of the number of pivotal players in the coalition. These sums for each player are then normalized to create an index. Felsenthal and Machover (1998) consider Johnston's intent to have been measuring I power.

The Burgin-Shapley (2001) index is like the Johnston index as being an intentional modification of Banzhaf's index, and also in its algorithm. For Johnston, a player is pivotal if it can break a winning coalition, and for Burgin-Shapley a player is pivotal if it can turn a losing coalition into a winner. The Burgin-Shapley index is apparently intended to measure I power.

### 3. EXPERIMENT DESIGN

Power indices estimate the power embodied in blocks of votes that allows voters to negotiate using those votes, with no shared interests between voters and no other differences between voters. Natural environments such as international organizations, legislatures, or board rooms may be unsuited for direct evaluations of power indices because voters in those environments differ in many dimensions and often share interests. The experiment we describe below approximates the conditions that underlie power indices.

These experiments consist of several subjects<sup>5</sup> taking the roles of players and meeting in an online chat room with two supervisors who assign some number of votes to each player. The players then divide a lump sum of money between themselves by simple majority-vote rule. This assignment of votes and distribution of money was repeated twelve or twenty-four times in each session of experiments. Each assignment of votes and distribution of money is called a round.

Consider an example. One player may receive eight votes while the others receive three each, with a majority of 12 required to determine the division of the money; {12;8,3,3,3,3,3}. For convenience identify the players with letters consecutively from A (with 8 votes). For example consider a division of \$10. Player A may propose to divide the money evenly with B and C. Players D, E and F could reply with a proposal to allow B to have four dollars and accept two dollars each. If each player has strictly less than half the votes, no matter how the players divide the money, there is always a majority that can benefit from a different division of the money.

Distribution of actual money was imperative in the experiments in order to maintain subjects' interest and motivate active participation. See Smith (1982) for more on the role of sufficient

---

<sup>5</sup> The term 'subject' applies to persons involved in our experiments while the term 'player' refers to an abstract role in a game.

returns for effort and other practical elements of social-scientific laboratory experiments. The Deakin University Human Research Ethics Committee, citing broader Australian standards, questioned the appropriateness of financial compensation for participants in experiments. Their concern centred on the possibility of undue coercion. Other standards, for example the American Anthropological Association (1986), state presumption in favour of compensation. We resolved the conflict by setting the amount of money to be divided at a level that would result in subjects receiving on average an hourly income approximating that of an undergraduate research assistant as the subjects were all undergraduates at our university. Pre-study debriefings and observations of play showed \$15 per round to motivate subjects and result in an appropriate hourly compensation.

Chat rooms permit political negotiation based upon logrolling (making deals based upon ones' votes or 'I agree to vote for something you want in return for you agreeing to vote for something I want Tullock 1976 in Johnston 1977) in an environment with greatly reduced effects from factors other than votes. Side deals and threats are obvious and avoidable in supervised chat rooms. Personality and charisma have much less potential for influence when deals must be made using brief formalised statements. Subjects may be anonymous in chat rooms. In these experiments, player's identities in the chat rooms consist of a number shared by all players in a particular game and a letter unique to each player (see the screen examples in Appendix 2).

During a pre-study and after several experiment sessions we debriefed subjects, proctors and recorders. They reported that the experiment procedure was easy to understand and apply, and that the chat room package was familiar. They reported that subjects were effectively anonymous through at least several rounds of play, which was also our experience when we supervised games. During the pre-study, one set of participants reported that they identified one of the participants during the game. This identification arose from that subject making an online comment in a characteristic, personal manner while idle between games. We instituted a policy of no chatting between games or rounds, discouraging side comments, and increasing the pace of the games. Proctors and recorders reported consistently during debriefings that they 'seldom knew' the identity of the players and were 'usually wrong' when they thought they knew. We conclude that the subjects were effectively anonymous. For further protection of anonymity, we changed the identity of the players every six rounds.

#### **4. PARTICIPANT CHARACTERISTICS**

We collected a variety of control data on the subjects in order to control for possible effects: gender, nationality, psychosocial and risk orientation (Table 1). Psychosocial orientation measures individuals' preferences for receiving payments in comparison with payments to others (Van Lange, Otten, De Bruin, and Joireman 1997). This Van Lange *et al* instrument divides subjects into three broad orientations<sup>6</sup>. Individualist subjects prefer to receive a higher payment for themselves without regard to payments received by others. Competitive subjects prefer to receive as much more than others as possible, even to the extent of accepting a lower payment for themselves to gain even lower payments to others. Pro-social subjects prefer the highest total for payments to themselves plus those to others; at least they will accept a somewhat decreased payment to themselves in order to gain more for others.

---

<sup>6</sup> Thanks to Dr. Janine Webb, School of Psychology, Deakin University for this advice and other guidance.

Based upon pre-study results we concluded that subjects with a competitive psychosocial orientation posed a threat to the experiment as they were so interested in reducing others' income that they would accept lower incomes for themselves in order to do so. Such behaviour is inconsistent with the self-serving behaviour assumed by power indices. Therefore, we excluded competitive subjects from the study. We also excluded those who gave inconsistent responses to the psychosocial instrument. Thus our subjects consisted of individualist and pro-social subjects. The potential magnanimity of pro-social subjects was potentially problematic for the study. We assured the subjects that all had an equal potential to earn money based upon their votes.

