When do conflicting parties share political power? An experimental study

Marco Battaglini
Department of Economics, 001 Fisher Hall,
Princeton University, Princeton, NJ, USA,
mbattagl@princeton.edu

Lydia Mechtenberg
Department of Economics, Von-Melle-Park 5,
University of Hamburg, D-20146 Hamburg,
lydia.mechtenberg@wiso.uni-hamburg.de

March 14, 2013

Abstract

When do privileged groups share their political power with other groups that have conflicting interests? We conducted a laboratory experiment to address this question. There are two groups of participants, the "yellows" and the "blues". The yellows collectively choose between the proportional voting rule and the simple-majorty rule, and all subjects participate in elections. In two control treatments, the blues can use a costly punishment option. The yellow group shares power voluntarily to only a small extent but is more inclined to do so if under threat of punishment. The blue group conditions punishment both on the voting rule and the electoral outcome. Under the proportional rule, they are more inclined to punish an unfavorable outcome.

Keywords: Behavioral Economics, Constitutional Design, Elections, Endogeneous Institutions, Experiment, Punishment, Voting Rule

JEL-Codes: C92, D02, D72, D79

1 Introduction

Do groups with constitutional authority share their political power with other groups that have conflicting interests? Do they share power voluntarily or only under threat? Recently, the literature in political economics has begun to address these important questions in a small number of ingenious papers. They find the following: On the one hand, a need for commitment power can motivate the elite to extend the suffrage (Acemoglu and Robinson 2000, Lizzeri and Persico 2004). Credible revolutionary threats from the disenfranchised have a similar effect (Acemoglu and Robinson 2001). On the other hand, the design of the electoral system - and thereby the distribution of political power within a given electorate - seems to be dictated by the self-interest of the elite that holds constitutional power (Aghion, Alesina and Trebbi 2004, Trebbi, Aghion and Alesina 2008). In particular, political majorities tend to choose voting rules that disempower political minorities (Trebbi, Aghion and Alesina 2008).[1] These insights have been obtained within the bounds of game-theoretic predictions under standard homo-oeconomicus assumptions. Our contribution is to complement this approach by a more comprehensive perspective that allows for behavioral determinants of power sharing. We know from experimental studies of bargaining games, like the dictator and ultimatum game, that people often deviate quite extensively and consistently from game-theoretic predictions if they are in sharing situations. Thus, the question suggests itself whether they deviate similarly in bargaining games in which groups can share political power. If so, existing theories of how political power is distributed should be complemented by a behavioral perspective.[2]

We conducted an experiment to investigate this issue. We study the collective choice of a voting rule by a group endowed with constitutional power when the entire population collectively decides between alternative

[1] Moreover, state officials in the U.S. often use their authority to enforce laws selectively, thereby reducing turnout in demographic minorities. For instance, they more often ask non-white than white voters to show their ID at the polling station. See, e.g., Bassi, Morton and Trounstine (2008). For further research on which electoral systems in the U.S. grant more political power to minorities, see, e.g., Gerber, Morton and Rietz (1998) and Morton and Rietz (2008).

[2] For a careful description of how the concept of power is defined in political science and closely related fields, see Sell et al. (2004). We use the concept "political power" to denote the probability with which a collective decision, e.g., an election, about the allocation of resources favors one's own group.

payoff allocations during an election. In particular, we investigate how one group selects the voting rule when another group with opposed interests has varying degrees of retaliatory power. In doing so, we firstly focus on the question whether behavioral determinants affect the extent of power sharing between groups. Secondly, we study how the chosen allocation of political power affects the acceptability of the resulting payoff allocation.

To our best knowledge there are as yet no expermental studies of collective power sharing. Thus, we designed a new and simple game that captures the essential features we are interested in. In the baseline treatment of our experiment (henceforth T1), we consider an environment in which the players are divided in two groups: the yellows, the majority; and the blues, the minority. Players have to collectively choose between one of two options: Y or B. Preferences are perfectly aligned within a group, and perfectly opposed across groups: the yellows prefer Y, the blues B. The game has two stages. In the first stage, the yellows deliberate on how the final outcome will be chosen. Two options are available: a majoritarian rule, that guarantees that the yellows can exclude the blues from the decision making[3]; or a proportional rule, that guarantees that for each group, the probability of choosing their preferred allocation is proportional to their respective size in the population. In the second stage the collective choice is taken according to the chosen rule. We then consider two additional treatments. In the second treatment (T2), the members of the minority can choose to "voice", i.e. directly punish the prospective majority at some individual cost. In the third treatment (T3), the members of the minority can choose to "exit," i.e., to stop playing the game after they learn the decision rule chosen by the yellows, so denying the majority the ability to benefit from the decision rule. All treatments have a unique Perfect Bayesian equilibrium in which the yellows choose the majoritarian rule in T1, the voters vote for their preferred policies in the second stage, and (in T2 and T3) the minority never finds it optimal to punish (neither at the final stage nor at the interim stage).

We find three main results. First, only a small (but significantly) positive fraction of members of the majority votes for a decision rule that favors the minority in the baseline treatment. The players are heterogeneous: 15% of the players vote altruistically more than 25% of the time (i.e., for

[3]Morton and Rietz (2008) show in an experimental study that a majority successfully disempowers the minority under a majoritarian voting rule in a much more complicated setting than ours. Thus, we feel safe to choose the simple majority rule as the disempowerment option in our experiment.

the proportional rule); 60% of the players, instead, vote opportunistically more than 95% of the time. Second, despite the fact that retaliation is never optimal in equilibrium, the possibility of punishment by the minority induces a very significant increase in fraction of majority members voting for a proportional rule. Indeed, it also induces the majority members to vote against their interest in the second stage when the majority rule is chosen. Third, the shift in the behavior of the majority is rational: punishments are often observed. The choice of the voting rule, however, changes the players' expectations and the way the minority punishes the majority: an unfavorable outcome is punished less under the majoritarian rule than under the proportional rule.

A conclusion that is suggested by the experiment is that the choice of the decision rule is heavily affected both by rational considerations and by behavioral responses that cannot be easily disentangled. The majority clearly shows opportunistic behavior (being overwhelmingly opportunistic in T1 and more "cooperative" in T2 and T3). Yet the majority is not responding to equilibrium expectations on the minority's behavior, since punishments are never optimal in equilibrium. Cooperation, therefore, seems to arise from the willingness of the minority to punish selfish behavior; the majority "rationally" reacts to this behavioral expectation.

Our paper contributes to three hitherto separated strands of the literature. The first is the literature on constitutional choice.[4] This literature naturally presupposes fully self-interested behavior consistent with game-theoretic predictions. Moreover, it assumes that all preferences are over absolute payoffs only, and that these preferences are not affected by institutional change. However, there are reasons why this standard approach should be complemented by behavioral research: People may be concerned with relative payoffs, or their utility from a given payoff might be reference-dependent. Moreover, people may have outcome-independent preferences over procedures, too, i.e., they may consider one procedure to be "fairer" than another. In addition, monetary outcomes might derive their acceptability from the way in which they were collectively chosen. Both the perceived intrinsic fairness of procedures and the way in which they affect the acceptability of resulting payoff allocations may have an influence on which procedures are chosen in the course of designing a constitution. Our experiment takes a first step toward including the behavioral perspective into the study of

[4]For an extensive survey on constitutional economics, see Voigt (1997, 2011).

endogenous constitutions by testing whether behavioral determinants affect constitutional choice in the lab.

