DISCUSSION PAPER SERIES

No. 9544

TAILORED BAYESIAN MECHANISMS: EXPERIMENTAL EVIDENCE FROM TWO-STAGE VOTING GAMES

Dirk Engelmann and Hans Peter Grüner

PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP9544.asp**

ISSN 0265-8003

TAILORED BAYESIAN MECHANISMS: EXPERIMENTAL EVIDENCE FROM TWO-STAGE VOTING GAMES

Dirk Engelmann, University of Mannheim, University of Copenhagen, CESifo and Economics Institute of the Academy of Sciences of the Czech Republic
Hans Peter Grüner, Universität Mannheim and CEPR

Discussion Paper No. 9544
July 2013

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 9544

July 2013

ABSTRACT

Tailored Bayesian Mechanisms: Experimental Evidence from Two-Stage Voting Games*

Optimal voting rules have to be adjusted to the underlying distribution of preferences. However, in practice there usually is no social planner who can perform this task. This paper shows that the introduction of a stage at which agents may themselves choose voting rules according to which they decide in a second stage may increase the sum of individuals' payoffs if players are not all completely selfish. We run three closely related experimental treatments (plus two control treatments) to understand how privately informed individuals decide when they choose voting rules and when they vote. Efficiency concerns play an important role on the rule choice stage whereas selfish behavior seems to dominate at the voting stage. Accordingly, in an asymmetric setting groups that can choose a voting rule do better than those who decide with a given simple majority voting rule.

JEL Classification: C91, D70 and D82
Keywords: Bayesian voting experiments, revelation principle, and two-stage voting

Dirk Engelmann
University of Mannheim
Department of Economics
L7, 3-5
68131 Mannheim
GERMANY

Email: dirk.engelmann@uni-mannheim.de

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=149489

Hans Peter Grüner
Lehrstuhl für Wirtschaftspolitik
Universität Mannheim
D-68131 Mannheim
GERMANY

Email: gruener@uni-mannheim.de

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=144072

*We thank seminar participants at the University of Zürich, WZB Berlin, DICE Düsseldorf, University of Mannheim, Utrecht University, Tilburg University, Tinbergen Institute/CREED-University of Amsterdam, Erasmus University Rotterdam, University of Erlangen-Nürnberg, University of Konstanz, Royal Holloway, University of Edinburgh, Maastricht University, Max-Planck Institute Bonn, CERGE-EI Prague, University of Mainz, the Schumpeter seminar at HU and FU Berlin, and participants at the International Economic Science Association Conference in New York 2012, the Economic Science Association European Conference in Cologne 2012, and the 2013 meeting of the economic theory group of the German Economic Association for their useful comments and Christoph Siemroth for excellent research assistance. Financial support from the German research foundation (DFG), SFB 884 is gratefully acknowledged.

Submitted 28 June 2013

1 Introduction

Voting mechanisms are widely used for binary decisions. Their major advantages are their simplicity and the fact that truthful voting is a dominant strategy in any private values setup. However, these two features come along with a cost because voting mechanisms can only imperfectly aggregate the much finer information about individuals' valuation of the relevant alternatives. Therefore, it may occur that many agents with a small positive valuation for some decision may trigger a decision which creates much larger losses for a small number of agents. Another problem is that, within the class of voting mechanisms, ex-ante efficiency is only achieved when the threshold for a majority vote is adjusted to the underlying distribution of preferences.[1] While there are theoretical papers that show how optimal majority rules have to be adjusted when the underlying distribution of preferences changes,[2] very little is known about how this could be achieved in practice. This paper addresses both problems with a series of Bayesian voting experiments.

In theory, it would be easy to adjust the voting mechanism when the underlying distribution of preferences is assumed to be common knowledge. In practice, it is difficult to adjust majority thresholds in referenda or voting rules in committees to every specific situation since this would require either a "complete" constitution that lists all possible cases or the existence of some benevolent planner who knows the underlying distribution of preferences and adjust the voting rules to that distribution. Both options are unrealistic in practice, the first because completeness is too costly, the second because perfectly benevolent individuals may be rare and, even if they exist, hard to identify.

Instead, one may consider to let voters themselves decide on the appropriate majority threshold whenever a new issue comes up. In theory, this will solve the problem at the (ex-ante) stage where individuals do not yet know their own preferences. Specifically, if all agents' preferences are drawn from the same distribution, they would suggest the rule that maximizes expected total

[1] More complex mechanisms such as the Vickrey Clarke Groves mechanisms do not require any such information at the design stage.

[2] Feddersen and Pesendorfer (1998) study the optimal majority rule in a Bayesian jury setup with identical preferences and heterogeneous information where strategic insincere voting arises from some voting rules. Tröger and Schmitz (2012) study optimal majority rules in a setup with heterogeneous preferences. The simple majority rule is efficient if preferences are stochastically independent and their distribution is not biased in favor of one alternative.

payoff. However, it is not straightforward that a rule choice stage at a point when agents are already informed about their preferences (i.e. at the interim stage) increases welfare. We consider this to be the more relevant case in practice because we expect that, when a new issue comes up, individuals form both their preferences and their beliefs about the distribution of other players' preferences.

This paper studies whether the introduction of a stage at which agents may choose the rules according to which they decide may help to solve the problem that voting inefficiently aggregates preferences and thus increase the sum of individuals' payoffs. We run three closely related experimental treatments to understand how individuals decide when they choose voting rules. In the context of a binary choice, choosing how to choose at the interim stage may seem useless because choosing the rules is like choosing the decision itself. However, our paper shows theoretically that if subjects exhibit efficiency concerns[3], then the interim rule choice stage is not redundant. Furthermore, our experiments show that efficiency concerns indeed play an important role on the rule choice stage and that this helps our experimental participants to choose better rules which fit to the underlying distribution of preferences.

More specifically, our paper experimentally studies the choice of a collective decision mechanism at the interim stage. We consider an environment in which agents can vote on a status quo with predetermined identical payoffs (alternative B) and an alternative with unequal stochastic payoffs (alternative A). The difference of both payoffs is hereafter called an agent's valuation of decision A. Agents privately learn about their randomly chosen valuation in the first stage of the game. In our setup, introducing a procedural choice stage is useless when agents act selfishly because they should then pick the rule which makes their preferred outcome most likely to obtain.[4]

[3]In line with the existing literature we use the term "efficiency concerned" for subjects who care about the - possibly weighted - sum of individual payoffs. Efficiency concerns in this sense should not be equated with a concern for Pareto-optimality. Obviously, in a setup with positive and negative valuations, all collective decisions are Pareto-optimal if one does not permit monetary transfers.

[4]More precisely, this holds if all are selfish and this is common knowledge. Otherwise, it is at least conceivable that even if all are selfish, but do not expect all others to be selfish, they could try to use a non-selfish vote at the rule choice stage to signal something about their valuation which might trigger non-selfish voting behavior in others, from which they would in turn benefit. However, for the specific rule choice procedure that we study, this is not a concern, because whenever one's own rule choice becomes relevant and known, one can unilaterally determine the outcome and there is thus no incentive for selfish players to try to signal a different type.

Consider, for example, the case where the rule itself is determined by a vote under majority rule. All second stage rules which favor alternative A in the sense that they require a supermajority for alternative B beat all other rules if outcome A makes the majority of agents better off. This is why the two-stage procedure should yield the same results as a simple one-stage vote. In our experiment, we implement the rule choice through a random dictator mechanism. Each player in a group suggests a threshold for the second-stage vote and then one of them is randomly chosen and this rule is implemented in the second stage. This allows a player whose rule is chosen to unilaterally implement the preferred outcome. If he prefers A, he picks as the rule that unanimity is required for B and votes for A and vice versa if he prefers B. Players are informed about the rule before the vote in the second stage.

The two-stage mechanism gives subjects an option to send finer signals about their valuation. In theory, they should not make use of this option and choose extreme voting rules whenever valuations differ from zero. In practice, however, a sizable share of subjects suggest voting rules that do react gradually to their valuation.

One important finding of our paper is that sequential voting mechanisms may increase social welfare. In our experiments, we find that agents often do not act selfishly at the rule choice stage, but instead propose rules which are biased towards the efficient (sum of payoffs maximizing) rule. We consider both symmetric (baseline treatment) and asymmetric distributions of types. With a symmetric distribution the simple majority rule is efficient and the bias at the rule choice stage is accordingly, that is, players with a small absolute valuation choosing a rule that deviates from their selfish rule towards majority voting. In that case, because majority voting is the ex-ante efficient rule, we cannot expect any efficiency gains from adding a rule choice stage. However, with an asymmetric distribution, the efficient majority threshold decreases and the bias adjusts accordingly. Participants with a small absolute valuation who do not choose the selfish rule split roughly equally between the rule that maximizes total welfare (if all vote selfishly in the voting stage) and all other non-selfish rules. This behavior leads to welfare that is slightly higher than in a control treatment that uses the same distribution of preferences but a pre-determined majority voting rule. Interestingly, though, in the control treatment where subjects cannot choose the rules, they vote more often non-selfishly in the second stage.

Results from our first two treatments suggest that some agents may follow different motivations at the rule choice stage than at the voting stage. Specifically, their behavior at the rule choice stage seems to be more efficiency driven than at the voting stage, which looks mostly selfish and only occasionally in line with inequality aversion or efficiency concerns. In contrast, at the rule choice stage we find no evidence for inequality aversion and substantial evidence for efficiency concerns. Testing for inconsistencies across stages is not straightforward since one has to control for subjects' beliefs regarding other players' motivations, valuations, and strategies. We develop two additional treatments that permit to rigorously check whether a subject's behavior is inconsistent across stages. We indeed find that this is the case for a substantial share of efficiency concerned subjects. Accordingly, the introduction of a rule choice stage turns efficiency concerns into a more important attractor. Therefore, adding a rule-choice stage can potentially increase efficiency for two reasons. On the one hand, it gives players the opportunity to more finely express their preferences than a binary vote does, on the other hand, it might make players actually act in a more efficiency minded way, apparently because choosing a rule suggests that one should do the right (i.e., efficient) thing. Our experimental design does not fully exploit the first aspect, because our random dictator mechanism, which we implemented in order not to make the design too complicated, makes only very coarse use of the information that players reveal in their rule choice since it uses only the information of one player. In case of asymmetric valuation distributions, where majority voting is not ex-ante efficient, adding a rule choice stage may have an additional advantage. Even if it is known that a biased rule is ex-ante efficient, implementing this may be politically difficult for a social planner. Allowing players to choose the rule themselves, may, however, lead them to choose the ex-ante efficient rule if they are sufficiently efficiency concerned, as frequently happens in our asymmetric treatment.

