When do conflicting parties share political power?*

Marco Battaglini
Department of Economics - Princeton University
mbattagl@princeton.edu

Lydia Mechtenberg
Department of Economics - University of Hamburg
lydia.mechtenberg@wiso.uni-hamburg.de

January 5, 2015

Abstract

We conduct a laboratory experiment to study the incentives of a privileged group (the "yellows") to share political power with another group (the "blues"). The yellows collectively choose the voting rule for a general election: a simple-majority rule that favors them, or a proportional rule. In two treatments the blues can use a costly punishment option. We find that the yellows share power voluntarily only to a small extent, but they are more inclined to do so under the threat of punishment, despite the fact that punishments are not subgame perfect. The blue group conditions punishments both on the voting rule and the electoral outcome.

Keywords: Constitutional Design, Endogenous Institutions.
JEL-Codes: C92, D02, D72, D79

*This revision: August 2014. We are grateful to Giri Parameswaran, Dan Myers and Julia Körner for excellent research assistance. This research has been supported by the DFG-Deutsche Forschungsgemeinschaft. This paper has been presented at the Annual Meeting of the Association for Public Economic Theory PET 2013, the Annual European Meeting of the Econometric Society ESEM 2013, and the Annual Meeting of the Economic Science Association ESA 2013, and the NYU Experimental Political Science Conference 2014.

1 Introduction

Institutions are one of the most important factors determining political outcomes. Institutions, however, are not exogenous: they are themselves political outcomes. Recognizing this fact, a recent but already significant literature, both in economics and in political science, has proposed positive theories of how political institutions are chosen.[1] At the core of most of these theories there is the implicit assumption that political actors are rational, selfish versions of the proverbial "homo economicus." On the other hand, history is full of examples in which groups entrusted with significant power choose to share it with others for what appear to be, perhaps superficially, purely idealistic reasons. Recent experimental evidence, moreover, shows that humans can behave in surprisingly altruistic ways; and that humans do not only care about *what* is decided, but also about *how* decision are made.[2] This suggests a number of open research questions: To what extent are institutions the result of Machiavellian calculations of convenience; and, conversely, to what extent are they shaped by other behavioral factors? In this paper, we make a first step in addressing these issues by presenting a laboratory experiment.

In our experiment, a population of citizens is divided in two groups, the "blues" and the "yellows." The entire population collectively decides between alternative payouts through an election. Before the election is held, the yellows are endowed with the constitutional power to choose the voting rule that determines the outcome of the election. They can choose between a majoritarian system and a system in which the policy outcome is determined by the groups in proportion to their size. We study the yellows' choice under conditions where the yellows are expected to be the majority and the blues

[1] Aghion, Alesina and Trebbi (2004) and Trebbi, Aghion and Alesina (2008) investigate constitutional choices within a modern democracy. For a theoretical study of the choice of disenfranchisement laws by states in the U.S., see Bassi, Morton and Trounstine (2008).

[2] Examples of prosocial behavior abound in the vast literature on the dictator game (first studied by Forsythe et al. (1994) and Hoffman, McCabe and Smith (1996)), the ultimatum game (introduced by Güth, Schmittberger and Schwarze (1982)), and the trust game (see Berg, Dickhaut and McCabe (1995)). See also the evidence on the problem of the commons, studied by Elinor Ostrom, and the literature on public good games. Procedural fairness concerns have been documented by, among others, Frey and Stutzer (2004), Bolton, Brandts and Ockenfels (2005), Aldashev, Kirchsteiger and Sebald (2010), Shor (2007), Mertins (2008), and Dickson, Gordon and Huber (2009).

have varying degrees of retaliatory power.

We find three results of note. First, a small but significant fraction of the yellows vote for a decision rule that favors the blues in the baseline treatment. Second, despite the fact that retaliation is never optimal in equilibrium, punishment occurs; the possibility of punishment by the minority, moreover, induces a significant increase in the fraction of majority members voting for the proportional rule. Third, the choice of the voting rule changes the way the minority punishes the majority: an unfavorable outcome is punished *more* under a rule that allocates proportional power to the two groups than under the majoritarian rule (that allocates all the power to the majority).