We also limited the subjects to those with apparent proficiency in English as pre-study results showed foreign nationality to affect earnings and debriefings suggested language ability to be the key issue in low foreigner performance.

We attempted to eliminate the effect of social orientation by structuring the experiments such that all subjects would receive approximately equal voting power over the course of the game. Thus, structural equity could replace the perceived need for subjects to pursue equity. Our subject instructions (Appendix 2) assured all subjects that the games were fair and that anyone temporarily in a weak position would be in a strong position later. The instructions also explicitly encouraged subjects to be self-centred or 'greedy'. This mechanism and encouragement was apparently effective as subjects with a pro-social psychosocial orientation received at least as much as did competitive subjects (Table 1).

We used a simple test (Appendix 1) to test for attitude toward risk: aversion, love, or neutrality. We found that gender, risk aversion, psychosocial orientation and national origin did not affect earnings significantly in either the statistical or practical sense. Coefficients on those variables were both small and insignificant (Table 1). For more detail of on the control of potential variation see (Geller and Mustard 2004)

Experience should matter in performance. In the pre-study, we tracked experience in order to document the length of the learning process, if experience continued to matter over time or if gains from experience were captured within a few trials of the game. Consistent with previous works (for example Kelly and Arrowood 1960; Komorita and Moore 1976), six rounds of play appeared to impart enough experience for proficient play. Inexperienced subjects sometimes focus upon other player's total votes as a proxy for their value in a game. However after only a few rounds, subjects focus upon sets of players that can form winning coalitions. They see beyond the veil of numbers of votes to the potential for alliance. Inexperienced subjects may consider the player with eight votes to be more powerful in  $\{12;8,3,3,3,3,3\}$ <sup>7</sup> than in  $\{22;8,7,7,7,7,7\}$ . After even a half a dozen rounds playing for money, subjects treat them as equivalent – the largest and any two players can win as can any four smaller players. In our analysis we limited our observations to those in which all subjects had already participated in a practice round and at least six rounds for money.

---

<sup>7</sup> Recall that this notation means that 12 votes are sufficient to win, one player has eight votes and five have three votes each.



## 5. EXPERIMENT

We distributed the subjects widely in a classroom style computer laboratory. We maintained at least one computer between every two subjects and seated subjects in the same experimental group more distantly. Each computer used by a subject had chat rooms for two player identities, permitting rapid change of identities between series of rounds. The proctor and recorder for each group of six players shared a computer, participating in the chats as a single individual. We copied files to each proctor's computer giving the listings of votes to be used each round and typical messages used during the rounds. We distributed to each recorder a hardcopy sheet giving the votes for each player and majorities required for each round as backup and verification against electronic records.

When the subjects arrived we sorted them into relatively homogeneous groups based on social orientation, risk orientation, gender, and nationality. We then assigned the subjects to computers without them knowing each other's player identities. When new subjects were participating, we provided instructions on the game and played a practice round without money. The instructions (Appendix 2) included procedures, rules of the game, suggestions on strategies, and that fifteen dollars would be divided each round. The subjects were students in the university and were familiar with the use of the chat rooms because the platform was used for educational purposes or student communication throughout the university. After we provided instructions, each group of six players ran independently.

Subjects within each group had the same information, communicating entirely through chat room windows shared by all group members. At the beginning of each round, the proctor submitted a message to one window, labelled 'Group Chat', on the subjects' monitors saying to wait and do nothing until further notice (Appendix 2 has screen examples). Second, the proctor sent a message to another window, labelled 'Vote Vector', on each subject's monitor giving the votes for each player and majority required for that round. All subjects in the group saw the same message and each knew the votes of all players in their group. This was the only message each round sent to the Vote Vector window. Third, the proctor sent a message to the Group Chat requesting that the subjects confirm their votes. Each subject responded with the number of votes assigned to them that round. The proctor and recorder confirmed each number of votes with the data file and hardcopy sheet, correcting any errors.

The proctors submitted a message to each Group Chat to begin the games. Subjects submitted messages proposing, rejecting, revising, or accepting various divisions of the money. Subjects could write plain English statements, use conventional chat room abbreviations, or use brief notation provided during the instructions for the game. Any subject proposing a division of money had to identify that proposal uniquely, using their player identification letter followed by a number. They were not permitted to use threats, deals for anything other than divisions of money that round, or statements that would reveal personal information. We authorised proctors to end a round without any payment to enforce the rules, but they never had to exercise such a punishment. Recorders and proctors watched the messages for the emergence of a consensus, a difficult task requiring two people. When subjects appeared to have reached a majority decision, the proctors would wait briefly and submit a call for votes. Each subject could then submit a

message stating which proposal he or she supported. The proctor and recorder then counted the votes. Subjects could change their acceptance of a proposal if they wished at any time until the proctor ended the round. If there was no winner, the proctor sent a message saying to continue negotiating. When a proposal received enough votes, the proctor sent a message saying that the round was over, which proposal won, and instructing the subjects to send no further messages until the next round started.

When each round ended the recorder wrote how much each player received on the hardcopy sheet and the proctor confirmed the record with the messages in the Group Chat. After each six rounds, the proctors submitted messages instructing subjects to close their Group Chat and Vote Vector and open the alternative version of each to proceed for six more rounds with a new identity.