Second, we contribute to the extensive experimental literature on simple bargaining games, like the dictator and ultimatum game. In the dictator game, a proposer - which can be an individual or, rarely, a group - has full authority to allocate a given amount of money between himself and a passive responder. Most experimental studies find that the average proposer in the games between individuals exhibits a considerably high willingness to share his endowment. This willingness significantly increases in the ultimatum game in which the responder has a costly punishment option. In both games, sharing is entirely driven by behavioral determinants; under standard assumptions, no money would be transferred to the responder in any of the two games. In the analoguous group bargaining games, proposer groups are in general closer to the game-theoretic predictions than individuals[5] but still do not match equilibrium predictions very well (Robert and Carnevale 1997, Bornstein and Yaniv 1998, Elbittar, Gomberg and Sour 2011, Luhan, Kocher and Sutter 2009). In the current study, we experimentally implement games that are similar in spirit to the dictator and ultimatum game between groups but differ in two important respects: First, payoffs cannot be shared directly in our experiment. Instead, (expected) payoffs can be shared only indirectly, by sharing power. Second, there is no "fair" ex-post allocation of payoffs: One group will always get considerably less than the other in the end. We investigate whether the well-known experimental results from the individual dictator and ultimatum game carry over "in spirit" to our games of power sharing between groups, i.e., whether groups share power contrary to equilibrium predictions. Importantly, we are not investigating *which* behavioral determinants lead to power sharing, but only *whether* groups share power when this would not be predicted under standard assumptions.

The third strand of the literature to which our paper relates is the new experimental literature on endogenous institutions. This literature predominantly looks at social dilemmata, in particular public-good games.[6] In these games, players have similar preferences: All prefer the cooperative outcome over the Nash equilibrium. Therefore, subjects vote for institutions that successfully foster cooperation. By contrast, we focus on constitutional choice

[5]One much-discussed excpetion to these findings is reported by Cason and Mui (1997).

[6]See Tyran and Feld (2006), Ertan, Page and Putterman (2009), and Sutter, Haigner and Kocher (2010), among others.

in a political-economy setting when players' preferences are opposed between groups. In our setting, no preference ranking over outcomes is common to everyone. Importantly, the majority group in our experiment is always already an interested party when selecting the voting rule. Such situations occur in reality when a collective cannot design its institutions behind the veil of ignorance and when specific classes have no access to constitutional power. As claimed by Aghion, Alesina and Trebbi (2004), any "political economy of constitutional design" must account for these two circumstances when providing a positive analysis of endogenous political institutions. Moreover, we contribute to this literature by studying how the minority reacts both to the degree of power sharing established by the endogenous voting rule and the electoral outcomes. A considerable amount of evidence provided by psychologists, sociologists and - more recently - economists supports the conjecture that procedures differ in both their perceived intrinsic fairness and the way in which they affect the acceptability of resulting payoff allocations. Within the economics literature, procedural fairness concerns have been documented by Frey and Stutzer (2004), Bolton, Brandts and Ockenfels (2005), Aldashev, Kirchsteiger and Sebald (2010) and Chlaß, Güth and Miettinen (2010), among others.[7] Generally, unbiased random procedures are considered to be "fair". To speak with Bolton, Brandts and Ockenfels (2005), "procedures deemed fair are typically those that create a 'level playing field', a place where the participants have equal opportunity even if the resulting allocation is not equal." (p. 1054) Thus, procedural fairness concerns may prompt the minority to punish the majority for disempowering them. Regarding the way in which procedures affect the acceptability of payoff allocations, Bolton, Brandts and Ockenfels (2005) find that subjects reject unfair offers more often when they resulted from biased procedures than when they resulted from unbiased procedures. Thus, "fair" (exogenous) institutions seem to make unfair outcomes more acceptable. This effect may interact in subtle ways with economic incentives. If fair procedures make unfair outcomes more acceptable, then the minority might "forgive" unfavorable electoral outcomes more easily the more the electoral system empowers them. Our findings, however, clearly contradict this conjecture. They are more in line with a somewhat contrary effect of procedures documented by Dal Bo, Foster and Putterman (2010) and Dickson, Gordon and Huber (2009). They conducted experiments in the context of public-good provision

[7] See also Shor (2007) and Mertins (2008).

to investigate how institutions legitimize certain kinds of behavior when they are endogenously chosen. In the experiment of Dal Bo, Foster and Putterman (2010), cooperation-enforcement regimes are more effective when chosen endogenously; and in the experiment of Dickson, Gordon and Huber (2009), imperfect monitoring regimes lower cooperation more when they are chosen by the monitor himself. Thus, in both experiments, endogenous procedures make the behavior that they incentivise more acceptable, too. A related effect arises in our political-economy context. To see this, consider a situation in which the majority refrains from sharing power. The allocation of expected payoffs between the two groups will be rather unequal, compared to when power is shared. Thus, the extent to which power is transferred to the minority affects everyone's expectations concerning the final payoff allocation. Individuals in the minority might therefore feel less disappointed about low payoffs when they already expected little due to their lack of power. Thus, the voting rule that favors the majority makes payoff allocations typically induced by it more acceptable in the eyes of the minority.

The remainder of this paper is structured as follows: In Section 2 and 3, we present the experimental design and procedure, respectively. Section 4 provides our hypotheses and Section 5 our results. We conclude and discuss our findings in Section 6. Appendix A contains all figures and tables and Appendix B a sample of the instructions handed out to our experimental subjects.

2 Experimental Design

Our experiment implements three different games of collective decision making. In all three games one of the two groups has exclusive constitutional power to choose the decision rule. In the next two sections we describe the game in greater details.

2.1 The Benchmark game

There are $N = 2n$ players. Nature assigns a color - yellow or blue - to each player. There are two possibilities. In the first, players extract (with replacement) their type from an urn in which there are $n+1$ yellows and $n-1$ blues; in the second they extract (with replacement) their type from an urn in which there are $n+1$ blues and $n-1$ yellows. The players observe their

type but they do not know from which urn it has been extracted. However, they know that the yellows are expected to be the majority since the first urn is $x \geq 1$ times more likely than the second.

After observing their type, the players play a game with two stages:

- In the constitutional stage the yellows - and only the yellows - privately choose between two voting rules, rule 1 and rule 2. One yellow's choice is randomly drawn and implemented; and all yellows are equally likely to be drawn. Then, all players are informed about which rule the yellows have implemented.

- In the voting stage all players participate in an election under the voting rule that was previously chosen. They privately vote for either alternative Y or alternative B. Alternative Y assigns a high payoff $\alpha + \delta$, with $\delta > 0$, to each yellow and a low payoff α to each blue player; and alternative B does the reverse.

Table 1 below depicts an example for this payoff structure. The outcome of the election is determined by the voting rule as follows: Under rule 1, one of the N players is randomly drawn and his decision implemented. All N players are equally likely to be drawn. Under rule 2, the alternative with the highest number of votes is implemented. If both alternatives attracted the same number of votes, a fair coin is flipped to decide which alternative to implement. Thus, rule 1 is the random dictator rule and rule 2 the simple majority rule. In the end, all players learn the outcome of the election and earn their resulting payoff.

The goal of the game described above (and of the variations described below) is to measure the extent to which a privileged group (the majority) shares power with an oppositional group (the minority). Before discussing the interpretation of the game in details, it is useful to characterize the associated equilibrium behavior. The Perfect Bayesian Equilibrium (PBE) of this game can be easily characterized by backward induction.[8] On the voting stage, yellows vote for Y and blues vote for B in all equibria. Anticipating this on the constitutional stage, the yellows choose the rule that makes a

[8]Since players do not know which color group is in the majority but get some information about it by learning their own type, the equilibrium concept that applies here is Perfect Bayesian equilibrium. As standard in voting games, we focus on equilibria with weakly undominated strategies.

final outcome of Y most likely. With probability $\frac{x}{x+1} \geq \frac{1}{2}$, the yellows are the larger group (of size $n+1$). Thus, a yellow assigns posterior probability

$$p_y = \frac{x\,(n+1)}{x\,(n+1)+n-1} > \frac{1}{2}$$

to the event that the yellows are the majority and probability $p_b = 1 - p_y$ to the opposite event. His expected payoff from selecting the random-dictator rule amounts to

$$\pi_y^1 = \alpha + \left(p_y \frac{n+1}{2n} + (1-p_y)\,\frac{n-1}{2n}\right)\delta, \tag{1}$$

while his expected payoff from the simple-majority rule is

$$\pi_y^2 = \alpha + p_y\delta.$$

Since $\pi_y^2 > \pi_y^1$, the yellows choose the simple majority rule. We conclude that in the unique PBE the simple majority rule is chosen with probability one and the outcome of the election is Y with probability $\frac{x}{1+x}$ and B with probability $\frac{1}{1+x}$ (depending on which group is majoritarian in the second stage).

