Meaningful learning in weighted voting games: An experiment*

Eric Guerci† Nobuyuki Hanaki‡ Naoki Watanabe§

March 1, 2015

Abstract

This paper experimentally investigates whether subjects learn, from their limited experiences, something about the underlying relationship between their nominal voting weights and expected payoffs in weighted voting games. In each session of the experiment, subjects are repeatedly asked to choose one of two 4-player weighted voting games where three out of four players are fictive. After each choice, payoffs for players are determined automatically according to a theory of voting power. We, however, vary the amount of feedback about the payoffs that subjects receive after each choice. The choice problem of games is changed to another one in the second part of the session. The main findings are as follows. (1) The fraction of subjects who chose the weighted voting games that generate higher expected payoffs increased even in the treatment without any immediate feedback on realized payoffs. (2) A statistically significant evidence of "meaningful learning" between two choice problems was observed only in the treatment without immediate feedback.

JEL Classification Numbers : C79, C92, D72, D83

Keywords : voting power, learning, experiment, two-armed bandit problem

*The authors wish to thank Gabriele Esposito and Xiaoyan Lu for stimulating discussion and collaboration at early stage of this project, and Yoichi Izunaga for his excellent research assistant. The ISER experimental economics laboratory at Osaka University, in particular, Keigo Inukai, Emi Kurimune, and Shigehiro Serizawa, has kindly hosted the authors for running this experiment. Financial support from MEXT Grants-in-Aid 24330078 and 25380222 (Watanabe), JSPS-ANR bilateral research grant "BECOA" (ANR-11-FRJA-0002) are gratefully acknowledged.

†Université de Nice Sophia Antipolis (I.S.E.M.), CNRS GREDEG (UMR 7321). E-mail: eric.guerci@gredeg.cnrs.fr

‡Corresponding author: Aix-Marseille University (Aix-Marseille School of Economics), CNRS, EHESS, and IUF. E-mail: nobuyuki.HANAKI@univ-amu.fr. Postal address: GREQAM, Aix-Marseille University. 3 Av. Robert Schuman, 13628, Aix-en-Provence, Cedex 01, France. Tel: +33-(0)442172982. Fax: +33-(0)491900227

§Faculty of Engineering, Information, and Systems. University of Tsukuba. E-mail: naoki50@sk.tsukuba.ac.jp

1 Introduction

This paper experimentally investigates whether subjects learn something about the underlying properties of a class of strategic situations (games) and whether they transfer what they have learned from their limited experiences in a set of games to another set of games within the same class. We examine weighted voting games as the class of games; thus our question is, more concretely, whether subjects can learn, from their limited experiences, about the underlying relationship between their nominal voting weights and actual expected payoffs derived from a theoretical power index in weighted voting games and generalize what they have learned in a game to similar but different one.

We begin with a brief survey of the literature on learning experiments. For many years, experimental studies on learning in games have been focused on whether subjects learn how to play games and how such learning proceed.[1] To better understand the dynamics of observed behavior of subjects, various models of learning have been proposed, such as, reinforcement learning (e.g., Erev and Roth, 1998), belief-based learning (e.g., Cheung and Friedman, 1997), and the experienced weighted attraction (EWA) learning (Camerer and Ho, 1999). Some researchers, however, reported that many learning models including the ones mentioned above failed to replicate observed behavior in repeated games.[2] New ideas are thus currently being examined among researchers.[3]

In comparison to the large amount of studies on learning how to play games and its process, only a few studies have been done to investigate "cross-game learning" (Cooper and Kagel, 2003) or "learning transfer" (Cooper and Kagel, 2008; Haruvy and Stahl, 2009), i.e., whether subjects learn about the underlying properties of games and generalize what they have learned in a situation to similar but different situations. In terms of the depth of what is learned, this higher-order concept of learning should be distinguished from learning to make choices that generate better outcomes in a given situation, and thus it is referred to as "meaningful learning" by Rick and Weber (2010).[4]

[1]This strand of literature is different in its objective from the literature on learning theory in games. See, e.g., Hart (2005) and the references therein for the theoretical literature, where authors mainly investigate the convergence properties of learning models to various equilibria.

[2]Arifovic et al. (2006) showed that many models of learning failed to replicate observed human behavior in such games as a repeated battle of sexes game. Erev et al. (2010) reported that the standard learning models based on the evolution of attraction did not perform well in predicting how people behave in market entry games.

[3]Marchiori and Warglien (2008) proposed to incorporate "regret" into their neural network-based learning model, showing that it better replicated the observed behavior than both EWA and neural network-based learning models without regret. Hanaki et al. (2005) and Ioannou and Romero (2014) extend reinforcement learning model and EWA learning model, respectively, to allow players to learn which strategies to use in repeated games. Arifovic and Ledyard (2012) report that their "individual evolutionary learning" model can capture most of the stylized results in public good game experiments.

[4]Dufwenberg et al. (2010) call this type of learning "epiphany," and investigate it in two race to X games (which they

Rick and Weber (2010) studied p-Beauty contest games (Ho et al., 1998) and found that withholding feedbacks, which do not produce any learning at all in many learning models, promotes meaningful learning in the sense that subjects learn to perform iterated dominance.[5] Neugebauer et al. (2009), on the other hand, reported that subjects did not learn to play their dominant strategy in a game of voluntary provision of public goods when they did not receive any feedback. Our aim is to study this deeper learning in weighted voting games.[6]

Weighted voting, which gives different numbers of votes to different voters, is one of the most popular collective decision-making systems in many contexts: stockholder voting in corporations, multi-party legislatures, and so on. The relationship between the nominal voting weight and real voting power, however, is often complex. In the study of the Council of Ministers in the European Economic Community, Felsenthal and Machover (1998, pp.164-165) reported that it must have been difficult for policy makers who designed the system to see through the underlying relationship between the nominal voting weights and actual voting powers.

To better understand this complex relationship inherent in weighted voting, recently, researchers have turned to experimental study to complement more traditional empirical analyses. This is because many features that are unobservable in actual practice can be controlled in experiments. Montero et al. (2008), Aleskerov et al. (2009), Esposito et al. (2012), and Guerci et al. (2014) conducted the experiments in which subjects decide how to allocate fixed amount of resources among them via weighted voting. These experiments are under a cooperative game environment because they did not specify explicitly the dynamic structures of negotiations among the subjects.[7] These studies

call "game of X"). In a race to X game, two players alternatively put 1 to M coins in the same hat which is initially empty. The game ends when there are X coins in the hat, and the player who has put the Xth coin into the hat is the winner. Both M and X are common knowledge. This game has a dominant strategy and can be solved by using backward induction. In the version studied by Dufwenberg et al. (2010), X were either 6 or 21, and M was 2. Dufwenberg et al. (2010) hypothesized that it is easier to realize the dominant strategy for this game when X is small, and once a subject realizes the dominant strategy, s/he will apply it every time playing the same class of games. They found, indeed, that those subjects who experienced playing race to 6 before playing race to 21 played the latter perfectly more often than those who started with race to 21. Learning to play race to X is also studied experimentally by Gneezy et al. (2010) who report that subjects learn to play the dominant strategy as they repeat the same game. Carpenter et al. (2013) report the correlation between subjects' cognitive ability, measured by Raven's Progressive Matrix test (described in Raven, 2008, for example), and their frequency of playing the dominant strategy in race to 5, 10, and 15 games with $M = 3$ against best-responding computers.

