Minority Protection in Voting Mechanisms – Experimental Evidence

Dirk Engelmann (HU zu Berlin)
Hans Peter Grüner (University of Mannheim)
Timo Hoffmann ()
Alex Possajenikov (University of Nottingham)

Discussion Paper No. 484

December 17, 2023

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

Minority Protection in Voting Mechanisms – Experimental Evidence*

Dirk Engelmann[†] Hans Peter Grüner[‡] Timo Hoffmann[§] Alex Possajennikov[¶]

December 17, 2023

Abstract

Under simple majority voting an absolute majority of voters may choose policies that are harmful to minorities. It is the purpose of sub- and super-majority rules to protect legitimate minority interests. We study how voting rules are chosen under the veil of ignorance and whether there are systematic biases in these choices. In our experiment, individuals choose voting rules for given distributions of gains and losses that can arise from a policy, but before learning their own valuation of the policy. We find that subjects on average adjust the voting rule in line with the skewness of the distribution. As a result, a higher share of the achievable surplus can be extracted with the suggested rules than with exogenously given simple majority voting. While the rule choices are not significantly biased towards under- or overprotection of the minority, towards majority voting or towards status-quo preserving rules, they only imperfectly reflect the distributions of benefits and costs. In expectation this leads to only 63% of the surplus being extracted. The participants are heterogeneous with respect to how well their rule choices adapt to the distribution of valuations, with a large share of the surplus loss caused by a small group of participants.

JEL-codes: D72; C91; *Keywords:* minority protection; voting; experiments

*We thank Matthias Lang, Cesar Martinelli, and participants at various seminars and conferences for helpful comments. Financial support from the German research foundation (DFG), SFB 884 and SFB TRR 190 is gratefully acknowledged. The experiment satisfied the rules of the experimental laboratory LERN at the University of Nuremberg where the experiment was conducted. Satisfying the rules, which ensure the usual standards in experimental economics, in particular informed consent of the participants and no deception, is a pre-condition for running experiments at LERN. These rules have been validated by an academic panel.

[†]Humboldt-Universität zu Berlin, School of Business and Economics; CESifo, Munich; CERGE-EI, Prague. E-mail: dirk.engelmann@hu-berlin.de

[‡]University of Mannheim, Department of Economics; CEPR, London. E-mail: gruener@uni-mannheim.de

[§]At the time the experiment reported in the paper was conducted Timo Hoffmann was affiliated with the Department of Economics at University Erlangen-Nuremberg. E-mail: Timo.Hoffmann@fau.de

[¶]University of Nottingham, School of Economics. Email: alex.possajennikov@nottingham.ac.uk

1 Introduction

When a simple majority of voters decides, it may choose policies that are harmful to the rest of the electorate. An outcome that is supported by only a narrow majority of voters can be socially undesirable when individual losses are particularly large.[1] In such cases, a super-majority rule can protect legitimate minority interests. But in order to support efficient decisions, the voting rule needs to fit the underlying problem. This choice is not a trivial task. It requires an evaluation of the expected gains and losses arising from a policy proposal. Reforms that are likely to affect individual losers more than individual winners should only be undertaken if a sufficiently large super-majority of voters benefits. Instead, when potential gains exceed potential losses, even a minority of winners should get its way. In the present paper, we study whether experimental participants manage to efficiently protect minorities or exhibit systematic biases in a setting where they can choose the voting rules at a stage when they are informed about the distribution of possible gains or losses but are not yet informed about their own net valuations of a proposal.

In practice, sub- or super-majority rules (including unanimity rule) are often used to protect minorities (Vermeule, 2005). Examples are particularly widespread in small decision bodies such as boards and committees including boards of sportsclubs[2], homeowner assemblies[3], national parliamentary committees and boards of international organizations[4]. While supermajority rules are frequently used in important decisions of relatively small decision bodies, simple majority rule is often considered the natural decision rule. It also plays a prominent role in large elections and referenda, including in cases where it is not unlikely that individual minority members may be more strongly affected by the outcome than those of the majority. Examples include the 2013 referendum in which a majority of Croatian voters decided that the constitutional definition of a marriage should apply exclusively to "a living union of a woman and a man"[5], the 2016 Brexit referendum, in which a narrow majority of British voters triggered the country's exit from the European Union, or parliamentary decisions about banning specific types of food that some

[1]The phenomenon that the majority takes decisions that are very harmful to a minority has been recognized early by John Adams who coined the term "tyranny of the majority" (Adams, 1788, p. 291).

[2]E.g., recalling a member of the supervisory board of soccer club SV Waldhof Mannheim during their tenure requires a 2/3-majority of the members present at the general meeting.

[3]According to the German "Wohneigentumsgesetz" (WEG), decisions in favor of maintenance repairs require a double qualified majority in the committee of apartment owners. Specifically, a majority consisting of two-thirds of the votes cast and half the co-ownership shares is required (§ 21(2) WEG).

[4]An example is the European Stability Mechanism (ESM) that supports European countries that are in financial difficulties. Voting rights of each ESM member state are proportional to the country's share of the capital stock of the ESM. Some ESM decisions require a supermajority of 80 percent, others one of 85%.

[5]The Guardian, 2013 - https://www.theguardian.com/commentisfree/2013/dec/04/croatia-gay-marriage-vote-europe-rotten-heart.

individuals may actually care about very strongly.[6] The widespread use of the simple majority rule raises the question whether those who select rules understand the decision problem that they are facing.

To see why the majority threshold should be adjusted to the underlying distribution of preferences, consider a binary voting decision between two alternatives, a change A and a status quo B. Intuitively, a vote in favor of one alternative should count more if those who benefit from that alternative on average benefit more than those who lose. Thus, the voting threshold for one alternative should generally decrease in the expected preference intensity of the supporters and increase in the expected preference intensity of the adversaries. Majority requirements that deviate from simple majority can thus effectively protect minorities in cases where preferences in favor of or against a decision may be particularly strong.[7]

The observation that institutions should fit citizens' preference intensities goes back to at least Buchanan and Tullock (1962). The formal analysis of voting setups with decentralized information about preferences was pioneered by Rae (1969), who analyzed voting problems with binary positive or negative valuations. Rae showed that in symmetric setups the optimal voting rule is the simple majority rule.[8] In later work, Schmitz and Tröger (2012), Azrieli and Kim (2014) and Drexl and Kleiner (2018) have shown that qualified majority rules maximize social welfare in the class of anonymous social choice mechanisms, also for asymmetric distributions of gains and losses.[9,10]

While super-majority rules are justified in theory and are used in practice, very little is known about what exactly motivates people to select them. Are super-majority rules chosen

[6]Simple majority rule is actually applied so often that sometimes democracy – the rule of the people – is identified with the rule of a simple majority. According to Rae (1969, p. 40), the widespread limitation to some prominent decision making rules "is illustrated by Abraham Lincoln's remark: 'Unanimity is impossible; the rule of a minority as a permanent arrangement, is wholly inadmissable; so that, rejecting the majority principle, anarchy or despotism in some form is all that is left.'"

[7]We make this point in more detail in Section 2. We note that we are concerned not with groups that are likely to be in the minority, but with actual realized minorities.

[8]Rae (1969)'s result was formally proven in Taylor (1969) and later generalized by Curtis (1972), Badger (1972), and Schofield (1972).

[9]The selection of voting rules and weighted voting have also been studied theoretically in Barberà and Jackson (2004, 2006). Grüner and Tröger (2019) study the optimal adjustment of voting rules in environments where voting is costly. See also Barberà (1979), Börgers and Postl (2009) and Gershkov et al. (2017) for related papers that study cases with more than two alternatives. Moldovanu and Rosar (2021) study optimal majority thresholds in a dynamic context when there is an irreversible option. A literature on strategic voting in committees beginning with Austen-Smith and Banks (1996) and Feddersen and Pesendorfer (1998) studies how voting rules should be adjusted to the distribution of information.

[10]Some novel voting mechanisms permit voters to express their preference intensities. One way to achieve this without the introduction of monetary transfers is to permit the storage of votes (Cassella, 2005).

to protect minority interests? If so, are the majority requirements strong enough to guarantee an efficient minority protection? Do individuals choose simple majority rule when gains and losses can be expected to be balanced? And do irrelevant aspects bias the choice of majority thresholds? The present paper addresses these questions with a controlled experiment to find out whether individuals choose appropriate majority thresholds when the rights of minorities should be protected. Since in many practically relevant cases the voting rules have to be chosen (long) before stakeholders' preferences materialize, our analysis focuses on these cases, i.e., subjects have to choose voting rules at a stage when their own preferences have not yet realized.[11] While such constitutional matters are often decided by experienced politicians after a long deliberation, the inexperience of experimental subjects in such matters is counterbalanced by the problem being much simpler than such problems are in practice. While noise naturally occurs in experiments, we deliberately chose a simple setting in order to detect possible biases in rule choice towards over- or under-protection of a minority, towards the simple majority rule or in favor of the status quo.

Our experiment implements the following two-step procedure about whether to enact a change to the status quo, such as whether to implement a public project. In the first step, individuals know possible benefits and costs that each group member can get from the project, and the probabilities of these gains and losses, but they do not yet know their own valuation of the project. Individuals suggest a voting rule (such as simple majority, super-majority, or even an extreme minority rule where one vote in favor is sufficient to enact the change). In the second step, individuals learn their own valuations of the project and vote about the implementation of the change according to one voting rule randomly selected from the suggested ones.[12] Thus, the experimental subjects vote on the outcome after receiving information about their own valuation but they have to decide on the voting rule before this uncertainty is lifted.

Theoretically, the expected-payoff maximizing number of votes required to enact the change increases in the expected cost of the opponents of the change and decreases in the expected gains of the supporters. We use payoff distributions that vary in their skewness so that the expected-payoff maximizing voting rule ranges across all possible thresholds from unanimity required to

[11]It is a key politico-economic insight that rules governing collective decision making should ideally be chosen before individual preferences about outcomes have realized. Rawls (1972) builds his theory of justice on the view that a fair system should maximize expected utility under a veil of ignorance, and Brennan and Buchanan (1985) argue that the establishment of an efficient system is more likely if the decision about the institution is taken *before* preferences have materialized. The present paper considers this case. The case where voters who are already informed about their preferences pick voting rules has been analyzed experimentally in Engelmann and Grüner (2017).

[12]Since we are interested mostly in the choice of the voting rules, individuals make rule suggestions for 21 distributions of benefits and costs, but vote for only 3 distributions, randomly selected.

enact the change to unanimity required to keep the status quo.[13]

Our main finding is that the suggested voting rules qualitatively follow the pattern of the payoff-maximizing rule and do not exhibit systematic bias. We find strong evidence for a monotonic relationship between the relative preference intensity of opponents and supporters and the chosen voting rule. The response of the chosen voting rule to the relative preference intensities is, on average, weaker than would be optimal, though. While many subjects respond in their rule choice to the underlying distribution of valuations by picking more extreme voting rules for more skewed distributions, fewer than half of the rule choices are for the payoff-maximizing rule. Deviations from the optimum reflect both under-protection and over-protections of minorities, and – sometimes – misguided protection of minorities opposite the ones that should be protected (i.e. those for whom gains or losses are weaker than that of a majority).

On average, though, there is no general tendency towards over-protecting or towards under-protecting the minority. Neither do we find a clear bias in our experiment in favor of the simple majority rule, which could have been expected if participants equated simple majority voting with democracy and valued democracy per se. There is only weak evidence that rule choices are overall biased towards conservative rules (i.e., those in favor of the status quo). Such a bias would be predicted by risk aversion or inequality aversion. We also find only weak support that increased variance in possible payoffs leads to more conservative rule choices, which would be in line with risk aversion. We do observe, however, that participants choose more conservative rules if the probability of a negative outcome increases, which is inconsistent with the theory, but arguably psychologically plausible. Overall, the absence of systematic aggregate bias suggests that the deviations from optimal rules are due to a combination of different individual biases and noise rather than any widely-shared bias in preferences.[14]

[13] We note that it is essential for the investigation of the selection of expected total-payoff maximizing rules that at the rule-choice stage only the distribution of possible valuations is known. An alternative design could at the rule-choice stage already fix the distribution of payoffs and only leave uncertain how these payoffs are matched to individuals. This procedure would have eliminated aggregate risk because it would have been known for sure which decision would increase total payoffs. The rule that maximizes an individual's expected payoff could still deviate from the simple majority. For example, if in a group of five, two winners each gain twice as much as three losers and assuming selfish voting in the second stage, one would maximize expected payoff (and total payoff) if two votes suffice to implement a decision. However, in this setting participants who are concerned with the maximization of total payoffs can also achieve this aim by non-selfish voting in the second stage, rendering the rule-choice stage potentially less important. Moreover, even though in our experiment participants chose rules for each choice problem, what we are really interested in and aim to study with this design is whether people chose the efficient rule for certain classes of problems where they have an idea about the relative gains and losses but not already of the realized share of winners and losers.

[14] As is typically the case when experimentally testing equilibrium predictions, our experiment amounts to a simultaneous test whether preferences are in line with the model's assumptions and whether participants are making optimal choices given their preferences. Our results suggest that deviations are rather due to individual

That the suggested rule choices frequently deviate from the optimal rule implies that on average more than one third of the expected total surplus would be lost.[15] Thus, our experimental participants largely understand and follow incentives to choose an appropriate voting rule but there are still noticeable surplus losses even in our "ideal" setting, where (i) the rule can be adjusted to the underlying problem and (ii) agents are ignorant about their own preferences when they choose voting rules.

We observe substantial heterogeneity across subjects. While more than half of the subjects show a strong positive correlation between the suggested and the payoff-maximizing rule (with one subject always suggesting the optimal rule), for other subjects the correlation is weaker. Some subjects choose more often conservative rules, some suggest rules closer to the simple majority rule, and some more often choose rules favoring change. For a few subjects the correlation between suggested and optimal rules is even negative, due to frequent suggestions of misguided minority protection. The rule suggestions of these subjects lead to particularly high surplus losses, sometimes higher than those from random rule suggestions.

Our experiment is a direct empirical test of subjects' ability to perform the task of rule selection at the ex-ante stage.[16] While there is a lot of research about what determines individuals' voting behavior under a given rule (see Martinelli and Palfrey, 2020, for a survey of experimental results), very little is known about how individuals choose rules that they would like to apply in the future. Specifically, very little is known about whether individuals are willing and able to properly adjust the majority threshold to the underlying distribution of voter preferences. One exception is the analysis by Engelmann and Grüner (2017) who study the choice of voting thresholds at the interim stage, i.e. when individual preferences have already realized. Their main finding is that efficiency concerns may make individuals choose rules that are not in their own favor, which can make a rule-choice stage welfare enhancing even if preferences have already realized. While interim rule choices are made in many practically important cases, it is equally important to understand whether individuals are capable of making choices that maximize expected total payoff under the closer to ideal conditions when they do not yet know their own preferences or whether they exhibit systematic bias in such choices.[17]

choices not being expected-payoff optimal than due to common preferences that aim at something other than expected-payoff maximization.