At the end of 12 or 24 rounds, we ended the game and we tallied each subject's winnings. We paid them precisely to five cents, the smallest denomination coin in Australia. We collected a receipt which included each subject's player identities during the game. While they waited for payment we listened to their conversation and sometimes inquired informally about the game. The subjects often asked each other who they had been during the game, receiving the reply as new information far more often than as confirmation of suspicions. Subjects, some of whom we knew well from classes and experiments, did note some problems. For example, we had to exclude a set of 24 rounds from our data because one subject stopped participating. After the final experiments when we could talk to subjects who would not play again, we asked seriously about identifying other subjects, manipulation of the games, and opinions about research objectives. No one reported confidence in identifying their fellow subjects. Proctors and recorders confirmed that they did not know subjects' identities with any reliability, and we were not able to identify subjects when we served as proctors. The only reported case of manipulation was that one a subject accepted a payment of one cent to see how we would pay him. No one expressed any understanding of our research objectives beyond the level presented in the instructions. (See appendix 2.)

## 6. RESULTS AND ANALYSIS

We ran 441 rounds of experiments for money, which is excluding the initial practice round whenever a new subject participated. As mentioned, four sessions of six payment rounds were excluded because one subject did not participate actively causing an ambiguity of whether the games were between five or six players. Of the remaining 417 rounds, there were 351 in which all subjects had participated in at least six previous payment rounds. In one round we accidentally used a quota of less than half the votes, leaving 350 valid observations. Fifty-nine of these form a pre-study for another investigation and do not relate to our current analysis. It was during one of these pre-study rounds that a player received a payment of one cent. Thus we address 291 observations in this paper, of which 288 have all control data.

Thirty-one subjects participated in the experiment rounds used in this analysis. Each was the largest player in terms of votes for between one and 26 valid observation rounds. We are missing demographic control data for one subject who was the largest player for three valid

observation rounds. Twenty subjects were male, 4 were foreign, 16 were individualist, 14 were pro-social, 4 risk averse, 4 risk loving, 15 risk neutral, and 7 gave inconsistent answers in the risk instrument.

We used 19 vote profiles that can be considered from several perspectives (see Table 2 and Figure 1). We sampled most profiles between 11 and 20 times. We sampled one profile 35 times in order to gain a ‘large’ sample size for one specific profile. Through copying errors we created two profiles accidentally and used one of them once and the other four times.

The profiles fall into 7 power-identical sets. Within each of these 7 sets, each player by rank has the same power as the same ranked player in each of the other profiles. That is, the largest player in each profile in each set has the same power index value for each given power index. Likewise, all the smallest players have the same power. Three of these sets are sampled in large enough numbers to allow for effective statistical analysis (see profiles  $p_i$ ,  $r_i$  and  $s_i$  in Table 2). A fourth may be aggregated with a very similar profile to create a sample size of 48 (see profiles  $t_i$  and  $u_1$  in Table 2).

Table 3 shows that the individual characteristics of gender, nationality, risk orientation and risk aversion had no significant effect on our experimental results. Not only were the coefficients statistically insignificant they were of low magnitude. Given their coefficients and means, taking any personal control variable to an extreme value of zero would have affected percentage earnings by less than one percent.

Tables 2 and 4 with figure 1 present our main results. The sample means for each profile are equal to or higher than the Shapley-Shubik and Banzhaf index values and well below the Johnston and Burgin-Shapley values. For the set of profiles with the largest sample size ( $p_i$ ), the sample mean (0.337) is arguably coincident with the value (0.333) for the former two indices<sup>8</sup> and widely divergent from the value (0.533) for the latter two indices. For the other three cases, the Shapley-Shubik and Banzhaf index values fall near the lower bound of 95% confidence interval<sup>9</sup> around the profile sample means. In all cases, the remaining two power indices fall far above the confidence interval.

These results lead to four conclusions.

1. Power indices can predict human voting outcomes in appropriate environments.
2. The Shapley-Shubik and Banzhaf indices are reasonably good predictors (at least).
3. These two power indices do not overstate the voting power of large voting blocks, and
4. the Johnston and Burgin-Shapley indices do overstate the voting power of large voting blocks.

---

<sup>8</sup> Profiles  $p_i$  also approximate the power distribution in 40 percent of top 400 UK publicly held corporations (Leach 2002).

<sup>9</sup> Note that data was effectively bimodal with values varying depending upon whether the largest player was in the winning coalition or not. As such, we used calculations similar to those used standard statistical test for stratified sampling to obtain the confidence interval estimates. The mean and standard error for the largest player’s earnings when not part of the winning coalition were of course zero.

Although the Shapley-Shubik index for P power coincides with the Banzhaf index for I power for the profiles we used (and rarely differed substantially for any profile we checked), conceptually P power and I power are different. Our experiment structurally addressed P power, and so the results apply to P power and conform fairly closely to the primary model of P power, the Shapley-Shubik index.

The significance of our results is that they provide empirical support for the first time that the Shapley-Shubik provides a solid foundation for the investigation of P type rivalrous voting power. Our empirical results suggest that previous studies showing weak applicability of power indices are driven by the environments of those studies differing substantially from the axiomatic foundation of power indices. Future research may focus upon the generalization of power indices to less idealized environments.