[5]Rick and Weber (2010) also found that asking subjects to explain the reason for their behavior promotes meaningful learning in the presence of feedback.

[6]Cooper and Kagel (2003, 2008) examined signaling games and found that letting subject to play in a team promotes "meaningful learning" or what they call "learning transfer." Haruvy and Stahl (2009, 2012) combined rule-based learning (Stahl, 1996) and EWA learning models to explain observed behaviors in the laboratory experiments in which subjects play a sequence of 4×4 symmetric normal form games that are dominance solvable. They found that a model that relabels the actions based on the steps of eliminating dominated strategies, which assumes that agents understand the basic property – dominance solvability – of the game, allows the model to better capture the observed behavior of subjects.

[7]Fréchette et al. (2005a,b), Drouvelis et al. (2010), and Kagel et al. (2010), on the other hand, conducted their ex-

found that an experimental measure of "a posteriori" voting power differs dramatically from the theoretical measures of "a priori" voting power such as Banzhaf (1965), Shapley and Shubik (1954), and Deegan and Packel (1978). The experiential measure of voting power is defined as the average payoff a voter has obtained during the experiment.

Those remarkable discrepancies between theoretical predictions and experimental observations calls for a better behavioral theory of weighted voting games including how and what subjects learn while playing these games. In particular, one can ask whether subjects learn something about complex underlying relationship between the nominal voting weights and voting power. In this paper, we take a small step toward answering this question. As shown by Esposito et al. (2012), trying to investigate what subjects learn when other subjects are also learning result in too complex a task for subjects to allow for very meaningful inferences. We thus simplify the experimental protocol drastically by abstracting away the issue of subjects learning how to play the game while other subjects are also doing the same. And we focus on whether subjects learn to become a member of a committee that generates a higher expected payoff for themselves. In our experiment, payoffs are automatically generated according to a theory of voting power as we describe below.

As a result of this drastic simplification, our subjects face two-armed bandit problems with contextual information related to payoffs, and simply choose a weighted voting game itself out of the two options. These two options share the number of voters, the total number of votes, the quota (the minimum number of votes required for a proposal to be adopted), and the number of their own votes. Only difference between the two is, thus, the votes apportioned among the other voters, although they are all fictive.[8] Payoffs for subjects from choosing one of the two options (committees) are determined stochastically based on the theory of voting power developed by Deegan and Packel (1978) which will be explained in the Section 2.

We have three treatments with different amounts of feedback subjects receive immediately after their choice; no feedback, partial feedback, and full feedback. In the no-feedback treatment, subjects are not given any information about the payoffs they obtain after each choice. Thus, only information they can use to make their choices is the description of how votes are apportioned across four players in two committees. In the partial-feedback treatment, after each choice, subjects are given information about their own payoffs for the option they have chosen (and not for the option

periments in non-cooperative game environments with explicit dynamic structures of negotiations based on variants of a legislative bargaining models.

[8]The standard two-armed bandit problems do not provide subjects with any contextual information that is related to the payoffs for the problem they face. See, for example, experiments by Meyer and Shi (1995), Banks et al. (1997), and Hu et al. (2013). Thus, our problems can be called as an extended two-armed bandit problem.

they have not chosen), which matches the feedback given in the standard two-armed bandit problems. In the full-feedback treatment, after each choice, subjects are given information about the payoffs for all the relevant voters again only for the option they have chosen and not for the option they have not chosen.

Standard models of learning based on the evolution of attraction (e.g., reinforcement learning and EWA learning models) cannot be used to make predictions about how subjects learn to choose the option with the higher expected payoff in no-feedback treatment. In fact, we are not aware of any models of learning that allow us to study learning without any explicit feedback.

We found that the fraction of subjects who chose the option with a higher expected payoff increased even without immediate feedback on payoffs. Further, we had a statistically significant evidence of "meaningful learning" between two choice problems in the treatment without immediate feedback but not in the treatments with immediate feedbacks on payoffs.

The rest of the paper is organized as follows. Section 2 describes our experimental design. Section 3 presents and discusses the experimental results. Conclusions are offered in Section 4.

2 Experimental design

Let $N = \{1, 2, 3, 4\}$ be the set of players. A four-player weighted voting game is represented by $[q; v_1, v_2, v_3, v_4]$, where q is the quota (the minimal number of votes required for a proposal to be adopted) and v_i is the voting weight (number of votes) player $i \in N$ has. We hereafter call a four-player weighted voting game simply a game. In each period, each subject, who acts as player 1 and is informed clearly that three other players are fictive, chooses one of the two games that have the same total number of votes, the same quota, and the same number of own votes. For example, in one period, a subject chooses between $[14; 5, 3, 7, 7]$ and $[14; 5, 4, 6, 7]$, in another period, the subject chooses between $[6; 1, 2, 3, 4]$ and $[6; 1, 1, 4, 4]$. In the former, the subject receives 5 votes whichever the choice he or she makes, and similarly in the latter, the subject receives 1 vote.

In our experiment, the payoff a subject obtains from his or her choice is stochastically determined according the idea behind an index of voting power proposed by Deegan and Packel (1978), and not according to the outcomes of four subjects actually negotiating how to divide the resources in the chosen committee. We have decided to rely on this clear and relatively simple theory to generate payoffs in order to avoid the complication resulting from subjects trying to learn which committee to choose while other subjects learning how to negotiate within each committee. Cason and Friedman (1997) point out a similar complication in their market experiments and introduce sessions in which subjects interact with "robots" that follow the equilibrium strategy. In addition to this complication, the existing experimental results (Montero et al., 2008; Aleskerov et al., 2009; Esposito et al., 2012; Guerci et al., 2014) provide us with little guidance as to what would be the expected payoffs for each member of a committee when we let subjects to play a weighted voting game. This lack of clear guidance makes it difficult to design binary choice problems where one of the options generates a higher expected payoff than the other without automatically generating payoffs based on a theory.

Let us now describe the payoff generating process we have employed. To do so, we first need to define a few terms. A nonempty subset S of N is called a coalition, and it is called a winning coalition if $\sum_{i \in S} v_i \geq q$; otherwise, it is called a losing coalition. A minimum winning coalition (MWC) is a winning coalition such that a deviation of any member of the coalition turns its status from winning to losing. In constructing an index of voting power, Deegan and Packel (1978) assumed that (1) a winning coalition will always be minimum, (2) all the MWCs are equally likely to be formed, and (3) within each MWC, the resources (120 points in our experiment) will be divided equally among its members. Thus, for example, in the case of $[14; 5, 3, 7, 7]$, possible MWCs expressed in terms of the number of votes each player has are $(5, 3, 7)$, $(5, 3, 7)$, and $(7, 7)$. Thus, player 1 may obtain 1/3 of the total resources 2/3 of the times (when a MWC $(5, 3, 7)$ is realized) and zero in remaining 1/3 of the times (when MWC $(7, 7)$ is realized).

Our choice of using the idea from Deegan and Packel (1978) in generating payoffs is mainly driven by the ease of its experimental implementation. Whether subjects can learn the relationship between the nominal voting weights and actual voting powers based on the other power indices defined by Banzhaf (1965) or Shapley and Shubik (1954) is an interesting question to be investigated in the future.

