Discussion Paper Series – CRC TR 224

Discussion Paper No. 002
Project B 01

Linear Voting Rules

Hans Peter Grüner [1]
Thomas Tröger [2]

January 2018

[1] Department of Economics, University of Mannheim; gruener@uni-mannheim.de
[2] Department of Economics, University of Mannheim, troeger@uni-mannheim.de

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Linear voting rules *

Hans Peter Grüner and Thomas Tröger

January 10, 2018

Abstract

How should a society choose between two social alternatives if participation in the decision process is voluntary and costly and monetary transfers are not feasible? Considering symmetric voters with private valuations, we show that it is utilitarian-optimal to use a linear voting rule: votes get alternative-dependent weights, and a default obtains if the weighted sum of votes stays below some threshold. Standard quorum rules are not optimal. We develop a perturbation method to characterize equilibria in the case of small participation costs and show that leaving participation voluntary increases welfare for linear rules that are optimal under compulsory participation.

1 Introduction

Participating in collective decision procedures is typically individually costly. Thus, if participation is voluntary then those for whom too little is on stake will abstain. This is a non-trivial issue in modern societies in which a vast range of decisions in public agencies, boards of companies, committees, parliaments, congressional and party caucuses and private and professional associations are reached democratically. From a welfare point of view, there is a trade-off between aggregating preferences while at the same time saving on participation costs. This raises important practical questions. Which decision rules should institutions use for different issues to maximize the welfare of their members? Should participation be voluntary or compulsory? Is it always desirable to lower participation costs when this

*We thank Felix Bierbrauer, Gorkem Celik, Luc Laeven, Stephan Lauermann, Andreas Kleiner, Benny Moldovanu, Igor Muraviev, Timofiy Mylovanov, Anne-Kathrin Roesler and anonymous referees for very helpful comments. This paper supersedes and corrects CEPR Discussion Paper No. 11127, "Optimal voting mechanisms with costly participation and abstention", February 2016.
Grüner, Tröger: University of Mannheim, Department of Economics, L7, 3-5 68131 Mannheim, Germany, gruener@uni-mannheim.de, troeger@uni-mannheim.de.
We gratefully acknowledge financial support from the German Science Foundation (DFG) through SFB 884 and SFB-TR 224.

is feasible? This paper delivers new insights on these questions using mechanism design theory.

We derive our results from a Bayesian model in which each individual is privately informed about her preferences over two social alternatives and about her preference intensity relative to the cost of participating in the decision. Such a private-values setup is well suited to capture many practical applications of voting which are dominated by material conflicts among the individuals. Assuming private values allows us to study the probability of being pivotal as the crucial determinant of an individual's participation decision. In a setting with common values, the event of being pivotal would in itself implies information that is relevant to an individual's participation decision, leading to well-known complications.[1]

We build a model that is symmetric across the individuals that constitute the electorate. Symmetry across the social alternatives, S(tatus quo) and R(eform), is not required. We allow for arbitrary anonymous mechanisms with voluntary participation, simultaneous moves, and no monetary transfers.[2] A mechanism together with a symmetric Bayes-Nash equilibrium is optimal if it maximizes the agents' ex-ante expected utility. Without loss of generality, we restrict attention to mechanisms that are voting rules, that is, each individual chooses among at most three actions that can be interpreted as voting for R, voting for S, and abstention.

We have two main contributions. First, we show that any optimal voting rule with voluntary participation is (up to outcome-equivalence) linear. A linear rule is a generalization of a majority rule. Under a linear rule, one alternative can be designated as a default such that, concerning the non-default alternative, only the votes that are cast *in addition* to some required minimum number are weighed against the votes in favor of the default (cf. Figure 1). An extreme type of linear rules are one-sided rules in which votes for the default alternative are not counted at all. We show that both qualified majority rules (i.e., majority thresholds different from 50% and no default) and all one-sided linear rules can be optimal for any population size. It can also happen that the optimal rule belongs to neither of these classes. We provide an example in a setting with 15 individuals in which the following rule is the unique optimum: implement reform if and only if there are at least two more R-votes than S-votes.

[1]Building on Condorcet's jury setting, there is an extensive literature on voting rules in settings with common values. Here, voters may abstain for informational reasons even when there is no participation cost (Feddersen and Pesendorfer, 1996). A major question addressed in this literature is whether in equilibrium the voters' private information is successfully aggregated (Feddersen and Pesendorfer, 1997).

[2]The analysis of compulsory participation is equivalent to the analysis of voting without a participation cost, which is well known for our setting; cf. Barbera and Jackson (2006), Schmitz and Tröger (2012).

Laruelle and Valenciano (2011) cover linear rules in their taxonomy of "weighted anonymous quaternary voting rules". They provide an extensive list of examples of one-sided linear rules and qualified majority rules used in political institutions, showing that a range of different majority requirements is applied. Thus, certain linear rules are commonly used. But since a linear rule that is neither a qualified majority rule nor a one-sided rule can be optimal as well, our results can also be used to provide novel recommendations for the practical design of voting rules.

It is important to distinguish the class of linear rules from the standard quorum rules that require a participation or approval quorum to upset the default alternative.[3] Overcoming the default requires both a minimum number of votes (in total or in favor of the non-default) and a qualified majority among *all* cast votes (cf. Figure 2). The minimum required number of votes can be seen as "a simple way of protecting the status quo" (Maniquet and Morelli, 2015). A similar protection is achieved in a linear rule with a default alternative. The difference is that the standard quorum rules have a kink in the cutoff line that separates the outcomes S and R, which makes these rules non-linear and thus, in our model, sub-optimal (unless complete abstention of one side of the electorate happens to be optimal).

An important special case of the costly-voting model is the neutral preference setting (cf. Börgers, 2004). Here, both alternatives are equally likely to be preferred with any given intensity. A rule is called neutral if it treats both alternatives identically. We show that there exist neutral settings in which no neutral rule is optimal (instead a one-sided rule is optimal). Thus, the ability to implement a non-neutral rule can be crucial.

The implementation of a non-neutral rule relies on the implicit assumption that the two social alternatives can be unambiguously mapped into the model variables R and S. This assumption fails in situations in which the available alternatives are fundamentally symmetric, like two non-incumbent candidates running for an office. In such situations, only rules that are neutral can be implemented. Specializing our analysis to neutral rules in a neutral preference setting, the standard voluntary majority rule is always optimal. This may be seen as a justification for the widespread use of the standard majority rule. At the same time, the result emphasizes the importance of the class of linear rules because the standard majority rule is the unique neutral linear rule.

A long-standing issue in voting theory is whether there are cases in which institutions should enforce participation in a vote. Our second main contribution

[3] From the outset of democracy, collective decisions were made both with and without quorum requirements. In the Athenian Assembly, most decisions were made by majority without a quorum requirement (i.e., a linear rule), but some decisions required a quorum of 6000 citizens and were made by secret ballots (cf. Blackwell, 2003). For quorum rules used in practice today, see Corte-Real and Pereira (2004) and Herrera and Matozzi (2010).

is a comparison of voluntary and compulsory participation in linear voting rules. We show that an institution designer with very limited information about the environment always prefers voluntary participation over compulsory participation. For technical reasons (see below), we restrict attention to settings with a small participation cost that can be obtained as perturbations of the setting without a participation cost. Given our model, the only information the designer needs about individuals' preferences in order to implement an optimal rule with compulsory participation is a single number: the ratio between the welfare weights put on the two alternatives (cf. Rae, 1969, Barbera and Jackson, 2006). This allows her to set the optimal qualified majority threshold. We show that leaving participation voluntary will increase equilibrium welfare in every deterministic linear rule with the optimal majority threshold. Thus, making participation voluntary is socially beneficial whether or not the design of the rule takes abstention into consideration optimally (note that the same conclusion is trivially true if the participation cost is sufficiently large).

On a technical level, our paper offers two contributions. First, we exploit a mismatch of dimensionalities. The symmetry across agents means that equilibria are mere two-dimensional objects, given by the expected participation rates of the agents preferring R and S, respectively. But the space of voting rules is of much higher dimensionality, given by the set of functions that map election results into probabilities of implementing R. The difference between the dimensionalities implies that any given equilibrium will prevail in a large set of voting rules. Thus, maximizing welfare across such a set of rules carries a long way. This maximization is easily tractable because both the welfare and the equilibrium conditions are linear functions of the probabilities that define the voting rules. Technically, this is the fundamental tool in our proof that optimal rules are linear.

Our second technical contribution is a perturbation method for establishing existence and analytical properties of equilibria in voting rules with voluntary participation if the participation cost is small. The equilibrium conditions establish a smooth relationship between three numbers: the (commonly known) participation cost and the participation rates of the R- and S-agents. An obvious idea is to start with full participation at 0 participation cost and use the implicit function theorem to establish an equilibrium with almost-full participation if the participation cost is small. For deterministic linear rules, however, this approach fails because either the R-participation rate or the S-participation rate changes at an infinite rate as the participation cost tends to 0. To overcome this problem, we switch the roles of dependent and independent variables. We show that for any S-participation rate, there exists an R-participation rate and a participation cost such that the equilibrium conditions hold. Finally, we invert functions back so that the equilibrium is described as a function of the participation cost.

Further results and relation to the literature

Ledyard (1984) and Palfrey and Rosenthal (1985) introduce the costly-voting setting with private values[4] on which we build our model.[5] They analyze the standard voluntary majority rule, mainly considering a setting in which the numbers of R-agents and S-agents are common knowledge. We can show that *in that setting optimal rules are again linear*, underlining the robustness of our main result (cf. Appendix B).