[15]This calculation is based on the assumption that all subjects would have voted sincerely (for the change if the subject's valuation is positive and against it if the valuation is negative) in the second stage, which they actually do about 85% of the time. Insincere voting leads to an additional 5-15% loss of surplus.

[16]In this, it is one of the few studies where subjects play the role of a mechanism designer. In our setup, subjects select a rule from a well-defined set of voting rules (mechanisms), before individual preferences are determined. Hoffmann and Renes (2022) study both ex-ante and interim choices from a set of four alternative mechanisms: the simple majority rule, the expected externality mechanism, the "no-change" rule and a random choice of outcome.

[17]Weber (2020) also experimentally studies the choice of voting rules (for representatives of differently sized

The next section outlines the theoretical argument how majority thresholds should be adjusted to the underlying distribution of valuations. Section 3 presents the experimental design and Section 4 the results. We conclude with a discussion in Section 5.

2 The voting problem

This section introduces the setup underlying our experiment. We also explain why and how a voting threshold should generally change with the underlying distribution of valuations for the alternatives.

Consider a population of finite size n that has to take a binary voting decision between two alternatives, A and B. Normalizing all players' payoffs resulting from alternative B to zero, we represent a player's preference by a payoff θ resulting from alternative A.[18] For each individual, the valuation θ is independently drawn from the same commonly known distribution on $\Theta \subset \mathbf{R}$ with zero mass on zero. Thus, indifference is excluded and there are strict winners or losers with potentially different preference intensities within and between these two groups.[19]

Voters must cast a vote for A or B. Abstentions are not allowed.[20] An anonymous voting mechanism maps the number of votes in favor of alternative A (which implies the number of votes in favor of alternative B) into a probability that A gets realized. An important subclass of such mechanisms consists of threshold voting rules which assign outcome A with probability 1 to any voting profile in which at least k voters voted in favor of A and outcome B with probability 1 to all other voting profiles.[21]

groups), but does so in a setup with binary valuations of identical absolute size. Both his paper and Hoffmann and Renes (2022) find that choices are generally different in ex-ante and interim stages. Similar findings appear in the literature on the choice of electoral rules (see, for example, Blais et al., 2015, and Bol et al., 2020), where supporters of a poitical party (or a candidate) are likely to prefer a rule that benefits their party but there is also support across the board for "fairer" rules.

[18]Of course, this is a simplification. Arguably, though, the status quo serves as a plausible reference point and alternative payoffs are likely evaluated by their deviation from this reference point, in line with our normalization. Furthermore, this simplification only affects the predictions based on inequality aversion.

[19]Excluding indifference is convenient for not dealing with multiple optimal choices in the voting stage, but the optimal mechanism does not depend on its exclusion (Schmitz and Tröger, 2012).

[20]The optimal voting mechanism below requires only two signals, A and B. If there are costs of voting, then a third signal – abstension – would be necessary for optimality (Grüner and Tröger, 2019). In our experiment voting is costless, thus only two signals are needed.

[21]These voting rules are often referred to as qualified majority rules. The importance of this class is that mechanisms from this class are optimal, in the sense of maximizing the expected total ex-ante payoff of n voters, among all anonymous strategy-proof mechanisms without transfers (Nehring, 2004; Schmitz and Tröger, 2012; Azrieli and Kim, 2014). Even with transfers, such mechanisms are optimal (among deterministic mechanisms) for "regular" distributions of valuations (Drexl and Kleiner, 2018).

An individual's voting strategy in a mechanism is a mapping $S : \Theta \to \{A, B\}$. Always voting in line with the sign of one's own valuation is a weakly dominant strategy for any possible threshold voting rule. Based on this voting behavior, the voting rule that maximizes the expected total payoff is as follows.[22] Denote conditional expected monetary gains of winners by $E^+ := E[\theta \,|\theta > 0]$ and conditional (absolute) losses by $E^- := |E[\theta \,|\theta < 0]\,|$.

Observation 1 *Among threshold voting rules, the rule that maximizes the expected (total and individual) payoff in the population of n voters is the rule with threshold k, where*

$$k = \left\lceil \frac{nE^-}{E^- + E^+} \right\rceil .$$

To see the intuition for this result, suppose that the realizations of valuations are such that the realized number of winners and losers are a and b. A payoff-maximizing rule specifies that a decision in favor of alternative A is made if $aE^+ > bE^-$ (and in favor of alternative B if $aE^+ < bE^-$; when $aE^+ = bE^-$, any outcome is optimal). Since zero valuations occur with probability zero,

$$aE^+ > bE^- \Leftrightarrow aE^+ > (n-a)E^- \Leftrightarrow a(E^+ + E^-) > nE^- \Leftrightarrow a > \frac{nE^-}{E^- + E^+}.$$

Hence, if the threshold is $k = \lceil nE^- / (E^- + E^+) \rceil$, then for all realizations of a and b for which $a > nE^- / (E^- + E^+)$, the rule selects A and for all other realization it selects B. Since rule k maximizes the total payoff for all realizations of a and b, it also maximizes the expected total payoff ex ante, before the realization of a and b.[23]

Note that the optimal choice of the voting mechanism does not depend on the probabilities of voters preferring A or B but only on the conditional gains and losses. This may appear counter-intuitive at a first glance, because one might think that if losses are more likely, one may want stronger protection of losers and hence a higher threshold. That intuition is false, however, because the number of voters in favor or against a policy change depends on the realized numbers of negative and positive valuations and not on the ex-ante expected numbers. Hence,

[22] For deterministic gains and losses, the result follows from the analysis in Curtis (1972) and Badger (1972). For expected gains and losses, it is stated in Nehring (2004), Azrieli and Kim (2014) and Drexl and Kleiner (2018) in the context of more general mechanisms. It also follows from Corollary 1 in Barberà and Jackson (2006). Note that the same rule maximizes the expected individual payoff of a player, since the total expected payoff is the sum of the expected payoffs of (ex-ante) symmetric players.

[23] On the other hand, a risk-averse individual, or an individual who cares about inequality in the distribution of payoffs, may prefer a different rule. Also, the optimal rule is derived under the assumption that voting follows the dominant strategy of voting for A if the realized valuation is positive and for B if the realized valuation is negative. If voting is not always consistent with realized valuations, then different rules may be optimal.

the latter are irrelevant for the determination of optimal voting rules. Also, the optimal rule does not depend on the whole distribution but only on expected gains and losses E^+ and E^-. For example, a distribution with $\theta = 2$ with probability p and another with $\theta = 1$ with probability $p/2$ and $\theta = 3$ with probability $p/2$ (and the rest of values and probabilities the same as in the first distribution) have the same optimal rule. We use the possible tension between the intuitive choice and the theoretical optimum to inform our choice of distributions for the experiment.

3 Experimental design and hypotheses

3.1 Design

Our experiment considers the following two-stage decision setup, consisting of a rule-choice stage and a voting stage. In the rule-choice stage, experimental participants are given a probability distribution of possible valuations (positive and negative) for alternative A. This distribution is the same for each member of the group of five individuals. The valuation for alternative B is always 0 € for all individuals. The distribution of valuations for alternative A differs across rounds. For example, it can be

- the valuation of A is 2 € with probability 2/3 and it is -5 € with probability 1/3.

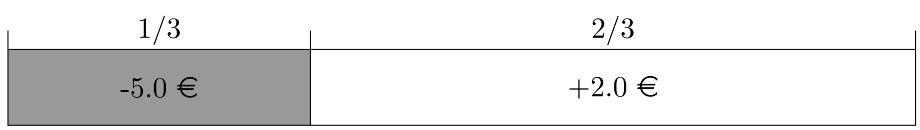


Figure 1: Illustration of a distribution of valuations

For a given distribution, each individual chooses one of the threshold voting rules, specifying how many individuals in the group of five need to vote for alternative A for it to be adopted. Since abstentions are not allowed, the available voting rules for groups of five voters are:

Rule I. At least 1 vote for alternative A is required for A to be chosen, thus 5 votes for alternative B are required for B to be chosen (unanimity for B);

Rule II. At least 2 votes for alternative A are required for A to be chosen, thus at least 4 votes for alternative B are required for B to be chosen (qualified majority for B);

Rule III. At least 3 votes are required for either A or B to be chosen (simple majority), that is, whichever alternative has more votes wins;

Rule IV. At least 4 votes for alternative A are required for A to be chosen (qualified majority for A), hence at least 2 votes for alternative B are required for B to be chosen;

Rule V. 5 votes for alternative A are required for A to be chosen (unanimity for A), thus at least 1 vote for alternative B is required for B to be chosen.

In the second stage of the experimental setup, the rule suggestion of one randomly chosen group member is taken to be the actual voting rule (a random dictator mechanism).[24] The subjects are informed about which rule was chosen, but not whose decision determined the voting rule nor the voting rule choices of the other four group members. The participants' valuations are then realized according to the given distribution and each participant learns his/her own valuation for alternative A. The participants then cast a binary vote (either for A or for B). The votes are tallied and the outcome (either A or B) is decided according to the chosen voting rule.

The two-stage procedure is designed to have individuals make decisions on the voting rule under the "veil of ignorance" (in the first stage, before they know their own valuation). The second stage is the one more commonly tested in the experimental literature on voting (see Martinelli and Palfrey, 2020), and we include it as a check on subjects' voting behavior and to make the rule-choice stage incentive compatible. However, our interest is mainly in the decisions in the first stage.

In the second stage of the procedure, it is a (weakly) dominant strategy to vote for A if one's realized valuation for A is positive and vote for B if the realized valuation is negative. If voting in the second stage is going to follow the dominant decisions, and if individuals maximize their expected payoff (or the expected payoff of the whole group), then the optimal rule is described in Observation 1. In particular, the optimal rule depends only on the ratio of the expected gains and losses in the distribution. For example, for the distribution shown above (2 € with probability 2/3 or −5 € with probability 1/3) the optimal rule Rule IV is skewed towards B: four votes for A are needed.[25] The benefits from A are lower than the losses that it causes and the players with losses need to be protected from an ex-ante point of view. In the example these players are likely to be in the minority (the probability of a negative value is 1/3), but it is actually irrelevant how likely negative (or positive) values are. Whenever two participants have the negative valuation −5 €, this outweighs three positive valuations 2 € and hence A should only be chosen if at least four participants support it. Since in the first stage individuals make a decision which rule to suggest before knowing their own valuation for alternative A, it is in

[24] We chose a random dictator mechanism because it is incentive compatible and moreover easy for participants to understand to be incentive compatible.

[25] According to the observation, the optimal threshold is $k = \left\lceil \frac{5 \cdot 5}{5+2} \right\rceil = \left\lceil \frac{25}{7} \right\rceil = 4$.

their own interest to suggest such a protection of the minority.

In the experiment, the two stages did not immediately follow each other. Instead, in the first part of an experimental session, the participants played the first stage (rule choices) for 21 rounds, with 21 different distributions of valuations for alternative A. These distributions are listed in Table 1. The order in which the distributions were shown to the subjects was randomly determined and thus varied between subjects; there was no feedback between rounds in the first part of the experiment.

The distributions are chosen to vary in their skewness such that the total-payoff maximizing voting rule ranges across all possible thresholds from unanimity required for alternative A (rule V) to unanimity required for alternative B (rule I). Distributions 1-5 are taken as the base; the rest of the distribution are derived from them. For example, distributions 6-11 are variants of distributions 1-5, but with different probabilities of each value. Distributions 12-16 are variants of distributions 1-5 but with increased variance, with one of the outcomes in distribution 1-5 being replaced by its mean-preserving spread. Finally, distributions 17-21 are derived from 12-16 by multiplying all valuations with -1.[26]

After the first part was finished, three of the 21 distributions were randomly selected, independently across sessions, for the second part of an experimental session.[27] In each round of the second part, the chosen distribution was shown to the participants and the valuation of each participant was drawn according to the distribution. The participants were informed about their own realized valuation but not about the realized valuations of others. Groups of five participants were randomly formed in each round and one voting rule among those suggested by the five group members for this distribution was randomly selected. The participants were informed about which rule was selected in their group and voted for alternative A or alternative B. The outcome of the voting in a group was then determined according to the voting rule. At the end of a round, the participants were informed about the outcome of the voting and their payoff. They were paid for all three group decisions from the second part.

The experiments were conducted at the Laboratory of Experimental Research Nuremberg (LERN) in December 2015. We ran 5 sessions, with the number of participants ranging between 15 and 30 in each. In total there were 130 participants. The experimental sessions were pro-

[26]Following this rule, distribution #17 should have been -1.5 with probability 1/3, -0.5 with probability 1/3, 5 with probability 1/3, with the optimal rule being rule I. Due to a copying error, 0.5 was entered instead of -0.5 in the experimental software, making rule II the optimal rule.

[27]To make sure that subjects could not lose money, we divided the distributions into three groups, according to the largest possible negative value (distributions 1, 2, 6, 7, 12, 20, 21 in group 1; distributions 3, 8, 9, 13, 14, 18, 19 in group 2; distributions 4, 5, 10, 11, 15, 16, 17 in group 3), and then randomly selected one distribution from each group. With this procedure, the probability that a given distribution is selected for the second part is the same as with unrestricted selection of triples of distributions.