Our results put into question the empirical relevance of the revelation principle. In theory, one could replace our two-stage voting mechanism by a simple one-stage mechanism and produce identical results. According to the revelation principle, this should actually hold for any multi-stage voting procedure. However, if different motivations play a role on different stages, the revelation principle would not apply. In our conclusion we discuss how this effect can be exploited in the design of collective decision mechanisms.

Our paper is related to a growing theoretical literature on optimal Bayesian voting mechanisms. One feature of most results from this literature is that the optimal design of the mechanism depends on the underlying distribution of preferences (e.g. Feddersen and Pesendorfer, 1998, Schmitz and Tröger, 2012). While it is known that mechanisms should adjust to the underlying problem, little is known about how this can be achieved in practice. Our paper is the first one to empirically address this problem.

Our paper also contributes to a literature that studies the choice of collective decision making mechanisms at the interim stage. This literature started with the seminal paper on private information in a bilateral trade setup by Myerson and Satterthwaite (1983). Further contributions on collective choice problems include Güth and Hellwig (1987), Cramton, Gibbons, and Klemperer (1987), Schmitz (2002), Segal and Whinston (2011), and Grüner and Koriyama (2011). In another closely related theoretical paper, Barbera and Jackson (2004) study the dynamic stability of pairs of voting rules (for constitutions and single issues) when single-issue rules can be changed over time. In our paper we empirically study one single decision problem and we use a random dictator mechanism to identify voters' preferred single issue rule. To the best of our knowledge, there is no experimental research on either the choice of mechanisms or the acceptance of mechanisms at the interim stage.

Furthermore, our paper contributes to the ongoing research about the role of different types of social preferences in economic experiments. Inequity averse individuals should favor the equal payoff option B unless their valuation for option A is high enough to offset the disutility from inequality. Agents who care about the maximization of the sum of individual monetary payoffs should instead be willing to accept outcome A (B) even if their valuation is slightly negative (positive) if this is welfare enhancing. Apart from the observation that we find evidence rather for efficiency concerns than for inequality aversion, one of our main results is that efficiency concerns may play a different role on different stages of a sequential game.

The rest of this paper is organized as follows. In Section 2 we explain our experimental design in detail, including the predictions based on different theories of subjects' behavior. We present the results of treatments with symmetric and asymmetric type distributions in Section 3. Results of two control treatments that test for the consistency of behavior across stages can be found in Section 4. We discuss implications of our results in Section 5. Theoretical predictions regarding the behavior of efficiency concerned subjects are derived in the Appendix.

2 Experimental design

In this section, we first describe the underlying games and derive predictions for different types of preferences for our two main treatments and a control treatment with pre-determined majority voting. Then we report the procedural details for these treatments. The design of two additional treatments, that are inspired by the results from the earlier treatments are reported subsequent to these results.

2.1 Treatment 1: Symmetric Valuation Distribution

Our first treatment is designed to study whether voting in the two-step procedure differs from rational maximization of one's own payoff and more specifically, to assess the relative importance of different motivations such as selfishness, inequity aversion, and efficiency concerns in our two-step choice procedure.

2.1.1 The Game

There are five players who collectively choose between two options, A and B. If they choose option B, all players receive a payoff of 0. If they choose option A, each player i receives a payoff θ_i. At the beginning of the game, the θ_i are drawn from an i.i.d. distribution on a given set of values. The distribution of types is common knowledge and uniform on the set $\{-7, -3, -1, -0.5, -0.2, -0.1, 0, 0.1, 0.2, 0.5, 1, 3, 7\}$. Each player only learns his own valuation θ_i. Next, players decide between options A and B in a two-stage voting process. In the first of these stages each player proposes a voting rule from the set $\{1, ..., 5\}$. If rule k is chosen, then in the second stage, k votes are required to choose option A. The actual rule is chosen by a random dictator mechanism, that is the rule that will be implemented is chosen randomly from the five voting rule that have been proposed by the five players.[5] In the second stage of the two-stage

[5] Due to a programming error, our randomization of whose rule was implemented did not work properly and thus the tie-breaking rule too often applied. Hence some subjects were far more and others far less often the dictator. However, subjects never learned that their personal choice was implemented, and even for those for whom this was never the case, the implemented rule often coincided with their choice. Hence subjects may have noticed that the chosen rule was surprisingly often or less often than expected the same as their own choice, but they should not have detected the systematic error. We also checked whether behavior differed with the frequency with which the subjects were chosen to be the rule dictator and did not find a significant difference.

voting process, players are informed about the chosen rule and then vote about the two options. Then either A or B is implemented depending on the rule chosen in the first stage and the votes from the second stage. If rule k is chosen and at least k players vote for A, then A is chosen and each player i receives θ_i. If fewer than k players vote for A, then B is chosen and all players receive payoff 0.

Players are informed about their own θ_i before the first stage but not about the complete vector of valuations. They are informed after each period of the game, however, about the realized distributions of valuations in their group. Note in particular that the choice of rules does thus not happen behind a veil of ignorance. Rules are thus to be understood here as ad-hoc rules for individual issues rather than as rules that are applied for a whole set of issues. Focusing on subjects' actions and therefore ignoring the two random moves, we will henceforth call the rule proposal stage the first stage and the voting stage the second stage of the game.

2.1.2 Theoretical Predictions

Selfish Subjects Selfish players have weakly dominant strategies. In those strategies, they must choose rule 1 (5) in case that they have positive (negative) valuations, and, in the second stage, vote for their preferred outcome if they have non-zero valuations. Voters can choose any rule/outcome if their valuation is zero, and hence the dominant strategy is not unique. All other strategies are weakly dominated by those strategies because independent of the strategies of the other players, these strategies maximize the probability for outcome A for $\theta > 0$ and maximize the probability for outcome B if $\theta < 0$.

Inequality averse subjects Predictions for inequality averse players in the two-stage game are as follows. We start with a derivation for the case that players assume others vote selfishly in the second stage, first to simplify the analysis and second, because this is what we primarily observe. We discuss deviations from this assumption below.

Players with valuations $\theta_i \leq 0$ choose rule 5 and vote in favor of alternative B. This is both maximizing their expected payoff and minimizing inequality amongst players. Thus, for negative valuations, they behave just like selfish players (but have an additional reason to prefer B). Players with valuations $\theta_i > 0$, however, face a trade-off. Alternative A yields a higher expected monetary payoff, but at the same time on expectation yields inequality. For small positive θ_i, an inequality averse player will thus still prefer B over A and choose rule 5 and vote for alternative B. For larger θ_i, however, concerns for own income dominate concerns for inequality and the player will prefer alternative A and thus choose rule 1 and vote for alternative A. For example, straightforward calculation shows that for $\alpha_i = 1$; $\beta_i = 0.3$ in the inequality-aversion model by Fehr and Schmidt (1999), i prefers B for $\theta_i \leq 1$. According to a classification made by Fehr and Schmidt based on ultimatum game data, about 40% of subjects exhibit inequality aversion at least as strong. According to an estimate of individual parameters by Blanco, Engelmann, and Normann (2011), it is satisfied by about 30% in their experiment. Note that $\alpha_i = 1$ implies that a subject would reject offers of less than 1/3 of the total pie in an ultimatum game. Even for substantially weaker inequality aversion, one would still get preferences for B for positive θ_i, e.g., for $\alpha_i = 0.25$; $\beta_i = 0$ in the Fehr-Schmidt model (implying rejecting offers of less than 1/5 in an ultimatum game), i prefers B for $\theta_i \leq 0.2$.

Also note that subjects who satisfy the assumption $\alpha_i \geq \beta_i$ as in the Fehr-Schmidt model would not choose intermediate rules unless they are strongly inequality averse and θ_i is large. If player i is considering A potentially attractive because θ_i is positive, then if i minds disadvantageous inequality (i.e., $\alpha_i > \beta_i$) more than advantageous inequality, she would rather have A when others have smaller valuations than her than if others have larger valuations than her. For θ_i relatively small, this is close to wanting A precisely when others vote for B but not if others vote for A.[6] This implies that A is most attractive precisely when all others vote against it and thus when A results from her own vote alone under rule 1. Every other rule results in A only when several other players have a positive valuation and thus when i likes A less. Therefore, i will switch from rule 5 straight to rule 1. For relatively large θ_i this argument only holds when α_i is substantially larger than β_i, because for large θ_i the expected advantageous inequality is larger than the expected disadvantageous inequality.[7] Note however, that in the estimation of the distribution of individual (α_i, β_i) pairs by Blanco, Engelmann, and Normann (2011), 38% among the subjects violate the $\alpha_i \geq \beta_i$ constraint.

To summarize, inequality averse players would show a similar choice pattern as selfish players,

[6] It matters here what i assumes about the other players' inequality aversion. If i considers others to be similarly inequality averse, then this relationship also holds for larger values of θ_i.

[7] One can show that for $\theta_i = 0.5$, i would not choose an intermediate rule if $\alpha_i \geq \frac{3}{2}\beta_i$ and for $\theta_i = 1$, i would never choose an intermediate rule for $\alpha_i \geq 2\beta_i$.

but their switch from rule 5 to rule 1 would not occur at $\theta_i = 0$, but at some $\theta_i > 0$. There are a number of alternative assumptions that one could make regarding the preferences and voting behavior of others. First, subjects could have rational expectations. Second, they could be subject to a so-called "false" consensus effect (Ross, Greene, and House, 1977, Engelmann and Strobel, 2000) and have believes that are biased towards their own type, in the extreme case believing that others behave exactly as they do. Given that subjects only observe actual votes of other participants, but not the relation between individual valuations and votes it is rather implausible that they have rational expectations regarding the others' voting behavior. Furthermore, actual voting behavior is very close to selfish voting, so for the actual behavior that we observe replacing the assumption that subjects expect selfish voting with the assumption that they have rational expectations regarding the other participants' voting behavior makes very little difference.