Concerning welfare comparisons of voting rules, a seminal contribution is Börgers (2004) who compares voluntary and compulsory participation in the standard majority rule.[6] He shows that voluntary participation yields a higher welfare if the preference setting is neutral across alternatives. Taken together with our result that the voluntary majority rule is optimal among all voluntary neutral rules, one obtains a fundamental conclusion: *the standard voluntary majority rule is optimal among all neutral rules, whether one assumes compulsory participation or voluntary participation*.

Krasa and Polborn (2009) point out that Börgers' result depends on the neutral preference setting. If ex-ante one alternative is sufficiently more likely to be preferred than the other, then participation will be inefficiently low in the standard voluntary majority rule so that a voting subsidy will increase welfare. The same conclusion applies if types are sufficiently strongly correlated, implying that enhancing the voluntary majority rule with a preelection opinion poll reduces welfare (Goeree and Grosser, 2007).

Further results on welfare have been obtained in the large-population limit. Campbell (1999) shows that, in general, a large majority rule with voluntary participation will be dominated by the most strongly affected voters, precluding efficiency. Krishna and Morgan (2015) provide a contrasting result by assuming heterogeneous participation costs with a positive density at 0: the standard majority rule is utilitarian optimal in the limit if the distribution of participation costs is identical on the two sides of the electorate.[7] This preference specification is a special case of our's.

An institution designer may be interested in promoting turnout in order to lend

[4]Palfrey and Rosenthal (1983) analyze costly voting with complete information. See Nöldeke and Pena (2016) for equilibrium existence and further comparative statics results in this setting.

[5]For costly participation in auctions, see Celik and Yilankaya (2009) and Cai et al. (2016).

[6]Further welfare comparisons can be found in the working paper Börgers (2000); e.g., an example in which a one-sided linear rule yields a higher welfare than a standard majority rule.

[7]This result is robust to various specifications of uncertainty of the population size, including Poisson uncertainty as introduced by Myerson (1998).

“legitimacy” to a collective decision (Qvortrup, 2005). The earlier literature has investigated turnout under specific voting rules. Palfrey and Rosenthal (1985) consider the standard majority rule and provide comparative statics of turnout with respect to participation cost and population size. Using numerical methods, Aguiar-Conraria and Magalhães (2010) give examples which show that augmenting a majority rule by a quorum requirement has an ambiguous effect on aggregate turnout. These examples also suggest that the problem of strategic abstention of the status-quo-supporters in quorum rules is most severe with a participation quorum. Using a large-population Poisson model, Maniquet and Morelli (2015) provide a formal result on this: due to strategic abstention, a planner prefers an approval quorum rule over a participation quorum rule even if with sincere voting she would be better off under the latter rule. Herrera et al. (2014), who's main interest is an experiment, compare turnout across the majority rule and the “proportional power sharing rule”.

An important theme in the literature on turnout is the “underdog compensation effect” in situations in which one alternative is supported by an expected minority. The effect refers to turnout among the underdog supporters being higher than turnout among the supporters of the opposite alternative, in a neutral voting rule. Ledyard (1984) demonstrates the underdog effect for the standard majority rule under the assumption that the conditional distributions of the R- and S-agents' preference intensities are equal. Making an equivalent assumption, Herrera et al. (2014) extend the underdog effect to proportional power-sharing rules, and Kartal (2015) to all “regular voting rules”.[8] Taylor and Yildirim (2010) show that the underdog effect is reversed if (i) each agent is equally likely to prefer S and R and (ii) the S-agents' conditional distribution of preference intensities stochastically dominates the R-agents' conditional distribution. Myatt (2015) shows that substantial turnout can occur in a large population if valuations are correlated, and under a weak condition the minority's preferred candidate can win the election.

Our linearity result can be extended to incorporate a concern for turnout: *maximizing a weighted average of R-agents' turnout, S-agents' turnout, and expected utility, again leads to a linear voting rule*.

There is a growing literature on the design of utilitarian-optimal voting rules when there is no participation cost (or, equivalently, when voting is compulsory).[9]

[8] Weak majority rules (cf. Schmitz and Tröger, 2012) are examples of non-regular neutral voting rules. To see that the underdog effect can fail without regularity, consider the weak majority rule that implements each alternative with probability $1/2$ unless both agents participate and vote unanimously for one alternative. If the participation cost is small then there exists an equilibrium with a positive participation rate among S-agents and zero participation of R-agents.

[9] A setting with two alternatives can be interpreted as a binary-public-good problem. Ledyard and Palfrey (2002) show that the welfare from any interim-efficient allocation of the public good

Rae (1969) considers the class of qualified majority rules. Barbera and Jackson (2006) consider the optimal two-stage aggregation of preferences and apply this to characterize the optimal voting rules at the level of the European Union. Schmitz and Tröger (2012) emphasize the potential importance of weak-majority rules when valuations are correlated across agents. Azrieli and Kim (2014) characterize the interim-efficient rules again with stochastically independent valuations. Drexl and Kleiner (forthcoming) show that a qualified majority rule is utilitarian optimal among anonymous deterministic dominant-strategy rules even when transfers are feasible, provided that any budget imbalance is subtracted from the welfare. Gershkov et al. (2017) characterize optimal voting rules in dominant strategies in a model with more than two social alternatives.

Concerning the design of optimal voting rules with costly participation, a dominant-strategy requirement is rather restrictive—not even majority rules have a symmetric dominant-strategy equilibrium. Relying on implementation in Bayes-Nash equilibrium, Kartal (2015) shows that in a preference setting that is neutral across alternatives, the standard majority rule is optimal among regular voting rules. In a setup with commonly known voter preferences and private costs, Faravelli and Sanchez-Pages (2014) compare all convex combinations of simple majority rule and proportional rules, finding that welfare increases in the weight of simple majority rule in neutral environments. Bognar et al. (2015) allow for asymmetric participation costs, but assume away any uncertainty about preference intensities. They show that, if both alternatives are equally attractive ex-ante, then the first best allocation can be implemented in a sequential variant of a one-sided voting rule in which the agents are invited to participate conditionally on the history of votes.

The rest of the paper organized as follows. After introducing the model in Section 2, we present our linearity results in Section 3. Section 4 identifies settings in which particular linear rules are optimal. Section 5 compares voluntary and compulsory participation and uses our techniques to draw further conclusions. Following the conclusion (Section 6), Appendix A contains the central proofs, and Appendix B summarizes the remaining proofs.

2 Model

Consider $n \geq 2$ individuals who have to implement one of two possible social alternatives, denoted S ("status quo") and R ("reform"). The decision about R

(with monetary transfers) can be achieved with a qualified majority rule in the large-population limit. Bierbrauer and Hellwig (2016) show that if one requires a robust version of coalition-proofness, then all mechanisms other than qualified majority rules become infeasible.

versus S is made via a mechanism that determines a social alternative depending on the participating individuals' actions. Each individual may abstain from the mechanism. We build a model that is symmetric across individuals.

Preferences

Each agent i cares about four outcomes R, iR, S, and iS, where R means that alternative R is implemented and agent i abstains from participating in the mechanism, iR means that R is implemented and agent i participates in the mechanism, and similarly for S and iS. Von-Neumann-Morgenstern preferences over outcome lotteries are represented in terms of the agent's *type* or *valuation* $v_i \in \mathbb{R}$ that is private information of the agent, and a *participation cost* $c > 0$ that is common to all agents. The agent's Bernoulli utilities are v_i for outcome R, $v_i - c$ for outcome iR, 0 for outcome S, and $-c$ for outcome iS. Assuming a common participation cost rather than private participation costs entails no loss of generality. Making c flexible is best for presenting our perturbation results. The absolute value $|v_i|$ can be viewed as the "intensity" of type v_i's preference.

Agents have i.i.d. beliefs about others' types; the c.d.f. is denoted F. We assume that F has no atoms and that there can be agents on both sides of the electorate,

$$F^R = 1 - F(0) > 0 \text{ and } F^S = F(0) > 0.$$

If $F(-v) = 1 - F(v)$ for all v, then we say the environment is *neutral*.

Voting mechanisms

A *(voluntary) mechanism* is a mapping $\Phi : \mathcal{A}^n \to [0, 1]$, where all individuals i simultaneously select actions $a_i \in \mathcal{A}$, where $\mathcal{A}$ includes a particular action A ("abstain"), and social alternative R is implemented with probability $\Phi(a_1, \ldots, a_n)$.[10] From individual i's point of view, the resulting lottery outcome is

$$\begin{pmatrix} R & S \\ \Phi(a_1, \ldots, a_n) & 1 - \Phi(a_1, \ldots, a_n) \end{pmatrix} \text{ if } a_i = A,$$

and

$$\begin{pmatrix} iR & iS \\ \Phi(a_1, \ldots, a_n) & 1 - \Phi(a_1, \ldots, a_n) \end{pmatrix} \text{ if } a_i \neq A.$$

[10] Sequential procedures cannot be studied in this framework. Transforming a sequential procedure into its normal form would mean that, from an individual's interim point of view, her participation can be uncertain and can depend on others' actions. Cf. Bognar et al. (2015).

We restrict attention to anonymous mechanisms, that is, mechanisms that treat all agents the same:[11] for all action profiles $(a_1, \dots, a_n)$ and all permutations ξ of $\{1, \dots, n\}$,

$$\Phi(a_{\xi(1)}, \dots, a_{\xi(n)}) \quad = \quad \Phi(a_1, \dots, a_n).$$

Each agent i employs a strategy that specifies an action in $\mathcal{A}$ for each type v_i. If all others employ the strategy σ and individual i of type v_i takes action a_i, then her (interim) expected utility is

$$v_i \rho^{\Phi,\sigma}(a_i) - c\mathbf{1}_{a_i \neq A},$$

where

$$\rho^{\Phi,\sigma}(a_i) = \int \Phi(a_i, (\sigma(v_j))_{j \neq i}) \prod_{j \neq i} \mathrm{d}F(v_j)$$

denotes the probability that the social alternative R is implemented, from the point of view of agent i.