This choice of payoff generating process led us to choose four-player weighted voting games, instead of simpler three-player weighted voting games. Because four-player games generate non-trivial sets of minimum winning coalitions, while three-player games do not. We do not explain this underlying payoff generating mechanism to our subjects, and simply inform them that payoffs are determined based on a theory of weighted voting games.[9]

Each session consists of 60 periods. Subjects face the same binary choice problem for the first 40 periods and then another binary choice for the next 20 periods. One can thus consider this

[9] See Appendix A for an English translation of the instruction.

Problem	Choice 1	(Expected payoff)	Choice 2	(Expected payoff)
A	$[14; \mathbf{5}, 3, 7, 7]$	(120 × 2/9)	$[14; \mathbf{5}, 4, 6, 7]$	(120 × 1/4)
B	$[6; \mathbf{1}, 2, 3, 4]$	(120 × 1/9)	$[6; \mathbf{1}, 1, 4, 4]$	(120 × 2/9)
C	$[14; \mathbf{3}, 5, 6, 8]$	(120 × 2/9)	$[14; \mathbf{3}, 6, 6, 7]$	(120 × 1/4)
D	$[9; \mathbf{1}, 3, 5, 6]$	(120 × 1/9)	$[9; \mathbf{1}, 2, 6, 6]$	(120 × 2/9)

Table 1: Four binary choice problems examined in the experiment. The number of votes allocated to our subjects are shown in bold. Note that for all the problems, Choice 2 has a higher expected payoff for our subjects.

problem as subjects facing two-armed bandit problems with a hint about hidden payoff generating mechanism. We are interested in (a) whether our subjects learn to choose the option with the higher expected payoff (henceforth, the better option), and, if they do so in the first 40 periods, (b) whether subjects who have experienced one choice problem choose the better option more frequently when they face a new choice problem than those subjects facing the same problem without such an experience.

We examined the set of binary choice problems shown in Table 1. In each committee, our subjects receive the number of votes assigned for player 1 that are shown in bold. Note that in terms of the payoffs (not only in terms of expected payoffs but also the set of all the possible payoffs) Problem A and C are the same. The same is true for Problem B and D. Our subjects face one of the following four sequences of choice problems in the order indicated by the arrow: $A \to B$, $B \to A$, $C \to D$, or $D \to C$. The first 40 periods are based on the first choice set, and the last 20 periods are based on the second choice set.

2.1 Three feedback treatments

We consider three feedback treatments: (1) no-feedback, (2) partial-feedback, and (3) full-feedback. In the no-feedback treatment, subjects are not given any information about realized payoffs after each choice. Thus, stimulus-response learning models that are based on the realized payoffs make no prediction about whether learning will take place in this treatment. In the partial-feedback treatment, after each choice, subjects are given information about the payoffs he or she receives but not the payoffs of the other (fictive) voters for the option they have chosen, but not for the one they have not chosen. This treatment is introduced because it is the closest to the feedback condition in the standard two-armed bandit problems in which subjects are informed only about the payoff they obtained after each choice. Our subjects will observe that the same choice generates varying payoffs, and the frequencies of receiving various payoffs differ between two choices. Stimulus-response learning models will predict that some learning would be observed in this treatment. In the full-feedback treatment, after each choice, subjects are given information about the payoffs for all the (fictive) voters. Just as in partial-feedback treatment this information is given only for the option they have chosen. Because under the full-feedback treatment, subjects observe that (a) payoffs are equally allocated among a subset of committee members and that (b) such subsets of committee members change from time to time, we expect the learning of the underlying relationship between the payoffs and votes to be much easier in the full-feedback treatment than in two other treatments.

In order to keep the amount of time given to subjects for (a) making choice, and (b) thinking with or without feedback constant across three feedback treatments, we assign 30 seconds time limit for the choice stage and 10 seconds in the feedback stage. If a subject does not make a choice within 30 seconds of the choice stage, he or she receives zero point for that period. In this case, the subject receives the special feedback, regardless of the feedback condition in the feedback stage, that he or she has obtained zero point for the period because he or she did not make a choice. The zero point in case of no choice was clearly written in the instruction.

If a subject makes a choice early, say within the first 10 seconds of the choice stage, a waiting screen is shown until all the subjects in the session make their decisions. If all the subjects make their choices before the end of the 30 seconds time limit, every subject enters the feedback stage. In the feedback stage, under the no-feedback condition, 10 seconds of waiting screen with a message that reads "Please wait until the experiment continues" is shown. In full or partial feedback treatments, the relevant payoff information is shown for 10 seconds.

In addition to the show-up fee of 1000 yen, subjects are paid according to the total points they receive over 60 periods with an exchange rate of 1 point = 1 yen. We pay our subjects in this way just as in other bandit experiments (see, e.g., Meyer and Shi, 1995; Hu et al., 2013), and not based on a randomly chosen period, because we want our subjects to learn to choose the option with a higher expected payoff, and not the option that can generate the highest possible payoff. The experiment lasted about 60 minutes including the instruction and post-experiment questionnaire. On average, our subjects earned about 2500 yen (about 25 U.S. dollars at the exchange rate around the date of the experiment.)

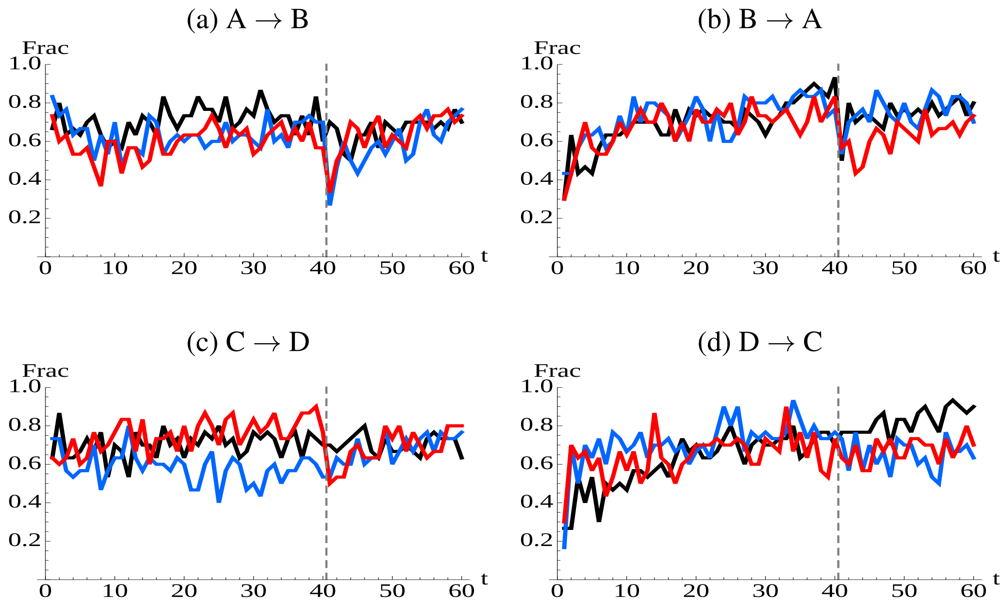

Figure 1: Time series of the fraction of subjects who have chosen "correct" option. Black: No-feedback, Red: Full-feedback, Blue: Partial-feedback.