This indicates that procedural factors are important: the choice of the voting rule affects players' expectations and hence their reference point with regard to the resulting payoff allocation. Based on these findings, we conclude that institutions are the result of both behavioral factors *and* self-serving calculations that interact in interesting ways.

2 The Experiment

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.

2.1 The Benchmark game (*TBase*)

There are six players. Nature randomly assigns the color yellow or blue to each player. With probability $\frac{9}{10}$, there are four yellow and two blue players in the group, and with the remaining probability $\frac{1}{10}$, the group consists of two yellow and four blue players. Players observe their type but they do not know whether their type constitutes the majority. However, they know the probabilities of that occurring.[3]

The players play a game with two stages:

- In the constitutional stage, the yellows, regardless of whether they hold the majority, privately choose between two voting rules, rule 1 and rule

[3] 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. One yellow's choice is randomly drawn with equal likelihood and implemented. Then, all players are informed about which rule the yellows have implemented.

- In the voting stage, all players participate in an election. They privately vote for either alternative Y or alternative B (abstention is not allowed). Alternative Y assigns a high payoff $\alpha + 5$ to each yellow and a low payoff α to each blue player; and alternative B does the reverse. The rule chosen in the previous stage will determine the result.

Under rule 1, one of the six players is randomly drawn with equal probability and his decision implemented. Under rule 2, the alternative with the highest number of votes is implemented.[4] Thus, rule 1 is the random-dictator rule and rule 2 the simple-majority rule.[5] As is well known, there is no widely accepted model of proportional electoral systems. An established literature has modelled power sharing typical of proportional systems by allowing the policy outcome to be determined by different constituencies in proportion to their size (see, for example, Sahuguet and Persico [2006]) as in rule 1.[6] Following this literature, we interpret rule 1 as representing a proportional system.

At the end of the game, all players learn whether alternative Y or B has been implemented and earn their resulting payoff. In the unique Perfect Bayesian equilibrium of this game, the yellows choose rule 2 and vote for Y, whereas the blues vote for B. Outcome Y is implemented whenever the yellows are the majority and B is implemented whenever the blues are the majority.

---

[4] If both alternatives attracted the same number of votes, a fair coin is flipped to decide which alternative to implement.

[5] Feddersen, Gailmard, and Sandroni (2009) use the random-dictator rule in an experimental context different from ours.

[6] See Sahuguet and Persico [2006] for an extensive discussion of this point and further references. As recognized by Sahuguet and Persico [2006], there are two ways to allow the outcome to reflect parties in proportion to their size: *probabilistic compromise*, in which there is a random dictator as in our model described above; and a *splitting-the-spoils-of-office* model, in which a divisible outcome is shared by the parties pro quota. While these two ways are conceptually different in general models, they are equivalent in models with risk neutrality as in our model.

2.2 Treatments with punishments (*TPunish*, *TExit*)

We consider two additional treatments in which the minority can punish the majority.

The ex-post punishment game (*TPunish*) The ex-post punishment game differs only in one respect from the benchmark game: It adds a punishment stage following the voting stage. On the punishment stage, each blue player privately chooses between accepting (*keep*) and changing (*change*) the final outcome of the election, i.e., Y or B. 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 7 points, the blue player whose choice was implemented loses 1 point, and the payoffs of the other blue player(s) remain unaffected. Hence, when the yellows are indeed punished, the blue who is responsible for it gets one point more than the yellows, while the other blue player(s) get(s) 2 points more.[7] All players are informed whether "change" or "keep" was implemented in the group. Punishments in this treatment are not sub-game perfect, so the unique equilibrium coincides with the equilibrium of *TBase*.

The interim punishment game (*TExit*) 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* in *TExit* are the same as those after *change* in *TPunish* when the electoral outcome has been Y. If *exit* is not implemented, the subjects get the same feedback as in *TBase*.

The difference between *TPunish* and *TExit* is that in *TPunish* the blues can condition their punishment behavior both on the voting rule and the

[7] This punishment scheme adopts the minimal amount of payoff reduction for the yellows and punishment costs for the blues that guarantee that, if punished, the yellows end up having less than any blue player.

outcome (Y or B), but in *TExit* punishment can only be conditioned on the voting rule. Here too punishments are not sub-game perfect, so the unique equilibrium remains unchanged from *TBase*.