To get a fully symmetric model, we assume that agents' behavior does not depend on their labels when playing the game, that is, we focus on *symmetric* (Bayesian) equilibria: all agents employ the same strategy σ, where

$$\sigma(v) \in \arg\max_{a \in \mathcal{A}} v\rho^{\Phi,\sigma}(a) - c\mathbf{1}_{a \neq A} \quad \text{for all } v \in \operatorname{supp}(F).$$

Given a mechanism-equilibrium pair $m = (\Phi, \sigma)$, let $U^m(v)$ denote the interim-expected utility of any type v.

Mechanism-equilibrium pairs (Φ, σ) and (Φ', σ') are called equivalent if they induce the same interim-expected utilities for all types.

Lemma 1 below shows that we can restrict attention to mechanisms with up to three actions, including A.[12] This holds because any participating R-agent chooses an action that maximizes the probability of the social alternative R, and any participating S-agent chooses an action that minimizes the probability of the social alternative R.[13]

[11] Committee rules in which only some players are allowed to participate, and anonymity is required among the participants (cf. Börgers, 2000) can be analyzed with similar methods.

[12] Thus, our analysis will be an instance of mechanism design with finite (specifically, three-elementary) action spaces. Another such exercise, in the different context of auctions and no participation cost, is Kos (2012).

[13] Lemma 1, while extending to asymmetric settings, is specific to the case of two social alternatives. Here, conditional on participation only two different preference relations exist—the agent prefers iR or iS. With a third social alternative T, post-participation preferences would concern three alternatives, iR, iS, and iT, yielding a continuum of post-participation preference relations and accordingly an *infinite* action set.

Lemma 1. *For any mechanism-equilibrium pair* (Φ, σ)*, there exists an equivalent mechanism-equilibrium pair* (Φ', σ') *such that* Φ' *allows at most three actions for each voter.*

Proof. An agent of type $v > 0$ will abstain or take an action $a \in \mathcal{A} \setminus \{A\}$ that maximizes $\rho^{\Phi,\sigma}(a)$; if instead $v < 0$, then she will abstain or take an action $a \in \mathcal{A} \setminus \{A\}$ that minimizes $\rho^{\Phi,\sigma}(a)$.

Consider (according to σ and F) the distribution $d_>$ over actions a across all types $v > 0$ that do not abstain; consider the distribution $d_<$ over actions a across all types $v < 0$ that do not abstain. We obtain a new, interim payoff equivalent, equilibrium in Φ by assuming that all types $v > 0$ that do not abstain randomize their action according to the distribution $d_>$, and all types $v < 0$ that do not abstain randomize their action according to the distribution $d_<$. Now define Φ' by restricting the set of actions to $d_>$, $d_<$, and A. This completes the proof.

By Lemma 1, it is sufficient to consider *voting* mechanisms (or *rules*) in which each agent chooses among three actions, denoted A, S, and R (for convenience we use the same notation S and R as for social alternatives). Any such mechanism can be described as a function $M : \{(r, s) | r \geq 0, s \geq 0, r + s \leq n\} \to [0, 1]$, where $M(r, s)$ denotes the probability that R is implemented if r agents play R and s agents play S. For convenience we will use the notation $M_{rs} = M(r, s)$ when the arguments r and s are simple enough expressions. Any outcome (r, s) is called a *tally*. A voting rule M is *deterministic* if $M(r, s) \in \{0, 1\}$ for all tallies (r, s).

A special class of voting rules are the *R-one-sided rules* that are defined by the property that $M_{r,s} = M_{r,0}$ for all (r, s), and the *S-one-sided rules* that are defined by the property that $M_{r,s} = M_{0,s}$ for all (r, s). In any equilibrium of a one-sided rule, at most one side of the electorate—either the S-agents or the R-agents—participate with positive probability. A voting rule that is not one-sided is called *two-sided*. A *constant* mechanism M has $M_{r,s} = M_{0,0}$ for all (r, s); in equilibrium, nobody participates.

A mechanism M is called *linear* with default bias ξ, R-weight ξ^R, and S-weight ξ^S if

$$\xi^R \geq 0, \ \ \xi^S \geq 0, \text{ and } (\xi^S > 0 \text{ or } \xi^R > 0 \text{ or } \xi \neq 0), \tag{1}$$

and, for all s and r,

$$M_{rs} \quad = \quad \begin{cases} 1 & \text{if } r\xi^R - s\xi^S - n\xi > 0, \\ 0 & \text{if } r\xi^R - s\xi^S - n\xi < 0. \end{cases} \tag{2}$$

(Cf. Figure 1.) Observe that linearity entails no condition on $M_{r,s}$ along the "cutoff line" where $r\xi^R - s\xi^S - n\xi = 0$. We say that (ξ, ξ^R, ξ^S) are *parameters* for M.

The class of linear rules is — up to the indeterminacy along the "cutoff line" where $r\xi^R - s\xi^S - n\xi = 0$ — a two-dimensional class of rules because only the relative size of the three parameters ξ, ξ^R, and ξ^S, is relevant.

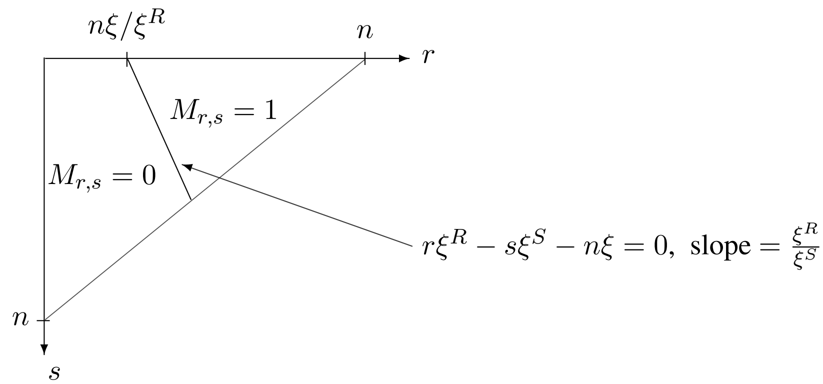

Figure 1: A voting rule maps every vote tally (r, s) into the probability of implementing alternative R. For each linear rule M, there is a cutoff line described by the equation $r\xi^R - s\xi^S - n\xi = 0$. Above the line, alternative R is chosen, below the line S is chosen. On the cutoff line there is no restriction.

While a linear rule can weigh votes for one alternative stronger than the votes for the other alternative (by having $\xi^R \neq \xi^S$), it can at the same time take either alternative as the "default" (by having $\xi \neq 0$). The linear rules with default bias $\xi = 0$ are the "qualified majority rules". An example of a qualified majority rule is the standard voluntary majority rule, which is defined by the property $M_{r,s} = 1$ if $r > s$, $M_{r,s} = 1/2$ if $r = s$, $M_{r,s} = 0$ otherwise.

Any R-one-sided linear rule can be represented with an S-weight $\xi^S = 0$; analogously for S-one-sided linear rules. Note that there exist linear rules with horizontal or vertical cutoff-lines that are not one-sided.

A quorum rule (see, e.g., Aguiar-Conraria and Magalhaes, 2010) reacts to a qualified majority among *all* cast votes, once a minimum number of votes for reform ("approval quorum") or minimum total number of votes ("participation quorum") is received; a default obtains if the quorum is not reached (cf. Figure 2). Thus, with the exception of pathological cases like with $n = 2$ agents, the standard quorum rules have kinks in the cutoff line that separates status quo and reform. In contrast to that, a linear rule with $\xi > 0$ and $\xi^R > 0$ reacts to a qualified majority among (i) the number of votes for the status quo and (ii) the number of votes that are cast for reform *in addition* to a minimum required number.

In many practical applications the available social alternatives can be unambiguously mapped into the model variables R and S (say, when there is a status quo and a reform proposal). In some situations, however, the alternatives are

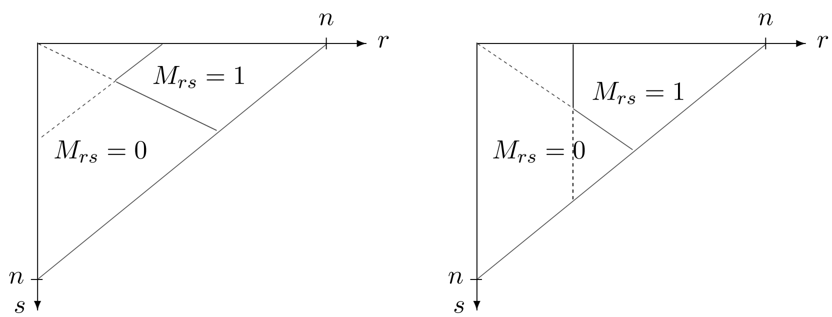

Figure 2: Examples of a participation quorum rule (left) and an approval quorum rule (right)

fundamentally symmetric, like two new candidates running for an office; in such applications, only rules that are neutral in the sense of treating both alternatives identically may be feasible. Formally, a voting rule M is *neutral* if

$$M(r, s) \quad = \quad 1 - M(s, r) \quad \text{for all } s \text{ and } r. \tag{3}$$

The unique neutral linear rule is the voluntary majority rule.