3 Results

A computerized experiment (developed with "z-Tree" Fischbacher, 2007) was conducted at the ISER experimental economics laboratory at Osaka University (Japan) and the experimental laboratory at University of Tsukuba (Japan) between June and November 2014. 360 subjects who have participated to our experiments were undergraduate students recruited from all over the campus. Neither third nor fourth-year students who major in economics participated in the experiment.[10]

3.1 Overview of the data

Figure 1 shows the time series of the fractions of subjects who chose the better option (i.e., the option that gives a higher expected payoff) in each of the four sequences of choice problems we examined. Three feedback treatments are shown in different colors: Black; no-feedback, Red: full-feedback, Blue: partial-feedback. The dotted vertical line in each panel separates the two choice problems, i.e., the one before period 40 and the another one after period 41.

[10] Out of 360 subjects, 240 subjects were at Osaka and remaining 120 subjects were at Tsukuba. 4 sequences of two problems $\times$ 3 feedback treatments $=$ 12 sessions in total. Each session consisted of 20 subjects in Osaka and 10 subjects in Tsukuba. Out of these 360 subjects, 54 of them failed to make a choice within the time limit at least once during 60 periods.

Comparing panel (a) and (b) as well as (c) and (d) of the figure, we observe that more subjects have chosen the better option from the beginning in Problem A and C than in B and D. The fractions of subjects who have chosen the better option in period 1 are 0.74 for Problem A and 0.67 for Problem C, while they are 0.34 for Problem B and 0.24 for Problem D. Here we are pooling the data from three feedback treatments because at the time of making the choice in period 1, no subject has received any feedback related to payoffs. In addition, we do not reject the null hypothesis that the fractions of subjects choosing the better option in the first period are the same across the three feedback treatments in none of the 4 problems. P-values are 0.338, 0.464, 0.646, and 0.466 for Problem A, B, C, and D, respectively, based on Kruskal-Wallis (KW) test. The null hypothesis that these fractions are the same across four problems is rejected ($p < 0.001$, KW). We do not, however, reject the null hypothesis that the fractions for Problem A and C are the same ($p = 0.255$, Mann-Whitney, MW, test, two-tailed). The same is true for the fractions for Problem B and D ($p = 0.146$, MW).

Observation 1. *Problem A and C were easier for subjects to figure out the better option than Problem B and D from the beginning of the experiment.*

A possible reason for this observation, based on the answers to the questionnaire about the reason for their choices at the end of the experiment, is that subjects quickly identified the crucial difference between the two options in Problem A and C. That is while the possibility of two large voters (voter 3 and 4 who have more votes than others) alone forming the winning coalition exists in one of the option ($[14; 5, 3, 7, 7]$ or $[14; 3, 5, 6, 8]$), the same is not true for the other option ($[14; 5, 4, 6, 7]$ or $[14; 3, 6, 6, 7]$). Subjects avoided the former option because they intuitively thought such a possibility must be dis-advantageous for themselves. For Problem B and D, on the contrary, there is no such an obvious difference in two choices because the two large voters can form a winning coalition by themselves in both choices. As a result, our subjects had to explore more to learn to make better choices.

Indeed, in Figure 1, we observe, for Problem B and D (shown in panels (b) and (d), respectively), an upward trend in the fraction of subjects who chose the better option in the first 40 periods. There is no such a trend in Problem A and C (shown in panel (a) and (c), respectively), however. This absence of upward trend in Problem A and C is probably due to many subjects having figured out the better option from the very beginning. It should also be noted that we do not observe much differences among three lines (corresponding to our three feedback treatments) in any of the four problems. Regardless of the differences in the payoff related feedbacks, the dynamics of the fractions of the subjects choosing the better options show the similar dynamics. Exceptions for this similarity can be identified in the following 4 cases. (1) Between period 41 and 60 in panel (b) where the line for the full-feedback treatment is clearly below the other two. (2) Between period 20 and 40 in panel (c) where the line for partial-feedback treatment is clearly below the other two. (3) Between period 5 and 15 of panel (d) where line for no-feedback treatment is below the the other two. And (4) between period 41 and 60 in panel (d) where line for no-feedback treatment is clearly *above* the other two. This final case is remarkable because it suggests that subjects who have received no payoff related feedback after each choice have learned to choose the better option more frequently than others who have received such feedback. Let us proceed to further investigate learning within and across problems.

3.2 Learning to choose the better option

To better analyze choices made by subjects and how their choices change over time, let us define, for each subject, a measure of his/her frequency of choosing the better option in various block of 5 periods. Namely, let FR^i_b be the fraction of bth block of 5 periods, in which subject i has chosen the better option. For example, FR^i_1 is the (normalized) number of times subject i has chosen the better option between period 1 and 5. FR^i_2 or FR^i_8 represent those between period 6 and 10 or period 36 and 40, respectively.

Based on FR^i_b, we can define the change in the relative frequency of choosing the better option between two blocks of 5 periods, f and g, as

$$\Delta FR^i_{f,g} = FR^i_f - FR^i_g$$

In this subsection, we focus on $\Delta FR^i_{2,1}$ and $\Delta FR^i_{8,1}$. The former is the change in the (normalized) number of periods in which a subject chose the better option in the second 5 periods (periods 6-10, $b = 2$) and the first 5 periods (periods 1-5, $b = 1$) of facing a problem, while the latter is the same change between the last 5 periods (period 36-40, $b = 8$) and the first 5 periods. We focus on these two changes to investigate whether differences in the amount of the feedback subjects receive have different impacts in the learning in a short ($\Delta FR^i_{2,1}$) and a long ($\Delta FR^i_{8,1}$) time horizons.

Figure 2 and 3 show the distribution of $\Delta FR^i_{2,1}$ and $\Delta FR^i_{8,1}$, respectively, for Problem A (panel a), B (panel b), C (panel c) and D (panel d) for three information treatments: Black: No-feedback, Red: Full-feedback, Blue Partial-feedback. For each game and feedback treatment, we test the