If an inequity averse player i expects other players to be equally inequity averse, as a strong consensus effect would predict, then i would expect others with positive valuation to occasionally vote for B. This increases i's estimate of the average valuation conditional on i being pivotal, which in turn would reduce the expected advantageous inequality but increase the expected disadvantageous inequality. Depending on i's inequality aversion parameters, this may make A more or less attractive and hence change i's voting behavior for $\theta_i > 0$. More importantly, however, it does not affect i's rule choice and voting behavior for $\theta_i \leq 0$. Outcome A then can only yield non-negative payoffs and inequality and is thus never preferred over B. The same holds also for rational expectations or indeed all others beliefs that i might have. So also for alternative assumptions regarding players' beliefs, the prediction for inequality averse players remains that they behave like selfish players for negative valuations, but may behave differently for positive valuations.

Efficiency concerned subjects According to Tröger and Schmitz (2012), with a symmetric distribution of types, simple majority rule maximizes the expected sum of individual payoffs. Suppose for a moment that all voters can be expected to vote "selfishly" in the second stage (we again address below the case that others do not vote selfishly in the second stage). In this case, players who consider the maximization of the sum of payoffs as their objective should propose rule 3 independently of their own valuation. In the second stage these voters must also reply by voting in favor of the alternative that maximizes their own payoff if indeed rule 3 is chosen. Therefore, suggesting rule 3 and voting selfishly is an equilibrium of the two-stage voting game amongst such players.

The choice of less than fully efficiency concerned subjects is more complex. Such agents care more about themselves than about others. In the appendix we assume that subjects maximize a convex combination of all agents' monetary payoffs (with a larger weight on their own monetary payoff) and show that when they expect selfish play in stage 2, they may propose rules 1, 2, 3, 4, or 5, depending on the value of θ_i. The proposed rule varies monotonically with the valuation. We note in particular that for the symmetric distribution of valuations in Treatment 1, the deviation from selfish rule choices is symmetric around 0 and that the suggested rule will always be between the selfish rule and majority rule. That is if i has efficiency concerns and $\theta_i < 0$, i may choose rule 3, 4, or 5, but not rules 1 or 2 and if $\theta_i > 0$, i may choose rule 1, 2, or 3, but not rules 4 or 5.

We can again consider alternative assumptions regarding the beliefs of subjects concerning the other participants' behavior. As argued above, rational expectations are neither a very plausible assumption, nor would that assumption make much of a difference. Assuming alternatively that others are equally efficiency concerned, as suggested by a strong consensus effect, leads to complications if efficiency concerns are relatively strong because players may then suggest a rule that is more strongly biased in one direction but vote against it if they take into account that choosing a different rule will change the voting behavior of others. However, in the experiment we observe essentially no evidence of strategic voting and therefore it does not appear to be a realistic assumption that players expect others to vote strategically.

We furthermore note that in Treatment 1 we cannot necessarily distinguish efficiency concerned subjects from those that have a preference for majority voting per se (for example because they equate that with democracy and have a preference for democracy). A relevant question is whether a participant with preferences for majority voting in this sense would also consider Rules 2 and 4 a reasonable compromise between selfishness and this concern for majority voting. In this case, such a participant would be indistinguishable from one who is efficiency concerned. It might be more plausible that a subject with preferences for majority voting would find Rules 2 and 4 poor compromises and thus only choose between Rules 1, 3 and 5, depending on θ_i and the intensity of her preferences for majority voting. This implies that at least subjects who choose all five rules in line with efficiency concerns cannot be explained by concerns for majority voting.

Furthermore, we note that selfish behavior with error where less costly errors are more likely to be made could produce patterns of behavior that look similar to that of efficiency concerned subjects. In the asymmetric Treatment 2, we can, however, distinguish between efficiency concerned behavior and both alternatives, preferences for majority voting per se and selfishness with error.

2.2 Treatment 2: Asymmetric Valuation Distribution

The main purpose of our second treatment is to establish whether the two-step voting procedure can increase efficiency over majority voting. We thus change the distribution of valuations such that majority voting is not ex-ante efficient. This treatment is intentionally designed as giving the two-step procedure a chance to provide efficiency gains compared to the simple majority rule. However, note also that the random dictator mechanism, which we chose primarily to make the design not too complex for the participants, is a rather inefficient way of aggregating preferences. Therefore, choosing a two-step procedure that makes better use of the preferences of participants regarding the voting rule may be more likely to produce efficiency gains.

2.2.1 The Game

In the second treatment, we have replaced the symmetric probability distribution from Treatment 1 by an asymmetric one with positive expected payoffs. The distribution of types is

$$\theta_i = \begin{cases} 14 & \text{with probability } 0.2 \\ -1 & \text{with probability } 0.8 \end{cases} .$$

This game has the same type of Bayesian Nash equilibrium for selfish players as the game that underlies Treatment 1. The efficient rule here is rule 1. Indeed the ex-post efficient outcome will always be reached if rule 1 is implemented and all players then vote according to their material preferences.

2.2.2 Predictions

A purely selfish subject should always vote in favor of his preferred alternative in stage 2. Moreover, he should propose rule 1 if $\theta_i = 14$ and rule 5 if $\theta_i = 14$. An inequality averse player i acts like a selfish subject if $\theta_i = -1$ and also if $\theta_i = 14$ unless i's aversion to advantageous inequality is extreme (that is for $\beta < \frac{14}{15}$ in the Fehr-Schmidt model, i would choose rule 1 if $\theta_i = 14$, and if i believes others vote in line with payoff maximization in second stage).

A sufficiently efficiency concerned subject (i.e., one who puts at least a third of the weight on the average payoffs of the others as on her own) would choose for $\theta_i = -1$ rule 1 and vote for B. Note in particular that efficiency concerns here impact on the rule choice, not on the voting behavior if indeed the efficient rule 1 is chosen. However, if the rule is rule 2, 3, 4, or 5, an efficiency concerned player with $\theta_i = -1$ would in equilibrium vote with positive probability for A. An efficiency concerned player with $\theta_i = 14$ would choose rule 1 and vote for A unless he actually cares substantially more for each other players earnings than his own, in which case he would vote for rule 2 and vote for A.

Given that inequality averse players should always vote in line with their monetary preferences, as should efficiency concerned players whenever the efficient rule 1 is chosen, considering alternative beliefs about others' behavior is not of much interest. If efficiency concerned players expect others to be efficiency concerned and thus vote for A with positive probability if rules 2, 3, 4, or 5 are chosen even if they have $\theta = -1$, this does still not make choosing these rules more attractive than choosing rule 1 and voting for B (assuming all others to do so when they have $\theta = -1$ and rule 1 is chosen) because only this choice ensures that A is chosen if and only if at least one player has $\theta = 14$.

2.3 Treatment 3: Asymmetric Valuation Distribution with Pre-determined Majority Voting

The purpose of our third treatment is to serve as a benchmark for Treatment 2 to assess whether the two-step procedure affects welfare compared to the most obvious simple alternative to decision making in the given context, namely simple majority voting.

2.3.1 The Game

Treatment 3 works with the same probability distribution as Treatment 2. However, it starts directly with the voting stage and the rule is always the majority rule. Hence players are informed about their valuations and then vote directly. Expected equilibrium payoffs are given by

$$0.2^5 \cdot 14 + 5 \cdot 0.8^1 \cdot 0.2^4 \cdot \frac{4 \cdot 14 - 1}{5} + 20 \cdot 0.8^2 \cdot 0.2^3 \cdot \frac{3 \cdot 14 - 2}{5} = 0.894$$

2.3.2 Predictions

Purely selfish subjects should vote for the alternative that maximizes their payoff. Efficiency concerned subjects with a positive valuation should always vote for alternative A. When all agents are sufficiently efficiency concerned, there is no sincere voting equilibrium. Instead, voters with a negative valuation play a mixed strategy in a symmetric equilibrium. Inequity averse subject with a negative valuation vote for alternative B, and with a positive valuation they vote for A, unless they are absurdly inequity averse (i.e., as long as others vote sincerely, they will do so as long as $\beta < \frac{14}{15}$ in the Fehr-Schmidt model.) Hence when all agents are either selfish or somewhat, but not excessively inequity averse, there is an equilibrium in which all agents vote sincerely.

To summarize, sincere voting can be motivated by selfishness and inequity aversion. Insincere voting of subjects with a negative valuation is compatible with efficiency aversion but not with selfishness or inequity aversion.

2.4 Procedures

The computerized experiments were run at the experimental laboratory mLab at the University of Mannheim with software programmed in z-Tree (Fischbacher 2007). Subjects were recruited using ORSEE (Greiner, 2004). In each session, 15 periods were played with random matching among subjects. In each session, between 10 and 20 subjects took part. We ran three sessions each for Treatments 1, 2, and 3 for a total of 45 subjects in Treatment 1 and 55 in both Treatment 2 and 3. To avoid any income effects we did not pay all periods, but at the end of each session, one period was chosen randomly, and that period determined the payoffs of all players. If in that period alternative A was chosen in subject i's group, then i's payoff was given by $\pi_i = \theta_i$, otherwise $\pi_i = 0$. Each subject i earned a show-up fee of 9 Euro plus the experimental payoff π_i, hence earnings could be between 2 and 16 Euro in Treatment 1 and either 8, 9 or 23 Euro in Treatments 2 and 3.