Equilibria of voting mechanisms

Consider a voting mechanism M and an equilibrium strategy σ. In the following we will use the shortcuts $\rho(a) = \rho^{M,\sigma}(a)$ for $a = A, S, R$. We can assume without loss of generality that $\rho(R) \geq \rho(S)$ (exchange the labels of the actions R and S if necessary). An agent of type v prefers action R over action A if

$$v(\rho(R) - \rho(A)) \quad \geq \quad c,$$

and prefers S over A if

$$v(\rho(S) - \rho(A)) \quad \geq \quad c.$$

For any mechanism-equilibrium pair there exists an equivalent pair such that

$$\rho(R) \geq \rho(A) \quad \text{and, analogously,} \quad \rho(S) \leq \rho(A). \tag{4}$$

To see this, note first that $\rho(R) = \rho(A)$ if M is S-one-sided. If $\rho(R) < \rho(A)$, then all types $v > 0$ will abstain; we can replace the mechanism M by an S-one-sided mechanism $\hat{M}$ (if $\rho(S) < \rho(R) < \rho(A)$, nobody will take action R, otherwise $\rho(S) = \rho(R) < \rho(A)$ and one first replaces S and R by a single action S by arguing as in the proof of Lemma 1).

Assuming (4) and using the fact that F has no atoms, any equilibrium strategy σ is characterized by the pivot probabilities $\Delta^R = \rho(R) - \rho(A) \geq 0$ and $\Delta^S = \rho(A) - \rho(S) \geq 0$, via

$$\sigma(v) = \begin{cases} R & \text{if } v\Delta^R > c, \\ S & \text{if } -v\Delta^S > c, \\ A & \text{otherwise.} \end{cases} \tag{5}$$

Thus, we can take the pair of pivot probabilities as independent variables and formulate the equilibrium conditions in terms of (Δ^R, Δ^S). An equilibrium requires that the strategy σ defined via (Δ^R, Δ^S) is a best response to itself.

To formulate the equilibrium conditions, we need additional notation. From the point of view of a given agent, the probability of any particular tally (r, s) ($r + s \leq n - 1$) of others' votes is given by a multinomial distribution,

$$\Pr_{\tau^R,\tau^S}(r, s) = \binom{n-1}{r\ s}(\tau^R)^r(\tau^S)^s(1 - \tau^R - \tau^S)^{n-1-r-s}, \tag{6}$$

where $\tau^R \in [0, F^R]$ ("R-participating rate") denotes the probability that a given other agent is a participating R-agent, $\tau^S \in [0, F^S]$ ("S-participating rate") denotes the probability that a given other agent is a participating S-agent, and $\binom{n-1}{r\ s} = \frac{(n-1)!}{r!s!(n-1-r-s)!}$. If an agent anticipates that the other agents use the strategy defined via (Δ^R, Δ^S), then from (5) she anticipates the participation rates

$$\begin{aligned} \tau^R &= l^R(\Delta^R) \stackrel{\text{def}}{=} 1 - F(c/\Delta^R), \\ \tau^S &= l^S(\Delta^S) \stackrel{\text{def}}{=} F(-c/\Delta^S). \end{aligned}$$

(Let $l^R(0) = 0$ and $l^S(0) = 0$.) An agent's best-response will be based on the pair of pivot probabilities (d^R, d^S) that she anticipates given the behavior of others,

$$\begin{aligned} d^R(M, \tau^R, \tau^S) &= \sum_{r+s\leq n-1} \Pr_{\tau^R,\tau^S}(r, s)(M_{r+1,s} - M_{r,s}), \\ d^S(M, \tau^R, \tau^S) &= \sum_{r+s\leq n-1} \Pr_{\tau^R,\tau^S}(r, s)(M_{r,s} - M_{r,s+1}). \end{aligned}$$

Pivot probabilities $(\Delta^R, \Delta^S) \in [0, 1]^2$ are called an *equilibrium* if[14]

$$\Delta^R = d^R(M, l^R(\Delta^R), l^S(\Delta^S)), \tag{7}$$
$$\Delta^S = d^S(M, l^R(\Delta^R), l^S(\Delta^S)). \tag{8}$$

[14]Note that, given an equilibrium $(\Delta^{R\prime}, \Delta^{S\prime})$ with $l^R(\Delta^{R\prime}) = 0$, the equilibrium strategy (5) could be represented via *any* pair (Δ^R, Δ^S) with $0 \leq \Delta^R \leq \Delta^{R\prime}$ and $\Delta^S = \Delta^{S\prime}$. Conditions (7) and (8) fix a unique representation. Similar for equilibria with $l^S(\Delta^{S\prime}) = 0$.

Observe that the equilibrium conditions are linear in M and are non-linear in Δ^R and Δ^S.

Given an equilibrium, any type or player who takes action R (resp., S) is called an *R-voter* (resp., *S-voter*).

Sometimes it will be convenient to describe equilibria not in terms of the pivot probabilities (Δ^R, Δ^S), but in terms of the participation rates $(\tau^R, \tau^S) \in [0, F^R] \times [0, F^S]$. There is a continuous one-to-one relationship between the equilibrium representations via pivot probabilities and via participation rates:[15]

$$(\Delta^R, \Delta^S) \mapsto (l^R(\Delta^R), l^S(\Delta^S)), \quad (\tau^R, \tau^S) \mapsto (d^R(M, \tau^R, \tau^S), d^S(M, \tau^R, \tau^S)). \tag{9}$$

Via this translation, the equilibrium conditions (7) and (8) can be applied to any pair (τ^R, τ^S).

We will refer to any (M, Δ^R, Δ^S) satisfying (7) and (8) as a *mechanism-equilibrium pair*; we say that (M, Δ^R, Δ^S) is *neutral* if M is neutral and $\Delta^R = \Delta^S$. Via the translation (9) we extend the same terminology to (M, τ^R, τ^S).

Optimal mechanisms

A mechanism-equilibrium pair $m = (M, \tau^R, \tau^S)$ yields, for each type v, the interim expected utility

$$U^m(v) \quad = \quad v\rho^m(A) + \begin{cases} \max\{vd^R(M, \tau^R, \tau^S) - c, 0\} & \text{if } v > 0, \\ \max\{-vd^S(M, \tau^R, \tau^S) - c, 0\} & \text{if } v < 0, \end{cases}$$

where, from the point of view of an abstaining voter, R gets implemented with probability

$$\rho^m(A) \quad = \quad \sum_{r+s \leq n-1} \Pr_{\tau^R, \tau^S}(r, s) M_{r,s}. \tag{10}$$

The objective is to maximize the welfare of each individual from an ex-ante point of view. For the equilibrium analysis we have normalized the Bernoulli utilities such that all types $v \in \text{supp}(F)$ have the same participation cost. Towards the welfare analysis, a weight function $g(v) > 0$ is required to capture cases with heterogenous participation costs (cf. Appendix B). We denote by the measure $G(v) = \int_{v' \leq v} g(v') \mathrm{d}F(v')$ the total weight assigned to types below v (where we

[15] Recall that, w.l.o.g., we are restricting attention to equilibria of the form (5) with $\Delta^R \geq 0$ and $\Delta^S \geq 0$. Equivalently, we are restricting attention to equilibria (τ^R, τ^S) with $d^R(M, \tau^R, \tau^S) \geq 0$ and $d^S(M, \tau^R, \tau^S) \geq 0$.

assume the first moment of G exists). The aggregate weight is denoted $\gamma = G(\infty)$. The welfare is defined as

$$W(m) \quad = \quad \int U^m(v) \mathrm{d}G(v) = E[U^m(\tilde{v})], \tag{11}$$

where we introduce the random-variable notation $\tilde{v}$ as a shortcut.

At a first reading, the reader may focus on the constant weight function $g(v) = 1$ for all v, that is, $G = F$. This means that changing the outcome from S to iS (or from R to iR) is equally costly for any two (equally likely) types. We find it important, however, to allow for a general g in order to cover various other welfare criteria that are used in the related literature. In particular, in much of the related literature, valuations are fixed on each side of the electorate and participation costs private. This leads to a different ex-ante welfare criterion according to which changing the outcome from S to R is equally beneficial for any two types who prefer R and equally costly for any two types who prefer S.[16]

Sometimes it will be useful to express the welfare in terms of pivot probabilities::

$$\begin{aligned} \check{W}(M, \Delta^R, \Delta^S) \quad = \quad & E[\tilde{v}]\rho^{(M, l^R(\Delta^R), l^S(\Delta^S))}(A) \\ & + E[(\tilde{v}\Delta^R - c)\mathbf{1}_{\tilde{v}\Delta^R > c}] + E[(-\tilde{v}\Delta^S - c)\mathbf{1}_{-\tilde{v}\Delta^S > c}]. \end{aligned} \tag{12}$$

We say that a mechanism-equilibrium pair is *optimal* if it solves the problem

$$\begin{array}{rll} \text{(opt)} & \max\limits_{(M, \Delta^R, \Delta^S)} & \check{W}(M, \Delta^R, \Delta^S) \\ & \text{s.t.} & (7),\ (8), \\ & & \Delta^R \geq 0, \quad \Delta^S \geq 0, \\ & & 0 \leq M_{rs} \leq 1 \quad \text{for all } (r, s). \end{array}$$

Implicit to our formulation is the classical mechanism-design doctrine: the designer selects an equilibrium with highest welfare if the optimal mechanism has multiple equilibria. The feasible set of problem (opt) is non-empty and compact. Hence, an optimum exists by Weierstraß' Maximum-Value Theorem.