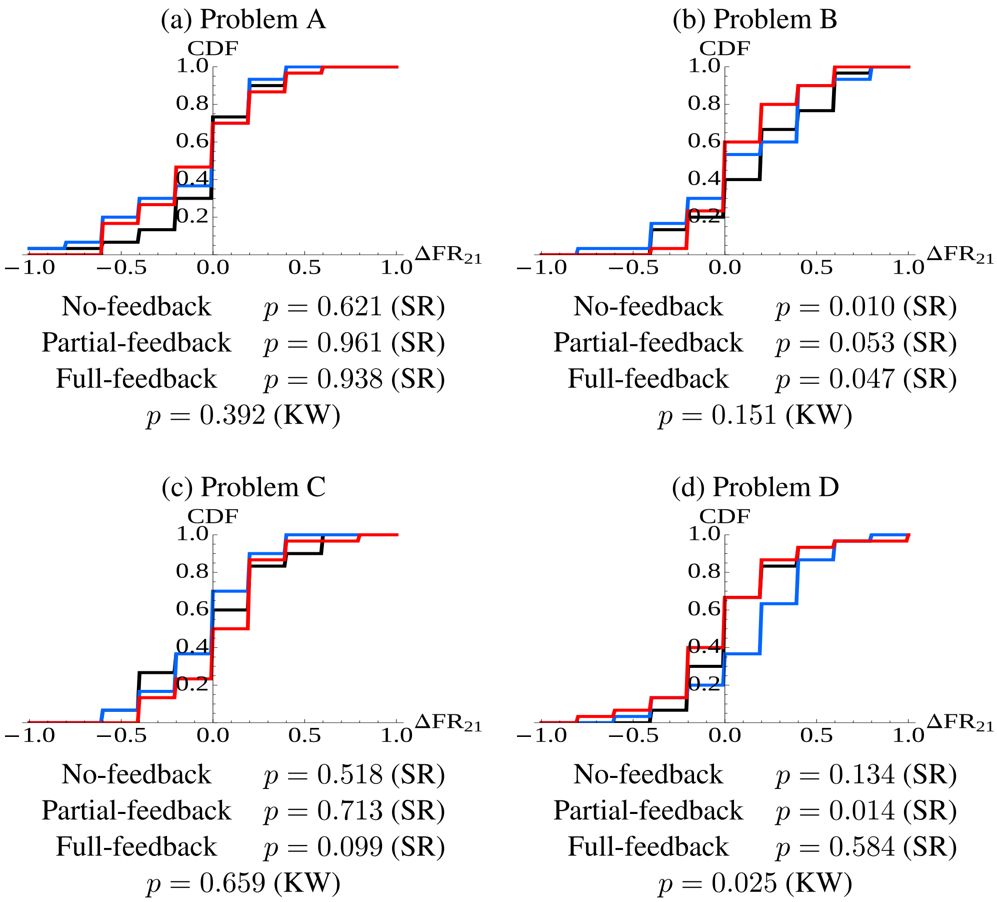

Figure 2: CDF of $\Delta FR^i_{2,1}$, the change in the (normalized) number of periods in which a subject chose the better option in the second 5 periods (periods 6-10, $b = 2$) and the first 5 periods (periods 1-5, $b = 1$) for Problem A (panel a), B (panel b), C (panel c) and D (panel d). Black: No-feedback, Red: Full-feedback, Blue: Partial-feedback. P-values for within treatment test (one-tailed, signed-rank test (SR)) and across treatments test (Kruskal-Wallis (KW)) are reported.

null hypothesis $\Delta FR_{2,1} = 0$ or $\Delta FR_{8,1} = 0$ (with the alternative hypothesis $\Delta FR_{2,1} > 0$ or $\Delta FR_{8,1} > 0$ with an assumption that learning will improve choices) using one-tailed-singed-rank test (SR). P-values are reported in the figure. We also test the null hypothesis that median $\Delta FR_{2,1}$s or $\Delta FR_{8,1}$s are the same across the three feedback treatments by Kruskal-Wallis test.

Panel (a) and (c) of both Figure 2 and Figure 3 show that for Problem A and C, which had higher FR_1 compared to Problem B or D, the distribution of $\Delta FR_{2,1}$ and $\Delta FR_{8,1}$ are not significantly greater than zero in most of the cases. Except for the full-feedback treatment of Problem C, we do not reject the null hypothesis that the median $\Delta FR_{2,1} = 0$ or $\Delta FR_{8,1} = 0$. The absence of the significant evidence on learning (to choose the better option) in Problem A and C is better under-

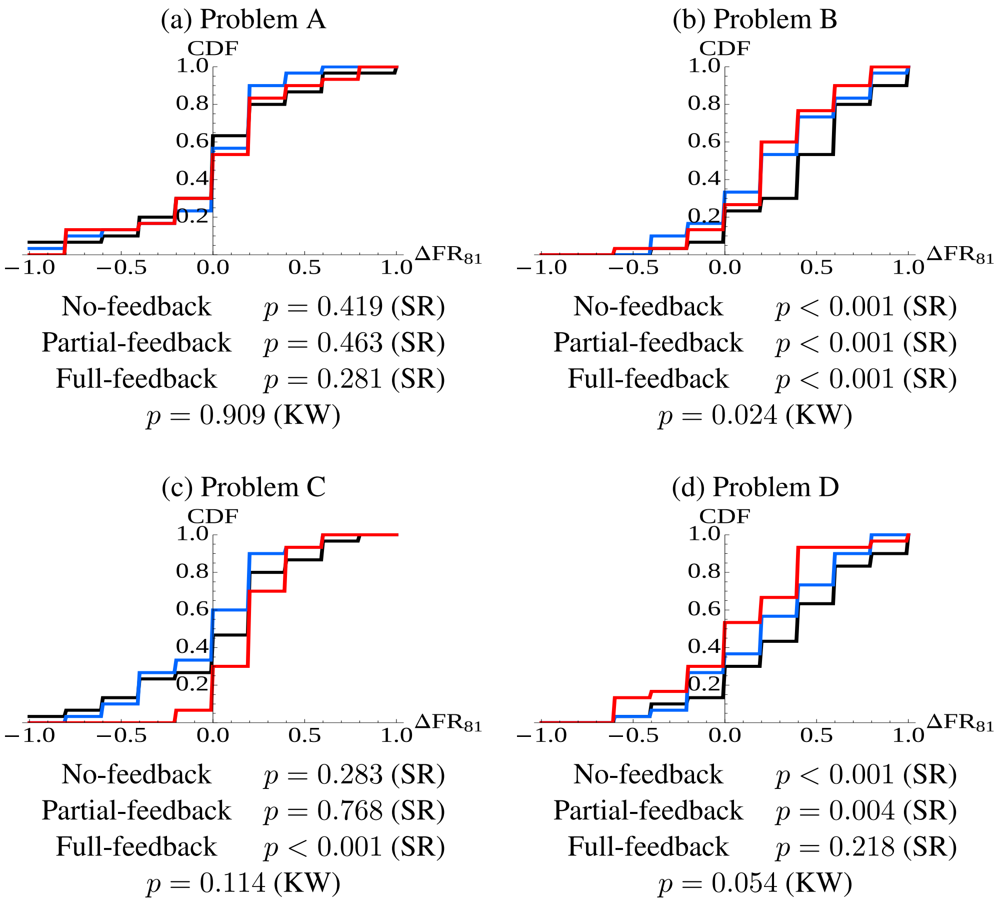

Figure 3: CDF of $\Delta FR^i_{8,1}$, the change in the (normalized) number of periods in which a subject chose the better option in the last 5 periods (periods 36-40, $b = 8$) and the first 5 periods (periods 1-5, $b = 1$) for Problem A (panel a), B (panel b), C (panel c) and D (panel d). Black: No-feedback, Red: Full-feedback, Blue: Partial-feedback. P-values for within treatment test (one-tailed, signed-rank test (SR)) and across treatments test (Kruskal-Wallis (KW)) are reported.

stood in light of Observation 1 above. Because subjects have some how figured out the better option without much trying, they did not change their choices significantly in the subsequent periods.

For Problem B, on the other hand, we reject the null hypotheses that $\Delta FR_{2,1} = 0$ (panel (b) of Figure 2) in all the three feedback treatments. The same is true for $\Delta FR_{8,1}$ (panel (b) of Figure 3). Thus, more subjects learn to choose the better option more frequently relatively quickly (even in the first 10 periods) regardless of the feedback treatments. Indeed, we do not reject the null hypothesis that $\Delta FR_{2,1}$ are significantly different across three feedback treatments ($p = 0.151, KW$). Comparing panel (b) of Figure 2 and Figure 3, however, one can notice that distributions of $\Delta FR_{8,1}$ lie towards the right of those for $\Delta FR_{2,1}$, suggesting that more subjects indeed learn to choose the better option more frequently when given more time to learn.