Table 1: Distributions used in the experiment

No.	V_1	V_2	V_3	$Pr(V_1)$	$Pr(V_2)$	$Pr(V_3)$	E^-	E^+	Optimal Rule
				Baseline Distributions					
1	−5	1		1/3	2/3		5	1	V
2	−4	1.5		1/3	2/3		4	1.5	IV
3	−2.5	2.5		1/2	1/2		2.5	2.5	III
4	−1.5	4		2/3	1/3		1.5	4	II
5	−1	5		2/3	1/3		1	5	I
				Modified Probabilities					
6	−5	1		1/2	1/2		5	1	V
7	−4	1.5		1/2	1/2		4	1.5	IV
8	−2.5	2.5		1/3	2/3		2.5	2.5	III
9	−2.5	2.5		2/3	1/3		2.5	2.5	III
10	−1.5	4		1/2	1/2		1.5	4	II
11	−1	5		1/2	1/2		1	5	I
				Mean-preserving Spreads					
12	−5	0.5	1.5	1/3	1/3	1/3	5	1	V
13	−4	1	2	1/3	1/3	1/3	4	1.5	IV
14	−2.5	2	3.5	1/2	1/3	1/6	2.5	2.5	III
15	−1.5	3	5	2/3	1/6	1/6	1.5	4	II
16	−1	3.5	6.5	2/3	1/6	1/6	1	5	I
				Inverted Distributions					
17	−1.5	0.5	5	1/3	1/3	1/3	1.5	2.75	II
18	−2	−1	4	1/3	1/3	1/3	1.5	4	II
19	−3.5	−2	2.5	1/6	1/3	1/2	2.5	2.5	III
20	−5	−3	1.5	1/6	1/6	2/3	4	1.5	IV
21	−6.5	−3.5	1	1/6	1/6	2/3	5	1	V

grammed using zTree (Fischbacher, 2007) and the recruitment of the participants was done with ORSEE (Greiner, 2015). Each participant was given a starting budget of 15 €. The valuations in Table 1 are in Euro; with the three distributions actually played out, the minimum amount a participant could earn was 3 € and the maximum amount was 27 €. The full experimental instructions are available in Online Appendix A.[28]

3.2 Hypotheses

Our experimental setup permits to test the following hypotheses about individuals' motives underlying the choice of voting rules:

1. **The distribution of gains and losses matters:** Voting rule choices take into account the skewness of the distributions towards larger positive or negative outcomes, thus reflecting which rules are optimal.

2. **Preference for simple majority rule:** Rule III is chosen more often than is warranted by the theoretical prediction of optimal rule choice. Such a preference could result from a preference for democracy and a perception that simple majority voting best represents democracy.

3. **Asymmetry:** There is a systematic bias towards rules IV and V as compared to rules I and II. This may reflect risk attitudes: since alternative A is more risky than alternative B, a risk-averse person would suggest rules IV and V more often. A person with maxmin preferences should always pick rule V, since for each distribution losses are possible. A bias towards conservative rules would also result from inequality aversion.

4. **Variance matters:** Although distributions that differ only by one outcome being replaced by a mean-preserving spread (for example, distributions 1 and 12 in Table 1) have the same theoretically optimal rule, decisions in the experiment may not reflect this. A distribution with a higher variance is less attractive for a risk-averse person, who should hence choose more conservative rules for a distribution derived by a mean-preserving spread. Inequality aversion also predicts a stronger bias towards conservative rules for more spread-out distributions.

5. **Probabilities matter:** Although distributions with the same outcomes but different probabilities of these (for example, distributions 1 and 6 in Table 1) have the same theoretically optimal rule, decisions in the experiment may not reflect this. Specifically, having

[28]Control questions about understanding of decisions in the experiment were asked via the experimental software. In one session, the control questions were accidentally skipped. The results from this session do not differ much from those of the others; we comment on the differences when they arise.

a higher probability of a loss may make a distribution appear to be more risky and hence lead to more conservative rule choices.

A strict form of Hypothesis 1 implies that participants always choose the optimal rule. A weaker form of this hypothesis, namely that the chosen rule correlates with the optimal rule, implies that participants take the need to protect minorities to some degree into account but that not all participants do this perfectly and at least some deviate from the optimal rule either due to preference-driven biases or errors. Hypotheses 2, 3, and 4 address the presence of a bias in the rule choice, either due to a generic preference for simple majority voting or due to risk aversion or inequality aversion. Hypothesis 5 addresses a reason for systematic errors. In addition, there can of course be other sources or error resulting, for example, from misunderstanding of the rules or miscalculation.

4 Experimental results

4.1 Rule choices - aggregate data

4.1.1 Minority protection

We have choices of voting rules of each of the 130 subjects for each of the 21 distributions of valuations, making a total of 2730 choices. Table 2 shows the summary of these rule choices, separated by which rule is optimal for the given distribution (under the assumption that voting in the second stage follows the dominant strategy of voting for A if the realized value is positive and voting for B if the realized value is negative). The last two rows of the table show the overall distribution of rule choices and the distribution of optimal rule choices. Figure 2 shows the overall distribution of rule choices by optimal choice graphically (along with a classification of choices, discussed further below).

To test Hypothesis 1 (distributions of gains and losses matter), we analyze how the rule choices of the subjects depend on the optimal rule. It is evident from both Table 2 and Figure 2 that subjects on average suggest higher majority thresholds when higher thresholds are optimal but suggested rules often deviate from the optimal one. The line in Figure 2 represents the result of a linear regression of the suggested rule on the optimal rule. The estimated relationship is

$$SuggestedRule = 1.783^{***} + 0.432^{***} OptimalRule \qquad (2730 \text{ obs}; R^2 = 0.167). \tag{1}$$

The regression coefficient is significantly different both from 0 and from 1.[29] Hence there is

[29] Since we did not provide any feedback in the first part of the experiment, each of our subjects represents an independent observation. The standard errors are clustered on the subject level to account for multiple observations per subject. Including session dummy variables does not identify significant differences across sessions.

Table 2: Aggregate rule choices

Optimal Rule	Obs	Median	Mean (St.Dev.)	Rule choices I	II	III	IV	V
I	390	2	2.254 (1.330)	39%	25%	17%	8%	11%
II	650	2	2.594 (1.261)	24%	27%	24%	15%	9%
III	650	3	3.080 (1.189)	11%	20%	35%	19%	15%
IV	520	4	3.598 (1.297)	9%	11%	22%	24%	33%
V	520	4	3.896 (1.348)	10%	9%	13%	21%	48%
Total	2730	3	3.100 (1.401)	17%	19%	23%	18%	23%
Optimal		3	3.048 (1.327)	14%	24%	24%	19%	19%

qualitative support for Hypothesis 1, but in aggregate subjects do not fully adjust their suggestions for the majority threshold to variations in the optimal threshold.[30] It also appears that the suggested rules are closer on average to rule III, the simple majority rule, than the optimal rules are.

Is the simple majority rule III then particularly preferred by our subjects? From comparing the last two lines in Table 2, there does not appear to be support in the aggregate data for Hypothesis 2 that subjects exhibit a specific preference for the simple majority rule III. In fact, rule III is chosen (slightly) *less* often than it should be according to the optimality criterion (23% instead of 24% of the times).[31]

Even though the simple majority rule does not have a clear bias in its favor, subjects might still favor majorities and do not give minorities the optimal protection. To get an idea whether this is the case, we distinguish the following categories of deviations from the optimal rule choice. Note that we do not distinguish here between minorities that prefer change (A) and those that prefer the status quo (B).

1. **Minority under-protection** ("*tyranny of the majority*"): a minority of m voters would deserve protection in the sense that the total surplus is maximized if these m voters were

[30] We observe a slight time trend towards stronger adjustment in line with the optimal rule. If we restrict the regression to the first ten periods, the coefficient is 0.396, while for the last eleven periods it is 0.459. Hence there appears to be some learning, even though we did not provide any feedback.

[31] This small difference is actually statistically significant according to a sign test whether the median value of the proportions of choices of rule III that each subject makes is equal to the one predicted by optimality (which is $5/21 = 0.238$): most subjects ($82/130 = 63\%$) chose rule III *fewer* than 5 times (p-value 0.004). The mean of the subjects' proportion of choices is not statistically significantly different from that predicted by optimality though, according to a t-test (p-value 0.692): there are a few subjects that chose rule III much more often than 5 times.

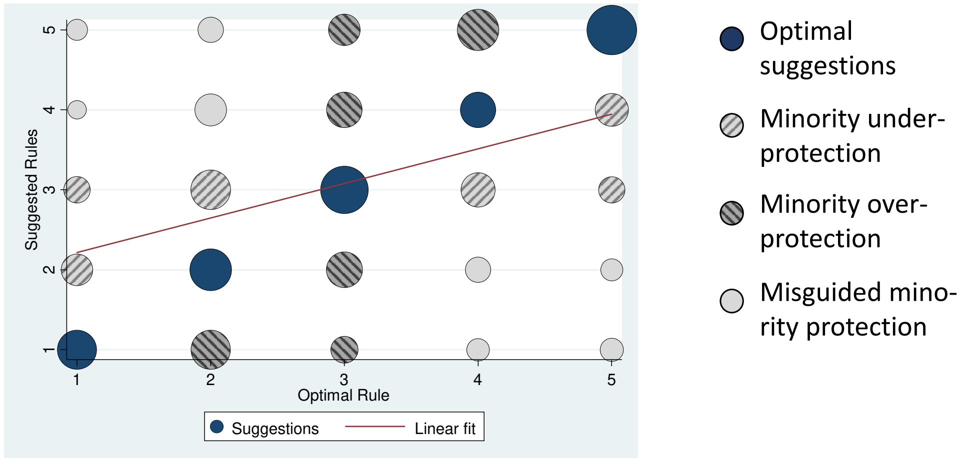


Figure 2: Rule choices in the experiment by optimal rule. The sizes of the circles are proportional to the number of observations for that combination of optimal rule and actually suggested rule.

allowed to determine the outcome, but the rule suggested is such that more than m votes are needed for the preferred outcome of this minority (but not so many that a minority preferring the other outcome would determine it: this would be the misguided minority protection below). In our experiment, this applies if the chosen rule is III while the expected-surplus maximizing rule is not, and if the chosen rule is II while the expected-surplus maximizing rule is I or if the chosen rule is IV while the expected-surplus maximizing rule is V. See Figure 2.

2. **Minority over-protection** ("*tyranny of the minority*"): either a rule is suggested where a minority can determine the outcome while a simple majority would be optimal, or the suggested rule is such that fewer than m votes are sufficient for the preferred outcome of a minority of m voters that deserves protection. In our experiment, this applies if the expected-surplus maximizing rule is III but the chosen rule is not, and if the expected-surplus maximizing rule is II but the chosen rule is I, or if the expected-surplus maximizing rule is IV, but the chosen rule is V. See Figure 2.

3. **Misguided minority protection**: the suggested rule is such that a minority of voters can determine the outcome, even though the maximal surplus is obtained if a minority of voters preferring the other outcome can determine it. In our experiment, this applies if the expected-surplus maximizing rule is I or II, but the chosen rule is IV or V, or if the expected-surplus maximizing rule is IV or V, but the chosen rule is I or II. See Figure 2.

In Figure 2, 23% of rule choices are classified as "tyranny of the majority" and 27% as "tyranny of the minority" (and 16% as misguided minority protection; the remaining 34% are optimal rule choices). These numbers can be misleading though since over-protection and under-protection of minorities are possible in different circumstances. Instead, we quantify the rates of decisions that are in line with under-protection and over-protection of minorities in the following way. For "tyranny of the majority", we consider all the observations for which it was possible to suggest a rule closer to the simple majority than the one maximizing the expected surplus (i.e., all the observations where the optimal rule was I, II, IV or V). Of these, we exclude observations in which misguided minority protection was suggested. Of the remaining observations, we count the proportion of observations involving minority under-protection. For "tyranny of the minority", we look only at those observations where it was possible (i.e. those with optimal rule II, III or IV), exclude observations with misguided minority protection and calculate among the remaining ones the proportion of those involving minority over-protection.

The results of the tyranny rates calculations are presented in Table 3. Overall, the "tyranny-of-the-minority" rate is higher than the "tyranny-of-the-majority" rate. Among the observations that do not suggest rules skewed towards a minority different from that implied by the optimal rule, rule choices on aggregate are actually biased towards stronger minority protection than necessary.[32] Note also that for rules II and IV, for which both under- and over-protection of minorities is possible, the under-protection rate is not much higher than the over-protection rate; in fact, over-protection is higher for rule IV. The apparent bias of average rule choices towards simple majority rule III in Figure 2 thus appears to be driven not by preferences for "more balanced" (closer to simple majority) rules but largely from the decisions we labelled as misguided minority protection.

While there is thus no bias towards rule III, is there a bias towards more conservative rules? Hypothesis 3 predicts asymmetry, i.e. that rules I and II would be chosen less often than rules IV and V. Although there is a slight shift to the rules with larger thresholds (rules IV and V are chosen 41% of the times while rules I and II only 36% of the times), this difference is not statistically significant.[33] There is thus no clear statistical support for Hypothesis 3 either.

Result 1: Aggregate rule choices and minority protection. Rule choices adapt to changes

[32] Even if the cases when the optimal rule is rule III and a different rule is suggested are excluded (which could also be interpreted as misguided minority protection because a minority can determine the outcome even though it should not), the overall "tyranny-of-the-minority" rate is 0.362, very similar to the "tyranny-of-the-majority" rate.

[33] Considering the difference between the numbers of choices of rules IV and V and the numbers of choices of rules I and II for each subject, neither the two-sided sign test (p-value 0.142) nor the t-test (p-value 0.113) find the median or mean of this variable to be significantly different from 0, even though for more subjects it is positive (68 subjects) than negative (51 subjects; for 11 subjects it is equal to 0).

Table 3: Tyranny rates by optimal rule

Optimal Rule	Tyranny of the majority rate	Tyranny of the minority rate
I	0.519	
II	0.323	0.317
III		0.648
IV	0.280	0.416
V	0.407	
Total	0.372	0.482

in the underlying distribution of valuations, although on average less than would be optimal. The aggregate data exhibit no evidence of a preference for simple majority rule III or for under-protection of minorities; over-protection of minorities is at least as common. There is also little conservatism bias.

4.1.2 Distributional changes and rule choices

The distributions of valuations in our experiment were chosen so that some of them are variants of others but with different probabilities of the same values, or with different variance. Theoretically, changing the probabilities or the variance while keeping the expected values of positive and negative valuations the same does not change which rule is optimal. Table 4 lists means of suggested rules for each distribution, organized in rows by the optimal rule (recall that distribution #17 was supposed to be a variant of distribution #5 but was not correctly implemented and is thus omitted from the table). The table also presents the results and p-values of two-sided sign-rank tests comparing the choices for each of the distributions to the corresponding base distribution in the left-most column. The distributions themselves can be found in Table 1. The first column of Table 4 shows the base distributions, the columns labeled Prob+ and Prob− present variants of the base distributions where the probabilities for positive or negative valuations, respectively, are increased. The columns labeled Var+ and Var− show results for valuation distributions with mean-preserving spreads, with increased variance of positive or negative valuations, respectively.