In a neutral environment, we may also consider a "neutral" social planner who maximizes the ex-ante welfare among all neutral mechanism-equilibrium pairs,

[16] For example, consider Börgers (2004), where, for all agents i, the valuations $v_i \in \{-1, 1\}$ are equally likely and the cost $c_i \in [\underline{c}, \overline{c}]$ is independently distributed according to a c.d.f. with density h. Translated to our preference representation, the distribution of types is F with density $f(v) = (1/2)h(c/|v|)c/v^2$ ($|v| \in [c/\overline{c}, c/\underline{c}]$). To make W equal to the ex-ante welfare of Börgers, define the weights $g(v) = 1/(|v|\overline{G})$, where $\overline{G} = 2\int_{c/\overline{c}}^{c/\underline{c}} f(v)/v$.

using neutral weights ($g(v) = g(-v)$ for all v). Describing any neutral equilibrium in terms of its pivot probability $\Delta^R = \Delta^S \overset{\text{def}}{=} \Delta$, the welfare with neutral rule M is

$$\check{W}(M, \Delta) \quad = \quad 2 \int \max\{v\Delta - c, 0\} \mathrm{d}G(v). \tag{13}$$

Thus, the welfare is maximal if Δ is maximal.

For a neutral pair (M, Δ), the equilibrium conditions (7) and (8) are identical to each other. Thus, the neutral planner solves

$$\begin{array}{lll} \text{(neutral opt)} & \max\limits_{(M,\Delta)} & \check{W}(M, \Delta) \\ & \text{s.t.} & (3), \\ & & (7) \text{ with } \Delta^R = \Delta^S = \Delta, \\ & & \Delta \geq 0, \\ & & 0 \leq M_{rs} \leq 1 \quad \text{for all } (r, s) \text{ with } r < s. \end{array}$$

Finally, we extend the welfare expression (11) to capture mechanism-participation-pair combinations $m = (M, \tau^R, \tau^S)$ in which (τ^R, τ^S) is not an equilibrium. This is achieved by first computing the aggregate welfare conditional on each tally. Expression (14) will allow us to consider the "first-best" benchmark of a hypothetical social planner who is able to impose arbitrary participation rates.

$$\begin{aligned} W(m) \quad = \quad & \frac{1}{n} \sum_{r+s \leq n} \binom{n}{r\ s} (\tau^R)^r (\tau^S)^s (1 - \tau^R - \tau^S)^{n-r-s} M_{rs} \omega_{rs}(\tau^R, \tau^S) \\ & -(\alpha^R(\tau^R) + \alpha^S(\tau^S))c, \end{aligned} \tag{14}$$

where

$$\alpha^R(\tau^R) = \gamma - G(F^{-1}(1 - \tau^R)) \ \text{ and } \ \alpha^S(\tau^S) = G(F^{-1}(\tau^S)),$$

and $\omega_{rs}(\tau^R, \tau^S)$ is the welfare conditional on the tally (r, s) and R being implemented. Thus,

$$\omega_{rs}(\tau^R, \tau^S) \quad = \quad r\eta^R(\tau^R) - s\eta^S(\tau^S) + (n - r - s)\eta^A(\tau^R, \tau^S), \tag{15}$$

where

$$\begin{aligned} \eta^R(\tau^R) \quad &= \quad \frac{E[\tilde{v}\mathbf{1}_{\tilde{v} > F^{-1}(1-\tau^R)}]}{\tau^R} \quad \text{if } \tau^R > 0, \\ \eta^S(\tau^S) \quad &= \quad \frac{E[-\tilde{v}\mathbf{1}_{\tilde{v} < F^{-1}(\tau^S)}]}{\tau^S} \quad \text{if } \tau^S > 0, \\ \eta^A(\tau^R, \tau^S) \quad &= \quad \frac{E[\tilde{v}\mathbf{1}_{F^{-1}(\tau^S) < \tilde{v} < F^{-1}(1-\tau^R)}]}{1 - \tau^R - \tau^S} \quad \text{if } \tau^R + \tau^S < 1 \end{aligned}$$

denote the G-conditional-expected valuation of an R-voter, S-voter, and abstaining type, respectively, each normalized such that the probability of the conditioning event if computed using F. (Note that (14) is continuous at the boundaries $\tau^R = 0$, $\tau^S = 0$, and $\tau^R + \tau^S = 1$ because the denominators in the definitions of η^R, η^S, and η^A get cancelled.) If the weight function g is identically equal to 1, the expressions above simplify to $\alpha^R = \tau^R$, $\alpha^S = \tau^S$, and $\eta^R = E[\tilde{v}|\tilde{v} > F^{-1}(1 - \tau^R)]$, $\eta^S = E[-\tilde{v}|\tilde{v} < F^{-1}(\tau^S)]$, $\eta^A = E[\tilde{v}|F^{-1}(\tau^S) < \tilde{v} < F^{-1}(1 - \tau^R)]$.

3 Optimality of linear rules

Our first contribution is to establish that always some linear mechanism is optimal. As an aside, we establish that full participation is never optimal (even when mechanism-equilibrium pairs with full participation exist). Moreover, linearity is essentially necessary for optimality: any optimal mechanism is linear on the set of tallies that occur with positive probability in equilibrium.

Proposition 1. *There exists an optimal mechanism-equilibrium pair in which the mechanism is a linear voting rule.*

Any optimal mechanism-equilibrium pair $(M^*, \tau^{R*}, \tau^{S*})$ is such that $\tau^{R*} + \tau^{S*} < 1$. Moreover, if $\tau^{R*} > 0$ and $\tau^{S*} > 0$, then M^* is linear; if $\tau^{R*} > 0$ and $\tau^{S*} = 0$, then M^*_{r0} ($r = 0, \dots, n$) is as an R-one-sided linear rule; if $\tau^{R*} = 0$ and $\tau^{S*} > 0$, then M^*_{0s} ($s = 0, \dots, n$) is as an S-one-sided linear rule.

To prove the optimality of partial abstention, we suppose to the contrary that there is an optimal rule such that in equilibrium all types participate. Then only the tallies with $r + s = n$ occur with positive probability. To obtain a contradiction, we replace the rule by an R-one-sided rule that yields the same collective decision at lower participation cost. The difficult part is to prove that the R-one-sided rule has an equilibrium in which all R-agent types participate; to this end, we first show that the optimal rule can be chosen as a linear rule.

Towards proving the linearity claims, an important initial insight is that the space of mechanisms M is $(n+1)(n+2)/2$-dimensional, whereas—by the symmetry across agents—equilibria are merely two-dimensional objects. Therefore, if a participation pair is an equilibrium in some mechanism, then it will be an equilibrium in a large set of mechanisms, $\mathcal{M}$. Thus, fixing an optimal participation pair and finding in $\mathcal{M}$ the mechanism that yields the highest welfare carries a long way.

Consider the "biased" case in which $E[\tilde{v}] > 0$ and neither side of the electorate abstains entirely, $\tau^R > 0$ and $\tau^S > 0$, so that each tally (r, s) occurs with positive probability. The welfare expression (12) reveals that a mechanism-equilibrium pair $m = (M, \tau^R, \tau^S)$ can be optimal only if M maximizes (10) subject to the equilibrium conditions (7) and (8). The objective (10) is increasing in each M_{rs}. But increasing an M_{rs} can imply that the planner loses some R-pivotality d^R on the right-hand side of (7) and/or S-pivotality d^S on the right-hand side of (8). For all (r,s) where losing pivotality is too expensive given the shadow prices μ^R and μ^S set by the Lagrangian, implementing S (i.e., $M_{rs}=0$) is optimal.

Imagine increasing some M_{rs} by ϵ. This changes the objective (10) by

$$\mathrm{d}\rho(A) = \Pr(r,s)\epsilon.$$

There are two countervailing effects on the R-pivotality in (7) because an R-vote can either bring about the election result (r,s) (with $\Pr(r-1,s)$) or avoid this election result (with $\Pr(r,s)$). The resulting gain of R-pivotality is

$$(\Pr(r-1,s) - \Pr(r,s))\epsilon = \left(r\frac{1-\tau^S}{\tau^R} + s - n\right)\frac{\mathrm{d}\rho(A)}{n-r-s}.$$

Similarly, in (8) there is an S-pivotality gain of

$$(\Pr(r,s) - \Pr(r,s-1))\epsilon = \left(-r - s\frac{1-\tau^R}{\tau^S} + n\right)\frac{\mathrm{d}\rho(A)}{n-r-s}.$$

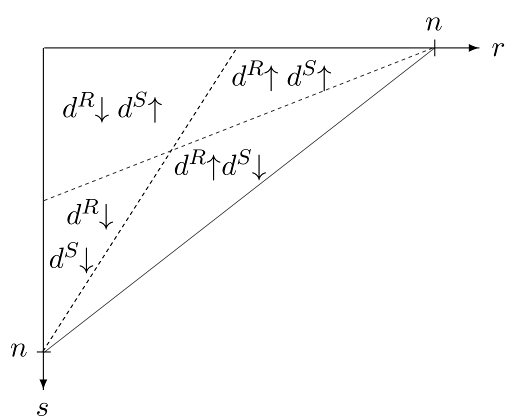

Figure 3: Effect of increasing any M_{rs} on the pivotalities. For all (r,s) below the steeper dashed line, increasing M_{rs} increases the R-pivotality d^R; vice versa below the line. Above the flatter dashed line, increasing M_{rs} increases the S-pivotality d^S; vice versa below the line. The intersection point of the dashed lines, $(n\tau^R, n\tau^S)$, is the stochastic expectation of the election result.

For all (r,s) such that increasing M_{rs} increases both pivotalities (the upper right triangle in Figure 3), it is optimal to implement R. For all other tallies there is a trade-off. Summarizing the expressions above, the value of the Lagrangian changes by

$$\left[n - r - s + \mu^R\left(r\frac{1-\tau^S}{\tau^R} + s - n\right) + \mu^S\left(-r - s\frac{1-\tau^R}{\tau^S} + n\right)\right]\frac{\mathrm{d}\rho(A)}{n-r-s}.$$

The expression in brackets $[\dots]$ is linear in r and s. Setting $M_{rs} = 1$ if $[\dots]$ positive and $= 0$ if it is negative yields the optimal rule.