For Problem D, $\Delta FR_{2,1}$ and $\Delta FR_{8,1}$ are both significantly larger than zero only in partial feedback treatments. For full-feedback treatments, both $\Delta FR_{2,1}$ and $\Delta FR_{8,1}$ are not significantly greater than zero. And, for no-feedback treatment, the latter is significantly greater than zero while the former is not. Although the reason for the absence of the significant evidence of learning under full-feedback treatment for this problem is not clear from our data, the significant difference of $\Delta FR_{2,1}$ between partial- and no-feedback treatments ($p = 0.082$, MW) suggests that, for this problem, receiving a feedback about own payoffs facilitated more subjects to learn to make the better choice more frequently from earlier rounds. However, there is no significant difference of $\Delta FR_{8,1}$ between no- and partial-feedback treatments ($p = 0.243$, MW). Thus, eventually, more subjects learn to choose the better option more frequently even without any feedback. Based on these results we can make the following observations.

Observation 2. *In Problem B and D, subjects learn to choose the option with a higher expected payoff even without any immediate feedback regarding the payoffs.*

3.3 Meaningful learning and learning transfer

Let us now investigate whether "meaningful learning" or "learning transfer" is observed in our data, and whether differences in the payoff related feedback had a significant effect on the outcomes. We first compare the initial fraction of subjects who have chosen the better choice when they have faced a problem for the first time without any prior experience in a similar problem (period 1) versus with an experience in a similar problem (period 41). For example, for Problem A, we are comparing the fraction of subjects who have chosen the better option in period 1 of sessions with $A \to B$ and period 41 of sessions $B \to A$.

Table 2 summarizes these fractions for four problems and three feedback treatments. Let us first compare the fraction of subjects with an experience in another problem choosing the better option in period 41 when they face a new problem. Except for Problem B ($p < 0.001$, KW), there was no significant difference across three feedback treatments (p-values are 0.878, 0.504, and 0.249 for Problem A, C and D, respectively, based on KW test). This suggests, as if, the difference in the feedback in the first 40 periods did not have significant effect on their first choice facing a different problem. This insignificant difference is a reason for us to present the result by pooling the outcome in the three treatments in Table 2 in addition to listing the result for each treatment separately.

Now let us compare the fractions of correct choices between those with experience (period

Table 2: Fraction of subjects choosing the better option.

Problem A	Pooled (90)	No-fb(30)	Partial-fb (30)	Full-fb (30)	P-value (KW)
Period 1	67	20	25	22	0.337
Period 41	48	15	16	17	0.878
p-value (MW)	0.003	0.197	**0.014**	0.183	

Problem B	Pooled (90)	No-fb(30)	Partial-fb (30)	Full-fb (30)	P-value (KW)
Period 1	31	9	13	9	0.464
Period 41	39	21	8	10	$p < 0.001$
p-value (MW)	0.223	**0.002**	0.183	0.776	

Problem C	Pooled (90)	No-fb(30)	Partial-fb (30)	Full-fb (30)	P-value (KW)
Period 1	60	19	22	19	0.466
Period 41	64	23	22	19	0.504
p-value (MW)	0.522	0.260	0.992	0.993	

Problem D	Pooled (90)	No-fb(30)	Partial-fb (30)	Full-fb (30)	P-value (KW)
Period 1	22	8	5	9	0.672
Period 41	52	21	16	15	0.249
p-value (MW)	<0.001	<**0.001**	**0.003**	0.115	

41) and without experience (period 1). For problem C, these fractions are very similar both at the aggregate (pooling three treatments) and for looking at each treatment separately. Two-tailed Mann-Whitney tests (MW) do not reject the null hypothesis that there is no difference between subjects with and without experiences with another problem.

For problem A, however, the fraction of subjects choosing the better option is significantly lower for the experienced subjects (period 41) compared to the inexperienced subjects (period 1). This is particularly so for the partial feedback case. It seems as if the experience in facing and learning to choose the better option in Problem B made it more difficult for these subjects to initially pick the better option in Problem A. Although we can only speculate about the possible reasons for this negative effect of experience, the following seems to be plausible. Subjects who faced problem B were choosing between $[6; 1, 2, 3, 4]$ and $[6; 1, 1, 4, 4]$ for 40 periods. In these 40 periods, they have learned that latter option gives them a higher payoff on average. Note that two large members of the committee have the same number of votes (4) in the latter option. Remembering this fact, when they faced problem A and asked to choose between $[14; 5, 3, 7, 7]$ and $[14; 5, 4, 6, 7]$, they have chosen the former because two large members of the committee have the same number of votes in this option. We, however, need to investigate, possibly by means of mouse-tracking or eye-tracking devise, which information subjects are paying attention to before making their choices in order to test this hypothesis. And thus, we leave this for future research.

The positive and significant effect of the experience (meaningful learning) is observed in Problem B and D. In these problems, especially under no-feedback treatment, significantly more experienced subjects (who have faced Problem A or C for 40 periods) have chosen the better option from the beginning compared to inexperienced subjects. Note that, unlike between Problem A and B discussed above, there is no obvious similarities in their options between Problem C and D. Furthermore, it seems that the positive effect of the prior experience is stronger when subjects received less payoff-related feedback. Namely, the effect of experience is statistically significant only under no- ($p = 0.002$ and $p < 0.001$ for problem B and D) and partial-feedback treatments ($p = 0.183$ and $p = 0.003$ for Problem B and D), and not under full-feedback treatment ($p = 0.776$ and $p = 0.115$ for Problem B and D). P-values are based on two-tailed Mann-Whitney test.

We complement the analysis by comparing the fraction of the first 5 periods of each problem in which subjects chose the better option between two groups of subjects: those who faced the problem for the first time (thus in periods 1-5, FR_1^i) versus those who have experienced a different problem for 40 times before facing the problem (thus in periods 41-45, FR_9^j). For example, for Problem A, we compare FR_1^i for those in $A \to C$ sessions and FR_9^j for those in $C \to A$ sessions. If there is a significant evidence on meaningful learning, we should observe FR_9^j to be significantly greater than FR_1^i.

The results are shown in Figure 4 in which for each choice problem and feedback treatment, the distributions FR_1^i (black) and FR_9^j (gray) are shown. We test the null hypothesis that the median $FR_1^i = FR_9^j$ against the alternative $FR_1^i \neq FR_9^j$ using the Mann-Whitney test (two-tailed). P-values for this tests are also reported in the figure.

In most of the panels of Figure 4, we do not observe much differences between the two distributions. There are three exceptions. For Problem B, C and D under no-feedback treatment, we observe that the distribution for FR_9^j (shown in gray) lies on the right of that for FR_1^i (shown in black). The two-tailed Mann-Whitney test shows that the median FR_9^j is significantly different (in fact, higher) from FR_1^i ($p = 0.049$ for Problem B, $p = 0.052$ for Problem C, and $p < 0.001$ for Problem D). Under partial- or full-feedback treatments of the same problem, however, such significant differences were not observed.