3 Results

3.1 Treatment 1

3.1.1 Second-stage Voting Behavior

We begin by discussing the behavior in the voting stage. Voting in the second stage is overwhelmingly in line with the maximization of the subjects' own payoff. Overall, only 4.3 percent of the votes cast are in favor of the alternative which does not maximize the subject's monetary payoff (i.e., are either for A when $\theta_i < 0$ or for B when $\theta_i > 0$). Table 1 shows the share of deviations for different valuations. We aggregate large negative, small negative, small positive and large positive valuations, respectively, and drop $\theta_i = 0$ because either vote is in line with payoff maximization in this case. As we can see in Table 1, deviations from payoff maximization occur most frequently for small positive valuations $0.1 \leq \theta_i \leq 0.5$. This is where inequality aversion would predict votes against own payoff maximization, so this suggests that inequality aversion may play a role for some subjects. We note, however, that the rate is still very small with 10.4%. Furthermore, voting for A when $\theta_i < 0$ is difficult to reconcile with any preference model, so seems likely to result from errors. Thus taking the 3.3 percent of votes against own payoff maximization for negative valuations as an estimate of the error rate, only about 7 % of the observations are in line with inequality aversion, where this is expected to matter most. Remember that only very small degrees of inequality aversion are needed in order to imply a preference for B when $0.1 \leq \theta_i \leq 0.5$. We also note, that, interestingly, we found no evidence of strategic voting, i.e., taking into account what being pivotal implies about the other's valuations. Votes are against own monetary preferences least often for rule 5 (2%) and most often for rule 4 (8%), which is only chosen 50 times altogether, so that this amounts to only 4 observations. Specifically, the rule that is in place does have much impact on the pattern presented in Table 1, with the interesting exception that of the 16 votes most indicative of inequality aversion (for B if $0.1 \leq \theta_i \leq 0.5$) 11 occur when rule 1 is implemented and thus when only one vote is needed to implement A. Note that in this *case*, a player is pivotal actually only if all others vote for B (including the rule chooser who thus must have made an error in one of the stages), and are thus likely to have negative valuations, and hence B is also on expectation maximizing efficiency.

$-7 \leq \theta_i \leq -1$	$-0.5 \leq \theta_i \leq -0.1$	$0.1 \leq \theta_i \leq 0.5$	$1 \leq \theta_i \leq 7$
3.5%	3.0%	10.4%	1.8%

Table 1: Deviations from selfish voting in the second stage

$-7 \leq \theta_i \leq -1$	$-0.5 \leq \theta_i \leq -0.1$	$0.1 \leq \theta_i \leq 0.5$	$1 \leq \theta_i \leq 7$
14.6%	39.3%	32.5%	18.8%

Table 2: Deviations from selfish voting in the first stage

3.1.2 First-stage Behavior

More interesting than the voting behavior are the rules that subjects choose. In contrast to the voting behavior in the second stage, subjects' proposals of rules in first stage frequently deviate from the payoff-maximizing rule. Overall, 25.4 percent of the proposals made do not coincide with a payoff maximizing voting rule. According to Table 2, which shows the rate of deviations from the selfish rule and again groups valuations into large negative, small negative, small positive and large positive valuations, there is a clear pattern. First, deviations from the payoff-maximizing rule are more frequent for small absolute valuations than for large absolute valuations, which is in line with any model that combines selfishness with some other motive. Second, deviations are about equally likely for negative valuations than for positive valuations, a pattern that is in line with efficiency concerns (or selfishness with error or preferences for majority voting). In the appendix of this paper we show that for our model of efficiency concerns, the proposed rule is monotonous in the agent's payoff and that the deviation from a selfish rule proposal is stronger, the smaller the absolute value of the valuation.

One obvious reason why there may be more deviations from selfishness in the first stage than in the second stage is that finding the selfish rule is more complex. Furthermore, for the choice of rules, there is only one way to be "right" (i.e., payoff maximizing) but four ways to be wrong. Both these arguments would suggest that pure confusion may contribute to the higher share of deviations from selfishness in the first stage. However, the smaller rate of deviations for larger absolute valuations suggests that the non-selfish rules are not chosen plainly out of confusion. Furthermore, some rules make sense given our model of efficiency concerns (or for preferences for majority voting), namely those that range from the selfish rule to majority voting whereas others

θ_i	1	2	3	4	5	average rule
-7	4	0	4	1	50	4.57
-3	2	2	3	0	44	4.61
-1	1	3	3	2	52	4.66
-0.5	1	2	3	8	28	4.43
-0.2	1	3	5	11	26	4.26
-0.1	2	0	11	6	28	4.23
0	7	1	22	2	13	3.29
0.1	32	3	7	3	4	1.86
0.2	29	9	4	0	4	1.72
0.5	43	5	10	0	1	1.49
1	47	2	5	3	2	1.49
3	46	3	7	1	0	1.35
7	45	1	3	4	1	1.43

Table 3: Distribution and average of the proposed rule for the different valuations in Treatment 1.

do not, namely those that are biased against the maximization of subjects' own payoff, i.e., the chosen rule favors alternative A (Rules 1 and 2) while the subject has a negative valuation or it favors B while the subject has a positive valuation (Rules 4 and 5). We find about 7 percent of the latter type of rule choices in each of the four groups of valuations. Interestingly, for $0.1 \leq \theta_i \leq 0.5$, where mild degrees of inequality aversion are consistent with the choice of rule 5 (and possibly 4), they are not chosen more frequently than the rules that are biased against the maximization of subjects' own payoff are chosen in any of the other groups of valuations. This suggests that inequality aversion does not play an important role on the rule choice stage. Given that 7 percent of rule proposals are for the two rules that make little sense given θ_i, a fair estimate of the error rate might be $2 * 7 = 14$ percent, given that there are four non-selfish rules. This would also explain most of the non-selfish choices for $|\theta_i| \geq 1$ as errors. In contrast, it would leave us with about 22 percent of non-selfish rule choices for small (≤ 0.5) absolute valuations.

A key prediction of our efficiency-concerns model and key distinction to the predictions by inequality aversion are that deviations from the selfish rule are not only decreasing in absolute valuation but are also symmetric around 0. In order to assess this prediction, we show the distribution of the chosen rules for each θ_i in Table 3. Looking at the disaggregated distribution, we see that for $|\theta_i| \geq 1$, the vast majority of choices are for the selfish rule, while the remaining choices are spread fairly evenly across the other rules, which would be in line with noise. In contrast, for $|\theta_i| < 1$, while still more than half of the rule choices are for the selfish rule, a substantial share is in each case in favor of the efficient rule 3 or the rule in between the selfish and the efficient rule. As we can see, the deviations from the selfish rule are quite symmetric. This pattern is in line with efficiency concerns. To test for the symmetry of the distribution, we compare for each valuation θ_i the distribution of chosen rules with the inverted distribution for $-\theta_i$. According to Fisher exact tests, the distributions are not significantly different for any $\theta_i \in \{0.1, ..., 7\}$ ($p > 10\%$ in each case).[8]

In contrast, rule choices that are biased against own monetary preferences are not chosen more often for $0 < \theta_i \leq 0.5$, where they would be in line with inequality aversion than for $\theta_i \geq 1$ or $\theta_i < 0$. We also see that for $\theta_i = 0$ the most frequently chosen rule is 3, as predicted by efficiency concerns, and not 5, which would be predicted by inequality aversion. Hence, Table 3 provides very little evidence in favor of inequality aversion. We also note that the median rule is always the selfish rule, so even for very small own monetary gains or losses, an absolute majority always goes for the selfish rule.

For a more thorough assessment of the rule choice, we run regressions of the rank number of the chosen rule on the valuation with individual random or fixed effects. For negative valuations, the chosen rule decreases only insignificantly in the valuation ($p > 0.3$), whereas, for positive valuations, the chosen rule decreases significantly in valuation ($p < 0.01$). Efficiency concerns would predict such an effect both for positive and negative valuations, inequality aversion only for positive valuations. Thus, these regressions suggest some role for inequality aversion. On the other hand, simply running probit regressions (again with individual random or fixed effects)

[8] The table also shows the average chosen rule. We note that one should not take these average rules too seriously as it is fair to argue that the rules are on an ordinal scale and not a cardinal scale. Thus, the averages serve primarily for illustration. For any given level of θ_i, comparing the absolute difference between the average rule and the selfish rule (1 or 5), with this difference for $-\theta_i$, we see that they never differ by more than 0.15.

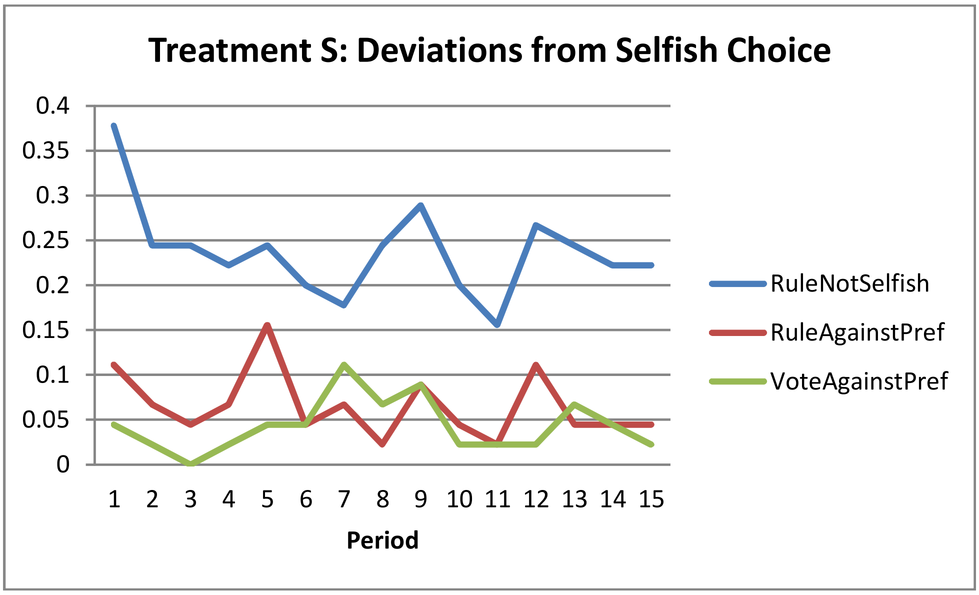

Figure 1: Share of choices that differ from the selfish prediction across periods. VoteAgainstPref corresponds to votes for A given $\theta < 0$ or votes for B given $\theta > 0$. RuleNotSelfish summarizes all rule choices that are not for rule 5 when $\theta < 0$ or not for rule 1 when $\theta > 0$. RuleAgainstPref is the subset where the rule is biased against the subject's own monetary preferences, i.e., for rule 1 or 2 when $\theta < 0$ or for rule 4 or 5 when $\theta > 0$.

for a dummy whether subjects are choosing a non-selfish rule find that this depends significantly ($p < 0.01$) on the valuation both for negative and positive valuations, supporting that efficiency concerns play a major role.