To interpret the game we can now make two comments. First, at the end of the game the allocation implies that one group receives δ more than the other no matter what the voting rule is. In the unique equilibrium, however, the voting rule has an important impact on the expected allocation. In a majoritarian allocation the yellows always impose their preferred choice with probability one. In the random dictator allocation, on the contrary, the probability that the minority chooses the policy is proportional to the number of yellows in the population. The second rule, therefore, is closer to a proportional rule. Although we model the decision rules in stylized form, they capture the essence of the problem we are interested in.

Second, the fact that we are assuming that the distribution of the population is random and that the players do not know it is not playing an important role in the game since x is large and the equilibrium coincides with the equilibrium we would have if the yellows were always the majority (i.e., x is infinity). We have adopted this design only to introduce enough uncertainty to keep the choice interesting.[9]

[9]If the yellows were the majority with certainty, the blues might play weakly dominated strategies under the majority rule in equilibrium. We exclude this by introducing a small probability that the blues are the majority.

2.2 Treatments with punishments

To explore the determinants of behavior in the previous game, we consider two additional treatments in which the minority can punish the majority. What varies between the three games is the strategy space of the group without constitutional power: It has no bargaining power whatsoever in the benchmark game. In the ex post punishment game, it can diminish the payoff of everyone in the privileged group when the election is over. In the interim punishment game, it can again reduce the payoffs of everyone in the privileged group, but prior to the election and with the additional consequence of ending the game immediately. Thus, ex post punishment can be conditioned both on the voting rule chosen by the privileged group and the election outcome. By contrast, interim punishment can only be conditioned on the voting rule. Thus, all three games help us to measure the extent to which a privileged group shares power with an oppositional group. These differences between the three games allow us to test whether - and how - outcome-related preferences interact with procedures.

The ex-post punishment game The ex-post punishment game differs only in one respect from the benchmark: It adds a punishing stage at the very end. On the punishment stage, each blue player privately chooses between accepting ("keep") and changing ("change") the final outcome of the election. One blue player in the group is randomly drawn, and his choice is implemented. (All blue players are equally likely to be selected.) If the selected choice is "keep", then the payoffs of this round remain unchanged. If it is "change", each yellow player loses $\delta+2$ points, the blue player whose choice was implemented loses 1 point, and the payoffs of the other blue player(s) remain unchanged. Again, all players are informed about the outcome of this stage.

It is easy to see that the Perfect Bayesian equilibrium with fully selfish payoff maximizers that has been characterized can be extended to this game: Since it becomes costly to choose "change" when one's choice is drawn, no blue should ever use the punishment option, and the final equilibrium outcome is the same as in the benchmark game.

The interim punishment game The interim punishment game moves the punishment stage up the game tree: Directly after the voting rule has been chosen and revealed to all players, the blues have to decide between

"continue" and "exit". Again, one blue player is randomly drawn and his choice implemented. If it is "continue", the game proceeds as in the benchmark game. But if it is "exit", the game ends immediately, without the voting stage being reached, and payoffs are as follows: Each yellow in the group earns $\alpha - 2$, the blue player whose choice has been implemented earns $\alpha - 1$, and the other blue player(s) earn α. Thus, the payoffs after "exit" are the same as those after "change" in the ex-post punishment game when the electoral outcome has been Y. The same argument as before applies: No blue should ever use the costly punishment option, and the final equilibrium outcome should be the same as in the two other games.

3 The experiment

3.1 Experimental Procedure

We conducted three experimental treatments, one for each game. The baseline treatment T1 implemented the benchmark game, the punishment treatment T2 the ex-post punishment game, and the exit treatment T3 the interim punishment game. The free software z-tree was used to computerize all three games. Each treatment had six sessions. In each session, 18 subjects participated anonymously in one of the three games. In each session, the game at hand was repeated over 30 rounds; and only one round per session was randomly drawn to be played out at the very end. At the beginning of each round, the computer randomly assigned the participants to three interaction groups of six: thus $n = 3$. Group membership changed randomly over different rounds. Moreover, the computer randomly determined whether the group consisted of $n + 1 = 4$ yellow and $n - 1 = 2$ blue players or instead of two yellows and four blues. Roles were assigned randomly, with random re-matching in every round. The matching group comprised all 18 participants in the session. In each round, it was nine times more likely that the yellows were in the majority; and subjects knew this. Thus, $x = 9$ and $p_y = \frac{18}{19}$. We chose a relatively high x because we wanted the yellows to be clearly perceived as the privileged group by all players. (The higher x is, the more the constitutional power of the yellows is of material consequence.) On the other hand, the small remaining uncertainty about the representation of one's role in the group guaranteed that even under the simple majority rule and fully self-interested rational voting behavior, a blue player would

see a small chance to get her favorite alternative and thus take the election seriously. The difference δ between the high and the low payoff was set to 5 in all sessions. In each session, we varied α randomly across rounds. It took the values 10, 15 and 20. We intentionally chose this payoff structure for the elections to keep the experiment simple.

At the beginning of the sessions and after reading Instructions, subjects had to solve a quiz in order to prove that they understood the game. They were allowed to ask questions about the Instructions in private to the experimenters. Only after each subject knew the correct answers to all quiz tasks and had no questions any more did the experiment start. The subjects earned an average of approximately 18 US-Dollars, including a show-up fee of 5 Dollars. Each subject participated in no more than one session.

Two sessions of the baseline treatment were conducted at Princeton University, USA, and all others at the Technical University of Berlin, Germany. All participants were undergraduate students from a variety of different subject fields, with a natural focus on technical fields in the sessions run in Germany. They were recruited by university web pages and signs posted in the university buildings. At the end of each session, the participants answered a few demographic questions. Then, the computer randomly drew one round that was played out to the participants privately and in cash. Appendix A contains the Instructions that we handed to participants at the beginning of the session.

3.2 Hypotheses

The Benchmark Game Both procedural and outcome-related fairness concerns can induce a yellow player to share power with the blues voluntarily, i.e., to choose the proportional rule with a significant probability in the baseline treatment. Placing a vote for the proportional rule would be a way for a yellow player to share the expected payoff of his color group with the group of blue players.[10] Thus, we state our first hypothesis:

[10]Note that voting behavior, by contrast, is unlikely to be affected by fairness concerns: First, a yellow player could not reduce inequality by voting non-selfishly, since the difference between the high and the low payoff would still be 5. Moreover, if his non-selfish vote was effective, it would reverse the relative standings of yellows and blues and reduce the expected total payoff of the entire interaction group. Neither existing fairness theories nor any well-known experimental results predict such behavior.

H1. *In the baseline treatment, yellow players choose the proportional rule with positive probability.*

Note that we do *not* intent to separate procedural from other fairness concerns. Rather, our intention is to test whether a privileged group - a group with constitutional power - voluntarily shares its power with an opposing group even if doing so is likely to have material consequences.