Hypothesis 4 states that the variance of distributions matters. From Table 4, support for this hypothesis is mixed, with tests finding significant differences for some comparisons but not for others. Differences are significant when the optimal rule is I and II and the variance of positive values is increased, or when the optimal rule is IV or V and the variance of negative values is increased. Moreover, in the former case the rule choices actually become less conser-

Table 4: Effects of changes in distributions on average chosen rules

Optimal Rule	Base	Prob+	Prob−	Var+	Var−
	D#5	D#11		D#16	
I	2.508	2.046***		2.208**	
		(0.001)		(0.038)	
	D#4	D#10		D#15	D#18
II	2.892	2.208***		2.677**	3.085
		(< 0.001)		(0.016)	(0.180)
	D#3	D#8	D#9	D#14	D#19
III	2.962	2.446***	3.746***	2.746	3.500***
		(< 0.001)	(< 0.001)	(0.114)	(< 0.001)
	D#2		D#7	D#13	D#20
IV	3.454		3.923***	3.308	3.708**
			(< 0.001)	(0.125)	(0.037)
	D#1		D#6	D#12	D#21
V	3.854		3.931	3.808	3.992*
			(0.198)	(0.576)	(0.063)

p-values in parentheses. *** - significant at 1%; ** - significant at 5%; * - significant at 10%.

vative, contrary to the hypothesis. It may be that not variance per se matters but rather the largest (in absolute terms) possible value. Distributions with optimal rules I and II have large positive valuations (relative to negative); increasing variance of them implies that the largest positive valuation is even larger (for example, D#5 has possible valuation 5; D#16 involves a mean-preserving spread of this valuation, with valuation 6.5 now possible). For distributions with optimal rules IV and V, negative valuations are large (relative to positive); increasing the variance of positive valuations still leaves them smaller in absolute terms than the negative valuations, and the effect on the choice of rules in the experiment is small. By contrast, increasing the variance of negative valuations increases the largest negative valuation (in absolute terms), making rule choices more conservative, with higher thresholds. The absence of clear support for Hypothesis 4 (together with little support for Hypothesis 3) suggests that neither risk aversion nor inequality aversion are important drivers of aggregate behavior in our experiment.

Hypothesis 5 states that probabilities of values matter although they should not in theory. Table 4 lends support to this hypothesis. Almost all tests in columns Prob+ and Prob− find significant differences in rule choices for the distributions with the same valuations but with different probabilities of those valuations (the only exception is the comparison of D#6 with D#1, for optimal rule V). The direction of differences is also clear: if the probability of positive values is increased, lower thresholds in favor of alternative A are suggested, while if the probability of negative values is increased, higher thresholds are suggested. Higher probabilities of positive values appear to make alternative A more attractive, while higher probabilities of negative values do so for the status-quo alternative B. While inconsistent with the theory, stressing gains or losses makes participants on average more or less daring, respectively, suggesting a source of systematic error.

Result 2: Aggregate effects of distributional changes. Mean-preserving spreads of large positive valuations lead to less conservative rule choices, and such spreads of large negative valuations lead to more conservative choices, but increased variance does not generally lead to more conservative rules. Changes in the probabilities of positive or negative valuations matter, making subjects suggest less conservative rules if the probability of positive valuations is increased and more conservative rules if the probability of negative valuations is increased.

4.1.3 Rule choices and expected surplus extraction

In our set-up there are two possible sources of inefficiency. First, realized average positive or negative valuations may deviate from the expected values, such that the ex-ante optimal voting rule is not optimal given the realized valuations.[34] Second, participants might not choose the ex-ante optimal rule, as indeed they frequently do, as shown in Figure 2. How much of the available surplus was lost due to these two distortions?

Supposing that in the voting stage the subjects would vote consistently with the sign of their valuation allows us to calculate, for each of the 21 distributions of valuations in our experiment, the expected surplus a group of voters would achieve for each voting rule. For each distribution, we calculate, averaging over all possible combinations of valuations arising in the distribution

- the "first-best" expected surplus S^{FB} as the sum of valuations if this sum is positive and 0 if the sum is negative,
- the "optimal" expected surplus S^o as the surplus if the voting rule were optimal and everyone voted according to their valuations,
- the "actual" expected surplus S as the expected surplus averaged over all rules actually suggested for this distribution and assuming again that everyone voted according to their

[34] The two-stage voting procedure can cause losses of expected payoffs if the realized positive or negative valuations differ from the respective conditional expectations. For example, for distribution #12, if there are four positive valuations 1.5 and one negative -5, then A would be the first-best choice, but the optimal voting rule V would lead to outcome B.

valuations.

The loss from the voting procedure per se is very small: if the optimal rule is used for each distribution, the average group expected surplus over the 21 distributions is $S^o = 3.11$. The average first-best expected surplus over the 21 distributions is $S^{FB} = 3.13$. Thus, with the voting procedure in our experiment, if subjects suggested optimal rules and voted according to the sign of realized valuation, they would in expectation only lose $1 - \frac{3.11}{3.13} = 0.006$, less than 1% of the first-best surplus. Since the difference between S^{FB} and S^o is very small, we take S^o as the appropriate benchmark for our experiment.[35]

Table 5 reports the surplus measures S^o and S averaged over distributions with the same optimal rule, as well as the aggregate measure for all 21 distributions (the last row Total). The table also reports how much of the expected surplus was lost because of the non-optimal rule suggestions. For comparison, the last column also reports the expected surplus S^{III} if the simple majority rule III were used for all distributions, and the expected loss arising from such a choice.

Table 5: Expected surpluses by optimal rule

Optimal Rule	Optimal S^o	Actual S	Actual loss $S^o - S$	Rule III S^{III}	Rule III loss $S^o - S^{III}$
I	7.16	5.54	1.61	4.62	2.54
II	4.71	3.61	1.10	3.77	0.94
III	2.55	1.90	0.65	2.55	0.00
IV	1.37	0.25	1.12	0.23	1.14
V	0.56	−0.93	1.47	−2.03	2.56
Total	3.11	1.97	1.14	1.82	1.29

As can be seen from the $S^o - S$ column, the loss is larger for the more extreme rules I and V, and is smallest for the distributions where rule III is the optimal rule. From the last two columns, if rule III were used for all distributions, it would lead to smaller losses for distributions with optimal rules II and III (and a very similar loss for IV), but would produce much larger losses for the more skewed distributions I and V.[36] Averaging over all distributions, the expected surplus for actually suggested rules is (slightly) higher than the surplus if rule III were always suggested: 1.97 compared with 1.82.

[35]Since many of the distributions of valuations in our experiment include only one positive and one negative valuation, situations where realized valuations differ from expected valuations (conditional on the number of positive and negative valuations) are indeed very rare in our data.

[36]This is of course to be expected, given that the chosen rules tend to adjust in the right direction and hence capture some of the surplus that the simple majority rule misses if the optimal rule is extreme, but also sometimes go in the wrong direction and hence lead to surplus losses not incurred with simple majority voting.

The ratio of the expected surplus with actually suggested rules to the potentially obtainable surplus can be called *surplus extraction rate* (SER).[37] For the expected surplus, the SER measure in our experiment is thus

$$SER^e = \frac{S}{S^o} = \frac{1.97}{3.11} = 0.634, \tag{2}$$

meaning that only 63% of the expected surplus would have been realized by our subjects if all of their suggestions were used in voting. These calculations assume voting according to the sign of the realized valuation in the second stage, and thus identify expected loss arising from suboptimal rule suggestions rather than from (possibly suboptimal) voting behavior.[38]

From the categories of rule suggestions represented in Figure 2, the largest expected surplus loss comes from misguided minority protection suggestions, as can be expected. The average loss from such suggestions is 4.23, much larger than the average loss 1.14 across all rule suggestions. Although cases of misguided minority protection constitute only 16% of all rule suggestions, they contribute 59% to the expected surplus loss. On the other hand, cases of minority under-protection make up 22% of surplus loss (average loss 1.12 for such suggestions), while cases of minority over-protection contribute only 18% of the expected surplus loss (average loss 0.77).[39]

Result 3: Expected surplus extraction. The (expected) surplus extraction rate from rule suggestions in the experiment is 0.634. Thus only 63.4% of the available surplus would have been realized in expected terms if all subjects' suggestions were used for voting rounds and voting followed the realized valuations. Most of the surplus loss arises from rules that protect a wrong minority that does not deserve protection.

4.2 Rule choices - individual data

4.2.1 Classification of subjects by rule choices

In order to obtain a deeper understanding of the aggregate patterns discussed above, we now consider individual behavior. To get a first idea, we begin with classifications of participants according to how frequently they choose the optimal rule, and according to the correlation of their choices with the optimal rule. Table 6 presents these two classifications.

[37] Gailmard and Palfrey (2005) introduced this measure, using it to compare mechanisms for public good provision.

[38] While dominant-strategy voting provides a reasonable benchmark, we also analyze actual behavior in the voting rounds of the experiment in Section 4.3.

[39] That minority under-protection leads to a larger surplus lost than its over-protection is a consequence of our choice of distributions: for many of them, deviating from an optimal rule towards rule III leads to a higher loss than deviating further away from rule III.

Table 6: Summary of rules choices by individuals

Number of optimal rule choices		Correlation with optimal rule	
	Subjects (Percent)		Subjects (Percent)
$x \leq 5$	43 (33%)	$r < 0$	25 (19%)
$5 < x \leq 10$	63 (48%)	$0 \leq r < 0.25$	13 (10%)
$10 < x \leq 15$	23 (18%)	$0.25 \leq r < 0.5$	22 (17%)
$x > 15$	1 (1%)	$r \geq 0.5$	70 (54%)
Total	130		130

In the first two columns of Table 6 we classify subjects by their frequency of optimal rule choices.[40] About half of the subjects made between 5 and 10 rule choices corresponding to the optimal rule. While there is variation in the number of rule choices that are in line with the theoretical optimum, frequent deviations occur for nearly all participants. The last two columns in Table 6 group subjects according to the correlation between the individually chosen rules and the optimal ones.[41] For most subjects the correlation is positive and for more than half it is above 0.5. These numbers indicate that a majority of subjects seem to take the expected payoff into account when choosing the group decision rule. However, about 20% of subjects exhibit a negative correlation. Again, there appears to be noticeable heterogeneity among subjects.[42]

The above classifications, although giving some indication of subjects' heterogeneity, are relatively arbitrary. For a better understanding of individual behavioral patterns, we conduct a hierarchical cluster analysis of subjects' rule choices. Since each of the 130 subjects made decisions for each of the 21 distributions, we have 130 vectors of individual rule choices, each with 21 elements. The cluster analysis aims to classify these 130 vectors into groups, thus helping to identify types of subjects according to their rule decisions.

The cluster analysis suggests to divide subjects into five clusters.[43] For each of the clusters,

[40] Remarkably, the one subject that chose an optimal rule more than 15 times chose it in all 21 distributions. In a post-experimental questionnaire, this subject explained to have calculated the optimal rule based on the conditional expected gains and losses.

[41] Two subjects always chose the same rule (one always rule I, the other always rule V). These subjects are classified as having correlation 0.

[42] In the post-experimental questionnaire, we also asked the subjects about their gender, risk and inequality attitudes, and their views on several policy issues as well as whether these should be decided by super-majorities or simple majorities. Notably, men adjust their suggested rule more strongly to the optimal rule and deviate more from simple majority rule than women. The political preferences do not have a clear effect on rule choices.

[43] As in Fallucchi et al. (2019), among many possible choices for measuring distances between and within clusters, we use Ward's linkage with L1 (sum of absolute distances) metric. We also use the Duda/Hart Je(2)/Je(1) index to determine the number of clusters that best resolves the trade-off between the similarity of observations within

we calculate several basic metrics summarizing rule choices of subjects within the cluster. The metrics are shown in Table 7 together with the number of subjects belonging to each cluster in column "Size". One metric is the average rule choice for the subjects in the cluster ("Average rule"). The other metrics are similar to the ones discussed in Table 6: the number of times the choice of a subject coincides with the optimal rule and the correlation of choices of a subject in the cluster with the optimal rule. The table reports, for each cluster, these metrics averaged over the subjects in the cluster. The last row ("Optimal") refers to what the value of the metric would be if the optimal rule were chosen for each distribution.

Table 7: Aggregate metrics for clusters

Cluster	Size (%)	Average rule	Average number of optimal rule choices	Average correlation with optimal rule
1	8 (6%)	1.96	3.25	−0.19
2	16 (12%)	2.95	3.88	−0.33
3	24 (18%)	2.74	5.63	0.36
4	60 (46%)	3.17	9.65	0.76
5	22 (17%)	3.84	6.05	0.29
Optimal		3.05	21	1

In addition to Table 7, Figure 3 shows, for each cluster, all the rule choices of subjects in this cluster. The line for each cluster represents the linear regression of choices on the optimal rule. From Table 7 and Figure 3, the five clusters can be described in the following way:

Cluster 1 (8 subjects) Mostly rules with a low majority threshold are chosen. The chosen rules are slightly negatively correlated with the optimal rule. While the tendency towards low majority thresholds could be interpreted as these subjects being risk-seeking, because of the negative correlation, they are better characterized as confused.

Cluster 2 (16 subjects) The chosen rules are negatively correlated with the optimal rule. Rule III is chosen more often than other rules. Rules with misguided minority protection are also chosen relatively frequently. These subjects can be characterized by a mixture of confusion and a preference for simple majority voting.

Cluster 3 (24 subjects) Extreme rules (I and V) are often avoided. The chosen rules are weakly positively correlated with the optimal rule. On average, rules are slightly biased towards

clusters and the dissimilarity of observations between clusters.