Deviations from selfish behavior does not decrease substantially over time. Figure 1 shows the trend over time of votes and rule choices that are not selfish. Except for a drop of non-selfish rule choices from the first to the second period, there is no clear discernible time trend. Hence, deviations from the selfish prediction are unlikely exclusively errors because we would then expect a decreasing time trend.

The analysis so far has only taken aggregate data into account and has found support that deviations from selfishness are largely in line with the predictions of our model of efficiency concerns. An important question is whether individual subjects behave according to a specific preference model across the periods where they draw different valuations. In order to do this, we classified the 45 subjects according to the following four criteria, which are key predictions of our efficiency model.

1. The subject always chooses rule 3 when $\theta_i = 0$.
2. The subject never choose a rule which is biased against selfish preferences.
3. The subject chooses at least once a rule that is not implementing selfish preferences.
4. The subject has a rule profile that is weakly monotone in the sense that the average rank number of the proposed rules is weakly decreasing in θ_i.

We call a subject consistently efficiency oriented if it satisfies all four criteria, with one deviation against (2) or (4) permitted. 11 out of 45 subjects are in this sense consistently efficiency oriented. We can also classify subjects as consistently selfish if they (i) propose rule 5 whenever $\theta_i < 0$ (ii) propose rule 1 whenever $\theta_i > 0$ (iii) vote for alternative B whenever $\theta_i < 0$ (iv) vote for alternative A whenever $\theta_i > 0$. We find 17 subjects who are consistently selfish. This number increases to 23 if we allow for one deviation on each stage. Finally, we classify subjects as inequality averse if they (i) propose rule 5 whenever $\theta_i \leq 0$ (ii) propose at least once rule 4 or 5 for $\theta_i > 0$ (iii) vote for alternative B whenever $\theta_i \leq 0$ (iv) vote for alternative B at least once when $\theta_i > 0$. We find no subject that satisfies these criteria perfectly, and three who satisfy it on the rule choice stage only, where two of them only once choose a non-selfish rule.

Remember that for the symmetric distribution of valuations, the simple majority rule is efficient and given an exogenously chosen majority rule, selfish voting is efficient. Thus, because the random dictator rule that we employ cannot result in a rule that is ex-ante more efficient than the majority rule (though at an interim-stage if the valuation of the dictator is $\theta = 7$ or $\theta = -7$, it might result in the more efficient rules 2 and 4, respectively) but often result in less efficient rules, it is not surprising that the total payoffs in Treatment 1 are about 35 percent lower than what would be obtained with pre-determined majority voting and selfish voting behavior given the draws of valuations we observe. However, given the rules of our experiment, the deviation from selfish behavior leads to improvements in welfare and the total payoffs of our game are about 8 percent higher than the payoff that selfish behavior on both stages would yield given the draws of valuations realized in the experiment.

3.2 Treatments 2 and 3

We again start by considering voting behavior in the second stage. In the asymmetric Treatment 2 with rule choice stage, the observed voting behavior of subjects with positive valuations ($\theta_i = 14$) is consistent with selfishness and efficiency concerns (as well as non-extreme form of inequality aversion). Only in 3 percent of cases, subjects with $\theta_i = 14$ voted for alternative B. This would correspond to the same error rate as observed in Treatment 1. Subjects with a negative valuation ($\theta_i = -1$) vote for alternative A in 13.6 percent of the cases. Their behavior is inconsistent with selfishness and inequity aversion, but is consistent with efficiency concerns if the rule is not rule 1 (because in that case, A would only maximize total payoffs if at least one player has $\theta_i = 14$, when one could expect that this player votes for A herself). Note, again, that the rate of non-selfish votes is relatively low, given that quite some efficiency gains are possible, so even moderately efficiency concerned players could be expected to vote for A if the rule is not rule 1.

Interestingly, subjects' voting behavior is statistically independent of the implemented voting rule. This is inconsistent with efficiency concerns that suggest a vote for A only for rules unequal to rule 1. Not very surprisingly, though, voting behavior is correlated with subjects' own rule choice. Only 3 percent among those subjects who choose the selfish rule vote for an outcome that does not maximize their own payoff (so they are consistently selfish), but between 19 and 33 percent among those who choose a non-selfish rule vote against their monetary preference in the second stage.

Regarding the rule choice stage, subjects with a valuation of $\theta_i = 14$ choose the selfish rule (rule 1) in 88 percent of cases (see Table 4). Among the 12 percent remaining proposals 6 percent are for the simple majority rule. The latter observation indicates that these subject may have a preference for democracy – interpreted as a preference for majority voting. In contrast, subjects with a valuation of $\theta_i = -1$ choose the selfish rule (rule 5) in only 56 percent of cases and the efficient rule (rule 1) in 19 percent of cases. The behavior of these 19 percent is inconsistent with selfishness and inequality aversion, but in line with efficiency concerns. Furthermore, 9 percent is for the simple majority rule, which could suggest a preference for majority voting per se or for mild efficiency concerns. While we thus observe that substantially more subjects deviate from the selfish rule if they have $\theta_i = -1$ than if they have $\theta_i = 14$, still a majority chooses the selfish rule 5 and in 97 percent of the cases than also vote for B and thus enforce B even if all other subjects have $\theta_i = 14$, which amounts to quite substantial selfishness.

θ_i	1	2	3	4	5
-1	123	72	58	38	365
14	149	5	10	4	1

Table 4: Distribution and average of the proposed rule for the different valuations in Treatment 2.

The actual average total payoff in Treatment 2 per period was $W_a = 6.52$. This value exceeds the hypothetical average total payoff with majority rule, selfish voting behavior and the observed valuation distribution of $W_s = 2.18$. Thus the combination of letting subjects choose the rule and their deviation from selfishness yields a quite substantial efficiency gain. Part of this gain stems from the fact that the efficient rule is sometimes chosen, part of it results from non-selfish voting when the rule is not rule 1. We thus also compare the total payoff to the hypothetical average total payoff that would result under exogenously chosen majority rule given the average voting behavior observed if the majority rule was chosen in the experiment and the observed valuation draws, which is $W_o = 4.51$. Since $W_o > W_s$, but $W_o < W_a$, the observed efficiency gains are derived indeed both from the choice of the efficient rule and from non-selfish voting. We also note that the efficient rule is sometimes chosen because a player with $\theta_i = 14$ is chosen as the dictator. Hence the choice of the efficient rule is not exclusively based on altruism, but sometimes also simply on allowing somebody with strong interests to have his will.

The comparison of Treatments 2 and 3 allows us to assess whether our two-stage decision process leads to efficiency gains compared to actual behavior given an exogenously given majority rule. Using an OLS regression with robust standard errors (controlling for the number of positive valuations in the group), we find that with rule choice stage, the average individual payoff is about 47 cents higher per period than under exogenously given majority voting. This difference is far from significant ($p > 0.5$). It is also smaller than what one would expect from looking at the behavior in Treatment 2. The reason is that for exogenously given majority rule, voting behavior is less selfish than observed with the two-stage procedure. Specifically, the probability to vote for A for $\theta_i = -1$ is by 11 percent-points lower in Treatment 2 with rule choice than in Treatment 3 and according to a linear probability model, highly statistically significant ($p < 0.01$, also if standard errors are clustered at group level). Hence in Treatment 3 with pre-determined majority voting, efficiency gains are rather obtained by non-selfish voting, in Treatment 2 with rule-choice stage by choosing the efficient rule. This suggests that subjects consider non-selfish behavior at the two stages partly as substitutes. They may engage in moral licensing, voting selfishly in the second stage if they have chosen a non-selfish rule, even if that rule was not chosen.

3.3 Interpretation

There is a non-trivial share of choice on the rule choice stage of Treatment 1 that is compatible with the assumption of efficiency concerns. Similarly, in the asymmetric Treatment 2, 19 percent of the subjects' rule proposals given a negative valuation can best be explained by efficiency concerns, and another 25 percent would be consistent with at least some degree of efficiency concerns. In contrast, in only 14 percent of cases, subjects with negative valuation vote for alternative B in Treatment 2. This share is, however, higher in Treatment 3, with exogenously given majority rule. Since efficiency concerned subjects should actually consider mixed strategies in Treatment 3, this may constitute a lower bound on the share of efficiency concerned individuals in Treatment 3. Overall, we observe non-trivial shares of rule choices as well as of votes that are in line with some degree of efficiency concerns.

Interestingly, in both Treatment 1 and 2, the share of own payoff maximizing votes on the second stage is higher than the share of own payoff maximizing choices on the rule choice stage. While this may reflect that efficiency concerns have little effect in the voting stage as argued above, it might also be that subjects approach the two stages differently and do not behave consistently with the same preference model across both stages. Partly, this may reflect some moral licensing, as subjects who have chosen a non-selfish rule feel justified to then cast a selfish vote, even if a different rule has been chosen. This is in line with the higher share of non-selfish votes in Treatment 3 then in Treatment 2 (if the majority rule is chosen in the latter) because in Treatment 3 such moral licensing does not apply because no rule is chosen. Alternatively, subjects may perceive the rule-choice stage as a more principled decision, where they should do the "right thing", whereas voting selfishly in the second stage is perhaps considered more acceptable. This would, however, not explain the higher share of non-selfish votes in Treatment 3. We address this apparent inconsistency in Treatments 4 and 5, where we eliminate uncertainty about what it means to be pivotal and can thus make clear predictions about consistent behavior across the two stages, independent of assumptions about beliefs.

Another interesting result in Treatments 1 and 2 is that some participants appear to have a preference for the majority rule (rule 3). It is chosen more often than both rule 2 and rule 4 in Treatment 1 and in Treatment 2 by those subjects with $\theta = 14$.

From the comparison of Treatments 1 and 2 we see that subjects very clearly react to the distribution of valuations. In particular, there are more choices for the efficient rule in Treatment 2 where efficiency gains are higher. This change in behavior is in line with subjects being concerned with the maximization of total payoffs.