2.3 Hypotheses

Our three games can be seen as collective-decision variations of the dictator and ultimatum game. Similar to the experimental literature on these games, it is reasonable to expect systematic deviations from the predictions of the Perfect Bayesian Equilibrium. First, we know from the literature on dictator games that under standard conditions subjects tend to share their payoffs to reduce inequality. In our setting, subjects can only reduce inequality in *expected* payoffs; this can be done by implementing the proportional rule on the constitutional stage. Second, the literature on ultimatum games strongly suggests that recipients react to disadvantagous inequality (if better options have been available) by chosing to punish the proposer even though this is not sub-game perfect. In our setting, the blues are those confronted with disadvantagous inequality, and they have access to punishment in the two control treatments. In particular, we expect that the yellows are willing to choose a less favorable voting system in order to leave some power to the blues; and that the blues are willing to pay a price to punish opportunistic behavior by the yellows. The willingness to chose a less favorable outcome, moreover, may depend on the possibility of punishments, despite the fact that punishments are never sub-game perfect. Similarly, the willingness to punish may depend on the policy outcome and voting rule with which it is chosen.

Motivated by these considerations, we test the following four hypotheses:

H1. *The yellows choose the proportional rule (rule 1) with positive probability in the benchmark game.*

H2. *The yellows (i) select the proportional rule more often in the two punishment games and (ii) vote for B more often in the ex-post punishment game than in the benchmark game.*

H3. *In the ex-post punishment game, the blues (i) punish more often after Y than after B; (ii) and they punish more under the simple-majority rule than under the proportional rule (both (a) in general and (b) conditional on a given outcome Y or B).*

H4. *In the interim punishment game, the blues punish more under the simple-majority rule than under the proportional rule.*

In contrast with dictator and ultimatum games, there is no obvious "fair" ex post allocation of payoffs in our model,[8] pro-social behavior at the constitutional stage cannot be exclusively attributed to concerns about fairness.[9] Indeed, our results suggest that perceived fairness of the final allocation depends on how it is achieved, in particular on the voting system chosen in the "constitutional" phase.

2.4 Experimental Procedure

We conducted three experimental treatments, one for each game.[10] Each treatment had six sessions. In each session, the game had 18 subjects. In each round, players were divided in 3 groups of 6 to play the game at hand. Subjects were randomly rematched into groups and roles between rounds.[11] Thirty rounds of the game at hand were played. We randomly varied α across rounds.[12] One round was randomly drawn and the payoffs were paid out to the participants in cash.

[8]For recent experiments on fairness concerns and voting, see, e.g., Gerber, Nicklisch, and Voigt (2013) and Hoechtl, Sausgruber, and Tyran (2012).

[9]Implementing the simple majority rule maximizes the aggregate payoffs in a given round and might hence be preferred by players with efficiency preferences. Taking into account that the roles of blues and yellows are randomly re-assigned each round, and that it is always more likely to become a member of the majority rather than the minority, one could also argue that implementing the simple majority rule and voting in one's self-interest is the ex-ante cooperative outcome for the experimental subjects. (For a similar argument, see Feddersen at al. 2009) However, considering the one-stage game that is played in one given round and abstracting from efficiency preferences, no prediction can be made on how the subjects evaluate the fairness of the (ex-post) allocation.

[10]A more detailed description of the experimental procedure is available in the on line appendix. The free software z-tree (Fischbacher (2007)) was used to computerize all three games. Two sessions of the baseline treatment were conducted at Princeton University, USA, and all others at the Technical University of Berlin, Germany.

[11]It would be interesting to see how the behavior of both yellows and blues would change if they remained in one group for at least several rounds. It would also be interesting to require the subjects to always play the same role in the game. For reason of space, we leave these interesting variations for future research.

[12]We intentionally chose this payoff structure for the elections to keep the experiment interesting for our subjects; but we did not expect changes of a to have any systematic effect on the choice of voting rule, the voting behavior or the punishing behavior. Since this has been confirmed by intense testing, we henceforth drop any reference to the variation

3 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.

3.1 When do subjects share power?

The upper part of Table 1 displays standard t-tests of whether the proportion of yellows voting for the proportional rule is significant in our three treatments.