The ex-post punishment game: punishment behavior Since the yellows have the opportunity to share power prior to the election, the blues might react differently to a given election outcome, depending on which voting rule the yellows implemented. There are four possible histories on which the blues can condition their punishment behavior, since any of the two rules can in principle be combined with any of the two outcomes. This considerations lead to the following hypothesis:

H2. *In the punishment treatment, blue players punish (i) more under the simple-majority rule than under the proportional rule, and (ii) more after Y than after B.*

For instance, inequity averse blues might choose to punish (i.e., "change") whenever Y is implemented, which implies the second part of the above hypothesis. Moreover, both inequity aversion and negative reciprocity are possible motivations to opt for "change" whenever the yellows implemented the simple majority rule, which implies the first part. Again, note that our intention here is not to test different theories of social preferences against each other. We aim at testing whether punishment incentives in bargaining games between individuals extend to games of collective decision making where groups instead of individuals are opposed to each other.

An especially interesting question concerns wether the blues might condition their reaction to the election outcome on the voting rule. Standard concepts of behavioral economics do not provide us with a unique prediction here. If, for instance, the electoral outcome is Y, then blues who exhibit negative reciprocity might want to punish the yellows more if they implemented the simple majority rule. The blues might ascribe unfriendlier intentions to the yellows if the latter refrained from sharing their power with them. However, the opposite might also be true: The implementation of the proportional rule might move the reference point of blue players more toward alternative B. Hence, if the electoral outcome is Y, blue players might experience stronger disappointment under the proportional rule than under the

simple majority rule, consequently feeling more inclined to punish. Similar arguments can be brought forward to show that no unique "behavioral" prediction can be derived regarding punishment after electoral outcome B. Thus, we have to consider all alternatives

H3(a). *In the ex-post punishment game, and conditional on getting alternative Y, blue players punish more under the simple-majority rule than under the proportional rule.*

H3(b). *In the ex-post punishment game, and conditional on getting alternative Y, blue players punish more under the proportional rule than under the simple-majority rule.*

H4(a). *In the ex-post punishment game, and conditional on getting alternative B, blue players punish more under the simple-majority rule than under the proportional rule.*

H4(b). *In the ex-post punishment game, and conditional on getting alternative B, blue players punish more under the proportional rule than under the simple-majority rule.*

The ex-post punishment game: Rule choice Given a specific punishment behavior of the blues, the question arises whether players learn to anticipate this behavior and best-respond accordingly when in the role of a yellow. If the expected probability of being punished under rule r is $q_Y(r)$ after alternative Y and $q_B(r)$ after alternative B, then a selfish yellow player weakly prefers alternative Y if and only if

$$\begin{aligned} 5 - q_Y(r)\,7 &\geq -q_B(r)\,7, \text{or} \\ q_Y(r) - q_B(r) &\leq \frac{5}{7}. \end{aligned}$$

Thus, if the difference between the two expected punishment probabilities becomes large enough, a selfish yellow player will try to make alternative B more likely to eschew punishment, either by choosing the proportional rule or by voting for B (or by both). Note that the difference in expected punishment probabilities under Y and B must exceed more than 70% in order to incentivise a deviation from the strategy to choose the simple-majority rule and vote for Y.

Suppose now that the blues condition their punishment only on the rule choice, not on the electoral outcome. In this case, $q_Y(r) = q_B(r)$, and no yellow will ever vote for B. Hence, if we observe a significant number of yellows frequently voting for B in the ex-post punishment game, then this will be evidence that the yellows implicitly believe punishment to be conditioned on electoral outcomes, too. In sum, we get the following two hypotheses:

H5. *The yellows select the proportional rule more often in the ex-post punishment game than in the benchmark game.*

H6. *The yellows vote for B more often in the ex-post punishment game than in the benchmark game.*

The interim punishment game Both procedural fairness concerns and inequity aversion can lead blue players to choose "exit" when the simple-majority rule has been implemented by the yellows. If the expected exit probability under the simple-majority rule exceeds the expected exit probability under the proportional rule by a sufficient amount, then a selfish yellow will prefer the proportional voting rule. In this case, we should observe a higher frequency of the proportional rule in the interim punishment game than in the benchmark game. Thus, if compared to the benchmark game, the interim punishment game serves as a robustness check of the ex-post punishment game.

However, there are also valuably insights to be gained from the comparison of the interim and the ex-post punishment games. Does the prospective minority (the blues) react less negatively to the choice of the simple-majority rule if the reaction has to be immediate, as required in the interim punishment game? How willing are the blues to end the game if it is not played according to their preferred rule, compared to their willingness to punish outcomes that they do not like? These questions are similar to the famous questions raised by Hirschman in his seminal book on Exit, Voice, and Loyalty: How willing are people to leave a system (an organization or state) if they do not like the way it is functioning, how willing are they to try and change it by staying inside and raising their "voice", and how willing are they to stay loyal without protest? And which of the two ways to protest, exit or voice, contributes more to changing the system? Prima facie, there are no clear and easy answers; hence, it might help to investigate these questions experimentally. In sum, we get the following three hypotheses:

H7. *In the interim punishment game, blue players choose "exit" more often under the simple-majority rule than under the proportional rule.*

H8. *There is no difference in the frequency with which the implementation of the voting rules makes the blue players choose "exit" in the interim and "change" in the ex-post punishment treatment.*

H9. *There is no difference in the frequency with which the yellow players choose the proportional rule in the interim and the ex-post punishment treatment.*

In the next section, we will report the results of our experiment and relate them to our hypotheses.

4 Results

We first report the results from the constitutional stage and the voting stage in all three games. We then present our results regarding the punishment behavior of the blues in the two punishment games. Hereafter, we will refer to treatments rather than games, i.e., the baseline treatment (benchmark game), punishment treatment (ex-post punishment game) and exit treatment (interim punishment game).

4.1 Do subjects share power voluntarily?

We first consider the constitutional stage in the baseline treatment. Do yellow players voluntarily share power with the blues by choosing the proportional rule? Table 3 displays standard tests of whether the proportion of yellows doing so is significant. As Table 3 shows:

Result 1. *On average 12% of the decisions made by yellow players are in favor of the proportional rule; this proportion is significanty positive at a 1% confidence level.*

To study how these choices are distributed over individual players, we separated our subjects into four categories, depending on how frequently they chose the proportional rule in the role of a yellow. The first category comprises all subjects who opt for the proportional rule in less than 5% of the rounds in which they are yellow. They behave as predicted by the equilibrium solution of the game, taking into account a small error rate. The

other three categories comprise players who choose the proportional rule in 5 to 15%, 15 to 25% and more than 25% of the time in which they are yellow, respectively. The first histogram in Figure 2 below depicts the distribution of subjects over these categories. As the first histogram in Figure 2 reveals:

Result 2. *In T1, most players are behaving in a selfish-rational way; there is however a significant minority who is not:*

- *More than 60% of players opt for the majoritarian rule in more than 95% of the rounds in which they are yellow.*
- *More than 15% of all players opt for the proportional rule more often than in 25% of the rounds in which they are yellow.*

We conclude that the players are heterogeneous: The willingness to share power seems to be much stronger in some players than in others.

To test whether the significant amount of yellow votes for the proportional rule is due to some lack of understanding of the game in the beginning of the sessions and vanishes when players learn by experience, we regressed the dummy for the choice of the voting rule (*rulechoice1*) on the round number (*period*). The dummy *rulechoice1* takes a higher value for a given player if he chose the simple-majority rule. Hence, a positive coefficient of the variable "period" indicates that a learning effect exists and reduces the deviation from selfishness on the constitutional stage. We ran this regression twice: Once without controlling for interdependency, and once with clustering on the highest possible level, i.e., on the session level, in order to control for any unspecified hierarchical interdependency in the data. Table 4 in Appendix A displays the regression results. Indeed, we find evidence of learning in the baseline treatment: it is highly significant, but small in magnitude. Taken together, these results suggest that a considerable number of players in the baseline treatment are willing to share power with the less privileged group. Hence, we find supporting evidence for our hypothesis H1.