4 Control Treatments: Identifying inconsistent behavior across stages

4.1 Motivation and Experimental Design

The results of Treatments 1 and 2 indicate that several agents behave differently across both stages. This holds for agents who consistently vote selfishly in the second stage and propose efficient rules in the first stage of the game. This raises the question whether subjects' behavior is consistent with utility maximization across stages. Related to that is the question whether each stage appeals to a different behavioral attractor in the sense that subjects act more altruistically in the rule-choice stage than in the voting stage. Our last two treatments are designed to rigorously test for inconsistencies across stages. In particular, in the first of these treatments, we do not need to worry about beliefs of subjects about the other subjects' rule choices and hence whether the chosen rule reveals something about what being pivotal implies for the other subjects' valuations. Treatment 4 implements the following game that is designed to test for any inconsistency across stages. Treatment 5 is a simplified variant that will be explained below.

In this game we consider three players. First nature determines players' valuations, where $\theta = -1$ or $\theta = 14$, both with probability $\frac{1}{2}$ (because the game is rather complicated, we decided to simplify the assignment of valuations and both outcomes being equally likely is easier to explain and understand than one occurring with probability $\frac{4}{5}$). Then as in the other treatments a two-stage voting procedure follows In the first stage a rule is suggested by each player. Each player's choice is limited to rules 1 and 2, i.e., rules that require either one or two votes to pass A.[9] Then either one player's rule choice is randomly selected to be implemented, or rule 2 is randomly chosen

[9] A similar argument could be made by restricting the rule set to rules 2 and 3.

to be implemented as the default. The probability for each player's rule to be implemented is $\frac{1}{4}$. Otherwise rule 2 is implemented for sure, which also happens with probability $\frac{1}{4}$. In the second stage of the voting procedure all agents cast their intended votes in favor of alternative A or B. However, only one randomly chosen subject can freely decide how to vote. The others are "forced" to vote in line with their monetary preferences, i.e., to vote for B if $\theta < 0$ and to vote for A if $\theta > 0$. (To gather more data, all vote, but all but one of the votes will be replaced with the money-maximizing vote).

We represent a strategy of a player who has a given valuation θ by a vector (R, v_1, v_2, v_c) , where $R \in \{1, 2\}$ denotes the rule proposed, and $v_1, v_2, v_c \in \{A, B\}$ denote the vote under rule 1, rule 2 proposed by a player, or rule 2 selected by the computer, respectively. Our test for consistency is based on the assumptions that (i) all subjects expect players with a positive valuation to always choose rule 1 and to vote in favor of alternative A for any implemented rule and (ii) all players with a negative valuation vote in favor of alternative B if rule 1 is implemented, and all players expect the others to do so.

We first show that a player i who proposes rule 1 even though $\theta_i < 0$ weakly prefers alternative A in case that exactly one other player has a positive valuation. A player i with $\theta_i < 0$ who proposes rule 1 chooses, under the above assumption (ii), strategy $(1, B, B, B)$, or $(1, B, A, A)$, or $(1, B, B, A)$, or $(1, B, A, B)$ (under the standard assumption that by the time they propose a rule the players actually have a complete plan how to vote contingent on all possible implemented rules, an assumption that may well be violated for experimental participants).

Consider first strategy $(1, B, B, B)$. Alternative A is chosen if (i) rule 1 is selected and exactly one other subject has a valuation $\theta > 0$ or (ii) two other subjects have a valuation $\theta > 0$ or (iii) exactly one other subject has $\theta > 0$ and one other subject with $\theta < 0$ votes for A and her vote is chosen to count. Compare this to the strategy $(2, B, B, B)$. Here, alternative A is chosen under the same conditions as above. Since rule 1 obtains with a lower probability under the second strategy and if exactly one other player has $\theta > 0$, A is more likely to result under rule 1 than under rule 2[10], i must be weakly in favor of alternative A, conditional on exactly one player having a positive valuation, if i chooses $(1, B, B, B)$ rather than $(2, B, B, B)$. The same reasoning holds

[10] Note that given that i votes for B under rule 2, and her own vote may be chosen to count, A will not result for sure if exactly one other player has $\theta > 0$, even if i assumes the remaining player would vote for A under rule 2 for sure.

for comparing strategy $(1, B, B, A)$ to $(2, B, B, A)$ because this only possibly leads to a different outcome if rule 2 was chosen by default, which is beyond the control of the players. For strategies $(1, B, A, A)$ and $(1, B, A, B)$ the same argument applies unless player i believes that both other players vote for A for sure if rule 2 is chosen by a player. For such a belief, the outcome would actually always be the same for $(1, B, A, v_c)$ and $(2, B, A, v_c)$ because in that case both when rule 1 is implemented and when rule 2, chosen by a player, is implemented, A will result whenever at least for one player $\theta > 0$. Thus we make the mild additional assumption that i does not assume that with certainty both other players are sufficiently altruistic such that they will vote for A even if they have $\theta < 0$ when rule 2 is implemented.

Now consider i in the voting stage when the rule has been determined to be rule 2, not chosen by any subject, so i has not learned anything about the other subjects' payoffs. Now if i's vote is relevant, then the other two players are forced to vote according to their monetary preferences. Thus voting for A implements A and i is pivotal if exactly one of the other players has $\theta > 0$. Thus if i prefers A when exactly one of the others has $\theta > 0$, then i should vote for A. Hence if i chooses rule 1, but votes for B if the rule is exogenously chosen to be rule 2, then this is inconsistent with our model, but more generally with any model that is consistent with choosing rule 1 in the first place (which must include altruism whereas social preferences such as inequality aversion do not yield voting against monetary preferences for anybody with $\theta < 0$).

Note that even if player i thinks the others are altruists who would want to vote for A given rule 2 to help one player with a positive payoff, i does not have an incentive to strategically choose rule 1 for $\theta_i < 0$. Given that only one player can freely choose in the voting stage, it cannot happen given rule 2 that A gets implemented even though both other players have a negative valuation but are altruistic. Thus given rule 2, A can only result if at least one of the others has a positive valuation. But if that is the case, A also results if i chooses rule 1 (unless the player with a positive valuation is chosen to be the one who can vote and votes for B, which is unlikely), and A will also result if rule 1 is chosen and a confused altruist's vote is chosen to count. Thus if $\theta_i < 0$ and i wants B, then i should choose rule 2.

Because the participants exhibited substantial problems in understanding the instructions in Treatment 4, which implemented this game, Treatment 5 was based on a simplified variant of this game. Specifically, the option that the computer automatically implements rule 2 with probability $\frac{1}{4}$ was eliminated. This simplifies the description of the first stage substantially. In the first stage now simply all three players choose a rule and one of them is then randomly selected and her chosen rule is implemented. The second stage remains unchanged, because for the test of inconsistency, it is essential that players know that if their vote counts, the other two players are voting in line with their monetary preferences. Simplifying the first stage now means that if a player learns that rule 2 has been implemented, he knows that this has been chosen by another player. Since it is highly implausible that a player with $\theta = 14$ chooses rule 2, the player should thus infer that another player has $\theta = -1$. Nevertheless, if rule 2 is implemented, a player thus still knows that he is only pivotal (in the sense that his vote is chosen to count and changes the outcome) if exactly one other player has $\theta = 14$. Hence if he prefers outcome A in that case, he should vote for A. Regarding the choice in the first stage, the same argument as above holds, that is a player with $\theta = -1$ should choose rule 1 only if he prefers A if exactly one player has $\theta = 14$. Hence a player with $\theta = -1$ choosing rule 1 but voting for B if rule 2 is implemented behaves inconsistently in the sense that his first-stage rule choice suggests a higher degree of altruism than his second-stage vote. The only difference between Treatments 4 and 5 is that in the latter, a player might (implausibly) assume that a player with $\theta = 14$ will choose rule 2 and hence conclude from the implementation of rule 2 that at least one other player has $\theta = 14$. However, this does not change anything about the conclusions from being pivotal. A wrong conditioning on being pivotal might possibly make such a player be more inclined to vote for A if rule 2 is implemented (because he now "knows" that one other player has $\theta = 14$ rather than inferring that from being pivotal). This, however, only strengthens the argument for voting A and thus still a player who chooses rule 1 but votes for B when rule 2 is implemented is behaving inconsistently.

Treatments 4 and 5 were implemented with 45 and 63 subjects, respectively, with random matching into groups of three in sessions with 12 to 24 participants. Otherwise, the same procedures were followed as for the other treatments.

4.2 Results

In Treatment 4, among the instances where inconsistent behavior could be detected, that is where a subject chooses rule 1 but then decides under rule 2, chosen by the computer, the subjects actually behave inconsistently in 25 percent of the cases. In Treatment 5, in 66 instances a player with $\theta = -1$ suggests rule 1 (only about 14 percent of the cases where $\theta = -1$). In 18 of these cases, rule 2 is chosen instead. In only 4 of these cases, i then votes for A. Thus in 14 out of

18 cases where an inconsistency could be observed, this is indeed observed. While these are few instances overall and the rate of inconsistency is smaller in Treatment 4, these results indicate that indeed inconsistencies in the sense of subjects being more concerned about efficiency in the rule choice stage than in the voting stage are a relevant phenomenon.

5 Discussion and concluding remarks

Our paper experimentally addresses a fundamental problem in political economics. Voting rules should adjust to the problem at hand while, frequently, they are fixed in practice. Our results indicate that many voters are willing to efficiently adjust the voting rule even when they are already privately informed about their preferences for the voting outcome. Choosing in a two-step procedure can yield higher aggregate payoffs than simple majority voting. On the one hand, voters with a weak preference can express this by supporting a rule not implementing their selfish preferences and then vote in line with their preference. On the other hand, people with strong preferences can implement these more often, so their preference carries more weight. If strong preferences are substantially stronger than weak ones and voters are sufficiently efficiency concerned, this can yield more efficient outcomes than simple majority voting.

Our results suggest that it may be useful to let voters fix or alter voting thresholds in referenda. However, drawing concrete welfare conclusions from our analysis requires additional research. In our analysis we have compared the aggregate payoffs that arise under different mechanisms, not welfare. This is not a problem if everybody is equally efficiency concerned or selfish, because then whatever maximizes payoffs maximizes welfare. However, if some voters are indeed inequality averse, our criterion would be paternalistic since it only takes into account the payoffs from the project itself.