## 7. CONCLUSION

The experiments work well procedurally as suggested by the match of experimental results with power indices and the insignificance of control variables.

The Shapley-Shubik and Banzhaf indices are empirically indistinguishable over most vote profiles, while they are easily distinguishable from Johnston and Burgin-Shapley indices. Our results support the Shapley-Shubik or Banzhaf indices or both as approximations of voting power in an environment with homogenous players with orthogonal interests. However, they may understate the power of large players in that environment. This result contrasts with other studies that suggest these indices overstate the power of large players in naturally occurring environments of heterogeneous players with shared interests. The next question is how shared interest and heterogeneity of players affects the application of power indices to non-experimental 'natural' environments with shared interests and heterogeneous players.

## REFERENCES

- American Anthropological Association 1986 Statements on Ethics: Principles of Professional Responsibility. Adopted by the Council of the American Anthropological Association May 1971 (As amended through November 1986) <http://www.aaanet.org/stmnts/ethstmnt.htm>. Accessed 10 July 2003.
- Banzhaf JFIII 1965 'Weighted Voting Doesn't Work: A Mathematical Analysis', *Rutgers Law Review* 19, 317-343.
- Burgin, M. and L.S. Shapley 2001 *Enhanced Banzhaf Power Index and Its Mathematical Properties*, UCLA Economics Working Papers, No 797, presented at the International Conference on Game Theory, Stony Brook, July 2000.
- Coleman JS (1971) 'Control of collectivities and the power of a collectivity to act' in *Social Choice*, ed B Lieberman Gordon and Breach: New York.
- Deegan J and Packel E 1978 'A new index for simple n-person games', *International Journal of Game Theory* 7, 113-123.
- Felsenthal D and Machover M 1995 'Postulates and paradoxes of relative voting power – a critical reappraisal' *Theory and Decision* 23, 195-229.
- Felsenthal D and Machover M 1998 *The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes*, Northampton, Mass: Edward Elgar.
- Geller, C and Mustard J 2004 'An experiment using ICT to investigate pure voting power', presented to Politics and Information Systems, Technologies and Applications, PISTA '04 conference, Orlando, Florida, 21-25 July.
- Gelman A, Katz J and Bafumi J Forthcoming 'Standard Voting Power Indexes Don't Work: An Empirical Analysis' *British Journal of Political Science*. (Available as Caltech Social Science Working Paper No. 1133.)
- Holler MJ and EW Packel 1983, 'Power, Luck, and the Right Index' *Journal of Economics (Zeitschrift für Nationalökonomie)*, 43, 21-29.
- Johnston RJ 1977 'National sovereignty and national power in European institutions', *Environment and Planning* 9, 569-577.
- Johnston RJ 1978 'On the measurement of power: Some reactions to Laver', *Environment and Planning* 10, 907-14.
- Kelly HH and Arrowood AJ 1960 'Coalitions in the triad: critique and experiment' *Sociometry*, 23, 231-244.

Komorita SS and Moore D 1976 'Theories and processes of coalition formation' *Journal of Personality and Social Psychology* 33, 371-381.

Leech D 2002 "An Empirical Comparison of the Performance of Classical Power Indices" *Political Studies*: 50, 1-22.

Penrose LS 1946 'The elementary statistics of majority voting', *Journal of the Royal Statistical Society* 109: 53-57.

Saari DG and Sierberg KK 1999 *Some Surprising Properties of Power Indices*, NU Center of Math Econ Discussion Paper # 1271, September 16, 1999.

Shapley LS (1953) 'A value for n-Person Games' in *Contributions to the Theory of Games*, eds HW Kuhn and AW Tucker, Princeton University Press, New Jersey Vol II: 307 – 317.

Shapley LS and Shubik M (1954) 'A method of evaluating the distribution of power in a committee system' *American Political Science Review*, 48, 787-92.

Smith V, 1982 'Microeconomic systems as an experimental science', *American Economic Review*, December 72, 923 – 955.

Taylor A and Zwicker W, 1997 'Interval measures of power', *Mathematical Social Sciences*, 33(1), 23-74.

Tullock G, 1976 *The Vote Motive : an essay in the economics of politics, with applications to the British economy*, Institute of Economic Affairs: London.

US Courts, *Baker v Carr*, 369 U.S. 186 (1962).

Van Lange PAM., Otten W, De Bruin EMN and Joireman JA 1997 'Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence', *Journal of Personality and Social Psychology* 73: 4, 733-746.

**TABLE 1: DATA DESCRIPTION**

Characteristics of the player with the most votes per round

	Number	Mean	St Dev	Min	Median	Max
% Earnings	291	0.354	0.228	0	0.400	0.733
SS, Shapley-Shubik	291	0.328	0.030	0.267	0.333	0.400
B, Banzhaf	291	0.327	0.028	0.267	0.333	0.393
J, Johnston	291	0.503	0.054	0.381	0.533	0.574
BS, Burgin-Shapley	291	0.508	0.051	0.389	0.533	0.574
Male	291	0.725	0.447	0	1	1
Foreign	291	0.076	0.265	0	0	1
PersI, Individualist	291	0.505	0.501	0	1	1
RiskA, Risk Averse	288	0.167	0.373	0	0	1
RiskL, Risk Loving	288	0.128	0.335	0	0	1
RiskN, Risk Neutral	288	0.569	0.496	0	1	1