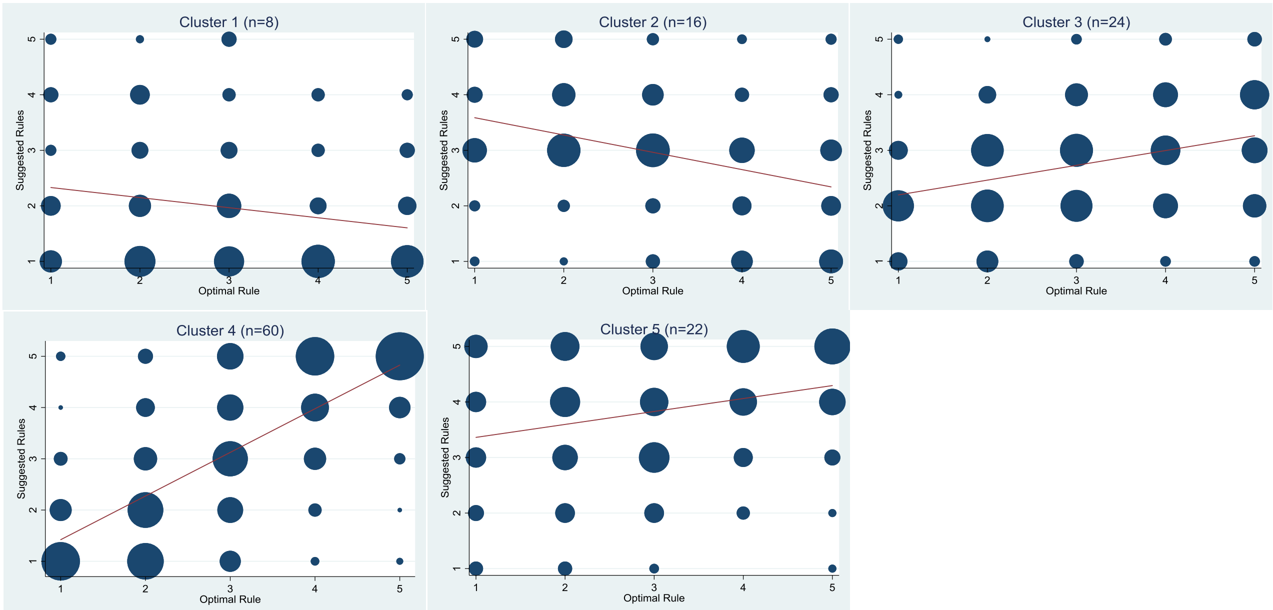


Figure 3: Rule choices by cluster. The sizes of the circles are proportional to the number of observations for that combination of optimal rule and suggested rule.

change. These subjects can by characterized by a mixture of a preference for moderate rules (i.e. minority under-protection) and slight risk-seeking.

Cluster 4 (60 subjects) The chosen rules are highly correlated with the optimal rule. The optimal rule is chosen more often than in the other clusters. If rule choices deviate, it is often in line with over-protection of minorities. These subjects can be characterized as being quite rational, with a strong sense of minority protection. This is the largest cluster, comprising almost half of all subjects.

Cluster 5 (22 subjects) Mostly rules with a high majority threshold are chosen. The chosen rules are weakly positively correlated with the optimal rule. Such subjects can be characterized as risk-averse or inequality-averse.

Although some subjects (mostly those in Clusters 1 and 2) seem confused and make choices that are difficult to explain, most subjects' choices reflect underlying incentives.[44] Of those, some subjects (Cluster 5) may be risk-averse or inequality-averse to some extent, choosing rules

[44]The coefficient in a regression of the suggested rules on the optimal rule is not significantly different from 0 in clusters 1 and 2, but is significantly different from both 0 and 1 in each of the other clusters.

with a high threshold. Some subjects (Cluster 3) avoid extreme rules, reflecting "tyranny-of-the-majority" (for subjects in this cluster, the tyranny-of-the-majority rate as calculated in Table 3 is 0.630 and the tyranny-of-the-minority rate is 0.340). Many subjects though (Cluster 4) frequently choose the optimal rule, but also often deviate to more extreme rules (thus leaning towards minority over-protection; for this cluster, the tyranny-of-the-majority rate is 0.202 and the tyranny-of-the-minority rate is 0.551).

The cluster analysis thus confirms that also on the individual level there is little evidence in favor of Hypothesis 2, a preference for simple majority voting. There are some subjects (Clusters 2 and 3) that choose rules closer to rule III more often but only 5 subjects suggested rule III at least 12 times: 4 in Cluster 2 and 1 in Cluster 3. But there are others (Cluster 4) that suggest more extreme rules: e.g., also 5 subjects suggested only rule I or rule V, of which 3 subjects are in Cluster 4, 1 in Cluster 1 (who suggested rule I always) and 1 in Cluster 5 (who suggested rule V always). It is also interesting to note that only 8 subjects (6%) never suggested either of the extreme rules, rule I and rule V. Therefore, most subjects did not even exhibit a preference for majority voting in the weaker sense that they completely shied away from the extreme rules. On the other hand, 41 subjects (32%) suggested one of the extreme rules but not the other.

Subjects in Cluster 4 also appear to react to – theoretically irrelevant – distribution changes in a (more) predictable way than subjects in the other clusters. The average number of times a subject in Cluster 4 suggests a lower voting rule when only the probability of the positive value is increased or suggests a higher voting rule when the probability of the negative value is increased is 2.55; for a subject in the other clusters this number is 0.89.[45] For the distributions in which only the variance of positive values is increased, the average number of times a subject in Cluster 4 suggests a lower voting rule or a higher rule if the variance of negative values is increased is 2.42, while for a subject in the other clusters this number is 0.20.

From the post-experimental questionnaire, subjects who stated to be "risk averse" or "somewhat risk averse" (89 subjects out of 129, 69%)[46] tend to choose "higher" rules than the remaining participants, but the difference is small and not significant (3.12 compared to 3.05, the p-value of a rank sum test is 0.521). Self-assessed risk attitude is therefore consistent with risk averse behavior, but does not provide much explanatory power.

[45]More precisely, for each subject, if the suggested rule is lower for the distribution in which the probability of the positive value is higher (such as distribution #11 relative to the base distribution #5, or distribution #8 relative to the base distribution #3), then 1 is added; if the suggested rule is higher, 1 is subtracted; if the suggested rule is the same, the count is not changed. For distributions in which the probability of the negative value is increased (such as distribution #9 relative to distribution #3), 1 is added is the suggested rule is higher and 1 is subtracted if the suggested rule is lower.

[46]One subject did not answer the risk attitude question and is omitted from the analysis in this paragraph.

Result 4: Individual rule choices summary. There is considerable heterogeneity in the rule choice behavior of subjects. Behavior of some subjects (Clusters 1 and 2, 18%) does not seem to reflect incentives in the distributions. There are sizable groups of subjects whose choices are somewhat biased towards simple majority voting (Cluster 3, 18%) or towards conservative rules (Cluster 5, 17%). About half of the subjects (Cluster 4, 46%), however, frequently choose the optimal rule, and also often overprotect minorities, counter-balancing the preference-driven biases of the other clusters. Sometimes these subjects react to changes in the distribution in a suboptimal (but predictable) way.

4.2.2 Individual rule choices and expected surplus

Given that there is a substantial heterogeneity in subjects' rule choices, we calculated, for each subject, the expected surplus lost if the subject's suggested rule was used for all 21 distributions. Figure 4 shows on the vertical axis the expected surplus lost by each individual participant, against the correlation between suggested and optimal rule for this participant on the horizontal axis. The figure also shows the expected surplus lost if the simple majority rule III were suggested for all 21 distributions (the horizontal line at 27.12), and the expected loss if each of the five rules was suggested with equal probability (the line at 41.23). For these rule suggestions the correlation with the optimal rule is 0, which is indicated by the vertical line in the figure.[47]

As can be expected, there is a clear negative relationship between the lost surplus and the correlation with the optimal rule: the higher the correlation, the less surplus is lost. There are several participants whose rule proposals are associated with larger losses than from a random choice of rules.[48] Most of those participants are in Clusters 1 and 2; their rule choices appear hardly affected by the optimal rules, often having in fact a negative correlation with the optimal rule. The average lost expected surplus in Clusters 1 and 2 is 47.71, higher than that with a random choice of rules.

However, most subjects have a positive correlation with the optimal rule and lose less surplus. While the mean expected lost surplus is 23.95, the median is 18.81. Across Clusters 3-5, the average lost surplus is 18.57. This loss is 28.4% of the optimal surplus, compared to 36.6% over all subjects. In the largest Cluster 4, which contains subjects that follow the optimal rule most closely, the average lost surplus is only 11.82. Recall that subjects in Cluster 3 tend to choose

[47] Three of the five subjects with the largest lost surplus were in the session in which control questions were skipped. The average and median lost surplus in this session is, however, not statistically different from that in the other sessions.

[48] For some of those participants, the expected loss is even larger than the total available surplus 65.31 (= $21 \cdot 3.11$ from Table 5) across all 21 distributions. Their suggested rules would on average lead to negative payoffs to the group because the group will often end up with A even if the sum of valuations is negative.

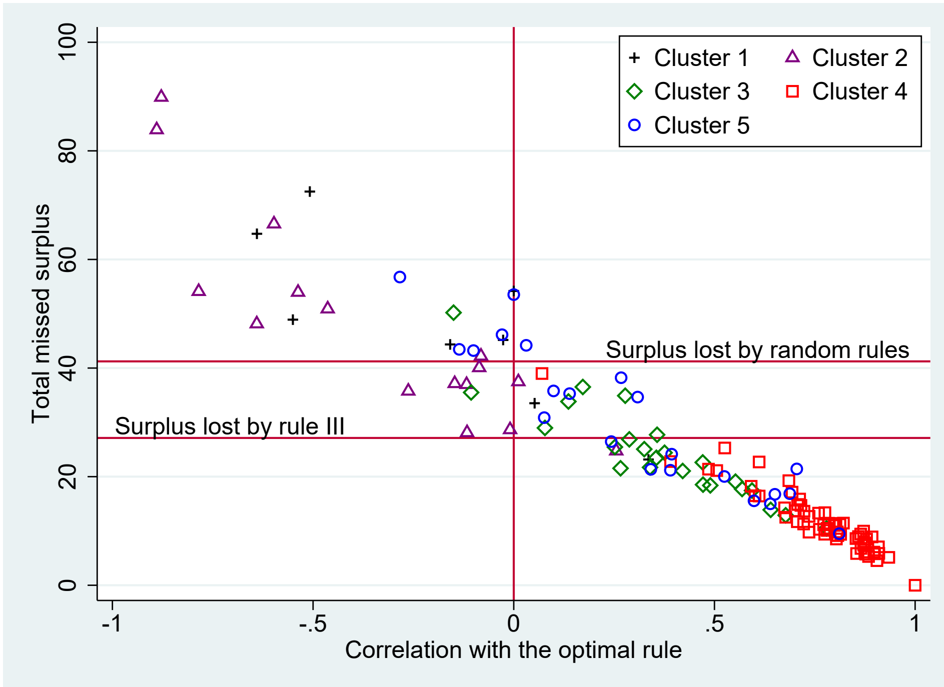


Figure 4: Missed surplus and correlation with the optimal rule

rules close to the simple majority rule III; for them the average lost surplus is 24.51, close to that if rule III were always suggested. Subjects in Cluster 5 tend to choose conservative rules; from Figure 4, there is a lot of variation within this cluster in terms of surplus lost. Those that had the highest correlation with the optimal rule would lose little surplus, but those that chose conservative rules even when rule I or II is optimal would lose quite a bit of the surplus. The average lost surplus for subjects in Cluster 5 is 30.47.

Result 5: Individual rule choices and surplus extraction. Most of the expected surplus loss comes from a minority of subjects who do not follow the incentives embedded in the distributions. For the majority of subjects (Clusters 3-5), whose rule choices do correlate with the optimal rules, the expected lost surplus is substantially smaller.

4.3 Voting behavior

4.3.1 Consistency of voting behavior

All of the analysis so far was based on the assumption that voting follows the realization of the valuations, i.e. subjects vote for A if their realized valuation is positive and for B if the valuation is negative. We call such voting *consistent*. To identify possible welfare effects of actual voting behavior, we also analyze the voting choices in the second stage. In total, we observe 390 individual voting decisions, since for each of the 130 experimental subjects we observe voting decisions for three distributions.

As a first observation, out of a total of 390 voting decisions, 332 (85%) are consistent. Out of 218 cases in which the valuation is positive, 175 (80%) are consistent; out of 172 negative valuation cases, 157 (91%) are consistent. While most voting decisions are indeed consistent, 10-20% of these decisions are not, and there is an apparent difference in consistency rates between positive and negative valuations.[49]

Table 8 reports the results of logit regressions of voting behavior, with standard errors clustered on the subject level, since each subject votes three times. The dependent variable is the binary variable of consistent voting (1 if voting is consistent and 0 otherwise). The first explanatory variable is the absolute value of the subject's valuation, since voting inconsistently is individually more costly if such a value is larger. Since subjects could also consider possible effects of their vote on others we also include the sign of subject's own valuation: voting for A when one's valuation is positive can lead to a loss to others, while voting for B when the valuation is negative can only prevent others from realizing their gain, an outcome possibly less damaging for others.[50]

We also consider variables that aim at measuring subjects' perception of the voting situation. The first variable is the difference between the rule suggested by the subject for the distribution in the current voting round and the implemented rule. Consider subjects who suggest a high threshold for which it is difficult for A to pass in a voting round where the implemented threshold is lower. They may then want to compensate for the lower threshold by voting for B even if their valuation is positive. Vice versa subjects who suggest a low threshold may compensate for a high threshold by voting for A even if they have a negative valuation. Thus for negative valuation realizations the variable is reversed: it is the difference between the rule actually used and the rule suggested by the subject. In addition, we consider for each subject the correlation

[49] We also checked whether inconsistency in voting relates to gender or the political attitudes elicited in the post-experimental questionnaire. We did not find reliable effects but there is evidence that more left-leaning participants are more likely to vote inconsistently if their valuation is positive (at 10% significance level). This would fit with left-leaning people being more inequality averse (even if the inequality is to their advantage) and is in line with results in Engelmann et al. (2020) who find in a survey experiment on voting that left-leaning "rich" participants vote for higher taxes than right-leaning "rich" participants.

[50] One could also include the maximum possible absolute value of a negative valuation for the distribution being used for voting, since in this case voting for A can lead to large negative payoff for others. However, our distributions are such that, e.g. small positive valuations imply large possible negative valuations for others, thus including such a variable can lead to multi-collinearity.

Table 8: Determinants of consistent voting

Dependent variable: Probability of consistent voting				
Variable	Coefficient	Average Marg. Eff.	Coefficient	Average Marg. Eff.
Absolute value of valuation	0.498***	0.058***	0.393***	0.047**
	(0.120)	(0.015)	(0.111)	(0.014)
Sign of valuation	−0.861***	−0.100***	−0.740***	−0.089***
	(0.260)	(0.030)	(0.268)	(0.032)
(Signed) difference between rule used and rule suggested	−0.293**	−0.034**		
	(0.116)	(0.013)		
Correlation with optimal rule	0.641*	0.075*		
	(0.344)	(0.039)		
Constant	1.046***		1.412***	
	(0.378)		(0.357)	

*** - significant at 1%; ** - significant at 5%; * - significant at 10%.

between subjects' suggestions in all 21 distributions in part I with the optimal rules, as a general measure of the subject's perception and understanding of the situation.

The first two columns of Table 8 include all variables defined above. The absolute value of a subject's own valuation has a significant positive effect on consistent voting as can be expected. The sign of the valuation also plays a role, however: consistent voting is less likely if the valuation is positive, which is in line with a concern for losses of others. The variables representing subjects' perception are also significant: subjects indeed less often vote consistently possibly trying to compensate for "too low" (or "too high") rules, and they more often vote consistently if their rule suggestions in Part I were more in line with optimal rules.