Another finding of our paper is that when choosing about rules, different types of social preferences may matter than when choosing about outcomes. Thus when predicting how experimental participants choose rules, one should be careful in taking typical experimental results as a guidance of the prevalence of preferences. Efficiency as a behavioral attractor seems to be more important when subjects decide on procedures than if they decide according to a given procedure. This effect could in principle be exploited to increase expected payoffs in democratic decisions or to end political gridlock in reform processes. Adding one or more pre-voting stages might result in more altruistic behavior and ultimately lead to an optimal outcome. However, again, it is difficult to draw conclusions regarding the optimal number of voting stages when preferences are not stable from one stage to another. Drawing institutional conclusions requires a theoretical framework that takes the instability of preferences into account. A potential complication of trying to raise efficiency gains through a multi-stage procedure arise from strategic voting given that the rule choices reveal something about other players' valuations. This in turn can change strategic rule choices. Such effects can both harm or enhance efficiency gains, depending on whether the dominating effect is altruists taking the information they gain into account in their attempts to maximize total welfare or selfish players strategically exploiting the altruists' reactions. In general, choosing about rules (possibly at multiple levels) allows people to condition voting behavior on information at various stages. In applications, people may know more or less about the others' preferences than in our experiment. For example, it is clearly not generally true that people know the underlying distribution of preferences. They may, however, often know crucial aspects of the distribution, such as that preferences against a project are much stronger than those in favor of it. Such information could often be enough to know that asymmetric voting rules are more efficient than majority voting in these cases.

How can we interpret that apparently subjects are more likely to deviate from selfish maximization of monetary payoffs in the rule-choice stage than in the voting stage? A first possible explanation is that at the rule choice stage, the final outcome is still somewhat remote and this may reduce the temptation to act selfishly and hence make subjects choose somewhat more altruistically. A second possibility is that at the rule choice stage, at least some participants believe they should choose what is the right thing to do, and choosing the efficient rule is likely perceived as the right thing to do. In spite of the fact that rules are chosen at an interim stage, choosing a rule may have a flavor of choosing what is best at an ex-ante stage, or put differently, even though rule choice does not happen behind a veil of ignorance, participant may think that it should. In contrast, in the voting stage, voting in line with one's own monetary preferences is likely to be perceived as legitimate. Voting selfishly is quite generally accepted, whereas tilting rules in one's favor is not. Therefore, a simple outcome-oriented preference model is unable to adequately capture behavior across both stages in our two-stage voting game.

Mechanisms where players first decide about the threshold are also related to the idea of "nudge" (Thaler and Sunstein, 2008). A social planner might know that given the distribution of valuations, the optimal voting threshold may deviate from simple majority rule. However, he may lack the political legitimacy to implement this rule. Instead, he may thus leave the decision regarding the threshold to the participants, who will, if they are sufficiently efficiency concerned, choose the optimal threshold. Naturally, this will not work perfectly because participants in this mechanism will not all be sufficiently efficiency concerned. However, it allows for some gains compared to exogenously assigned majority voting. These gains can be increased, if the rule choice optimally exploits the efficiency concerns of those who have any.

We note that the specific two-stage procedure we have implemented is not well suited at capturing efficiency gains because the random-dictator mechanism at the rule-choice stage is very inefficient since it does not aggregate information about preferences and because it allows a randomly chosen person to implement her preferred choice independent of the preferences and behavior of the other participants. This leads in many individual decisions problems to worse outcomes than majority voting. We had chosen this rule to obtain evidence regarding subjects' individually preferred voting rule that is not affected by first-stage strategic considerations. A more subtle way of aggregating preferences for rules makes use of more than just one subject's proposal. In our asymmetric treatment we found that in about 44 percent of the cases subjects suggested a rule that was not completely biased in favor of alternative B. This indicates that an appropriately designed majority rule for the choice of the voting rule may produce better results than our simple random dictator mechanism. Testing and comparing such more complex multi-stage mechanisms is worth additional research.

Overall, we know rather little about how people behave in two-stage voting procedures. Our experiment provides only a few early steps. More research is required to understand whether, when and how one would want to implement such procedures in practice. There are actually many instances in which committees first decide about procedural rules before deciding about the issue itself. Frequently, decision mechanisms are the issue of a debate (though not necessarily of a formal vote) at a stage at which voters' privately known valuations have already realized. One important example is the United Nations General Assembly which occasionally chooses to proceed under consensus only when the choice is particularly controversial.[11][12] A challenging task is to develop and experimentally test more formal real-world decision procedures. One first step could be to experiment with hybrid voting mechanisms in local referenda. Such hybrid mechanisms would ask voters to simultaneously approve both a decision rule and, contingent on the realized rule, a decision outcome.

More generally, other mechanisms such as school choice and matching mechanisms may also benefit from the introduction of a pre-play stage that enables efficiency concerned participants to express their beliefs about the underlying distribution of types as well as their own intensity of preferences and, related to that, their preferred mechanisms.

[11] The Encyclopedia of the Nations notes that "In effect, each member state of the League (of Nations) had the power of the veto, and, except for procedural matters and a few specified topics, a single "nay" killed any resolution. Learning from this mistake, the founders of the UN decided that all its organs and subsidiary bodies should make decisions by some type of majority vote (though, on occasion, committees dealing with a particularly controversial issue have been known to proceed by consensus)." (http://www.nationsencyclopedia.com/United-Nations/Comparison-with-the-League-of-Nations-VOTING.html).

[12] Another example are families that may decide to take a detour for some ice-cream if at least one person votes in favor when it is known that the cost of the detour are small for those who are not interested in ice-cream.

A Appendix

A.1 Modelling efficiency concerns

In this section we study the optimal rule choice of an agent who cares about a weighted sum of all $2n + 1$ players' monetary payoffs. Assume that this player has altruistic preferences captured by a utility function $u_i(\theta) = \gamma_i\theta_i + (1 - \gamma_i)\overline{\theta}_{-i}$, where $\overline{\theta}_{-i} = \frac{1}{2n}\sum_{j\neq i}\theta_j$. We assume that agent i expects that, for any rule, all players (including player i himself) vote symmetrically and strictly selfishly in the second stage.[13] That is, player j votes for A if $\theta_j > 0$, votes for B if $\theta_j < 0$ and flips a fair coin if $\theta_j = 0$. Given that the rule that player i chooses only affects outcomes if that rule is chosen, we can treat the rule choice as if player i could decide the rule for sure.

It is useful to first introduce some additional notation. Denote with p_k^+ the probability that alternative A is chosen if the rule is k and player i votes for A and p_k^- the probability that A is chosen if the rule is k and i votes for B.

Let E_k^+ be the expected average payoff of the $2n$ other players given rule k and player i voting for A and E_k^- the expected average payoff of the other players given rule k and player i voting for B. Let ω_j be the expected average of the valuations of the other players conditional on at least j among these players voting for A. Then $E_k^+ = p_k^+\omega_{k-1}$ (because if at least $k-1$ of the other players vote for A and player i votes for A, then at least k vote for A, so A is chosen) and $E_k^- = p_k^-\omega_k$ (because if at least k of the other players vote for A and player i votes for B, then at least k vote for A, so A is chosen). Moreover $E_k^+ = p_k^+\omega_{k-1} = p_{k-1}^-\omega_{k-1} = E_{k-1}^-$ for all $k \in \{2, .., n\}$

Finally, denote by

$$u_i^+(\theta_i, k) := \gamma p_k^+\theta_i + (1 - \gamma)\, E_k^+$$

the expected utility of player i under rule k conditional on all other players voting selfishly in the above defined sense and player i voting for alternative A. We define $u_i^-(\theta_i, k)$ and $u_i^0(\theta_i, k)$, the expected utility when player i votes for B or flips a fair coin, respectively, in an analogous way.

Note that for all $k = 1, .., n$ the functions $u_i^+(\theta_i, k)$ and $u_i^-(\theta_i, k)$ are linear in the valuation θ_i. Moreover $u_i^+(\theta_i, 1) = \gamma\theta_i$ because if player i votes for A if rule 1 is chosen, A is implemented for sure and hence nothing can be inferred about the valuations of the other players and hence their conditional expected valuation is 0 given the symmetry of the underlying distribution of

[13] The optimality of the assumed second stage voting behavior will be discussed below.

valuations, $u_i^-(\theta_i, 2n+1) = 0$ (because in that case, B is chosen for sure since unanimity would be required for A but is prevented by i), $u_i^+(0,k) = (1-\gamma)\,E_k^+ \geq 0$ because for rule 1, nothing is learned about the other players' valuations from the implementation of A, whereas for the other rules it can be inferred that at least one of the other players has a positive valuation, and hence in that case the inequality is actually strict, and $u_i^-(0,k) = (1-\gamma)\,E_k^- > 0$ because for any rule it can be inferred that at least one of the other players has a positive valuation.

It is a direct consequence of Proposition 2 in Schmitz and Tröger (2012) that the simple majority rule is optimal for agent i when his valuation is zero. This, in combination with the strict monotonicity of the slopes of $u_i^+(\theta_i,k)$ and $u_i^-(\theta_i,k)$ (p_k^+ and p_k^-) in k implies that rules $k > n+1$ ($k < n+1$) are suboptimal for agents with strictly positive (negative) valuations. Therefore, for agents with strictly positive (negative) valuations, it remains to compare all rules k with $k \leq n+1$ ($k \geq n+1$). For $k \geq n+1$, $u_i^-(0,k)$ is strictly monotonously decreasing in k and for $k \leq n+1$, $u_i^+(0,k)$ is strictly monotonously increasing in k.[14]

This and the strict monotonicity of the slopes of $u_i^+(\theta_i,k)$ and $u_i^-(\theta_i,k)$ in k implies that there are unique valuations $\theta_{kk'} < 0$ at which players are indifferent between any two rules $k \neq k'$ with $k, k' \geq n+1$. For the same reason, there are unique valuations $\theta_{kk'} > 0$ at which players are indifferent between any two rules $k \neq k'$ with $k, k' \leq n+1$. The following symmetry property is useful to prove our main result.