4.2 Power-sharing and threats

We now compare behavior on the constitutional stage across treatments in order to identify the effects of anticipated punishment on the choice of the voting rule. Within a probit model, we regressed *rulechoice1* on the treatment dummies for the punishment and exit treatment. *T2* denotes the punishment treatment and *T3* the exit treatment. To control for unspecified

hierarchical dependencies of individual data in a given session, we clustered on the session level. Table 5a in Appendix A displays the results. We find that:

Result 3. *The average number of yellow players choosing the proportional rule increases significantly under the threat of punishment, both when punishment occurs ex post and when it occurs on the interim stage.*

In order to test whether this result is robust, we took the average of *rulechoice1*, obtaining one observation per session, and regressed it on the treatment dummies. The results are displayed in Table 5c below. Evidently, our findings are robust. This strongly supports our hypothesis H5. Moreover, the yellow players in the two Princeton sessions, indicated by the dummy D_C, chose the proportional rule less often than the yellow players in the four Berlin sessions; this difference is significant. This latter result indicates that subjects who live in a proportional system, like Germans, are more willing to share power with a political minority.

Again, we took a closer look at the distribution of constitutional choices over individual players.[11] The second and third histogram in Figure 2 refer to the punishment and exit treatment, respectively. As Figure 2 shows, the behavioural effect of the threat of punishment is very significant.

Result 4. *Both in T2 and T3, we find a strong effect of anticipated punishment on rule choice:*

- *The share of players who choose the majoritarian rule more often than in 95% of the rounds in which they are yellow is more than halved under the threat of punishment.*
- *The share of players who choose the proportional rule more often than in 25% of the rounds in which they are yellow is more than doubled.*

As Table 6 shows, these effects are highly significant. Moreover, the main effect seems to be on the selfish players who would have played the equilibrium strategy in the absence of punishment: Most of them seem to react to the risk of punishment mainly by choosing the proportional rule

[11] As Eckel, Johnson and Wilson (2002) claim, we think it is important to go beyond simple treament comparisons and account for the heterogeneity of subject samples. We did so without, however, eliciting information about individual characteristics or beliefs.

more than five times more often than they would have done without the risk of being punished.

As both Table 5b and 5c reveal: anticipation of punishment seems to have similar effects when the punishment looms directly after the rule choice and when it may hit at the end of the game. If at all, the difference in effects is in favor of the exit treatment; there, the number of players who choose the proportional rule in more than 25% of the time is largest. However, the difference in average rule choices becomes significant only if session dummies are included. Moreover, the similarity of the two control treatments is not restricted to the constitutional stage. Punishment behavior is also very similar.

Result 5. *Behavior on the constitutional stage does not differ much between the two control treatments.*

We used a Man-Whitney-U test to investigate whether the frequency of "change" differs from the frequency of "exit". As Table 11 in Appendix A shows, there is no significant difference. Taken together, these findings support our hypotheses H9 and H10.

4.3 Do subjects vote for giving others more than themselves if threatened?

In the previous subsection, we have seen that there are players willing to share power, and that even selfish players choose to do so if they anticipate being punished - directly or indirectly - for doing otherwise. Before turning to the actual punishment behavior of the blues, we want to consider the voting stage. Do yellows try to eschew punishment by voting for alternative B? Table 7 displays standard tests whether a significant number of players do so.

On average, the share of yellow votes for alternative B (B_Y_vote) is slightly below 2% and 5% in the baseline and the exit treatment, respectively; but much higher, namely 13,7%, in the punishment treatment. All three means are significantly different from zero.

We tested for treatment differences by running a probit regression. The dependent variable, D_Y_vote, is a dummy that encodes the yellows' choices on the voting stage. It takes negative values if the player is blue or if the voting stage was not reached, which happened in the exit treatment when "exit" was implemented. It takes a higher value for a given subject if that

subject placed a vote for alternative B. Excluding the negative values, we regressed *D_Y_vote* on the dummies for the punishment and the exit treatment, D_T5 and D_T6, where D_T5 is the dummy for T2 and D_T6 is the dummy for T3.[12] Again, we controlled for hierarchical within-session dependencies in the data by clustering on the session level. Table 8a in Appendix A shows the results. The coefficients of the treatment dummies are positive and highly significant. Thus, we have that:

Result 6. *The share of yellow votes for alternative B is significantly higher in both the punishment and exit treatment than in the baseline treatment.*

The (small) increase in the exit treatment could be due to reciprocity: If the blues refrain from ending the game, the yellows seem to reward them by voting for alternative B with a higher probability than the blues in the baseline treatment. By contrast, the large increase in yellow votes for B in the punishment treatment, compared to the baseline treatment, seems to be due to the effort of yellows to avoid punishment: If they are in the role of a yellow, subjects fear to be punished if the outcome of elections is Y and vote for B to avoid this.

As a robustness check, we took the average of *D_Y_vote,* obtaining one observation per session, and regressed it on the treatment dummies, naturally excluding all rounds in the exit treatment that ended after the constitutional stage due to an exit choice. Table 8b below displays the results.

The difference between the baseline and the punishment treatment remains highly significant. Thus, the effect of an ex-post punishment threat on the yellows' willingness to vote for B is robust. However, the difference between the baseline and the exit treatment becomes insignificant. Overall, we can conclude that the yellow subjects believe their blue opponents to condition punishment on the electoral outcome Y and not only on the choice of the simple-majority rule. The strong effect of anticipated punishment on the yellows' voting behavior is in line with our hypothesis H6.

4.4 (How) do subjects punish the privileged group?

In the previous section, we have seen that both threats, the threat of "exit" and the threat of "change", affect the behavior of the yellow players. On the

[12] Orginally, we planned six different treatments, but in the end, we conducted only three, so that T5 and T6 became T2 and T3. However, we did not correct the treatment names when doing the empirical analysis.

constitutional stage, the threat of "exit" in the exit treatment has a slightly larger effect than the threat of "change" in the punishment treatment. On the voting stage, the latter threat has a much larger effect. In sum, the fear of being punished affects the behavior of yellows more than positive reciprocity. A natural question now arising is whether the punishment and exit behavior of the blues is - at least roughly - in line with the implicit beliefs of the yellows. Do the blues indeed punish the yellows more under the simple-majority rule and after outcome Y? Moreover, only the punishment behavior of the blues can inform us about the way in which the voting rule affects acceptability of the electoral outcome: Does the benign choice of the proportional rule make the unfavorable outcome Y more acceptable for the blues? Or is outcome Y easier to accept under a voting rule that makes this outcome more likely, like the simple-majority rule?

Consider first the punishment treatment. To test whether the blues conditioned their decision between "change" and "continue" in the punishment treatment on the voting rule alone, on the outcome of the elections, or on a combination of both, we ran probit regressions of "change" (the dummy encoding the individual punishment decision of the blues) on "rule" (the dummy that encodes the implemented voting rule), and on "decide" (the dummy encoding the outcome of the elections). For a given blue subject, "change" took a higher value if the subject chose to punish. "Rule" takes a higher value if the simple-majority rule was implemented; and "decide" takes a higher value if Y was the outcome. We excluded yellow subjects from the regression and clustered on the session level. Table 9a reveals the results.