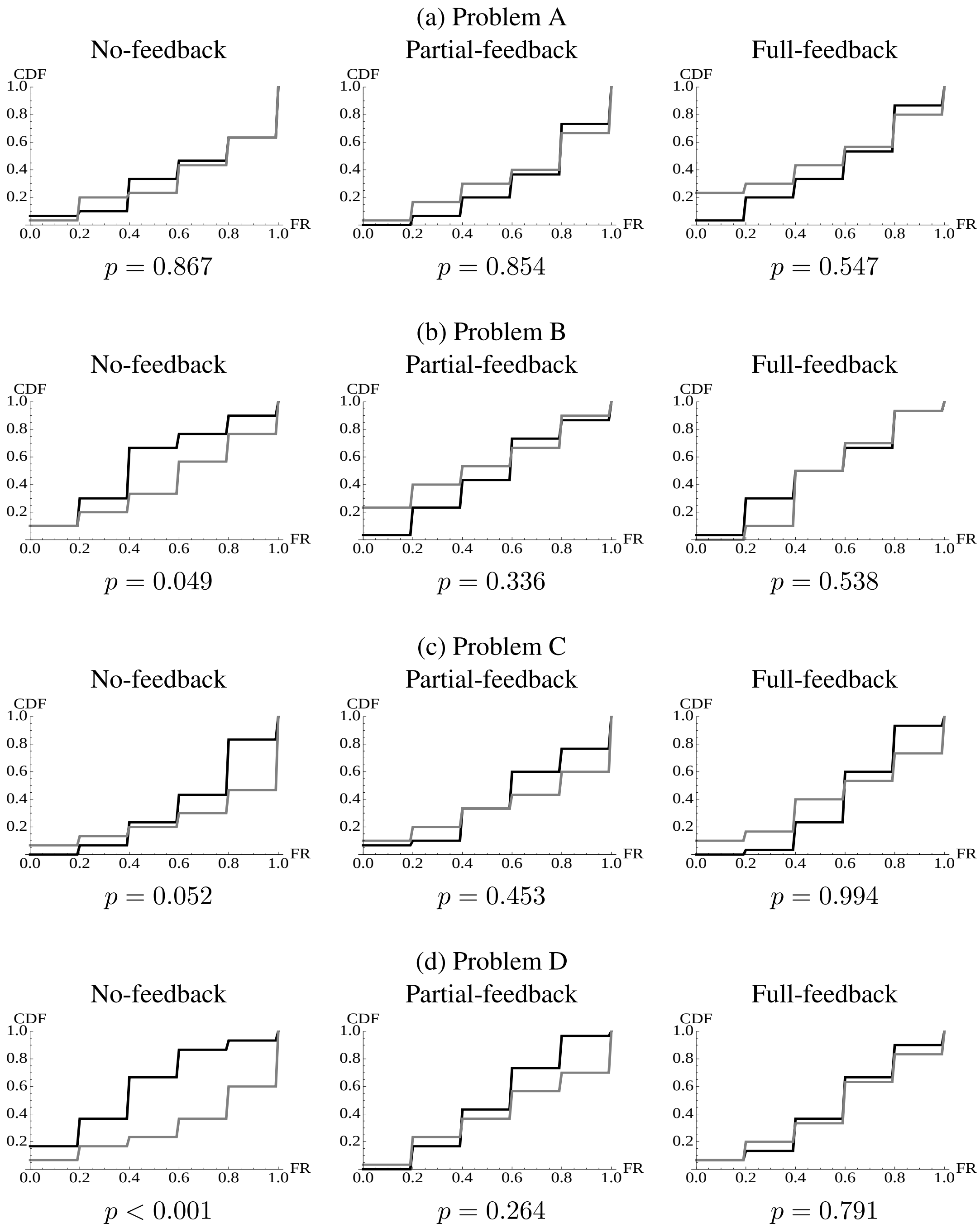

Figure 4: CDF of FR_1^i (black) and FR_9^j (gray) for four problems in three feedback treatments: No-feedback (left), Partial-feedback (middle), and Full-feedback (right). Note that FR_1^i is the fraction of the correct choice in the first 5 periods when subject i faced the choice problem for the first time, and FR_9^j is that when the subject j faced the choice problem after experiencing another choice problem. P-values for two-tailed Mann-Whitney test are reported.

These results suggest that, in two of our extended binary choice problems, withholding immediate payoff related feedbacks stimulated “meaningful learning” just as Rick and Weber (2010) found for dominance solvable games. Together with the result based on the comparison of choices in period 1 and 41, we can summarize the results of this subsection as follows.

Observation 3. *For Problem B and D, subjects who have experienced Problem A and C, respectively, before facing the problem under no-feedback treatment have chosen the option with a higher expected payoff much more frequently in the first five periods than those without such an experience.*

4 Conclusion

In this paper, we experimentally investigated whether subjects learn, from their limited experiences, something about the underlying relationship between their nominal voting weights and actual voting powers (expected payoffs) in weighted voting games. To simplify the learning problem as much as possible, we converted the problem into an extended two-armed bandit problem in which subjects are asked to choose between two committees that divide the resources among four (hypothetical) members.

We found that, in some problems, our subjects could learn to choose the committee that results in a higher expected payoff for themselves more frequently even without immediate feedback on payoffs. In addition, we have a statistically significant evidence of meaningful learning or learning transfer between two choice problems only in the treatment without immediate feedback. Our results replicate, at least in some of the choice problem we have experimented with, the results obtained by Rick and Weber (2010) regarding the learning in the absence of immediate payoff related feedbacks.

In our experiment, the payoffs were generated based on the assumption made by Deegan and Packel (1978) in creating their index of voting power. As is well known, there are other indices on voting power such as Banzhaf (1965) and Shapley and Shubik (1954). It will be an interesting future research to experimentally investigate which of these indices are more intuitive for subjects to learn.

References

ALESKEROV, F., A. BELIANIN, AND K. POGORELSKIY (2009): "Power and preferences: an experimental approach," Available at SSRN: http://ssrn.com/abstract=1574777.

ARIFOVIC, J. AND J. LEDYARD (2012): "Individual evolutionary learning, other-regarding preferences, and the voluntary contributions mechanism," *Journal of Public Economics*, 96, 808–823.

ARIFOVIC, J., R. D. MCKELVEY, AND S. PEVNITSKAYA (2006): "An initial implementation of the Turing tournament to learning in two person games," *Games and Economic Behavior*, 57, 93–122.

BANKS, J., M. OLSON, AND D. PORTER (1997): "An experimental analysis of the bandit problem," *Economic Theory*, 10, 55–77.

BANZHAF, J. F. (1965): "Weighted voting doesn't work: a mathematical analysis," *Rutgers Law Review*, 19, 317–343.

CAMERER, C. AND T.-H. HO (1999): "Experience-weighted attraction learning in normal form games," *Econometrica*, 67, 827–874.

CARPENTER, J., M. GRAHAM, AND J. WOLF (2013): "Cognitive ability and strategic sophistication," *Games and Economic Behavior*, 80, 115–130.

CASON, T. N. AND D. FRIEDMAN (1997): "Price formation in single call markets," *Econometrica*, 65, 311–345.

CHEUNG, Y.-W. AND D. FRIEDMAN (1997): "Individual learning in normal form games: some laboratory results," *Games and Economic Behavior*, 19, 46–76.

COOPER, D. J. AND J. H. KAGEL (2003): "Lessons learned: generalized learning across games," *American Economic Review, Papers and Proceedings*, 93, 202–207.

——— (2008): "Learning and transfer in signaling games," *Economic Theory*, 34, 415–439.