Result 1. *(i) Supporting hypothesis H1, on average 12% of the decisions made by yellow players in TBase are in favor of the proportional rule; this proportion is highly significant. (ii) A higher average of roughly 26% (TPunish) and 29% (TExit) of the decisions made by yellow players in the two punishment treatments are in favor of the proportional rule; again, these proportions are highly significant.*

Comparing *TBase* with *TPunish* and *TExit*, we find: (1) the share of players who always or nearly always choose the majoritarian rule (95% or more of their opportunities to) is more than halved under the threat of punishment; and (2) the share of players who often choose the proportional rule (95% or more of their opportunities to) is more than doubled (not shown in the tables).[13]

As a first step toward testing whether these differences are significant, we used a Random Effect logit model as displayed in Table 3. To visualize this result, we took a closer look at the distribution of constitutional choices over individual players and estimated the cumulative distribution functions of yellows choosing the proportional rule in our three treatments (Figure 1). Interestingly the distributions in *TPunish* and *TExit* are similar and they both first-order stochastically dominate the corresponding distribution in *TBase*. This difference is statistically significant.[14] We conclude:

of a.

[13]In Figure A3 in the on line Appendix, we provide a histogram depicting this result.

[14]In Table A1 in the on line appendix we present the results of a Two Sample Kolmogorov-Smirnov test to support this claim.

Result 2. *Supporting our hypothesis H2(i), the average number of yellow players choosing the proportional rule increases significantly under the threat of punishment, both when punishment occurs ex post (TPunish) and when it occurs on the interim stage (TExit).*

To test for learning effects, we ran logit regressions as displayed in Table 2.[15] We find that learning in *TBase* is highly significant, but small in magnitude.[16]

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

As shown in the lower part of Table 1, the average frequency of voting for alternative B when yellow (Y_vote_B) is slightly below 2% and 5% in *TBase* and *TExit*, respectively; but much higher, 13.8%, in *TPunish*. This difference is intuitive: it is only in *TPunish* that the blues can condition their punishment on the alternative, which the yellows seem to anticipate. All three means are significantly different from zero.

We tested for treatment differences by running a logit regression (Table 3(3)); and we conclude:

Result 3. *Supporting hypothesis H2(ii), the share of yellow votes for alternative B is significantly higher in TPunish than in TBase.*

3.3 (How much) do the blues punish?

Consider TPunish first. The upper panel of Table 4 reveals that the blues punished significantly more after outcome Y; they also punish less frequently under the simple-majority rule, although this latter effect is not statistically robust.[17]

[15] The dummy *rulechoice* 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.

[16] We have not studied how the outcome in previous rounds affect the players' behavior. This seems and important question: for reasons of space, we leave it for future research.

[17] To interpret Table 4 note that, for a given blue subject, *change* takes a higher value if the subject chose to punish. The variable *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 controlled for interdependency of the data.

To test whether the blues punish more when they get outcome Y under the proportional rule and more when they get outcome B under the majoritarian rule, as suggested by the descriptive statistics,[18] we split the sample into two subsamples, one with outcome B ((3) in Table 4) and one with outcome Y ((4) in Table 4). There seems to be a significant effect of the voting rule on the acceptability of outcome Y: 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 simple-majority rule. Overall, we can conclude that:

Result 4. *Supporting part (i) of H3 the blues punish more after Y than after B. There is however weak evidence supporting the hypothesis that the blues generally punish less under the simple-majority rule than under the proportional rule (contrary to (ii.a) of H3). Moreover, conditioning on a Y outcome, the blues punish more under the proportional rule, contrary to (ii.b) of H3.*

A plausible explanation for the fact that the blues are more inclined to punish for a Y outcome under rule 1 is that the choice of the voting rule affects the reference point according to which the blues evaluate the outcome of the vote: the blues feel more entitled to expect a more favorable outcome under the proportional rule than under the majoritarian rule.

Consider now the treatment *TExit*. To test whether the blues punish the yellows for choosing the simple-majority rule, we ran a logit regression of the dummy *exit* on *rule*. We obtain:[19]

Result 5. *Supporting our hypothesis H4, the blues choose "exit" significantly more often under the simple-majority rule than under the proportional rule.*

Is the behavior of the yellows optimal given the actual punishing behavior of the blues? In Table 5, we report the average frequency of punishment after the four possible combinations of voting rule and voting outcome. These

[18] See Table 5.