As Table 9a reveals, the blues did not condition their punishment decisions on the choice of the voting rule but rather on the electoral outcome: They punished the yellows when alternative Y resulted from the elections, but not necessarily when a specific voting rule was implemented. Thus, only the second part of our hypothesis *H2* is supported by evidence. Importantly, however, there seems to be a significant effect of the voting rule on the acceptability of outcome Y: Perhaps surprisingly, the blues punished the yellows less often for alternative Y when the latter chose the simple-majority rule. This is also supported by a sub-sample regression of "change" on "rule" (Table 9b): In the subsample in which B was the electoral outcome, the blues do not punish the yellows significantly more under any of the two rules. By contrast, in the subsample in which Y was the outcome of elections, the blues punish the yellows significantly *less* under the proportional rule. Overall, we can conclude that:

Result 7. *We find evidence for the hypothesis that the unfavorable electoral outcome becomes more acceptable when it has been expected anyway under the chosen voting rule.*

Hence, our findings regarding the behavior of the blues in the punishment treatment support hypothesis *H3b* rather than *H3a*. We do not find any evidence for either *H4a* or *H4b*.

Consider now the exit treatment. Here, the punishment choice of the blues is between ending the game at terms that are unfavorable to the yellows or continuing the game as in the baseline treatment. Due to the timing of the game, the blues have no possibility of conditioning their punishment choice on anything else than on the voting rule. To test whether the blues punish the yellows for choosing the simple-majority rule, we ran a probit regression of the dummy "exit" on "rule", again clustering on the session level. "Exit" encodes the individual decision of the blues and takes a higher value if the blue subject concerned chose "exit". Again, we excluded the yellows from the regression. Table 10 displays the results.

Unsurprisingly, and supporting our hypothesis *H8*, we find that the blues choose "exit" significantly more often under the simple-majority rule than under the proportional rule. This can be due to, for instance, procedural fairness concerns or inequity aversion combined with the anticipation of outcome Y.

5 Conclusion and Discussion

While there is a vast experimental literature on sharing money, our paper presents the first experimental study on constitutional power sharing between groups. Although the games that we implemented in our experiment are similar in some respects to the familiar dictator and ultimatum game (for instance, we also compare voluntary sharing with sharing under off-equilibrium punishment threats), our games differ in too many respects from these well-known bargaining games to be directly comparable to them. For instance, we do not only look at group interaction, but also include the endogenous choice of the rule that is supposed to govern this interaction. Moreover, payoffs are not shared directly, but indirectly by power sharing, and the "size of the cake" varies with the voting rule. Nonetheless, our results clearly show that most insights from the experimental literature on the dictator and ultimatum game carry over "in spirit" to our games of power sharing between

groups: The privileged group (i.e., the majority) shares power voluntarily to only a small extent. Thus, the prevalent determinant of their constitutional choices is self-interest. They become more inclined to share power if under threat of punishment, since they "rationally" react to an anticipated punishment threat which, in itself, is not consistent with equilibrium predictions.[13] However, it is not empowerment but only a favorable electoral outcome that can appease the minority in our experiment: An election outcome that favors the majority is punished more often from the minority if the latter has been empowered by the proportional rule.

To see what can be learned from these findings, note firstly that our study is the first that complements the theoretical literature on endogenous constitutional design by a behavioral perspective. For instance, Acemoglu and Robinson (2000, 2001) argue that throughout history, political elites tended to extend the suffrage to disadvantaged classes if the latter can credibly threat them with a revolution or other forms of social unrest, in particular when the opportunity costs of participating in a riot are sufficiently low. While this is certainly an important insight, our experiment suggests that even if no credible threat is possible under standard assumptions, i.e., when a cool-minded cost-benefit analysis would prevent everyone from rioting, behavioral motives like negative reciprocity can nonetheless trigger a revolution, and members of the elite are quick to understand this. Thus, consider situations in which the elite cannot be credibly threatened by the underprivileged. This case has been theoretically analyzed in an ingenious paper by Lizzeri and Persico (2004). They show that extending the franchise to the poor can function as a commitment device for the elite to refrain from ad personam redistribution and move on to more beneficial policies. Again, this is an important and convincing argument. However, the results from the baseline treatment in our experimental study suggest that a privileged group is sometimes (although rather rearely) willing to share its political power with the underprivileged in the absence of threats even if self-interest would dictate not to do so. This opens up a new line of research: *Which* behavioral determinants lead to power sharing in the absence of threats? How strong are procedural fairness concerns, compared to other determinants like inequity aversion or social norms? How large can the material costs of sharing one's power become

[13] Our findings are roughly in line with results from experiments on collective versus individual decisions on transfers: Groups seem to be less generous (or trusting) but as reciprocating as individuals. See, e.g., Wilson and Eckel (2009) on trust games.

before the privileged start acting selfishly again? And, in relation to this, how large can the costs of participating in a social upheaval become before a threat of revolution directed against the elite stops being credible? This is only a collection of possible questions to be addressed by further research.

A second contribution of our paper lies in its clear prediction concerning the effect of endogenous voting rules on the acceptability of electoral outcomes, a nexus completely ignored by the existing literature. We show that punishment is conditioned both on the voting rule and the electoral outcome, and that the privileged group understands this and often votes against their immediate interests in order to appease the other group. The underprivileged group is more inclined to accept an unfavorable electoral outcome (i.e., to refrain from retaliation) if they were disempowered by the electoral rule. From this we can conclude that there is a trade-off: On the one hand, sharing power increases the likelihood of electoral outcomes in favor of the underprivileged, thus decreasing the risk of punishment. On the other hand, sharing power increases the risk of punishment for electoral outcomes that do not favor the underprivileged.

Overall, our experimental study suggests that behavioral determinants both directly affect endogenous constitutional design and influence the way in which endogenous constitutions shape collective decision making.

6 References

1. Acemoglu, D. and J. Robinson (2000): Why did the West Extend the Franchise: Democracy, Inequality and Growth in Historical Perspective. Quarterly Journal of Economics 115, 1167-1199.

2. Acemoglu, D. and J. Robinson (2001): A Theory of Political Transition. American Economic Review 91, 938-963.

3. Aghion, P., Alesina, A. and F. Trebbi (2004): Endogenous Political Institutions. Quarterly Journal of Economics, May 2004, 565-611.

4. Barbera, S. and M. O. Jackson (2004): Choosing how to Choose: Self-Stable Majority Rules and Constitutions. Quarterly Journal of Economics 119, 3, 1011-1048.

5. Bassi, A., Morton, R., and J. Trounstine (2008): Delegating Disenfranchisement Decisions. Unpublished Working Paper.

6. Bornstein, G., and I. Yaniv (1998): Individual and Group Behavior in the Ultimatum Game: Are Groups More "Rational" Players? Experimental Economics 1, 101-108.

7. Cason, T.N., and V.L. Mui (1997): A Laboratory Study of Group Polarisation in the Team Dictator Game. The Economic Journal 107, 444, 1465-1483.

8. Dal Bo, P., Foster, A. and L. Putterman (2010): Institutions and Behavior: Experimental Evidence on the Effects of Democracy. American Economic Review 100, 5, 2205-2229.

9. Dickson, E.S., Gordon, S.C. and G.A. Huber (2009): Enforcement and Compliance in an Uncertain World: An Experimental Investigation. Journal of Politics 71, 4, 1357-1378.

10. Eckel, C., Johnson, M., and R.K. Wilson (2002): Fairness and Rejection in the Ultimatum Bargaining Game. Political Analysis 10,4, 376-393.

11. Elbittar, A., Gomberg, A., and L. Sour (2011): Group-Decision Making and Voting in Ultimatum Bargaining: An Experimental Study. The B.E. Journal of Economic Analysis and Policy 11, 1, 1-31.

12. Ertan, A., Page, T., and L. Putterman (2009): Who to punish? Individual decisions and majority rule in mitigating the free rider problem. European Economic Review 53, 495–511.

13. Gerber, E.R., Morton, R.B., and T.A. Rietz (1998): Minority Representation in Multimember Districts. The American Political Science Review 92, 1, 127-144.