The cases in which $E[\tilde{v}] < 0$ and/or one side of the electorate abstains, $\tau^R = 0$ or τ^S, are similar.

Finally, consider the remaining "unbiased" case $E[\tilde{v}] = 0$. Here, the welfare (12) is independent of $\rho^m(A)$ and is increasing in Δ^R and Δ^S. Thus, the optimal mechanism makes each side, the R-voters and the S-voters, "as pivotal as possible" given the equilibrium conditions. Put differently, among the pairs (Δ^R, Δ^S) that are equilibrium feasible, the optimal mechanism selects a Pareto point.

The problem of characterizing Pareto points (Δ^R, Δ^S) can be looked at from the point of view of any player, given the participation decisions of the other players. More formally, we first show that the equilibrium conditions (7) and (8) can be relaxed to become inequality ($\leq$) constraints. Consequently, at an optimum one cannot change the mechanism to increase both the R-pivotality and the S-pivotality because otherwise one could increase Δ^R and Δ^S without violating the constraints. Given that, the separating-hyperplane theorem implies that any optimal rule maximizes a weighted sum of the voter's R-pivotality and her S-pivotality. This implies linearity of the optimal rule by arguments analogous to the biased case above.

In summary, the essential reasons behind the optimality of a linear rule are as follows. First, the objective and the constraints are linear in the probabilities M_{rs}. Second, we rely on the particular form of the relative probabilities of neighboring tallies; this form follows from the stochastic independence of the voters' types.

The proof of Proposition 1 can be found in the Appendix, except for the unbiased case which is relegated to Appendix B.

By varying the arguments from the proof of Proposition 1, its conclusions can be extended to a variety of important related settings. Firstly, a linear rule is optimal if there is a positive mass of R-types that do not strictly prefer outcome R over iR, that is, who have no participation cost, and a positive mass of S-types with analogous preferences (cf. Appendix B). Secondly, a linear rule is optimal if the designer's objective is any convex combination of expected utility, turnout of R-agents (τ^R), and turnout of S-agents (τ^S); here, full participation may be optimal. Thirdly, a linear rule is optimal if there is a fixed population of R-agents and a fixed population of S-agents, as in Palfrey and Rosenthal (1985) (cf. Appendix B). This last example constitutes one case where the agents' types are not i.i.d. distributed and the optimal rule is still linear.[17]

[17] Another possible deviation from the i.i.d. case is that the vector of valuations has an exchange-

Quorum rules are, with the exception of pathological cases, non-linear, due to the kink in the cutoff line (cf. Figure 2). By Proposition 1, such quorum rules can be optimal only with an equilibrium in which one side of the electorate abstains; in this case the outcome is the same as with a one-sided linear rule. Accordingly, our results suggest that linear rules rather than quorum rules should be used.

It is instructive to compare the "second-best" considered in Proposition 1 with a hypothetical planner who chooses a "first-best" mechanism-participation-rates pair (M, τ^R, τ^S) that maximizes (14) not subject to the equilibrium conditions. It is immediate that any first-best M is linear, with parameters

$$\xi^R = \eta^R - \eta^A, \quad \xi^S = \eta^S + \eta^A, \quad \xi = -\eta^A. \tag{16}$$

Proposition 1 shows that the "second-best" is linear as well, but with generally different parameters of the optimal rule.

For the neutral social planner, a result similar to Proposition 1 holds. The standard voluntary majority rule—the unique voting rule that is both linear and neutral—is always optimal. In contrast to Proposition 1, however, full participation may be optimal.

Proposition 2. *Consider a neutral environment.*

There exists a neutrally optimal mechanism-equilibrium pair in which the mechanism is the standard voluntary majority rule. Moreover, in any neutrally optimal mechanism-equilibrium pair with non-zero participation and non-zero abstention, the mechanism is identical to the standard voluntary majority rule.

The proof of Proposition 2 is analogous to the proof of the unbiased case in Proposition 1: one first shows that (7) can be relaxed into an inequality and then maximizes the right-hand-side of (7) across all neutral M; we omit the details.[18]

On first sight, Proposition 2 may appear un-surprising. But other neutral settings are known in which the standard majority rule is suboptimal, even among neutral rules and neutral equilibria: Schmitz and Tröger (2012) show that, in the absence of voting costs, but with correlated valuations, a weak majority rule can yield a higher welfare. In a weak majority rule, a lottery is used if neither alternative has a sufficiently strong majority over the other alternative.

---

able distribution, as in Schmitz and Tröger (2012), where the participation cost is zero. In some cases such as when the valuations are affiliated, Schmitz and Tröger obtain the optimality of the standard majority rule among dominant-strategy rules. With a participation cost, an agent's participation decision depends crucially on her believed pivotality, so dominant-strategy implementation is rather restrictive. Turning to Bayesian implementation, characterizing optimal rules with affiliated valuations is an open problem even when there is no participation cost.

[18] The essence of the arguments goes back to Rae (1969).

It is interesting to relate the solution to problem (neutral opt) to Börger's (2004) fundamental insight that *too many* agents participate in a voluntary majority rule compared to the 'first-best" participation rate selected by a social planner who is not constrained by equilibrium. In the "second-best", it is optimal to use a mechanism that induces participation of *as many agents as possible*. The voluntary majority rule is optimal because it induces maximum participation among neutral rules and neutral equilibria. Analogously, in the unbiased case $E[\tilde{v}] = 0$ of Proposition 1, optimality requires a "Pareto-optimal" participation pair.

4 Which linear rules can be optimal?

Which linear rules can be optimal if one goes beyond the case of a neutral social planner operating in a neutral environment? Information about the distribution of types is needed to answer this question. We consider low-cost environments and show that a one-sided rule is optimal if uncertainty concerns mainly an agent's preferred alternative, but not her preference intensity. In this case, it is enough to only count one side of the electorate because almost all the relevant information about the other side is already contained. A qualified majority rule is optimal if preference intensities play a role in an all-or-nothing manner, that is, if each agent is either significantly affected by the public decision or is almost indifferent. In this case, the participation decision screens out the significantly affected agents and it is important to distinguish types of agents on both sides of the electorate. Both results hold for arbitrary continuous welfare-weight functions. A general linear rule that belongs to neither of these classes can also be optimal.

The constructions in this section rely on type distributions that approximate two- or three-point distributions. Using these to build examples is natural because with three actions a distribution with three types corresponds to the maximum amount of separation that technically can be achieved in a voluntary voting mechanism. In addition, our examples are such that the equilibrium in the optimal mechanism is straightforward. Consider a distribution of valuations, $\hat{F}$:

$$\begin{pmatrix} v^S & v^0 & v^R \\ p^S & p^0 & p^R \end{pmatrix},$$

with the realizations $v^S < v^0 \leq 0 < v^R$ occuring with the probabilities $p^S > 0$, $p^0 \geq 0$, and $p^R > 0$, respectively. Thus, $\hat{F}$ is a three-point distribution if $p^0 \neq 0$ and is a two-point distribution otherwise.

We fix a continuous weight function $g(v) > 0$ for all v in a neighborhood of supp$(\hat{F})$.

While the equilibrium conditions (7) and (8) are not formulated for a discrete type distribution such as $\hat{F}$, they do apply to approximations. Let $\epsilon > 0$. In an ϵ-approximation, each realization of $\hat{F}$ is replaced by a continuum of types in its ϵ-neighborhood (for example, the ϵ-approximation may have a piecewise flat density).

Formally, a continuous type distribution F is an ϵ-approximation of $\hat{F}$ if

$$\text{supp}F \quad \subseteq \quad [v^S - \epsilon, v^S + \epsilon] \cup [v^0 - \epsilon, v^0 + \epsilon] \cup [v^R - \epsilon, v^R + \epsilon],$$

where $p^R = 1 - F(v^R - \epsilon)$ and $p^S = F(v^S + \epsilon)$. We assume ϵ is small enough so that g is defined on supp(F).

In the discrete setting, similar to (15), the welfare from implementing R conditional on r voters having type v^R, s voters having type v^S, and $n - r - s$ voters having type v^0, is

$$w(r, s) \quad = \quad rv^R g(v^R) + sv^S g(v^S) + (n - r - s)v^0 g(v^0), \tag{17}$$

ignoring the participation cost.

An important benchmark for our analysis are environments in which the valuation distribution F approximates a two-point distribution, that is $p^0 = 0$. Here, uncertainty concerns mainly an agent's preferred alternative, but not her preference intensity. Assuming the participation cost is small, we show that the optimal rule is one-sided, where the side of the electorate with lower (weighted) participation is called to vote. Remark 1 implies that one-sided rules with arbitrary parameters can be optimal for any population size.

Remark 1. (Optimality of one-sided rules.) *Consider a two-point distribution $\hat{F}$ satisfying*

$$w(r, n - r) \neq 0 \quad \textit{for all } r. \tag{18}$$

Let the participation cost c be sufficiently small. If F is an ϵ-approximation of $\hat{F}$ where ϵ is sufficiently small (given c), then a one-sided voting rule is optimal. Specifically, if $p^R g(v^R) < p^S g(v^S)$, then the mechanism-equilibrium pair $(\mathbf{1}_{w(r,n-r)>0}, p^R, 0)$ is the unique optimum; if $p^R g(v^R) > p^S g(v^S)$, then $(\mathbf{1}_{w(n-s,s)>0}, 0, p^S)$ is the unique optimum.

To prove this (for details see the Appendix), observe that, if c is small and F a close approximation of $\hat{F}$, then the R-one-sided mechanism $\mathbf{1}_{w(r,n-r)>0}$ has an equilibrium with full participation of the R-agents, and the S-one-sided mechanism $\mathbf{1}_{w(n-s,s)>0}$ has an equilibrium with full participation of the S-agents.