	% Earnings	SS	B	J	BS	Gender	Foreign	PersI	Risk Defined	RiskA	RiskN	RiskL
% Earnings	1.000											
SS	0.166	1.000										
B	0.161	0.999	1.000									
J	0.104	0.921	0.936	1.000								
BS	0.122	0.932	0.947	0.994	1.000							
Male	-0.016	0.021	0.018	-0.005	-0.004	1.000						
Foreign	0.007	-0.061	-0.061	-0.055	-0.047	-0.464	1.000					
PersI	-0.058	-0.037	-0.037	-0.029	-0.036	0.130	0.205	1.000				
Risk Defined	-0.001	-0.028	-0.025	-0.012	-0.012	0.004	-0.001	0.071	1.000			
RiskA	-0.097	-0.122	-0.125	-0.130	-0.128	-0.305	0.503	0.447	0.177	1.000		
RiskN	0.085	0.075	0.079	0.088	0.081	0.071	-0.304	-0.042	0.455	-0.514	1.000	
RiskL	-0.019	-0.004	-0.004	0.001	0.011	0.238	-0.110	-0.363	0.152	-0.172	-0.442	1

**TABLE 2: VOTE PROFILES**

Profile	S.S.	Banz.	B.S.	John.	<i>n</i>	$\bar{w}$	$\overline{\bar{w}}$
$p_1 = \{31:11,10,10,10,10,10\}$	0.333	0.333	0.533	0.533	18	0.381	
$p_2 = \{36:14,12,12,11,11,11\}$	0.333	0.333	0.533	0.533	20	0.423	
$p_3 = \{22:8,7,7,7,7,7\}$	0.333	0.333	0.533	0.533	17	0.351	
$p_4 = \{16:11,4,4,4,4,4\}$	0.333	0.333	0.533	0.533	35	0.298	0.337
$p_5 = \{20:14,5,5,5,5,5\}$	0.333	0.333	0.533	0.533	20	0.296	
$p_6 = \{12:8,3,3,3,3,3\}$	0.333	0.333	0.533	0.533	17	0.280	
$p_7 = \{18:9,7,5,5,5,4\}$	0.333	0.333	0.533	0.533	1	0.667	
$r_1 = \{13:8,4,3,3,3,3\}$	0.367	0.362	0.558	0.551	14	0.300	
$r_2 = \{18:10,7,5,4,4,4\}$	0.367	0.362	0.558	0.551	15	0.484	0.395
$r_3 = \{9:5,3,2,2,2,2\}$	0.367	0.362	0.558	0.551	14	0.393	
$s_1 = \{10:5,3,3,3,2,2\}$	0.317	0.317	0.493	0.479	12	0.289	
$s_2 = \{25:10,8,8,8,7,7\}$	0.317	0.317	0.493	0.479	15	0.369	0.351
$s_3 = \{13:6,4,4,4,3,3\}$	0.317	0.317	0.493	0.479	11	0.394	
$t_1 = \{24:10,8,8,7,7,6\}$	0.300	0.300	0.453	0.435	12	0.433	
$t_2 = \{18:8,6,6,5,5,4\}$	0.300	0.300	0.453	0.435	12	0.406	0.317
$t_3 = \{36:14,12,12,11,11,10\}$	0.300	0.300	0.453	0.435	4	0.400	
$u_1 = \{18:7,6,6,6,5,5\}$	0.300	0.300	0.438	0.438	20	0.178	
$v_1 = \{22:8,7,7,7,7,6\}$	0.267	0.267	0.389	0.381	22	0.336	
$w_1 = \{13:8,5,3,3,3,3\}$	0.400	0.393	0.574	0.574	12	0.581	
<p>S.S. = Shapley-Shubik Index                      Banz. = Banzhaf Index                      B.S. = Burgan Shapley Index                      John. = Johnston Index  <i>n</i> = Number of rounds played  <math>\bar{w}</math> = Average percentage won by largest player  <math>\overline{\bar{w}}</math> = Average of <math>\bar{w}</math> across a number of power-identical sets                      Mean for all relatively equal <math>p_i</math> profiles (<math>p_1</math>-<math>p_3</math>) is 0.387 with <i>n</i> of 55.                      Mean for all relatively unequal <math>p_i</math> profiles (<math>p_4</math>-<math>p_7</math>) is 0.298 with <i>n</i> of 73.</p>							