[19] The detailed results are presented in Table A2 in the on-line appendix. The variable *exit* encodes the individual decision of the blues and takes a higher value if the blue subject concerned chose to exit the game. Again, we excluded the yellows from the regression and controlled for interdependency of the data.

frequencies allow us to verify that it is not worth it for the yellows to "appease" the blues by choosing the proportional rule or by voting for B. Given the empirical frequencies of Y, B, and punishment, and conditioning on pivotality, the yellows' expected utility of voting for Y is $E[\alpha] + 2.641$ after rule 2 and $E[\alpha] + 1.759$ after rule 1; the expected utility of voting for B is $E[\alpha] - 0.833$ after rule 2 and $E[\alpha] - 0.399$ after rule 1; the expected utility of voting for rule 1 is approximately $E[\alpha] + 0.86$ and the expected utility of voting for rule 2 is approximately $E[\alpha] + 1.60$.[20]

4 Conclusion

While there is a vast experimental literature on sharing money, our paper presents the first experimental study on sharing constitutional power between groups. Our results clearly show that some (but not all) insights from the experimental literature on the dictator and ultimatum game carry over to our games of power sharing between groups: The privileged group (i.e., the majority) shares power voluntarily only to a small extent. Thus, the prevalent determinant of their constitutional choices is self-interest. They react - or even overreact - to an anticipated punishment threat by becoming more inclined to share power, even though the punishment is not consistent with equilibrium predictions.

To see what can be learned from our findings, note that our study is the first that complements the theoretical literature on endogenous constitutional design with a behavioral perspective. Our experiment suggests that even if a cool-minded cost-benefit analysis would prevent structural minorities from using costly punishment against the structural majority, behavioral motives can nonetheless trigger them to do so, and the majority is quick to understand and even overreact to this. Moreover, the results from the baseline treatment in our experimental study suggest that the majority is sometimes (although rather rarely) willing to share its political power with the minority in the absence of threats even if self-interest would dictate not to. Another contribution of our paper is the 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

[20] To validate this, see Table 5.

both on the voting rule and the electoral outcome, and the group with constitutional power understands this and often votes against their immediate interests in order to appease the other group. Importantly, the group without constitutional power 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 conclude that the majority faces 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 if the electoral outcome does not favor the underprivileged.

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

5 References

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

2. Aldashev, G., Kirchsteiger, G., and A. Sebald (2010): How (not) to decide: Procedural Games. ECARES working papers.

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

4. Berg, J., Dickhaut, J., and K. McCabe (1995): Trust, Reciprocity, and Social History. Games and Economic Behavior 10, 122-142.

5. Bolton, G.E., Brandts, J., and A. Ockenfels (2005): Fair Procedures: Evidence from Games Involving Lotteries. The Economic Journal 115, 506, 1054-1076.

6. 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.

7. Feddersen, T., Gailmard, S., and A. Sandroni (2009): Moral Bias in Large Elections: Theory and Experimental Evidence. American Political Science Review 103, 2, 175-192.

8. Fischbacher, U. (2007): z-Tree: Zurich toolbox for ready-made economic experiments. Experimental Economics 10, 2, 171-178.

9. Forsythe, R., Horowitz, J., Savin, N. E., and M. Sefton (1994): Replicability, Fairness and Pay in Experiments with Simple Bargaining Games. Games and Economic Behavior 6, 3, 347-69.

10. Frey, B.S., and A. Stutzer (2004): Introducing Procedural Utility. Not only What, but also How Matters. Journal of Institutional and Theoretical Economics 150, 3, 377-401.

11. Gerber, A., Nicklisch, A. and S. Voigt (2013): Strategic Choices for Redistribution and the Veil of Ignorance: Theory and Experimental Evidence. Working paper, October 2013.

12. Hoechtl, W., Sausgruber, R., and J.-R. Tyran (2012): Inequality Aversion and Voting on Redistribution. European Economic Review 56, 1406-1421.