Lemma 1: For all $n \geq 1$, for all $j \in \{-n, ..., n\}$ and for all $\theta_i > 0$:

$$u_i^+(\theta_i, n+1+j) - \frac{\gamma}{2}\theta_i = u_i^-(-\theta_i, n+1-j) - \frac{\gamma}{2}(-\theta_i)\,.$$

Proof

$$u_i^+(\theta_i, n+1+j) = p_{n+1+j}^+\,(\gamma\theta_i + (1-\gamma)\omega_{n+j})$$

$$u_i^-(-\theta_i, n+1-j) = p_{n+1-j}^-\,(-\gamma\theta_i + (1-\gamma)\omega_{n+1-j})$$

[14] When $k \leq n+1$, E_k^+ is increasing in k, because then cases where fewer than half of the other players have a positive valuation are dropped in the calculation of E_k^+. Similarly, when $k \geq n+1, E_k^-$ is decreasing in k, because then cases where more than half of the other players have a positive valuation are dropped in the calculation of E_k^-.

Hence the claim is that

$$p^{+}_{n+1+j}\left(\gamma\theta_i+(1-\gamma)\omega_{n+j}\right)=p^{-}_{n+1-j}\left(-\gamma\theta_i+(1-\gamma)\omega_{n+1-j}\right)+\gamma\theta_i.$$

We know that $p^{+}_{n+1+j}=1-p^{-}_{n+1-j}$ (if the rule is $n+1+j$ and i votes for A, then A is implemented if at least $n+j$ others vote for A and thus if they have a positive valuation, while if the rule is $n+1-j$ and i votes for B, then B is implemented if at least $2n-(n+1-j)+1=n+j$ others vote for B and thus if they have a negative valuation, and at least $n+j$ of the others having a positive valuation is equally likely as at least $n+j$ of them having a negative valuation). Hence the previous equality follows from

$$p^{+}_{n+1+j}\omega_{n+j}=p^{-}_{n+1-j}\omega_{n+1-j}.$$

Call $p_{m,2n}$ the probability that m out of $2n$ agents vote yes. We have, with $\bar{\theta}^{+}$ denoting the expected valuation conditional on voting for A:

$$\begin{aligned}p^{+}_{n+1+j}\omega_{n+j} &= \sum_{m=n+j}^{2n} p_{m,2n}\left(m-(2n-m)\right)\bar{\theta}^{+}\\ &= 2\bar{\theta}^{+}\sum_{m=n+j}^{2n} p_{m,2n}\left(m-n\right),\end{aligned}$$

and

$$\begin{aligned}p^{-}_{n+1-j}\omega_{n+1-j} &= \sum_{m=n+1-j}^{2n} p_{m,2n}\left(m-(2n-m)\right)\bar{\theta}^{+}\\ &= 2\bar{\theta}^{+}\sum_{m=n+1-j}^{2n} p_{m,2n}\left(m-n\right).\end{aligned}$$

Therefore, our claim can be rewritten as:

$$
\begin{aligned}
2\bar{\theta}^{+} \sum_{m=n+j}^{2n} p_{m,2n}(m-n) &= 2\bar{\theta}^{+} \sum_{m=n+1-j}^{2n} p_{m,2n}(m-n) \\
\sum_{m=n+j}^{2n} p_{m,2n}(m-n) &= \sum_{m=n+1-j}^{2n} p_{m,2n}(m-n) \\
\sum_{m=n+1-j}^{n+j-1} p_{m,2n}(m-n) &= 0 \\
\sum_{m=n+1-j}^{n-1} p_{m,2n}(m-n) + \sum_{m=n+1}^{n+j-1} p_{m,2n}(m-n) &= 0 \\
\sum_{m=n+1}^{n+j-1} p_{m,2n}(m-n) &= \sum_{m=n+1-j}^{n-1} p_{m,2n}(n-m) .
\end{aligned}
$$

The last inequality holds because $p_{k,2n} = p_{2n-k,2n}$.

Q.E.D.

Based on these preliminaries, we can prove the main result for the case of five players, i.e. $n = 2$.

Proposition 1: For $n = 2$, there exist $\underline{\theta}_i < \overline{\theta}_i < 0$ such that player i's preferred rule is rule 5 for $\theta_i < \underline{\theta}_i$, rule 4 for $\underline{\theta}_i < \theta_i < \overline{\theta}_i$ rule 3 (majority rule) for $\overline{\theta}_i < \theta_i < -\overline{\theta}_i$, rule 2 for $-\overline{\theta}_i < \theta_i < -\underline{\theta}_i$ and rule 1 for $\theta_i > -\underline{\theta}_i$.

Proof Consider first $\theta_i < 0$. If player i chooses rule 5, then her utility is 0 for sure (because she will vote for B in the second stage). If she chooses rule 3 (majority rule) and her choice is implemented, then her expected utility is $p_3^-(\gamma\theta_i + (1-\gamma)\omega_3)$. So she is indifferent between rule 5, which implements B for sure and the majority rule, if $p_3^-(\gamma\theta_i + (1-\gamma)\omega_3) = 0$ or $\theta_i = -\frac{(1-\gamma)}{\gamma}\omega_3$.

Now consider the intermediate rule 4. Player i's expected utility for $\theta_i = -\frac{(1-\gamma)}{\gamma}\omega_3$ and rule 4 is

$$p_4^-(\gamma\theta_i + (1-\gamma)\omega_4) = p_4^-(-(1-\gamma)\omega_3 + (1-\gamma)\omega_4) > 0,$$

because $\omega_4 > \omega_3$, since the expected average payoff ω_k increases strictly in k. Thus at the point where i is indifferent between rule 5 and the majority rule, i strictly prefers the intermediate rule.

The first part of the proposition follows from this and from (i) the linearity of all conditional utility functions (ii) the strict monotonicity of their slopes in k and (iii) the fact that rule 3 strictly maximizes welfare conditional on $\theta_i = 0$. The second part of the proposition follows from the symmetry property shown in Lemma 1, since, for all $\theta_i > 0$ and $j \in \{0, 2\}$,

$$\begin{aligned}
u_i^+(\theta_i, 2) &= u_i^+(\theta_i, 3 - j) \Rightarrow \\
u_i^+(\theta_i, 2) - \frac{\gamma}{2}\theta_i &= u_i^+(\theta_i, 3 - j) - \frac{\gamma}{2}\theta_i \Rightarrow \\
u_i^-(-\theta_i, 4) - \frac{\gamma}{2}(-\theta_i) &= u_i^-(-\theta_i, 3 + j) - \frac{\gamma}{2}(-\theta_i) \Rightarrow \\
u_i^-(-\theta_i, 4) &= u_i^-(-\theta_i, 3 + j).
\end{aligned}$$

Hence, indifference between rules 2 and 3 (2 and 1) for players with some type $\theta_i > 0$ implies indifference of a player with type $-\theta_i$ between rules 4 and 3 (4 and 5). Q.E.D.

Hence, assuming selfish voting behavior, the optimal rule choice of efficiency concerned subjects is monotonous and symmetric and, when the type space is large enough, it covers the entire set of available rules.

Moreover, one can easily show that, when player i has no additional knowledge about other players' types, one can rule out that a player prefers to vote non-selfishly under any voting rule that he has proposed himself. This follows because $u_i^+(\theta_i, k) = u_i^-(\theta_i, k - 1)$ for all $k = 2, .., n$. Hence, if $\theta_i > 0$ and $u_i^+(\theta_i, k) \geq u_i^+(\theta_i, k')$ for all $k' = 1, .., 5$ then $u_i^+(\theta_i, k) \geq u_i^-(\theta_i, k' - 1)$ for all $k' = 2, .., 5$. Furthermore, choosing rule 5 and voting for B is worse than choosing rule k and voting for A., i.e., if $\theta_i > 0$, $u_i^+(\theta_i, k) \geq u_i^-(\theta_i, 5) = 0$. The same type of argument can be made to rule out that players with negative valuations are better off by voting insincerely.

Optimal voting behavior may instead not be selfish when another player's rule proposal has been successful. In such cases, player i can infer something about other players' types from being pivotal in the voting stage. Moreover, player i may learn something about other player's expected valuations from the rule chosen in stage 1. This is why non-selfish voting may be optimal for small positive or negative values of i's willingness to pay. Similarly, if an efficiency concerned player expects other players to vote non-selfishly if certain rules are chosen, the rule choice may deviate from the one derived in Proposition 1.

References

[1] Barbera, S. and M.O. Jackson (2004) "Choosing How to Choose: Self-Stable Majority Rules and Constitutions." *Quarterly Journal of Economics* 119, 1011-1048.

[2] Blanco, M., D. Engelmann, and H.-T. Normann (2011) "A Within-Subject Analysis of Other-Regarding Preferences." *Games and Economics Behavior* 72, 321-338.

[3] Cramton, P., R. Gibbons, and P. Klemperer (1987) "Dissolving a Partnership Efficiently." *Econometrica* 55, 615-32.

[4] Engelmann, D. and M. Strobel (2000). "The False Consensus Effect Disappears if Representative Information and Monetary Incentives are Given" *Experimental Economics* 3, 241-260.

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

[6] Fehr, E. and K. M. Schmidt (1999) "A Theory Of Fairness, Competition, And Cooperation." *Quarterly Journal of Economics* 114, 817-868.

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

[8] Güth, W. and M.F. Hellwig (1986) "The Private Supply of a Public Good." *Journal of Economics* 5, 121-159.

[9] Grüner, H.P. and Y. Koriyama (2012) "Public Goods, Participation Constraints, and Democracy: A Possibility Theorem." *Games and Economic Behavior* 75, 152-167.

[10] Myerson, R. and M. Satterthwaite (1983) "Efficient Mechanisms for Bilateral Trading." *Journal of Economic Theory* 29, 265-81.

[11] Ross, L., D. Greene, and P. House, (1977) "The 'False Consensus Effect': An Egocentric Bias in Social Perception and Attribution Processes." *Journal of Experimental Social Psychology* 13, 279-301.

[12] Segal, I. and M. Whinston (2011) "A Simple Status Quo that Ensures Participation (with Application to Efficient Bargaining)." *Theoretical Economics* 6, 109-125 .

[13] Schmitz, P. W. (2002) "Simple Contracts, Renegotiation under Asymmetric Information, and the Hold-up Problem." *European Economic Review* 46, 169-188.

[14] Schmitz, P.W. and T. Tröger (2012) "The (Sub-)Optimality of the Majority Rule." *Games and Economic Behavior* 74, 651-665.

[15] Thaler. R.H. and C.R. Sunstein (2008) "Nudge: Improving Decisions about Health, Wealth, and Happiness." Yale University Press.