Since the two last variables are subject-specific, using them for predictions would require adding beliefs about the behavior of subjects in the same voting group. Instead, the last two columns suggest a minimal model of voting behavior based only on the distribution being used for voting.[51] This model can be seen as incorporating the possibility of inconsistent voting as

[51] If session dummies are added to the regressions in Table 8, the significance of the coefficients of the variables in the table remains the same, but there is also some difference in consistent voting across sessions (in the session in which control questions were omitted voting is less consistent than in three of the other four sessions but it is

an error, with more costly errors (i.e. when the absolute value of the valuation is high) being less likely. Additionally, it incorporates the inequality concern via the sign of the valuation. The probability of consistent voting can thus be predicted as

$$\Pr(\text{Voting is consistent}) = \frac{e^x}{1+e^x}, \text{ where } x = 1.412 + 0.393 \cdot |Valuation| - 0.740 \cdot I_{\{Valuation>0\}}, \tag{3}$$

and $I_{\{Valuation>0\}}$ is the indicator function taking value 1 if the valuation is positive and 0 if negative. With this model, the probability of consistent voting in our distributions ranges from 0.704 (when the valuation is 0.5) to 0.981 (when the valuation is -6.5). The average predicted probability of consistent voting using this model is 0.85, the same as the average actual consistent voting. Separated by observations with positive and negative valuations, the average predicted probability also fits the actual one: 0.80 for positive valuations and 0.91 for negative ones.

Result 6: Consistency of voting behavior. Subjects vote in accordance with their valuation 85% of the time. They vote consistently more often the higher is the absolute value of the realized valuation, and when this valuation is negative. A model of voting based on the realized valuation but with error predicts the average probability of consistent voting well.

4.3.2 (In)consistent voting and expected and realized surplus

The possibility of inconsistent voting changes the expected surplus under a given voting rule, and can change which rule maximizes expected surplus for a given distribution. The estimated model of voting (3) changes the optimal rule for 5 of our 21 distributions: for distributions #1, #12, #21 the optimal rule becomes rule IV instead of rule V, for distribution #8 the optimal rule becomes rule II instead of rule III, and for distribution #17 the optimal rule becomes rule I instead of rule II. Note that the model of voting leads to lower rules becoming optimal: the higher probability of inconsistent voting by voters with positive valuations means that it may be better to have a lower threshold, if the expected total surplus of the group is to be maximized. Thus anticipating inconsistencies in voting according to the model would lead to suggestions of lower rules than considering only consistent voting. By contrast, the suggested average rule is 3.10 in our experiment, higher than the average optimal rule assuming consistent voting (3.05), which in turn is higher than the average optimal rule based on the actual partly inconsistent voting behavior (2.81). It does not therefore look like that anticipating voting inconsistency explains the deviations of the rule choice from the predicted optimal rule.

not different from the remaining session). The minimal model of voting focuses on the prediction of consistent voting for the whole population and does not consider the session effect.

As in Section 4.1.3, we can define various measures of expected surplus, but based on the empirical model of voting (3) instead of always consistent voting. Using this model, the expected surplus if rule choices were optimal for always consistent voting, averaged over 21 distribution, is $S_v^{o,cons} = 2.63$. If rule choices are optimally adjusted to the model of voting, the expected surplus is $S_v^{o,adj} = 2.66$. Thus not adjusting the rule to inconsistent voting leads to hardly any losses; it is inconsistency itself that is responsible for much of the loss: compared with the surplus $S^o = 3.11$ from Table 5, 15% ($= 1 - \frac{S_v^{o,cons}}{S^o} = 1 - \frac{2.63}{3.11}$) would be lost due to inconsistent voting. With the actual choice of voting rules, the expected surplus using the voting model (3), averaged over all distributions and subjects, is $S_v = 1.80$. Compared with the surplus $S = 1.97$ from Table 5, 9% ($= 1 - \frac{S_v}{S_{cons}} = 1 - \frac{1.80}{1.97}$) of the expected surplus would be lost because of inconsistencies in voting. The larger part of the expected surplus loss again comes from the suboptimal suggestion of voting rules because with the suggested voting rules only 68% ($= \frac{S_v}{S_v^{o,cons}} = \frac{1.80}{2.63}$) of the possible surplus given the empirical model of voting would be realized. The surplus extraction rate measure from equation (2), taking into account the inconsistencies in voting, can be adjusted to

$$SER_v^e = \frac{S_v}{S^o} = \frac{1.80}{3.11} = 0.579,$$

implying that only 58% of the surplus achievable with optimal rules and consistent voting would have been realized.

The analysis in the previous paragraph is based on all 21 distributions, extrapolating voting from the valuations and rules that were actually used in the voting rounds. If we look only at the voting rounds, we can also add the actually realized surplus S (and compare it with S_v, predicted by the empirical model of voting). Table 9 shows the surplus measures averaged over voting instances only (based on actually chosen voting rules and actually realized valuations in each instance), in all 78 instances and separated by optimal rule.[52]

Again, there is evidence that the empirical model of voting predicts expected surplus quite well: in total the difference between predicted surplus $S_v = 1.16$ and actual surplus $S = 1.18$ is small, and the difference is also relatively small for each group of distributions. The total average surplus obtainable in voting rounds is 2.44, smaller than 3.11 from Table 5: the distributions in voting rounds and the realizations of valuations were less favorable than what could be expected on average. There were no instances of distributions with optimal rule I, which have the highest obtainable surplus. For the distributions with optimal rule II, subjects overall manage to realize a larger surplus than they would with consistent voting. Despite this, the picture emerging from Table 5 is similar to that from the previous paragraph: the inconsistency of voting loses some

[52] Given that the loss from not adjusting the optimal rule to the empirical model of voting is small, unadjusted optimal rules are considered in the table.

Table 9: Average surpluses in voting rounds

Optimal rule	Number of instances	Actual voting and rules S	Empirical model of voting (3) Actual rules S_v	Optimal rules $S_v^{o,cons}$	Always consistent voting Actual rules S_{cons}	Optimal rules S_{cons}^o
I	0	-	-	-	-	-
II	26	3.33	3.27	3.77	3.19	4.13
III	26	0.48	0.41	1.74	0.65	2.35
IV	11	-0.14	-0.14	0.55	0.23	1.05
V	15	-0.37	-0.26	0.15	0.03	0.67
Total	78	1.18	1.16	1.94	1.32	2.44

surplus (in total, $11\% = 1 - \frac{S}{S_{cons}} = 1 - \frac{1.18}{1.32})$[53], but the larger loss in surplus comes from not using the optimal rules (even the ones unadjusted for inconsistencies in voting): 39% in total $(= 1 - \frac{S}{S_v^{o,cons}} = 1 - \frac{1.18}{1.94})$. The surplus extraction rate in voting rounds is

$$SER_v^a = \frac{S}{S_{cons}^o} = \frac{1.18}{2.44} = 0.484,$$

with only 48.4% of the surplus available with optimally chosen rules and always consistent voting realized.[54]

As with the distribution of the expected surplus loss across subjects, the distribution of losses across voting instances is also skewed: while the average such loss is 1.26, the median is 0 (in 53 out of 78 voting instances the surplus is the same as under optimal rule and consistent voting). Interestingly, it is the "tyranny-of-the-minority" rules that happen to lead to the largest lost surplus. These rules were used in 28 of voting instances (36%) but lead to lost surplus 2.18 on average (and 62% of total lost surplus). Even the misguided minority protection rules were not so detrimental: they were used in 7 instances (9%) and lost 2.14 on average (15% of total lost surplus). "Tyranny-of-the-majority" rules were used in 15 instances (19%) and lost 0.73 of surplus on average (11% of lost surplus), while optimal rules were used in 28 instances (36%) and lost 0.39 (also 11% of total lost surplus). The number of instances is, however, small, thus we do not put too much stake on these numbers.

Result 7: Voting and surplus. Inconsistent voting leads to a further loss (9-15%) in obtain-

[53]Note also that inconsistency in voting did not "correct" for the "wrong" rules: the actual surplus 1.18 is lower than the surplus 1.32 that would have been achieved if the subjects had voted consistently using the actual rule in the group.

[54]On the other hand, the absolute surplus lost in voting rounds $(2.44 - 1.18 = 1.26)$ is similar to the one lost in expected terms over all 21 distributions: $(3.11 - 1.80 = 1.31)$.

able surplus, but the main source of inefficiency is suboptimal rule choices, with only 61-68% of surplus realized (if voting follows the estimated model). With actual voting and rule suggestions, the surplus extraction rate is only 48-58% of the surplus achievable with optimal rules and consistent voting.

5 Conclusion

A wide range of institutions rely on voting procedures to aggregate members' preferences. Often the very same institution applies different majority thresholds to decide on different classes of issues. While many decisions are determined by a simple majority of participants as a default, decisions on more sensitive issues require super-majority or even unanimous support.[55] This is compatible with the view that institutions should protect minorities with strong preferences against particularly unfavorable majority decisions. The observed variety of voting rules in practice is in line with our experimental finding that individuals adjust the voting rule to the distribution of preference intensities. However, we also find that subjects on average fail to adjust the rules strongly enough, missing, in expected terms, a substantial proportion of available surplus. We find that sometimes minorities are not protected enough when they deserve being protected; but we also find frequent cases of over-protection of minorities, in particular of conservative minorities. For example, in a third of the cases where it would be efficient to implement change as long as at most one person objects, participants actually suggest unanimity.

Both under-protection and over-protection of minorities occur to similar degree in our experiment – an observation that partly results from heterogeneous preferences among the participants but also from errors. Overall, we do not find evidence for systematic under-protection of minorities that would result from a widely-held preference for simple majority voting. Neither do we find systematic biases in favor of conservative rules that would result from risk-averse or inequality-averse preferences. This absence of a general preference-driven bias suggests that endogenous rule choice has the potential to lead to substantial efficiency gains relative to the use of any given fixed majority rule. The sizable surplus losses relative to the first best that we observe appear to be driven primarily by confusion of some subjects,[56] suggesting that the surplus losses could be smaller for more experienced participants, for tasks that are easier, or for compromise rules based on suggestions from the entire group of voters. If, instead of a random dictator rule, the *median* rule among those suggested in the group were used in the 78 voting instances, the

[55] See Barberà and Jackson (2004, 2006) and Grüner and Tröger (2019) for examples and further references.

[56] Although the aggregate pattern of choices also appears consistent with a model incorporating (costly) mistakes, we argue in Online Appendix B that such a model cannot explain all features of the data. Choices are thus better explained by a combination of some subjects having preferences aiming at something else than expected-payoff maximization or a wrong understanding of the problem and noise than by trembles alone.

surplus loss would be approximately only half as large under the assumption that first-stage rule choices would be no different if the median instead of the random dictator rule were applied. The latter assumption is justified since in both mechanisms participants should state their preferred rule. For practical implementation of a rule-choice stage, one should therefore rather choose the median rule to aggregate stated preferences over rule choices, as this partly corrects for erroneous individual rule choices. In our experiments, we choose the random-dictator rule primarily because it is easier to explain that it is incentive compatible.

For practical reasons, our experiment draws on student subjects and considers an election with a relatively small number of voters. Thus, taken literally, it evaluates inexperienced individuals' ability to choose rules for small decision making bodies (such as e.g. boards of clubs or professional associations, parliamentary committees, or central bank councils) in situations where the set of possible valuations and their probabilities can be, at least in principle, calculated (or more precisely, where the expected valuations conditional on being positive and negative are known). Providing this information in our experiment can balance out subjects' inexperience. It is an important open question to what degree the benefits of an endogenous choice of voting rules also apply to large elections.[57]

Political institutions often protect minorities against unfavorable changes of policy but seldom permit minorities to trigger reforms which lead to large but concentrated welfare gains, sometimes making it difficult to implement efficiency-enhancing reforms. An example are those constitutions that require either a simple or a super-majority in parliamentary decisions but never base decisions just on a sub-majority. Related to this, in our experiment, minorities favoring a change are slightly less likely to be protected by the voting rule than minorities favoring a status quo, but the bias towards rules protecting the status quo is small.

Our paper shows that letting people choose democratic rules depending on the issue at hand, while improving on the exogenously imposed simple majority rule, does not always lead to the choice of the efficient rule. Even in a relatively simple environment with pre-specified valuations, some subjects fail to properly choose majority rules and aspects that are irrelevant, in our case the probabilities of gains and losses, can affect the chosen rules. The observation that minorities with strong preferences can be under-protected raises the question how to identify real world institutions that suffer from this problem. While such claims are often made, e.g., by minorities after a referendum, as has been the case with Brexit, a systematic analysis of whether more under- or over-protection exists in practice requires the measurement of preference intensities

[57]A different theoretical framework is needed to fully understand the underlying tradeoffs. By the law of large numbers, in a large election with uncorrelated valuations the optimal decision is known if the distribution of valuations is known. Correlated types make these elections more meaningful. Furthermore, if the possible valuations are known but not their associated probabilities, voting is informative.

for various institutions and decision problems and a comparison of optimal and actual rules. Furthermore, in our setting individuals choose rules in isolation whereas real decision-making bodies, such as constitutional assemblies allow for communication and collective deliberation before reaching a decision, raising the question by how much this can improve rule choices.

References

[1] Adams, J. (1788). "A Defence of the Constitutions of Government of the United States of America." London: C. Dilly, Vol. III.

[2] Austen-Smith, D. and J.S. Banks (1996). "Information Aggregation, Rationality, and the Condorcet Jury Theorem," American Political Science Review, 90, 34–45.

[3] Azrieli, Y. and S. Kim (2014). "Pareto Efficiency and Weighted Majority Rules," International Economic Review 55, 1067–1088.

[4] Badger, W.W. (1972). "Political Individualism, Positional Preferences and Optimal Decision Rules," in Niemi, R. and H. Weisberg (eds) "Probability Models of Collective Decision Making." Columbus: Charles E. Merril, 34–59.

[5] Barberà, S. (1979). "Majority and Positional Voting in a Probabilistic Framework," Review of Economic Studies 46, 379–389.

[6] Barberà, S. and M.O. Jackson (2004). "Choosing How to Choose: Self-Stable Majority Rules and Constitution," Quarterly Journal of Economics 119, 1011–1048.

[7] Barberà, S. and M.O. Jackson (2006). "On the Weights of Nations: Assigning Voting Weights in a Heterogeneous Union," Journal of Political Economy 114, 317–339.