13. Hoffman, E., McCabe, K., and V. L. Smith (1996): Social Distance and Other-Regarding Behavior in Dictator Games. American Economic Review 86,3, 653-660.

14. Mertins, V. (2008): Procedural Satisfaction Matters: Procedural Fairness Does Not. No. 07/2008. IAAG Discussion Paper.

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

16. Sahuguet N. and N. Persico (2006): Campaign Spending Regulation in a Model of Redistributive Politics, *Economic Theory*, 28, 95-124.

17. Shor, M. (2007): Rethinking the Fairness Hypothesis: Procedural Justice in Simple Bargaining Games. SSRN 1073885.

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

19. 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.

Table 1: Behaviour of the Yellows

	Rule Choice		
	TBase	TPunish	TExit
	Frequency of choosing the proportional rule when Yellow H_0: proportional = 0 H_1: proportional > 0		
	Clustered on Session Level		
Mean	0.124 ***	0.267***	0.293***
T–stat	5.39	10.09	10.29
Number of Sessions	6	6	6
	Clustered on Group Level		
Mean	0.122 ***	0.267***	0.294***
T–stat	16.85	26.58	29.00
Number of Groups	540	540	540
	Voting for Outcome		
	TBase	TPunish	TExit
	Frequency of voting for the B Outcome when Yellow H_0: Y_vote_B = 0 H_1: Y_vote_B > 0		
	Clustered on Individual Level		
Mean	0.019 ***	0.138***	0.048***
T-stat	3.95	6.50	4.41
Number of individuals	108	108	108

*** p<0.01, ** p<0.05, * p<0.1, 95% Confidence Interval

Note– *proportional* is a dummy variable and equals 1 if the proportional rule is chosen.
A one sample t-test against a one-sided alternative is used.
To account for the clustered data structure, the tests are conducted on the proportion of Yellows choosing the proportional rule. In the upper part of the table, the average over groups is calculated first. Then the average over sessions is calculated. Hence, we are left with 6 independent data points per treatment. In the middle part of the table, the average over groups is tested.
In the lower part of the table, the proportion of how often an individual votes for Outcome B, when she is Yellow is tested. With 6 sessions per treatment and 18 individuals per session, we are left with 108 individuals per treatment

Table 2: Logit Regression of Rulechoice in TBase

	(1)		(2)	
	Pooled Estimator		Fixed Effects Estimator	
VARIABLES	rulechoice	tstat	rulechoice	tstat
period	0.0268***	2.787	0.0437***	3.758
Constant	1.563***	7.507		
Observations	2,046		848	
Number of individuals			45	

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

Note– *rulechoice* is the dependent dummy variable and equals 1 if the simple majority rule is chosen.
Coefficients estimated using logit regressions.
In model (1) Standard Errors are clustered on session level.

Figure 1: Treatment Effects on Rulechoice

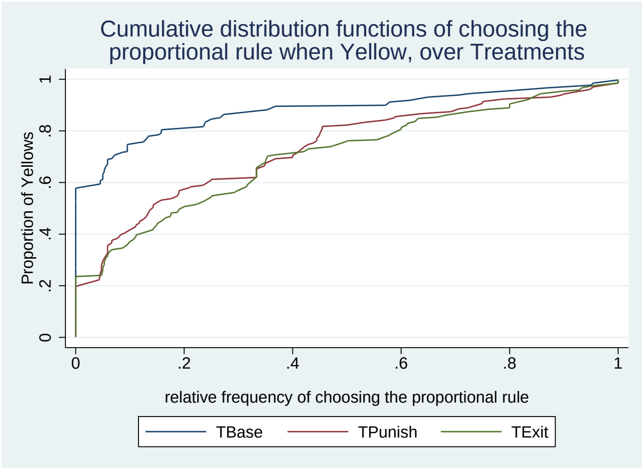

Table 3: Treatment Effects on the Behaviour of the Yellows

	(1) All Treatments		(2) TPunish and TExit only		(3) All Treatments	
VARIABLES	rulechoice	tstat	rulechoice	tstat	Y_vote_B	tstat
Germany	-0.681*	-1.76	-0.974**	-2.05	1.799***	3.66
TPunish	-1.898***	-5.18			2.620***	5.76
TExit	-1.858***	-4.82	0.132	0.38	0.606	1.24
Constant	4.085***	10.49	2.334***	5.90	-7.16***	-12.04
Observations	6,138		4,092		5,578	
Number of individuals	324		216		324	

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

Note– Regression estimated using a Random Effect Logit model

rulechoice is the dependent dummy variable in model (1) and (2) and equals 1 if the simple majority rule is chosen.