DEEGAN, J. AND E. PACKEL (1978): "A new index of power for simple n-person games," *International Journal of Game Thoery*, 7, 113–123.

DROUVELIS, M., M. MONTERO, AND M. SEFTON (2010): "The paradox of new members: strategic foundations and experimental evidence," *Games and Economic Behavior*, 69, 274–292.

DUFWENBERG, M., R. SUNDARAM, AND D. J. BUTLER (2010): "Ep*i*ph*a*ny *i*n t*h*e G*a*me *o*f 2*1*," *Journal of Economic Behavior and Organization*, 75, 132–143.

EREV, I., E. ERT, AND A. E. ROTH (2010): "A choice prediction competition for market entry games: an introduction," *Games*, 2, 117–136.

EREV, I. AND A. E. ROTH (1998): "Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria," *American Economic Review*, 88, 848–81.

ESPOSITO, G., E. GUERCI, X. LU, N. HANAKI, AND N. WATANABE (2012): "An experimental study on "meaningful learning" in weighted voting games," Mimeo, Aix-Marseille University.

FELSENTHAL, D. S. AND M. MACHOVER (1998): *The Measurement of Voting Power : Theory and Practice, Problems and Paradoxes*, London: Edward Elgar.

FISCHBACHER, U. (2007): "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10, 171–178.

FRÉCHETTE, G. R., J. H. KAGEL, AND M. MORELLI (2005a): "Gamson's law versus non-cooperative bargaining theory," *Games and Economic Behavior*, 51, 365–390.

——— (2005b): "Nominal bargaining power, selection protocol, and discounting in legislative bargaining," *Journal of Public Economics*, 89, 1497–1517.

GNEEZY, U., A. RUSTICHINI, AND A. VOSTROKNUTOV (2010): "Experience and insight in the Race game," *Journal of Economic Behavior and Organization*, 144–155.

GUERCI, E., N. HANAKI, N. WATANABE, X. LU, AND G. ESPOSITO (2014): "A methodological note on a weighted voting experiment," *Social Choice and Welfare*, 43, 827–850.

HANAKI, N., R. SETHI, I. EREV, AND A. PETERHANSL (2005): "Learning strategy," *Journal of Economic Behavior and Organization*, 56, 523–542.

HART, S. (2005): "Adaptive heuristics," *Econometrica*, 73, 1401–1430.

HARUVY, E. AND D. O. STAHL (2009): "Learning transference between dissimilar symmetric normal-form games," Mimeo, University of Texas at Dallas.

——— (2012): "Between-game rule learning in dissimilar symmetric normal-form games," *Games and Economic Behavior*, 74, 208–221.

HO, T.-H., C. CAMERER, AND K. WEIGELT (1998): "Iterated dominance and iterated best response in experimental "p-beauty contests"," *American Economic Review*, 88, 947–969.

HU, Y., Y. KAYABA, AND M. SHUM (2013): "Nonparametric learning rules from bandit experiments: The eyes have it!" *Games and Economic Behavior*, 81, 215–231.

IOANNOU, C. A. AND J. ROMERO (2014): "A generalized approach to belief learning in repeated games," *Games and Economic Behavior*, 87, 178–203.

KAGEL, J. H., H. SUNG, AND E. WINTER (2010): "Veto power in committees: an experimental study," *Experimental Economics*, 13, 167–188.

MARCHIORI, D. AND M. WARGLIEN (2008): "Predicting human interactive learning by regret-driven neural networks," *Science*, 319, 1111–1113.

MEYER, R. J. AND Y. SHI (1995): "Sequential choice under ambiguity: intuitive solutions to the armed-bandit problem," *Management Science*, 41, 817–834.

MONTERO, M., M. SEFTON, AND P. ZHANG (2008): "Enlargement and the balance of power: an experimental study," *Social Choice and Welfare*, 30, 69–87.

NEUGEBAUER, T., J. PEROTE, U. SCHMIDT, AND M. LOOS (2009): "Selfish-baised conditional cooperation: on the decline of contributions in repeated public goods experiments," *Journal of Economic Psychology*, 30, 52–60.

RAVEN, J. (2008): "General introduction and overview: the raven progressive matrices tests: their theoretical basis and measurement model," in *Uses and abuses of intelligence*, ed. by John and J. Raven, Edinburgh, Scotland: Competency Motivation Project, chap. 1, 17–68.

RICK, S. AND R. A. WEBER (2010): "Meaningful learning and transfer of learning in games played repeatedly without feedback," *Games and Economic Behavior*, 68, 716–730.

SHAPLEY, L. S. AND M. SHUBIK (1954): "A method for evaluating the distribution of power in a committee system," *American Political Science Review*, 48, 787–792.

STAHL, D. O. (1996): "Boundedly rational rule learning in a guessing game," *Games and Economic Behavior*, 16, 303–330.

Appendix

A Instructions

You will be repeatedly asked to make a simple choice between two options.

Imagine you need to represent your interests in a voting committee. This committee decides how to divide 120 points among its members. The committee has three other members, and each member has a predetermined number of votes, which may be different from one another. A decision will be made in this committee only when a proposal receives the pre-determined required number of votes. This required number of votes will be indicated to you. If more than one proposals are made in the committee, each member cannot vote for multiple proposals by dividing the votes she or he has. A member can vote for only one proposal, and when s/he votes, all the votes s/he has will be cast for the proposal.

The choice you are asked to make is, between two possible committees, which committee you prefer to belong to. The number of votes each of the four member of the committee has (including you) as well as the required number of votes for a proposal to be approved will be shown to you. The number of votes you have will always be indicated with the label YOU.

A.1 Full-feedback treatment

There will be total of 60 rounds. In each round, you have 30 seconds to make your choice between the two committees. If you do not make any choice during the 30 second in one round, you will obtain zero point for the round. Once choice is made, for each round of choice, the chosen committee will automatically allocate 120 points among the four members. The outcomes may vary from one period to another, but they are based on a theory of decision making in a committee. Once allocation is made, you will immediately see the resulting allocation. At the end of the experiment, you will be paid according to your total earnings during the 60 rounds, with an exchange rate of 1 Yen = 1 point.

If you have any question, please raise your hand.

A.2 No-feedback treatment

There will be total of 60 rounds. In each round, you have 30 seconds to make your choice between the two committees. If you do not make any choice during the 30 second in one round, you will obtain zero point for the round. For each round of choice, the chosen committee will automatically allocate 120 points between the four members. The outcomes may vary from one period to another, but they are based on a theory of decision making in a committee. You will not see the resulting allocation after each round. You will, however, will be informed about the total point you have obtained during the 60 rounds at the end of the experiment. At the end of the experiment, you will be paid according to your total earnings during the 60 rounds, with an exchange rate of 1 Yen = 1 point.

If you have any question, please raise your hand.

A.3 Partial-feedback treatment

There will be total of 60 rounds. In each round, you have 30 seconds to make your choice between the two committees. If you do not make any choice during the 30 second in one round, you will obtain zero point for the round. For each round of choice, the chosen committee will automatically allocate 120 points among the four members. The outcomes may vary from one period to another, but they are based on a theory of decision making in a committee. Once allocation is made, you will immediately see the number of points which has been allocated to you. The allocations for the rest of the members will not be notified to you. At the end of the experiment, you will be paid according to your total score, with an exchange rate of 1 Yen = 1 point.

If you have any question, please raise your hand.