In any optimum, at least one side of the electorate will fully participate: if both a type close to v^R and a type close to v^S did not participate, then the equilibrium conditions would yield upper bounds for both the R-pivotality and the S-pivotality such that the welfare achieved would be close to the welfare of a constant rule, a contradiction to optimality because c is small.

Say that all R-agents participate. Then, given any election result (r, s), it is ex-post optimal to implement R if and only if $w(r, n-r) > 0$, independently of s. The least participation-intense way to achieve this implementation is with the R-one-sided rule $\mathbf{1}_{w(r,n-r)>0}$. Similarly, we would have arrived at the S-one-sided candidate $\mathbf{1}_{w(n-s,s)>0}$ if above we had assumed that all S-agents participated. Among the two candidates, the one with less participation (as measured with the weights g) gives higher welfare because both candidates implement the same collective decision.

In order to make the case for qualified majority rules, we consider a family of type distributions $\hat{F}^{v^0}$ where v^0 is treated as a parameter, while v^S, v^R, and the probability distribution (p^S, p^0, p^R) are kept fixed. By considering v^0 close to 0 we capture environments in which each agent is either significantly affected by the public decision or is almost indifferent. We use the notation $w^{v^0}(r, s)$ for the function (17). We fix a continuous weight function g in a neighborhood of supp$(\hat{F}^0)$.

Assuming a small participation cost and excluding, in the limit $v^0 \to 0$, non-generic cases (19) as well as extreme cases in which a single agent's preferences can outweigh everybody else (20), we show that the optimal rule is a qualified majority rule in which only the significantly affected agents participate. Remark 2 implies that qualified majority rules with arbitrary parameters can be optimal for any population size.

Remark 2. (Optimality of qualified majority rules.) *Consider a three-point type distribution $\hat{F}^{v^0}$ for each v^0 in a left-neighborhood of 0. Assume that*

$$w^0(r, s) \neq 0 \quad \textit{for all } (r, s) \neq (0, 0), \tag{19}$$

$$w^0(n-1, 1) > 0, \quad \textit{and} \quad w^0(1, n-1) < 0. \tag{20}$$

Let the participation cost c be sufficiently small. If F is an ϵ-approximation of $\hat{F}^{v^0}$ where $v^0 < 0$ is sufficiently close to 0 (given c) and ϵ is sufficiently small (given c and v^0), then the mechanism-equilibrium pair $(\mathbf{1}_{w^0(r,s)>0}, p^R, p^S)$ is the unique optimum.

The proof is relegated to the Appendix. From the start we can restrict attention to v^0 so small that

$$w^{v^0}(r, s) \text{ has the same sign as } w^0(r, s) \text{ for all } (r, s) \neq (0, 0). \tag{21}$$

Note that $w^{v^0}(0,0) < 0$ if $v^0 < 0$. We first consider the limit setting in which types are distributed according to the three-point distribution $\hat{F}^{v^0}$ and the participation cost equals 0. We consider the "first-best" problem (*): maximize the welfare (14) across all mechanisms and all feasible participation pairs, ignoring the equilibrium conditions. Clearly, $(\overline{M}, p^R, p^S)$, with $\overline{M}_{rs} = \mathbf{1}_{w^0(r,s)>0}$, is a solution because by (21) it is welfare-maximizing tally by tally. Moreover, by (19) and (20) it is the unique maximizer if $v^0 < 0$; if $v^0 = 0$, then there are other maximizers that differ at $(r,s) = (0,0)$ so that the set of maximizers is

$$\overline{\mathcal{M}} \quad = \quad \{M \mid M(r,s) = \mathbf{1}_{w^0(r,s)>0} \text{ for all } (r,s) \neq (0,0)\}. \tag{22}$$

This implies (by a subsequence argument) that any maximizer of any variant of (*) with a slightly perturbed objective must be close to the set $\overline{\mathcal{M}} \times \{p^R\} \times \{p^S\}$.

With this background, we move away from the limit setting. Since we choose c small, v^0 close to 0 given c and ϵ small given c and v^0, the pair (p^R, p^S) becomes an equilibrium in $\overline{M}$ and, moreover, the welfare objective is only slightly different from the objective in (*). Thus, any optimal mechanism-equilibrium pair is close to $\overline{\mathcal{M}} \times \{p^R\} \times \{p^S\}$.

Using the gaps in the support of F, the equilibrium conditions imply that in any mechanism sufficiently close to $\overline{\mathcal{M}}$, any equilibrium sufficiently close to (p^R, p^S) is in fact equal to (p^R, p^S). Thus, in any optimal mechanism-equilibrium pair the equilibrium (p^R, p^S) is played. Given these participation rates, the welfare conditional on any tally (r,s) approximates $w^{v^0}(r,s)$ so that by (19) no mechanism other than $\overline{M}$ can be optimal.[19]

Remark 2 relies on the assumption that v^0 is small given c. It is instructive to also consider the opposite case in which c is small given v^0. Here, every "first-best" solution still approximates a point in the set $\overline{\mathcal{M}} \times \{p^R\} \times \{p^S\}$. However, it is straighforward to see that no mechanism close to $\overline{\mathcal{M}}$ has an equilibrium close to (p^R, p^S) because the types close to $v^0 < 0$ would not abstain. Thus, the second-best mechanism differs from the first-best mechanism.

Finally, we provide an example of an environment in which neither a one-sided rule nor a qualified majority rule are optimal. Instead, optimality requires that not only the ratio of R-votes versus S-votes is taken into account, but also the number of abstentions is. Consider the following three-point distribution:

$$\begin{pmatrix} v^R & v^S & v^0 \\ p^R & p^S & p^0 \end{pmatrix} \quad = \quad \begin{pmatrix} 54 & -66 & -6 \\ 0.3 & 0.2 & 0.5 \end{pmatrix}. \tag{23}$$

[19]Remark 2 does not cover the welfare criterion that is obtained when the participation cost is private and the valuation is fixed on each side of the electorate as, e.g., in Palfrey and Rosenthal (1985) or Börgers (2004). The corresponding welfare weights $g(v)$ are proportional to $1/|v|$ which cannot be extended continuously to $v = 0$. A valuation approaching 0 in our setting corresponds to a participation cost approaching infinity in Palfrey and Rosenthal (1985) or Börgers (2004).

It is optimal to implement R if and only if there are at least two more votes in favor of R than votes in favor of S; in equilibrium, the strongly affected types 54 and -66 participate and the less affected type -6 abstains.

Remark 3. (Optimality of a general linear rule.) *Consider a population of $n = 15$ agents with participation cost $c = 1$. Assume the weight function $g(v) = 1$ for all v. If the type distribution F is an ϵ-approximation of* (23)*, where ϵ is sufficiently small, then the mechanism-equilibrium pair $(\mathbf{1}_{r>s+1}, p^R, p^S)$ is the unique optimum.*

The proof is relegated to Appendix B. We first show that all one-sided rules yield a lower welfare than the candidate optimum. This implies that in any optimum the types around v^R and around v^S participate. If $n = 15$, and $c = 1$, and F is an ϵ-approximation with sufficiently small ϵ, then the participation pair $(\tau^R, \tau^S) = (p^R, p^S)$ is an equilibrium in the linear mechanism $\overline{M}$,

$$\overline{M}_{rs} \quad = \quad \mathbf{1}_{r>s+1}. \tag{24}$$

It is straightforward to verify that the equilibrium maximizes the sum of the payoffs ex post:

$$r > s + 1 \quad \Leftrightarrow \quad w(r, s) > 0.$$

Given that, we can establish optimality of $(\mathbf{1}_{r>s+1}, p^R, p^S)$ by ignoring the equilibrium conditions, similar to the proof of Remark 2.

5 Voluntary versus compulsory participation

An important question is whether the designer of a voting rule should make participation compulsory. Börgers (2004) answers this to the negative for the standard majority rule in neutral environments. On the other hand, a subsidy for voting can improve welfare if the standard majority rule with voluntary participation is used in a non-neutral environment (Krasa and Polborn, 2009). This suggests that compulsory participation can be better than voluntary participation. However, given any non-neutral environment and assuming compulsory participation, the standard majority rule is not an optimal rule under the ex-ante welfare criterion. This raises a question: if one starts with a rule that is optimal with compulsory participation, what is the welfare effect of leaving participation to be voluntary?

To simplify, we make a genericity assumption:

$$\overline{y}(r) \stackrel{\text{def}}{=} r\frac{E[\tilde{v}\mathbf{1}_{\tilde{v}>0}]}{F^R} + (n-r)\frac{E[\tilde{v}\mathbf{1}_{\tilde{v}<0}]}{F^S} \neq 0 \quad \text{for all } r = 0, \ldots, n, \tag{25}$$

where $\overline{y}(r)$ is the welfare of implementing R conditional on r agents prefering R. Let $t^* = \min_{\overline{y}(r)>0} r$ denote the minimum number of agents preferring R such that alternative R is welfare-maximizing.

Compulsory participation in a voting rule M corresponds to the assumption that agents cannot take the action A, that is, full participation is enforced. A voting rule M with compulsory participation is welfare-maximizing if and only if

$$M_{r,n-r} = \mathbf{1}_{r \geq t^*}, \tag{26}$$

where the R-agents take action R and the S-agents take action S (cf. Rae, 1969, Barbera and Jackson, 2006). Thus, very limited information about the environment is required in order to implement an optimal rule with compulsory participation: knowing the ratio between the weighted conditional valuations on the two sides of the electorate is enough for a planner to set the optimal qualified majority threshold t^*. Many linear rules satisfy (26) because it entails no restriction of M_{rs} if $r + s < n$.