Y_vote_B is the dependent dummy variable in model (3) and equals 1 if the Yellow votes for outcome B.

Germany is a dummy variable indicating that the sessions were conducted in Berlin rather than in Princeton. *TPunish* and *TExit* are dummy variables as well, indicating the Treatment.

Table 4: Logit Regression of Change

	Full Sample			
	(1)		(2)	
	Pooled Estimator		Fixed Effect Estimator	
VARIABLES	change	tstat	change	tstat
decide	2.166***	4.056	4.245***	11.03
rule	-0.401**	-2.468	-0.289	-1.142
Constant	-1.945***	-5.483	-4.961***	-7.628
Observations	1,200		811	
Number of individuals			71	
	Sub Sample			
	(3)		(4)	
	Group voting B		Group voting Y	
VARIABLES	change	tstat	change	tstat
rule	0.225	0.803	-0.519***	-2.787
Constant	-2.997***	-3.775	0.428	1.194
Observations	453		747	

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

NOTE– *decide* is a dummy variable and is 1 if the outcome Y. *rule* is equal to 1 if the simple majority rule is chosen and 0 otherwise.

Coefficients estimated using logit regressions. Model (3) and (4) estimated using pooled data.

Standard Errors are clustered on session level.

Table 5: Frequency of Punishment

Rule		PROPORTIONAL	SIMPLE MAJORITY
		H_0 : "change"= 0 in TPunish	
OUTCOME Y	Mean	0.463***	0.337***
	T–stat	12.16	15.87
	Observations	80	283
	Rel. Frequency	0.584	0.702
OUTCOME B	Mean	0.057***	0.119***
	T–stat	2.75	5.80
	Observations	57	120
	Rel. Frequency	0.416	0.298
		H_0 : "exit"= 0 in TExit	
EXIT TREATMENT	Mean	0.160***	0.288***
	T–stat	8.20	17.28
	Observations	160	380

*** p<0.01, ** p<0.05, * p<0.1, 95% Confidence Interval

NOTE – A one-sample t-test against a one-sided alternative is conducted.

To account for the clustered data structure, the test is conducted on the calculated mean over groups.

Relative Frequency refers to the relative frequency of the outcome after the Yellows chose the rule.

For example, we observed overall 540 group decisions and from that 137 groups chose the proportional rule. When the proportional rule was chosen, 80 group decisions had Y as the outcome, so the relative frequency of outcome Y under the proportional rule is 58.4%.

Appendix B: Instructions

Thank you for agreeing to participate in this experiment. During the experiment we require your complete, undistracted attention and ask that you follow instructions carefully. Please turn off your cell phones. Do not open other applications on your computer, chat with other students, or engage in other distracting activities, such as reading books, doing homework, etc. You will be paid for your participation in cash, at the end of the experiment. Different participants may earn different amounts. What you earn depends partly on your decisions, partly on the decisions of others, and partly on chance. It is important that you not talk or in any way try to communicate with other participants during the experiments.

Following the instructions, there will be a practice session and a short comprehension quiz. All questions on the quiz must be answered correctly before continuing to the paid session.

At the end you will be paid in private and you are under no obligation to tell others how much you earned. Note that we are bound not to use deception, so all information in these instructions is true. If something is unclear to you while reading, or if you have other questions, please let us know by raising your hand. We will then answer your questions individually.

As a matter of course, your anonymity and the anonymity of the other participants will be guaranteed throughout the entire experiment. You will neither learn about the identity of the other participants, nor will they learn about your identity.

1. General

This is an experiment in decision-making. It consists of **thirty rounds**. At the end of the experiment, the computer randomly draws one of the rounds; and the dollars that you have earned in this round will be paid out to you in cash. The exact sequence of the stages of the experiment is explained in detail in the following.