14. Lizzeri, A. and N. Persico (2004): Why did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain's Age of Reform. Quarterly Journal of Economics 119, 707–765.

15. Luhan, W.J., Kocher, M.G., and M. Sutter (2009): Group Polarization in the Team Dictator Game Reconsidered. Experimental Economics 12, 26-41.

16. Morton, R.B., and T.A. Rietz (2008): Majority Requirements and Minority Representation. NYU Annual Survey of American Law 705, 63, 691-726.

17. Persson, T. and G. Tabellini (2000): Political Economics: Explaining Economic Policy. Cambridge, MA, MIT Press.

18. Persson, T. and G. Tabellini (2003): The Economic Effects of Constitutions. Cambridge, MA, MIT Press.

19. Persson, T. (2002): Do Political Institutions Shape Economic Policy? Econometrica 70, 883-906.

20. Robert, C. and PJ. Carnevale (1997): Group choice in ultimatum bargaining: an experimental study. Organizational Behavior and Human Decision Processes 72, 2, 256-279.

21. Sell, J., Lovaglia, M.J., Mannix, E.A., Samuelson, C.D., and R.K. Wilson (2004): Investigating Conflict, Power, and Status Within and Among Groups. Small Group Research 35, 44-72.

22. Sutter, M., Haigner, S. and M.G. Kocher (2010): Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations. Review of Economic Studies 77, 1540-1566.

23. Ticchi, D. and A. Vindigni (2010): Endogenous Constitutions. The Economic Journal 120, 543, 1-39.

24. Trebbi, F., Aghion, P. and A. Alesina (2008): Electoral Rules and Minority Representation in U.S. Cities. Quarterly Journal of Economics 123, 1, 325-357.

25. Tyran, J.-R. and L.P. Feld (2006): Achieving Compliance when Legal Sanctions are Non-deterrent. Scandinavian Journal of Economics 108, 1, 135-156.

26. Voigt, S. (1997): Positive Constitutional Economics: A Survey. Public Choice 90, 11–53.

27. Voigt, S. (2011): Positive constitutional economics II—a survey of recent developments. Public Choice 146, 205–256.

28. Wilson, R.K., and C.C. Eckel (2009): Trust and Social Exchange. Working Paper, prepared for the Handbook of Experimental Political Science.

Appendix

Figure 2: Graphic Representation of the three Games

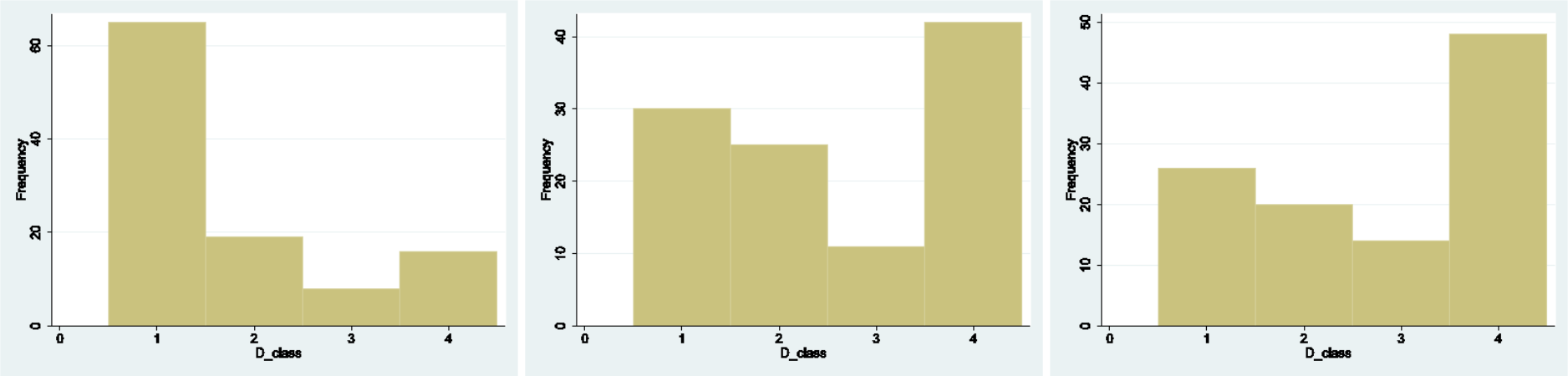

Table 3: One-sample T-test - Treatment 1

Variable	Observation	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
C_prop~l	6	0.1226852	0.0564215	0.0564215	0.0634745	0.1818959

mean = mean(C_proportional) t = 5.3263

Ho: mean = 0 degrees of freedom = 5

Ha: mean < 0	Ha: mean != 0	Ha: mean > 0
Pr(T < t) = 0.9984	Pr(T > t) = 0.0031	Pr(T > t) = 0.0016

Wilcoxon Sign test - Treatment 1

Sign	Observed	Expected
positive	6	3
negative	0	3
zero	0	0
all	6	6

One-sided tests:

Ho: median of C_propor~l - 0 = 0 vs. Ha: median of C_propor~l - 0 > 0

Pr(#positive >= 6) = Binomial(n = 6, x >= 6, p = 0.5) = 0.0156

Ho: median of C_propor~l - 0 = 0 vs. Ha: median of C_propor~l - 0 < 0

Pr(#negative >= 0) = Binomial(n = 6, x >= 0, p = 0.5) = 1.0000

Two-sided test:

Ho: median of C_propor~l - 0 = 0 vs. Ha: median of C_propor~l - 0 != 0

Pr(#positive >= 6 or #negative >= 6) = min(1, 2*Binomial(n = 6, x >= 6, p = 0.5)) = 0.0313

Table 4: Probit Regression of Rulechoice on Period and Session (Treatment 1)

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		No Clustering				With Clustering			
Equation	Variables	rulechoice1	tstat	rulechoice1	tstat	rulechoice1	tstat	rulechoice1	tstat
rulechoice1	period	0.0140***	3.380	0.0145***	3.431	0.0140***	2.888	0.0145***	2.764
	_Isession_i_2			-0.111	-0.853			-0.111***	-462.4
	_Isession_i_5			-0.387***	-3.091			-0.387***	-270.8
	_Isession_i_6			-0.512***	-4.152			-0.512***	-920.9
	_Isession_i_16			-0.135	-1.037			-0.135***	-37.09
	_Isession_i_17			0.328**	2.190			0.328***	47.17
	Constant	0.948***	13.50	1.115***	9.766	0.948***	8.233	1.115***	14.29
	Observations	2,046		2,046		2,046		2,046	

Table 5a: Probit Regression of Rulechoice, Controlling for Country, Treatment and Session

		(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
		All Treatments (1, 5 and 6)							
		No Clustering				**With Clustering**			
Equation	Variables	rulechoice1	tstat	rulechoice1	tstat	rulechoice1	tstat	rulechoice1	tstat
rulechoice1	D_C	-0.287***	-5.726	0.311**	2.089	-0.287***	-3.779	0.311***	1.437e+08
	D_T5	-0.533***	-11.42	-1.018***	-7.518	-0.533***	-5.062	-1.018***	-4.703e+08
	D_T6	-0.525***	-10.85	-1.192***	-8.850	-0.525***	-3.771	-1.192***	-5.509e+08
	_Isession_i_2			-0.111	-0.856			-0.111***	-7.387e+13
	_Isession_i_3			0.498***	2.736			0.498***	2.302e+08
	_Isession_i_4			0.615***	3.360			0.615***	2.842e+08
	_Isession_i_5			-0.704***	-5.025			-0.704***	-3.252e+08
	_Isession_i_6			-0.828***	-5.971			-0.828***	-3.824e+08
	_Isession_i_7			-0.111	-1.074			-0.111***	-2.016e+14
	_Isession_i_8			-0.124	-1.221			-0.124***	-5.892e+14
	_Isession_i_9			-0.232**	-2.297			-0.232***	-4.138e+14
	_Isession_i_10			-0.171*	-1.734			-0.171***	-1.679e+14
	_Isession_i_11			0.329***	3.166			0.329***	6.054e+14
	_Isession_i_12			0.267***	2.577			0.267***	1.341e+15
	_Isession_i_13			-0.0646	-0.649			-0.0646***	-2.410e+14
	_Isession_i_14			0.228**	2.237			0.228***	4.715e+14
	_Isession_i_16			-0.457***	-3.161			-0.457***	-2.113e+08
	Constant	1.355***	26.85	1.337***	14.06	1.355***	17.37	1.337***	9.007e+14
	Observations	6,138		6,138		6,138		6,138	