[8] Blais, A., Laslier, J.-F., Poinas, F. and K. Van Der Straeten (2015). "Citizens' Preferences about Voting Rules: Self-Interest, Ideology, and Sincerity," Public Choice 164, 423–442.

[9] Bol, D., Blais, A., Coulombe, M., Laslier, J.-F. and J.-B. Pilet (2020). "Choosing an Electoral Rule," working paper.

[10] Börgers, T. and P. Postl (2009). "Efficient Compromising," Journal of Economic Theory 144, 2057–2076.

[11] Brennan, G. and J.M. Buchanan (1985). "The Reason of Rules: Constitutional Political Economy." Cambridge: Cambridge University Press.

[12] Buchanan, J.M. and G. Tullock (1962). "The Calculus of Consent." Ann Arbor: University of Michigan Press.

[13] Casella, A. (2005). "Storable votes," Games and Economic Behavior 51, 391–419.

[14] Curtis, R.B. (1972). "Decision Rules and Collective Values in Constitutional Choice," in Niemi, R. and H. Weisberg (eds) "Probability Models of Collective Decision Making." Columbus: Charles E. Merril, 23–33.

[15] Drexl, M. and A. Kleiner (2018). "Why Voting? A Welfare Analysis," American Economic Journal: Microeconomics 10, 253–271.

[16] Engelmann, D. and H.P. Grüner (2017). "Tailored Bayesian Mechanisms: Experimental Evidence from Two-Stage Voting Games." CESifo Working Paper Series 6405, CESifo Group Munich.

[17] Engelmann, D., E. Janeba, L. Mechtenberg, and N. Wehrhöfer (2020). "Preferences over Taxation of High-Income Individuals: Evidence from a Survey Experiment." CESifo Working Paper Series 8595, CESifo Group Munich.

[18] Fallucchi, F., R.A. Luccasen III, and T.L. Turocy (2019). "Identifying Discrete Behavioural Types: A Re-analysis of Public Goods Game Contributions by Hierarchical Clustering," Journal of the Economic Science Association 5, 238–254.

[19] Feddersen, T. and W. Pesendorfer (1998). "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting," American Political Science Review 92, 23–35.

[20] Fischbacher, U. (2007). "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," Experimental Economics 10, 171–178.

[21] Gailmard, S. and T.R. Palfrey (2005). "An Experimental Comparison of Collective Choice Procedures for Excludable Public Goods," Journal of Public Economics 89, 1361–1398.

[22] Gershkov, A., Moldovanu, B. and X. Shi (2017). "Optimal Voting Rules," Review of Economic Studies 84, 688–717.

[23] Greiner, B. (2015). "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE," Journal of the Economic Science Association 1, 114–125.

[24] Grüner, H.P. and T. Tröger (2019). "Linear Voting Rules," Econometrica 87, 2037–2077.

[25] Hoffmann, T. and S. Renes (2022). “Flip a Coin or Vote? An Experiment on the Implemenation and Efficiency of Social Choice Mechanisms,” Experimental Economics 25, 624–655.

[26] Martinelli, C. and T.R. Palfrey (2020). “Communication and Information in Games of Collective Decision,” in Capra, C.M., Croson, R., Rigdon, M. and T. Rosenblat (eds) “Handbook of Experimental Game Theory”, Edward Elgar, 348–375.

[27] Moldovanu, B. and F. Rosar (2021). “Brexit: A comparison of dynamic voting games with irreversible options,” Games and Economic Behavior 130, 85–108.

[28] Nehring, K. (2004) “The Veil of Public Ignorance,” Journal of Economic Theory 119, 247–270.

[29] Rae, D.W. (1969). “Decision Rules and Individual Values in Constitutional Choice,” American Political Science Review 63, 40–56.

[30] Rawls, J. (1972). “A Theory of Justice.” Oxford: Clarendon Press.

[31] Schmitz, P.W. and T. Tröger (2012). “The (Sub-) Optimality of the Majority Rule,” Games and Economic Behavior 74, 651–665.

[32] Schofield, N. (1972). “Ethical Decision Rules for Uncertain Voters,”British Journal of Political Science 2, 193–207.

[33] Taylor, M. (1969). “Proof of a Theorem on Majority Rule,” Behavioral Science 14, 228–31.

[34] Vermeule, A. (2005). “Submajority Rules: Forcing Accountability upon Majorities”, The Journal of Political Philosophy 13, 74–98.

[35] Weber, M. (2020). “Choosing the Rules: Preferences over Voting Systems for Assemblies of Representatives,” Journal of Economic Behavior & Organization 174, 420–434.

Online Appendix to "Minority Protection in Voting Mechanisms – Experimental Evidence"

Dirk Engelmann* Hans Peter Grüner† Timo Hoffmann‡ Alex Possajennikov§

March 24, 2023

Abstract

This online appendix contains the experimental instructions (Appendix A) and the quantal response analysis (Appendix B).

*Humboldt-Universität zu Berlin, School of Business and Economics; CERGE-EI, a joint workplace of Charles University and the Czech Academy of Sciences, Prague, Czech Republic; CESifo, Munich. E-mail: dirk.engelmann@hu-berlin.de

†University of Mannheim, Department of Economics; CEPR, London. E-mail: gruener@uni-mannheim.de

‡At the time the experiment reported in the paper was conducted Timo Hoffmann was affiliated with the Department of Economics at University Erlangen-Nuremberg. E-mail: Timo.Hoffmann@fau.de

§University of Nottingham, School of Economics. Email: alex.possajennikov@nottingham.ac.uk

Online Appendix A - Experimental Instructions

You are now taking part in an experiment. The amount of money you earn depends on your choices and the choices of the other participants. It is therefore important that you understand the instructions. Please do not communicate with the other participants during the experiment. If you have any questions, please raise your hand and we will come to your seat.

All the information you provide will be treated anonymously. The experiment is run through a computer program, which determines the resolution of all random events during the experiment.

You will begin the experiment with a starting budget of *15€*. This amount can be increased or decreased depending on all participants' choices in the experiment, as explained below. Your final earnings, however, cannot be negative, that is, there is no risk that you will have to pay us. For each participant, the minimum possible earnings of the entire experiment are *3€* and the maximum possible earnings of the entire experiment are *27€*. The earnings of all participants will be paid out privately in cash after the experiment.

Thank you for participating.

THE EXPERIMENT

The experiment consists of two parts, each of which consists of a number of rounds.

In each round you will be asked to consider a problem of making a choice between two alternatives, called A and B, by a group consisting of 5 members, you and four other participants. Your payoff will depend on which alternative is ultimately chosen. If alternative B is chosen, your (and the other group members') payoff is 0. If alternative A is chosen, the payoff of each group member (including you) depends on a randomly assigned valuation, which can be positive or negative. Each group member has the same possible valuations for alternative A and the same corresponding probabilities. The description of the problem in each round will consist of a list of the possible valuations and the probabilities of their realization as illustrated in the example below.

Example*: For the current round, the valuation for alternative A of each group member can be either −5€ with probability 1/3 or +2€ with probability 2/3. The valuation will be randomly assigned to each participant by the computer using the given distribution. In this example this is equivalent to each participant rolling a 6-sided dice. If the outcome is a 1 or 2 the valuation for alternative A of the participant is −5€. If the dice outcome is a 3, 4, 5 or 6, the valuation for alternative A of the participant is +2€.*

1	2	3	4	5	6
-5.0 €		+2.0 €			

PART I: RULE CHOICE ROUNDS

Part I of the experiment consists of 21 rounds. In each round, a different collective decision problem like the one above will be presented to all participants. You (and each of the other participants) will be asked to choose one of five group decision rules for this problem, listed below. The rules determine how the group decision about alternative A or B is derived from the individual votes of all group members. The actual voting, according to one of these rules, selected as explained below, will take place in Part II of the experiment if this round is selected for it.

AVAILABLE GROUP DECISION RULES

Five group decision rules are available. All five rules are voting rules where voters have to vote for A or B. Hence, no-one can abstain from voting, so, for example, only two votes for A automatically means that there are three votes for B.

Rule I. At least 1 vote for alternative A is required for A to be chosen, thus 5 votes for alternative B are required for B to be chosen (unanimity for B).

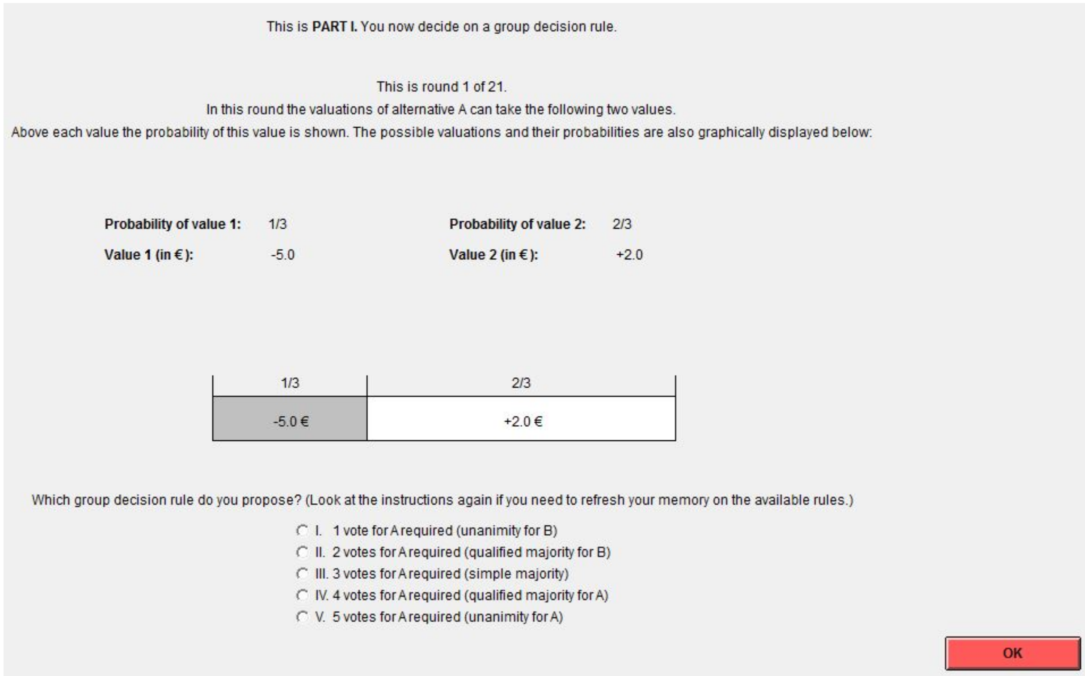
Rule II. At least 2 votes for alternative A are required for A to be chosen, thus at least 4 votes for alternative B are required for B to be chosen (qualified majority for B).

Rule III. At least 3 votes are required for either A or B to be chosen (simple majority), that is, whichever has more votes wins.

Rule IV. At least 4 votes for alternative A are required for A to be chosen (qualified majority for A), hence at least 2 votes for alternative B are required for B to be chosen.

Rule V. 5 votes for alternative A are required for A to be chosen (unanimity for A), thus at least 1 vote for alternative B is required for B to be chosen.

Note that at the stage when you propose a voting rule you know neither your own valuation for alternative A (which can be positive or negative), nor the valuations of the other participants. You know only the possible valuations for alternative A and their probabilities, as in the example above. The decision screen for Part I looks like this, using the valuations and probabilities from the example above.



Your choice in each of the rounds may be selected as one of the rules that will be used to determine the actual voting outcomes in Part II, as explained below. After all participants have made their choices for the 21 decision problems, Part II begins.

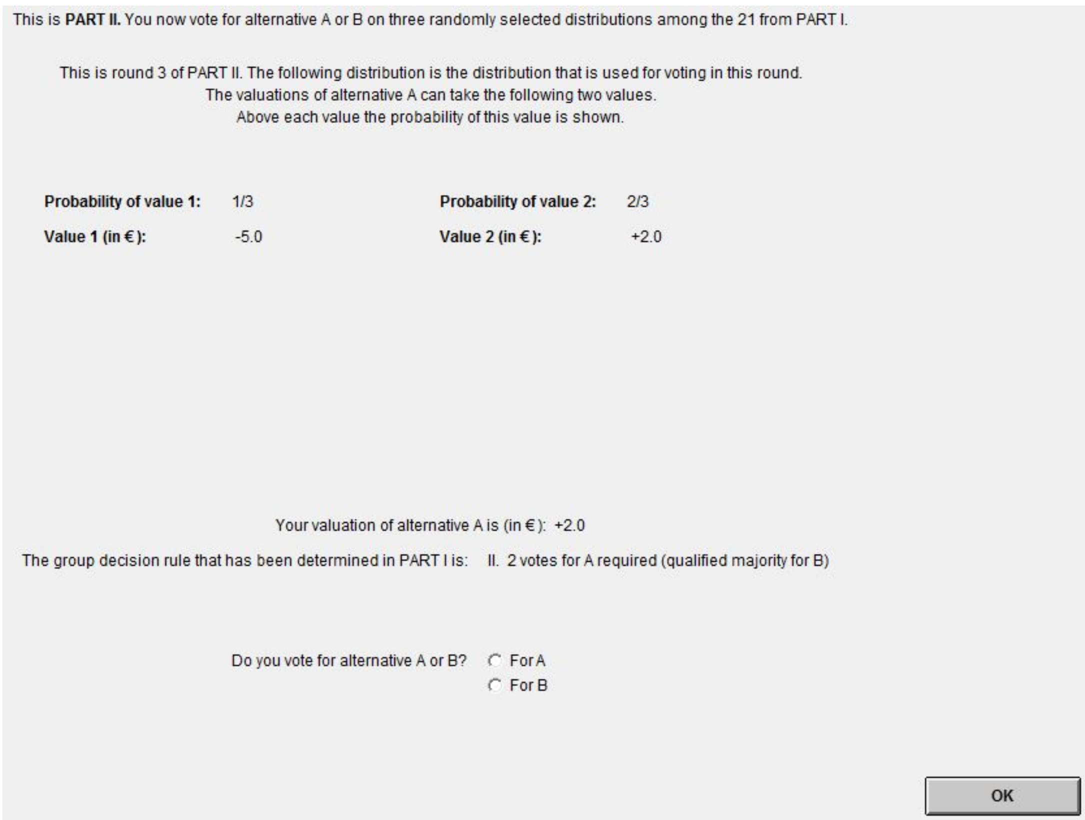
PART II: VOTING ROUNDS

In Part II, each player will participate in three different collective decisions. All three decision problems will be selected randomly from the 21 problems of Part I, with each problem being equally likely to be selected. In each round, groups of 5 participants will be randomly formed. Note that group members cannot recognize each other, so even if you should encounter the same participant in different rounds, you will not be able to identify her or him.