2. Players

There are 18 participants in total. At the beginning of each round, the computer randomly assigns the participants to 3 groups of 6. During each round, you interact exclusively with the participants of the group you are assigned to. No participant knows the identity of the other members of his or her group. Group membership changes randomly over different rounds.

There are two possible states, state 1 and state 2. In each round, the state is implemented as follows: After participants have been matched into groups, the computer randomly draws a number from 1 to 10. All numbers are equally likely. If the number is less than 10, then the state is state 1. If the number is 10, then the state is state 2. You are not informed about the state. The state remains the same during a round but may change over different rounds.

1. In state 1, the computer randomly assigns the role of a "yellow player" to 4 out of the 6 participants in a group and the role of a "blue player" to the remaining 2 participants.

2. In state 2, the computer randomly assigns the role of a "blue player" to 4 out of the 6 participants in a group and the role of a "yellow player" to the remaining 2 participants.

Although nobody observes the state, each participant is informed about his (her) own role, i.e. whether he (she) is a yellow or a blue player. Nobody observes the role of any other participant. Note that since the assignment of roles is random and is repeated each round, roles change across rounds. In each round your role is indicated on the upper left of your screen.

3. Payoffs Structure

Your group will be asked to collectively choose one of two **alternatives**: **Y, B**. Your payoff will depend on your assigned role and the alternative (Y or B) that is chosen.

The following table presents an example of a possible payoffs structure. The columns designate the roles, and the rows designate the alternatives:

Payoffs

	yellow	**blue**
Y	20	15
B	15	20

In this example, if you are a yellow player and alternative Y is chosen, or if you are a blue player and alternative B is chosen, your payoff is 20 $. However, if you are a yellow player and alternative B is chosen, or if you are a blue player and alternative Y is chosen, your payoff is 15 $.

The payoffs may change from round to round.

In addition, and independently of roles and decisions during the experiment, you earn a show-up fee of 10 $.

4. The choice of the alternative Y or B

The group's decision between Y and B is made in two stages.

In the first stage, the **yellow players** collectively choose a **decision rule**. There are two possible decision rules:

Rule 1 Both the Yellow and the Blue players individually vote for either Y or B. The **computer randomly draws one of the players in the group and implements this player's choice**. All players in the group are equally likely to be drawn.

Rule 2 Both the Yellow and the Blue players individually vote for either Y or B. **The alternative with the largest number of votes is implemented.** In case of a tie, the computer selects the alternative randomly, with both alternatives being equally likely to be implemented.

The rule choice is made as follows. Each yellow player individually votes for Rule 1 or Rule 2. The computer randomly draws one of the yellow players and implements this player's choice. All yellow players in the group are equally likely to be drawn.

In the second stage, the final alternative, either Y or B, is chosen according to the decision rule that was selected in the first stage.

This procedure is repeated 30 times.

5. Summary

1. The computer randomly matches the 18 participants in 3 groups of 6 players.
2. The computer randomly draws a number from 1 to 10. All numbers are equally likely. If the number is below 10, then the state is state 1. If the number is ten, then the state is state 2. You are not informed about the state.
3. The computer randomly allocates the roles of a yellow player and a blue player. In state 1, 4 out of 6 players in the group are yellow, and the remaining 2 players are blue. In state 2, 4 out of 6 players are blue, and the remaining 2 players are yellow. You are informed about your own role. You are not informed about the roles of the other players.
4. The computer randomly selects a payoff structure. This payoff structure is displayed on the screen.
5. In each group, the yellow players collectively select Rule 1 or Rule 2.
6. All players in the group are informed about the rule.
7. Alternative Y or B is chosen according to the rule.
8. Your payoff in the round is determined by the rule, the votes and the payoff structure.
9. All players in the group are informed about the final outcome.

6. The end of the experiment

At the end of the 30th round, the computer randomly draws a round that determines your earnings from the experiment. All rounds are equally likely of being selected. You will be informed of your final earnings. You will not learn anything about the earnings of the other participants.

Please remain seated and wait quietly until we call you by your identification number (the number of your computer place). Please come when called, and you will be paid out your total earnings in private.

If there was anything you did not understand, please let us know by raising your hand. We will answer your questions on an individual basis.

Thank you for participating!