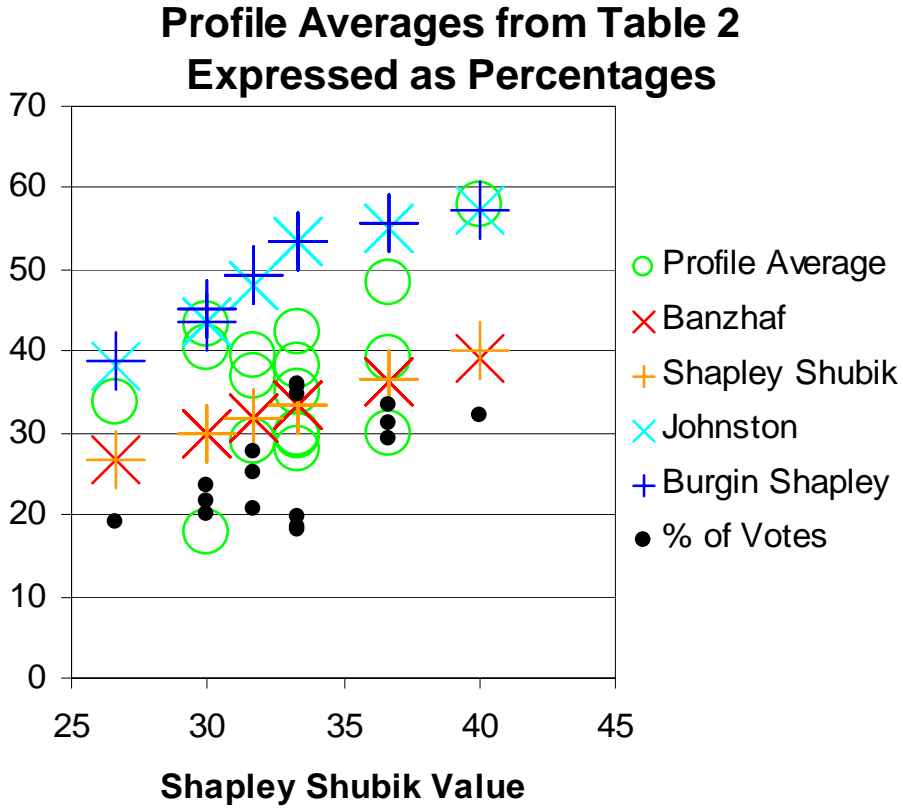
TABLE 3: REGRESSION OF CONTROL VARIABLES

Linear Regression Analysis						
n	288	s	0.22679	d		
R <sup>2</sup>	0.039727	Adj R <sup>2</sup>	0.019223	Raw R <sup>2</sup>	0.716735	
Tests, etc, for Individual Coefficients						
Variable	Coeff. (est.)	Std Err	H <sub>0</sub> : Coefficient = 0 H <sub>1</sub> : Coefficient ≠ 0		Confidence Ints.	
			T	p-value	Level	0.95
			Lower	Upper		
1	-0.01913	0.152059	-0.12578	0.899994	-0.31843	0.28018
SS	1.195785	0.45004	2.657063	0.008329	0.309948	2.081622
Male	-0.00323	0.037845	-0.08547	0.931951	-0.07773	0.071257
Foreign	0.06142	0.063227	0.971429	0.332161	-0.06303	0.185872
PersI	-0.01576	0.034636	-0.45491	0.64952	-0.08393	0.05242
RiskA	-0.0658	0.046732	-1.40792	0.160248	-0.15778	0.02619
RiskL	-0.027	0.045432	-0.59421	0.552845	-0.11642	0.06243

TABLE 4: CONFIDENCE INTERVALS

Profile	n	$\bar{w}$	St.Err	95% C.I	SS, B	J, BS Lowest
$p_i$	128	0.337	0.007	0.322-0.351	0.333	0.533
$r_i$	43	0.395	0.017	0.360-0.429	0.362	0.551
$s_i$	38	0.351	0.013	0.324-0.377	0.317	0.479
$t_i$ and $u_1$	48	0.317	0.012	0.294-0.341	0.300	0.435

**FIGURE 1**



From left to right, the profiles for each column of values are:  $v_1$  at SS=26.7%,  $t_i$  and  $u_1$  at SS=30.0%,  $s_i$  at SS=31.7%,  $p_i$  at SS=33.3%,  $r_i$  at SS=36.7%,  $w_1$  at SS=40.0%.

**APPENDIX 1: RISK AVERSION INSTRUMENT**

In this section the alternatives are different chances that you will receive different amounts of money, given that we select your reply. We will select twenty replies for this section and each of these twenty will have the following chances to win the following rewards.

Consider this example:

	A	B	C
(10) Chance:	100%	50%	10%
Winnings	\$5.00	\$10.00	\$50.00

Suppose that this reply was one of the twenty we chose for payment. If you selected alternative A, we would give you five dollars. IF you chose B we would flip a coin, and if it came up 'heads' we would give you ten dollars. If you chose C we would roll a ten sided dice, and if it came up 10, we would give you fifty dollars.

Choose and circle A, B or C for each of the following alternatives.

	A	B	C
(10) Chance:	100%	50%	10%
Winnings	\$5.00	\$10.00	\$50.00

	A	B	C
(11) Chance:	100%	50%	10%
Winnings	\$4.00	\$12.00	\$40.00

	A	B	C
(12) Chance:	100%	50%	10%
Winnings	\$4.50	\$10.00	\$45.00

	A	B	C
(13) Chance:	100%	50%	10%
Winnings	\$4.00	\$10.00	\$40.00

Note: This is not part of questionnaire. Risk loving answers are: C,C,C,C; C,B,C,B; and C,B,C,C. Risk neutral answers are: B,B,B,B; A,B,B,B; and C,B,B,B. Risk averse answers are: A,A,A,A; A,B,A,B; and A,B,A,A. Other combinations are intransitive.

## APPENDIX2: SUBJECT INSTRUCTIONS

Note that this appendix will be available from the authors by website or email request for the proceedings of the conference.