Table 5b: Probit Regression of Rulechoice, Controlling for Country, Treatment and Session

		(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
		Treatment 5 and 6 only							
		No Clustering				**With Clustering**			
Equation	Variables	rulechoice1	tstat	rulechoice1	tstat	rulechoice1	tstat	rulechoice1	tstat
rulechoice1	D_C	-0.362***	-5.548	-0.304***	-2.854	-0.362***	-5.903	-0.304***	-6.446e+07
	D_T5								
	D_T6	0.0312	0.673	-0.175*	-1.723	0.0312	0.344	-0.175***	-9.384e+12
	_Isession_i_2								
	_Isession_i_3			-0.117	-1.070			-0.117***	-2.475e+07
	_Isession_i_4								
	_Isession_i_5								
	_Isession_i_6								
	_Isession_i_7			-0.111	-1.074			-0.111***	-4.668e+14
	_Isession_i_8			-0.124	-1.221			-0.124***	-1.502e+14
	_Isession_i_9			-0.232**	-2.297			-0.232***	-1.476e+12
	_Isession_i_10			-0.171*	-1.734			-0.171***	-1.659e+11
	_Isession_i_11			0.329***	3.166			0.329***	1.843e+10
	_Isession_i_12			0.267***	2.577			0.267***	5.063e+11
	_Isession_i_13			-0.0646	-0.649			-0.0646***	-3.353e+11
	_Isession_i_14			0.228**	2.237			0.228***	7.001e+12
	_Isession_i_16								
	Constant	0.875***	16.05	0.934***	11.92	0.875***	20.30	0.934***	1.980e+08
	Observations	4,092		4,092		4,092		4,092	

Table 5c: OLS Regression of C_RuleChoice (on collapsed data set)

	(1)	(2)	(3)	(4)	(5)	(6)
	T1 and T5 only		**All Treatments**		**T5 and T6 only**	
Variables	C_rulechoice	tstat	C_rulechoice	tstat	C_rulechoice	tstat
D_C	-0.0726**	-2.341	-0.0726*	-2.014	-0.111*	-2.218
D_T5	-0.144***	-4.942	-0.144***	-4.252	-0.00926	-0.248
D_T6			-0.148***	-4.109		
Constant	0.926***	31.67	0.926***	27.25	0.816***	14.73
Observations	12		18		12	
R-squared	0.769		0.706		0.384	

Table 6: OLS Class

	-1	-2	-3	-4	-5	-6	-7	-8
	OLSClass1		**OLSClass2**		**OLSClass3**		**OLSClass4**	
Variables	B_class_1	tstat	B_class_2	tstat	B_class_3	tstat	B_class_4	tstat
D_T5	-1,050***	-2.52E+13	180***	1.91E+13	90***	2.08E+13	780.0***	2.63E+13
D_T6	-1,170***	-3.59E+13	30***	5.06E+12	180***	5.20E+13	960.0***	3.98E+13
Constant	1,950***	6.08E+13	570***	1.27E+14	240***	7.08E+13	480.0***	2.04E+13
Observations	9,720		9,720		9,720		9,720	
R-squared	1		1		1		1	

Table 7: T-test of B_Y_vote_B

One Sample T-test of B_Y_vote_B == 0 if (subject == 1) & (condition == 1)

Variable	Observation	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
B_Y_vo~B	180	0.0180556	0.0030987	0.0415735	0.0119409	0.0241702

mean = mean(B_Y_vote_B) t = 5.8268
Ho: mean = 0 degrees of freedom = 179

Ha: mean < 0 Ha: mean != 0 Ha: mean > 0
Pr(T < t) = 1.0000 Pr(T > t) = 0.0000 Pr(T > t) = 0.0000

One Sample T-test of B_Y_vote_B == 0 if (subject == 1) & (condition == 5)

Variable	Observation	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
B_Y_vo~B	180	0.137037	0.0092297	0.1238298	0.118824	0.1552501

mean = mean(B_Y_vote_B) t = 14.8474
Ho: mean = 0 degrees of freedom = 179

Ha: mean < 0 Ha: mean != 0 Ha: mean > 0
Pr(T < t) = 1.0000 Pr(T > t) = 0.0000 Pr(T > t) = 0.0000

One Sample T-test of B_Y_vote_B == 0 if (subject == 1) & (condition == 6)

Variable	Observation	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
B_Y_vo~B	177	0.0499058	0.0060652	0.080692	0.037936	0.0618757

mean = mean(B_Y_vote_B) t = 8.2283
Ho: mean = 0 degrees of freedom = 176

Ha: mean < 0 Ha: mean != 0 Ha: mean > 0
Pr(T < t) = 1.0000 Pr(T > t) = 0.0000 Pr(T > t) = 0.0000

Table 8a: Output of the Probit Regression for Treatment Differences

		(1)	(2)
		Probvote1	
Equation	Variables	D_Y_vote_B	tstat
D_Y_vote_B	D_T5	0.972***	6.004
	D_T6	0.425***	3.183
	Constant	-2.074***	-25.01
	Observations	5,578	

Table 8b: Output of the OLS Regression OLSvote1

	(1)	(2)
	OLSvote1	
Variables	D_Y_vote_B	tstat
D_T5	0.119***	4.226
D_T6	0.0319	1.133
Constant	0.0181	0.907
Observations	18	
R-squared	0.561	

*** p<0.01, ** p<0.05, * p<0.1

Table 9a: Output of the Probit Regression of Change on"rule" and "decide"

		(1)	(2)	(3)	(4)	(5)	(6)
		ProbChangeNoDC1		**ProbChangeNoDC1**		**ProbChangeNoDC1**	
Equation	Variables	change	tstat	change	tstat	change	tstat
change	rule	-0.128	-0.970			-0.221**	-2.393
	decide			1.186***	3.882	1.202***	4.020
	Constant	-0.411*	-1.666	-1.488***	-9.048	-1.115***	-5.773
	Observations	1,200		1,200		1,200	

Table 9b: Sub-Sample Regression of "change" on "rule"

		(1)	(2)	(3)	(4)
		ProbChangeRuleSubB		ProbChangeRuleSubY	
Equation	Variables	change	tstat	change	tstat
change	rule	0.108	0.837	-0.323***	-2.832
	Constant	-1.672***	-4.509	0.266	1.170
	Observations	453		747	

Table 10: Output of the Probit Regression of the Dummy "exit" on "rule"

Equation	Variables	ProbExit1 exit
exit	rule	0.428***
	Constant	-1.415***
	Observations	1,188

Table 11: Mann-Whitney-U-test on PunishRule2 in Treatments 5 and 6

Rank-Sum PunishRule2 if (PunishRule2 > -1), by(condition)

Condition	Observation	Rank Sum	Expected
5	888	752528	765900
6	836	734422	721050
combined	1724	1486950	1486950
unadjusted variance	1.07E+08		
adjustment for ties	-43655590		
adjusted variance	63059810		
Ho: Punish~2(condit~n==5) = Punish~2(condit~n==6)			z = -1.684
			Prob > |z| = 0.0922