After the groups are formed, the group decision rule is determined as the choice in Part I for this round of one of the five group members, selected randomly with equal probability for each group member. The selected group decision rule is announced to all members of the group. In addition, each group member is privately informed about his or her valuation for alternative A. No participant can see another participant's valuation for A at this stage or at any later point of time. The random draws of the valuations for the group members are independent of each other; thus, learning your own valuation does not change the possible valuations and their probabilities for each of the other members of your group.

Example: *The round in which the valuation for alternative A of each group member is −5€ with probability 1/3 and +2€ with probability 2/3 is selected for Part II. The selected voting rule is Rule II (At least 2 votes are required for alternative A to be chosen, hence at least 4 votes for alternative B are required for it to be chosen). You are further informed that your valuation for alternative A is +2€; from your point of view, the valuation for alternative A of each of the other group members is still −5€ with probability 1/3 and +2€ with probability 2/3.*

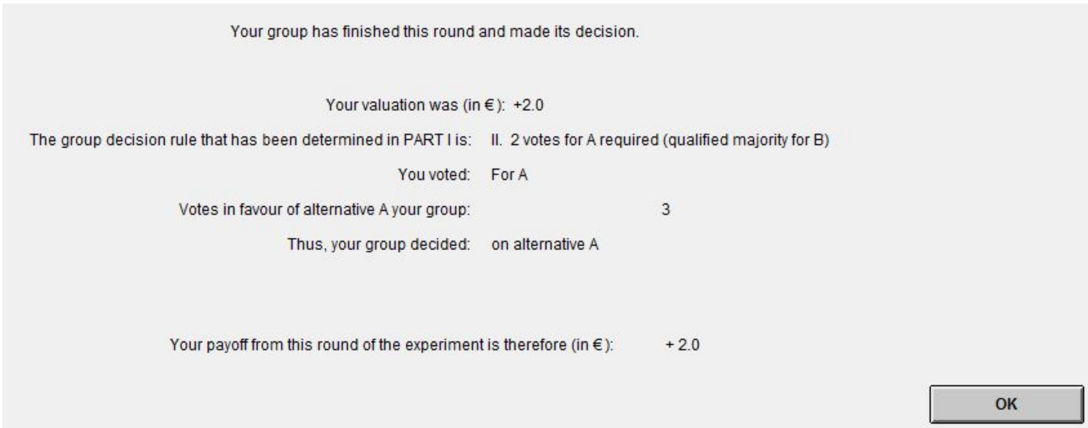
After being informed about the selected rule and your valuation, you and the other group members vote for alternative A or B. The decision screen for Part II looks like this.



After all group members made their choice, the group decision automatically results from the individual votes according to the selected group decision rule. The decision is announced and you get the payoff equal to your valuation of the chosen alternative.

Example: *Suppose that the voting rule is Rule II and that your valuation for alternative A is +2€. Suppose further that there are 3 votes for alternative A and 2 votes for alternative B. According to Rule II, alternative A is chosen. Your payoff for this round is your valuation for alternative A, that is, +2€.*

At the end of each round of Part II, you will be informed about the outcome of the round. The screen looks like this.



PAYOFFS

Your payment from the experiment is the sum of payoffs you get in the three rounds of Part II, plus the starting budget of *15€*. Recall that your payoff in a round is your valuation of the alternative chosen by your group in that round. Recall that the valuation of alternative B is 0 for all participants, while your (and the other participants') valuation for alternative A is determined anew in each round and can be positive or negative.

Payment rule example: *Assume that your valuation for alternative A in round 1 of Part II was +5€, in round 2 it was −3€, and in round 3 it was +1€. If your group voted for alternative A in each round, then your payoff from the three rounds would be 3€ (=+5€+(−3€)+1€) and your final earnings would be 18€ (=15€+3€). If your group chose alternative B in all three rounds, you would have final earnings 15€, equal to the starting budget. If your group chose alternative A in the second round and alternative B in the other two rounds, then your payoff from the three rounds would be −3€ (=0€+(−3€)+0€) and your final earnings would be 12€ (=15€+(−3€)).*

OVERVIEW OF THE EXPERIMENT

Here is the structure of the experiment again in a short overview:

- There are 2 parts;
- Part I consists of 21 rounds. In each round:
  - Possible valuations for alternative A and their probabilities are announced;
  - Each participant selects one of the five group decision rules.
- Part II consists of 3 rounds. In each round:
  - One round from Part I is randomly chosen and the corresponding possible valuations for alternative A and their probabilities are announced;
  - Groups of 5 participants are randomly formed;
  - The group decision rule chosen in Part I by one randomly chosen participant in the group of 5 is selected and announced to all group members;
  - Each group member privately learns his or her valuation for alternative A;
  - Each group member votes for alternative A or B;
  - The group decision is taken based on the selected group decision rule and the votes;
  - All group members are informed about the outcome of the vote and their payoff.
- The sum of payoffs from the 3 rounds of Part II, added to the starting budget, is paid out privately.

Online Appendix B - Quantal response model of rule choice

B.1 The model and its parameter estimation

Suppose that all participants have preferences for expected payoff maximization but make mistakes in their choices of voting rules. Suppose that mistakes that are more costly are less likely than mistakes that are less costly, as in logit choice (used, for example, in the quantal response equilibrium (QRE), McKelvey and Palfrey, 1995).

For each of the 21 distributions in the experiment and each of the five voting rules, let E_d^k be the expected (with respect to the realization of voters' valuations, assuming own payoff-maximizing voting in the second stage) sum of voters' payoffs for rule k for distribution d (one voter's expected payoff is then $E_d^k/5$). The probability of choosing voting rule k is then

$$Pr_d^k = \frac{e^{\lambda E_d^k}}{\sum_{i=1}^5 e^{\lambda E_d^i}},$$

where λ is a parameter measuring the noise in choices. If λ is close to 0, choices are close to uniformly random; if $\lambda \to \infty$, the probability of the rule that maximized the expected payoff approaches 1.

We estimate the value of λ that fits the data best by two methods. The first method minimizes squared deviations between the predicted probabilities of rule choices Pr_k^d and the actual frequencies in the experiment; we find $\lambda = 0.304$. The second method used Pr_k^d to maximize the log-likelihood of the entire experiment data sample, with $\lambda = 0.267$ as the result. We choose $\lambda = 0.3$ for the analysis below.

We simulate 200 samples of 2730 observations (130 individuals making choices for 21 distributions, each behaving according to this model of noisy choice). The analysis below focuses on 95% intervals for the values of variables from these 200 samples (2.5-97.5% ranges).

B.2 Average rule choices

Table B.2 shows the 95% intervals for the proportions of rule choices, grouped by distributions with a given optimal rule, as in Table 2 in the paper. Bold entries indicate an interval end beyond which the corresponding proportion from Table 2 lies (intervals without bold entry therefore indicate intervals encompassing the proportion observed in the experiment).

The medians for each optimal rule choice in the simulated samples are the same as in the actual sample. The mean in the actual data (3.100) is slightly higher than in the simulated samples.

The regression of actually chosen rules on optimal rules (equation (1) in the paper) has intercept 1.783 and slope 0.432, with $R^2 = 0.167$. For the same regressions in 200 simulated

Table B.2: Simulated rule choices

Optimal Rule	Median	Mean	Simulated rule choices I	II	III	IV	V
I	2-2	2.021-**2.221**	31-40%	**29**-38%	13-21%	5-11%	3-**7**%
II	2-2	2.409-**2.586**	22-29%	26-33%	18-25%	11-16%	7-12%
III	3-3	2.886-3.102	**12**-19%	19-25%	23-**30**%	18-25%	12-18%
IV	4-4	3.390-**3.587**	7-12%	10-15%	18-25%	**27**-34%	22-**29**%
V	4-4	3.787-3.960	4-**8**%	6-10%	**14**-20%	**29**-38%	33-**40**%
Total	3-3	2.977-**3.061**	16-19%	**20**-22%	20-23%	**20**-23%	17-**20**%

samples, for the intercept the 95% confidence interval is 1.528-**1.735**. For the slope, the 95% interval is 0.420-0.490. For R^2, the 95% confidence interval is 0.166-0.225. The regressions for the simulated samples have a similar slope and R^2 but a lower intercept than the regression based on the actual data.

Overall, the simulated rule choices are thus on average slightly lower than the actual choices. The simulated choices also tend to be less extreme. With one exception (optimal rule III and chosen rule I), whenever the experimental data lies outside the confidence interval of the simulated data, extreme rules (I and V) are chosen more often in the experimental data, but less extreme rules (II and IV) are chosen less often.

B.3 Rates of over- and under-protection of minorities

The 95% intervals of the proportions of choices in 200 simulated samples are 12-**15**% for misguided minority protection, **25**-28% for under-protection of minorities ("tyranny-of-the-majority"), 27-30% for over-protection of minorities ("tyranny-of-the-minority"), and 30-**33**% for optimal choices. There are more optimal and misguided choices in the actual data than in the simulated data, but fewer choices in line with underprotection of minorities.

The tyranny rates for the simulated samples, calculated in the same way as Table 3 in the paper, are given in Table B.3. In the simulated data there is more majority tyranny because rules I and V are played less often than in the actual data (where the overall majority tyranny rate is 0.372). Rule III is also played less often in the simulated data than in the actual data, thus there is an increase in the minority tyranny when rule III is optimal, but it is compensated by a decrease in minority tyranny because rules I and V are played less often in the simulated data than in the actual data when rules II and (especially) IV are optimal.

Table B.3: Tyranny rates in simulated samples

Optimal Rule	Tyranny of the majority rate	Tyranny of the minority rate
I	**0.529**-0.632	
II	0.242-0.328	0.295-0.369
III		**0.692**-0.771
IV	0.233-0.323	0.283-**0.374**
V	**0.533**-0.624	
Total	**0.399**-0.443	0.472-0.520

B.4 Distributional changes

Our distributions are chosen so that some modifications of them (e.g. in the probabilities of particular values, or in mean-preserving spreads) do not change which rule is optimal. These changes, however, affect the expected payoff of all rules and thus affect the quantal response choices.

The 95% confidence intervals for average rule choices from 200 samples are given in Table B.4 (recall that the bold entries indicate the end of the interval beyond which the statistics from the actual data lies, from Table 4 in the paper). The table also shows the proportion of samples in which the median test detects a significant difference from the base distribution at the 1% level (for Prob+ and Prob− columns) and at the 5% level (for Var+ and Var− columns).[1]

When the probabilities change, then the choices in the quantal model do change in the same direction as in the data: if the probability of the positive value increases, then lower rules are chosen more often; if the probability of the negative value increases, then higher rules are chosen more often, as in the experimental data. The effect is, though, smaller than in the data: for only 8-20% of samples the difference is significant at the 1% level, and for some distributions (D#8, D#9 and D#7), the actual data lie beyond the interval of simulated data.

What the quantal model cannot explain is the change in choices when the variance of the distribution changes. By construction, with a mean-preserving spread, all rules have the same expected payoffs as in the base distribution, thus the model predicts no changes in rule choices. (There are, of course, some differences appearing by chance: in 3-9% of the simulated samples the difference is significant at the 5% level.) In the experimental data, however, we see a certain pattern of changes: if the variance of positive payoff is increased, then lower rules are chosen

[1]We chose the different levels of significance to obtain meaningful comparisons in both cases. In Table 4 in the paper, most of the comparisons for probabilities are significant at 1% level and for variances at 5% level.

Table B.4: Effect of changes in distributions in simulated data

Optimal Rule	Base	Prob+	Prob−	Var+	Var−
I	D#5 2.008-**2.392**	D#11 1.838-2.169 (8%)		D#16 1.938-2.408 (7%)	
II	D#4 2.369-**2.792**	D#10 2.069-2.454 (20%)		D#15 2.385-2.777 (3%)	D#18 2.369-**2.785** (4%)
III	D#3 2.762-3.177	D#8 **2.523**-3.008 (15%)	D#9 3.069-**3.438** (15%)	D#14 **2.746**-3.246 (7%)	D#19 2.800-**3.215** (6%)
IV	D#2 3.185-3.623		D#7 3.500-**3.877** (20%)	D#13 3.192-3.646 (7%)	D#20 3.200-**3.631** (6%)
V	D#1 3.585-3.992		D#6 3.831-4.185 (9%)	D#12 3.615-4.054 (9%)	D#21 3.608-4.008 (7%)

In parentheses: proportions of samples for which the difference with the base distribution is significant: at 1% (Prob+ and Prob−) or at 5% (Var+ and Var−).

more often (sometimes significantly so); if the variance of negative payoff is increased, then higher rules are chosen more often (again, sometimes significantly more so).

B.5 Surplus extraction

Table B.5 replicates Table 5 in the paper, but instead of the expected surplus and the loss for rule III, the last two columns present the 95% intervals of surpluses and losses from the 200 samples of the quantal model.

The simulated quantal choice model has a higher surplus (loses less surplus) than the data when optimal rules are I and V (mostly because it leads to choices of "misguided minority protection" less often than in the data). On the other hand, it leads to a lower surplus (more loss) when the optimal rule is III (because it leads to fewer choices of rule III then). Overall, the simulated quantal response model has a higher surplus than in the experimental data.

Table B.5: Expected surpluses in actual and simulated data

Optimal Rule	Optimal S^o	Actual S	Actual loss $S^o - S$	Simulated S^q	Simulated loss $S^o - S^q$
I	7.16	5.54	1.61	**5.69**-6.12	1.02-**1.44**
II	4.71	3.61	1.10	3.60-3.81	0.90-1.12
III	2.55	1.90	0.65	1.64-**1.76**	**0.79**-0.90
IV	1.37	0.25	1.12	0.15-0.37	1.00-1.22
V	0.56	-0.93	1.47	$\mathbf{-0.90}$--0.57	1.10-**1.43**
Total	3.11	1.97	1.14	**1.99**-2.09	1.02-**1.12**

B.6 Summary

Overall, there are many patterns (more choices of rules I, V, and III in the data than in the quantal-response model; tyranny rates for distributions with these optimal rules; the significance of the effects of probability changes and the effects of mean-preserving spreads; surplus amount) that are quite different in the experiment data than in the quantal-response model. It is therefore unlikely that the observed choices come purely from the model of expected payoff maximization with noise. Choices are more likely to be explained by a combination of individual preferences that differ from expected payoff maximization (some of which identified in the cluster analysis in the paper) or misinterpretation of the decision problem and its incentives and noise.

References

[1] McKelvey, R.D. and T.R. Palfrey (1995) "Quantal Response Equilibria for Normal Form Games," Games and Economic Behavior 10, 6–38.