### Player Instructions for the Power Index Game

You're about to become a 'player' in an experimental game that investigates the influence that voting power of electoral groups has on democratic decisions. Confusing? A player could represent a political party and all its voters or a corporate board member where votes could be the number of your shares. The procedures are really quite simple and you get to make some big money!

#### The Game is as Follows:

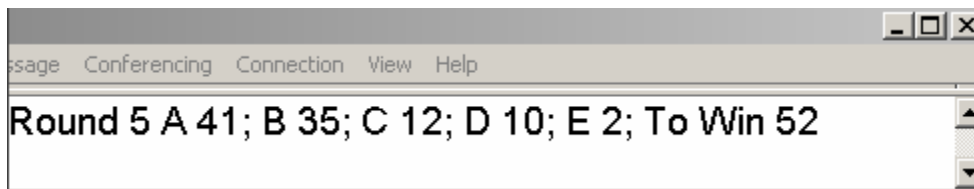
- You will be divided into groups, but you won't know whom you're playing against.
- You will communicate only through open First Class chat rooms (not private).
- You will be given a username such as player **A**, or **B**.
- A game consists of twelve rounds.
- In each round you are allocated a different number of votes, e.g. 12 or 2 votes.
- In each round your group has between \$15 to divide democratically.
- Division of money is done by offering a proposal, and accepting the proposal.
- Each round goes for a maximum of 5 minutes.
- At the end of all the rounds you get the sum of what you earned in CASH.

Your job is to be GREEDY and get as much money as possible!

#### Game Notation:

It is difficult to grasp the relationship between votes and power and so is getting used to the games notation.

You will see a 'vote vector' like this:



This tells us that:

- We are in round 5
- Player A has 41 votes
- Player B has 35 votes etc...
- To win this game the players who agree on a proposal must have at least 52 votes combined to win.

Lets say you are playing for \$10 in this round. To offer a proposal in the open chat room you could use the following notation:

*Prop B1: A5 B5*

This tells us that:

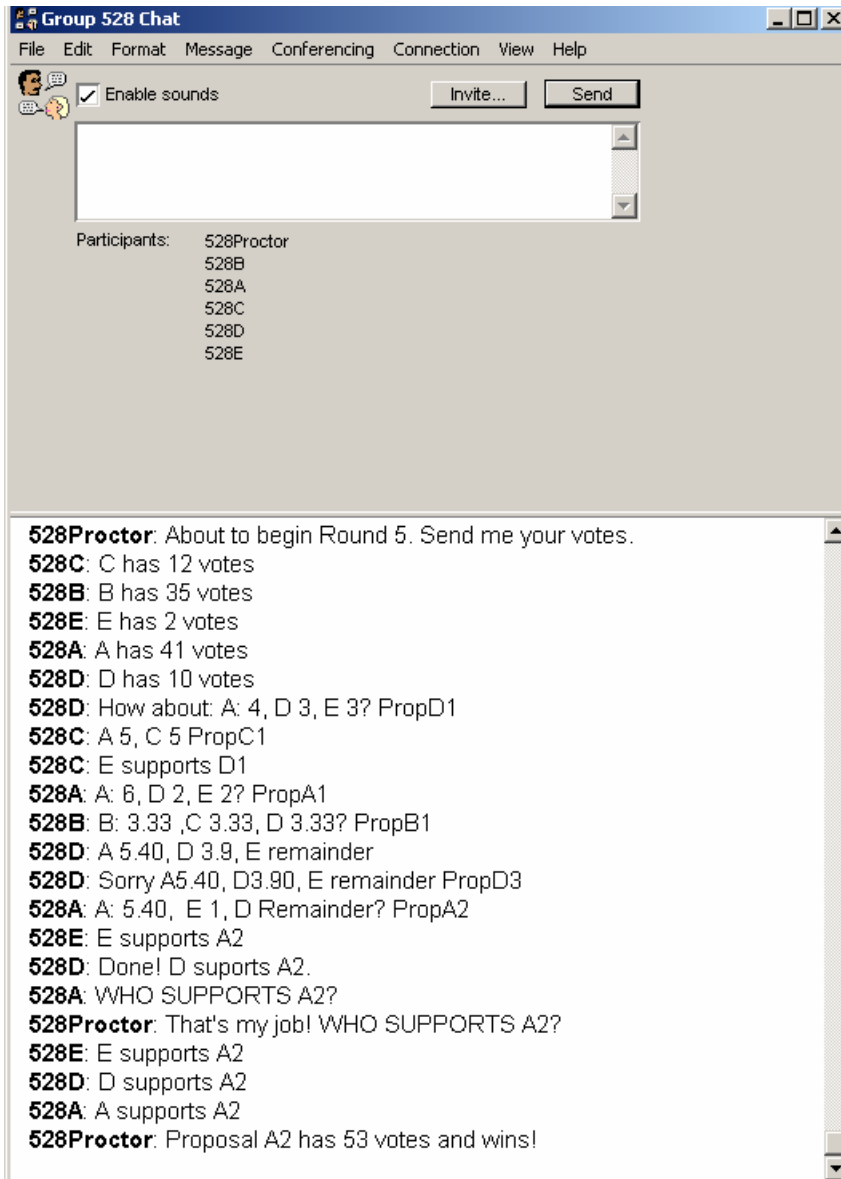
- Player B is making his/her first proposal (prop B1)
- Player B offers A \$5 and B \$5
- The combined votes (if A accepts) is 76, this proposal has majority votes and is hence accepted.

If Player A doesn't accept and player B wishes to make a second proposal it should be titled **Prop B2**, or just **B2**.

To accept a proposal Player A could write:

**Supports B1**

An example of a game for the above vote vector is as follows:



Players in the group # 528

Enter the number of votes you have so proctor knows there are no misunderstandings

Negotiations begin; present your proposals here, a maximum of 5 minutes before

Enter your support of a proposal. Proctor announces winner for this round

Note: Example is from the pre-study with a purse of \$10.

Here you can see that the ‘proctor’ announces which proposal wins. Proposal A2 has since players A, D, and E all support A2 giving a total of exactly 53 votes. Whereby, **A** receives \$5.40, **E** \$1 and **D** \$3.60. i.e. A 5.40, E 1, D remainder Prop A2

It is also interesting to note that just because you’re allocated are large amount of votes doesn’t necessarily mean you will get the biggest earnings! Player B with 35 votes completely missed out! On one hand, Player B should have (once he realised that he may not get anything) offered a proposal like; Prop B2: A 8, B 2. He could then get \$2 instead of **0**. In this case player A could be greedier and accept this proposal, forgetting any proposals that included players C, D, or E!

### *Playing the Game*

At the start of the game, the proctor will send a message:

**“About to begin round 1. Send me your votes.”**

Reply stating your letter and number of votes:

**“A has 41”** or **“B has 35”** etc.

Then the proctor will announce the beginning of the round:

**“Start round 1”**

And the end of the round or when sufficient support is shown before a proposal:

**“End of round Prop ## wins. Stop voting”** At this point no more proposals can be offered or accepted

If a proposal appears to have enough votes to win, the proctor will call a vote. (If the proctor has not noticed support for a proposal, you may send a message pointing out that a proposal has received enough support.) The proctor will say, for example, ‘Call for votes.’ At that time you cast a vote for one proposal: ‘I vote for PropE89’. There is no need to vote against a proposal because the Proctor will count no reply as a vote against. So, if you want a proposal to pass, vote for it even if it is your proposal. Proctors will declare if a proposal has passed or not.

- It is also not imperative that you stick to the notation examples, plain English is fine, however we have found that short hand is easily understood.
- A good message could be ‘Prop F43 is nearly winning. Hey F, C&D and I will give you more. B 3.25 C 2.00 D 2.00 F 2.75: PropB46.’
- You can argue about a proposal, try to persuade other players to support your proposal...whatever gets you the most money!

You will want to find other players whose votes can be added to yours to equal or exceed the minimum votes required to divide the money. You get them to cooperate with you by offering them some of the money. Of course, you want to keep as much as you can for yourself. Other players will propose deals that give you nothing. Perhaps you can get some of them to share with you by offering them more than their current deal.

Once the game is over, wait for the proctor to give you instructions to begin a new game.

### **Game Rules:**

- No Private Chat rooms are allowed!
- You are not allowed to make references to the world outside of the game. You may not use your names or any other real world identifier. You may not make side deals such as offers to do homework for or threats of violence to other players in the game.

Players who violate these rules will forfeit their earnings and be removed from the game

- Be ***Greedy!*** Try to get as much money as possible. In previous experiments players have earned \$60 after a couple of rounds! Don't feel sorry for players with a small amount of votes. You never know in the next round they could have the majority!
- If you identify two proposals with the same number and one of them passes, we will give you the lower amount of money.

### **Summary:**

Let us review. In each round of this game you will divide \$10 between yourself and other players. You will get some votes which may change between every round. You will use a Group chat room to discuss how the money should be divided. You will vote on who gets the money in the Group chat. Whatever you earn in this game is yours to keep.

Consider some possible strategies. You could:

- o Look at every player's votes to see what combinations can win.
- o Who can help you? Who can you help?
- o Are any players dependent on your votes in particular?
- o Make offers quickly hoping to make big earnings before others catch on.
- o If you are left out of a proposal, think of a way to divide the money so that some of the players in the proposal get more by voting with you and your votes are enough with theirs to create a majority. That is, break up coalitions that leave you out.

### **Please keep several points in mind:**

- You are playing for real money.
- You may earn nothing in this experiment.
- You don't know who is in your group and you cannot speak during rounds
- Check the Votes Vector area because votes may change between rounds.
- Be careful not to close the group chat room or the vote vector.



- It is a majority of votes, not of players that decides how the money gets divided.
- You may not offer any deal that involves anything from outside of the game. That is, you cannot offer to do someone's homework or threaten them to make them cooperate with a deal.
- Every majority coalition of players can be broken.
- This game is divided into rounds. In each round, your group of players will divide \$10 among yourselves.
- Each round will last up to five minutes, but may be shorter if players reach an agreement sooner. Any later and no one gets any money
- Keep an eye out for bogus proposals. Player B for example names a proposal 'D2'. This is technically possible, against the rules, probably pointless and silly, but it could cause confusion.
- Make sure that your proposals contain players that will give you the minimum amount of votes 'To win'. For example if B has 35 and C has 2 votes, To win is 52, then Prop B1 B5, C5 cannot work since the summation of the votes is only 37
- Make sure your proposals add up! If you're dividing \$10, make sure that you don't offer B1 B9, C5. This adds up to \$14...not \$10