Proposition 3 below is the central result in this section. Starting with any deterministic linear voting rule that is optimal with compulsory participation, leaving participation voluntary improves welfare if the participation cost is sufficiently small. Proposition 3 does *not* rely on the comparison with the optimal rule with voluntary participation. Thus, any designer whose it able to compute the optimal compulsory majority threshold t^* (and knows that the participation cost is small) prefers voluntary over compulsory participation.

Proposition 3. *Assume (25). Suppose that F has a strictly positive and continuous density in a neighborhood of 0.*

Consider any deterministic two-sided linear voting rule M that is optimal if participation is compulsory. If participation is made voluntary and the participation cost is sufficiently small, then M has an equilibrium that yields a higher welfare than compulsory participation.

The main difficulty towards proving this is to find and characterize the claimed equilibrium in M. We cannot use the techniques from the examples in Section 4 where we established equilibria in particular voting rules using the gaps in the support of F. Also, we need to go beyond the known literature in which equilibrium analysis is largely restricted to qualified majority rules.

Our approach to equilibrium analysis in voting rules is presented in Lemma 2. For a two-sided rule, we can assume full participation at 0 participation cost. The neighboring small-cost cases are interesting due to the assumption that the support of F includes 0; otherwise a small participation cost would keep full participation

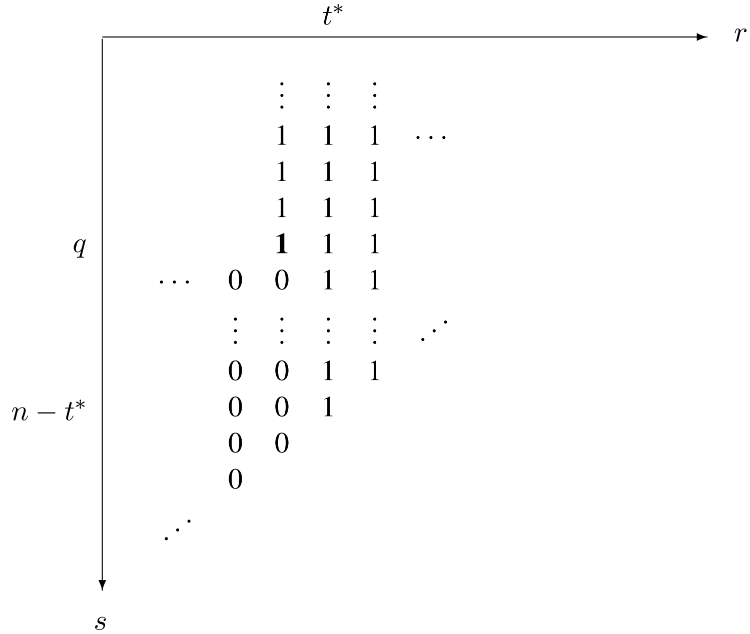

Figure 4: Some entries of a two-sided deterministic linear rule M. Here, M_{rs} is the entry at the intersection of the column labelled r and the row labelled s.

intact so that there would be no difference between voluntary and compulsory participation.[20] Using the implicit function theorem, we establish an equilibrium with almost-full participation if the participation cost is sufficiently small. We obtain the conclusion of Proposition 3 via comparing, across equilibria with voluntary and compulsory participation, the first-order welfare effect of changing c at $c = 0$.[21]

Additional notation is needed to lay out the details. Without loss of generality, assume that (cf. Figure 4)

$$M_{t^*,n-t^*} = 1, \quad M_{t^*-1,n-t^*+1} = 0, \quad M_{t^*-1,n-t^*} = 0. \tag{27}$$

(Assuming the mirror $M_{t^*-1,n-t^*} = 1$ of the third condition yields an analogous analysis with the roles of R and S reversed.)

Let $q \leq n - t^* - 1$ be maximal with the property $M_{t^*-1,q} = 1$ (such a q exists because M is two-sided).

If participation in M is compulsory, the resulting welfare is $W^* - c$, where

$$W^* \quad = \quad \frac{1}{n} \sum_{r=t^*}^{n} \binom{n}{r} (F^R)^r (F^S)^{n-r} \overline{y}(r). \tag{28}$$

Consider voluntary participation. We begin by showing the existence of an equilibrium with almost-full participation. A pair (τ^R, τ^S) with $\tau^R > 0$ and $\tau^S >$

[20]Consider for example the preference representation of Börgers (2004) where each agent i has a private participation cost c_i distributed according to a c.d.f. with compact support $[\underline{c}, \overline{c}]$ and a valuation $\in \{-1, 1\}$. Translating this into our preference representation yields a type distribution with support $[-1/\underline{c}, -1/\overline{c}] \cup [1/\overline{c}, 1/\underline{c}]$, that is, with a gap around 0.

[21]For some non-deterministic two-sided linear rules, the first-order welfare effect is the same with voluntary and compulsory participation so that one would have to compare higher-order effects.

0 is an equilibrium in a voting rule M if and only if both type $F^{-1}(1-\tau^R) > 0$ and type $F^{-1}(\tau^S) < 0$ are indifferent between participating and abstaining, that is,

$$\phi(c, M, \tau^R, \tau^S) \quad = \quad \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{29}$$

where

$$\phi(c, M, \tau^R, \tau^S) \quad = \quad \begin{pmatrix} F^{-1}(1-\tau^R) d^R(M, \tau^R, \tau^S) - c \\ F^{-1}(\tau^S) d^S(M, \tau^R, \tau^S) + c \end{pmatrix}.$$

We are now ready to present Lemma 2: if c is close to 0, then M has an equilibrium with almost full participation, and there is an explicit formula for the marginal welfare effect of introducing a participation cost.[22] Note that Lemma 2 in fact holds for all $t^* = 1, \ldots, n-1$. Using similar methods, first-order welfare effects for non-deterministic rules, for other equilibria, and for one-sided rules can be computed.[23] In the welfare effect (30), the term proportional to the density $F'(0)$ stems from types around 0 beginning to abstain if a participation cost is introduced. The other term, $-\gamma$, is the direct cost effect that stems from increasing c in (14) when the mechanism-equilibrium pair is kept fixed: in equilibrium essentially everybody pays the participation cost.

Lemma 2. *Make the assumptions of Proposition 3 and consider a deterministic two-sided linear voting rule M satisfying (27). Then, for all $c > 0$ sufficiently close to 0, there exists an equilibrium $(\tilde{\tau}^R(c), \tilde{\tau}^S(c))$ ($\to$ (F^R, F^S) as $c \to 0$) such that*

$$\lim_{c \to 0} W(M, \tilde{\tau}^R(c), \tilde{\tau}^S(c)) = W^*.$$

Moreover,

$$\frac{d}{dc} W(M, \tilde{\tau}^R(c), \tilde{\tau}^S(c)) \bigg|_{c=0} \quad = \quad F'(0) \left(1 - \frac{1}{n - t^* + 1 - q}\right) (-\overline{y}(t^* - 1)) - \gamma. \tag{30}$$

The proof of Lemma 2 is relegated to the Appendix. A major difficulty is that we *cannot* apply the implicit function theorem directly to describe the equilibrium

[22] Formula (30) shows that the first-order welfare effect is strictly decreasing in q. Thus, M with $q = 0$ ("almost one-sided rule") beats every other rule considered in the lemma in pairwise welfare comparison if c is small.

[23] Yet, a comparison of first-order welfare effects is not necessarily sufficient to find the optimal voting rule for small c because the set of (deterministic and non-deterministic) linear mechanisms is not finite, and the first-order welfare effect is not a continuous function of the mechanism.

(τ^R, τ^S) as a function of the cost c because the full-rank condition for the relevant Jacobi matrix is violated. Indeed,

$$\lim_{c \to 0} \frac{\mathrm{d}\tilde{\tau}^S}{\mathrm{d}c} \quad = \quad -\infty.$$

To overcome this problem, we first describe τ^R and c as implicit functions of τ^S. This allows us to describe the welfare as a function of τ^S. We show that the relation between τ^S and c is monotonic, so that a (locally unique) equilibrium in fact exists for every small c. Finally, we use L'Hospital's rule to compute the marginal welfare effect.

Proof of Proposition 3. In the limit $c \to 0$, both voluntary and compulsory participation yield the welfare W^*. The marginal welfare effect of introducing c equals $-\gamma$ if participation is compulsory. If participation is voluntary, the welfare effect (30) is $> -\gamma$ because $\overline{y}(t^* - 1) < 0$ by definition of t^*. This completes the proof.

The basic intuition behind Proposition 3 is that, with voluntary participation and a small participation cost, equilibrium turnout $\tau^R + \tau^S$ is already "too high" relative to the turnout that would be enforced by a planner who is not constrained by equilibrium: while in equilibrium $\tau^R + \tau^S \approx 1$, the unconstrained planner would require a non-negligible abstention rate.[24] This suggests that making participation compulsory lowers welfare. This intuition is similar to Börgers (2004), who's setting is indeed one-dimensional. But in our setting welfare is a function of the turnout pair (τ^R, τ^S) rather than a function of the turnout $\tau^R + \tau^S$—we have to deal with two dimensions rather than one. We address this problem with a different proof strategy that relies on local variations of equilibria in low cost environments.[25]

To conclude the analysis of voluntary versus compulsory voting, consider again a neutral social planner in a neutral environment. With compulsory participation, she finds the compulsory majority rule optimal (cf. Schmitz and Tröger, 2012). Börgers' (2004) analysis implies that the voluntary majority rule yields a higher welfare than the compulsory majority rule. Thus, from Proposition 2 we have:

[24]Suppose otherwise; that is, suppose the planner's optimum is $(\tau^R, \tau^S) = (F^R, F^S)$ at $c = 0$. Then, by the envelope theorem, the "first-best" welfare (14) achieved by the planner would change at the rate $\mathrm{d}W/\mathrm{d}c|_{c=0} = -\gamma$. This contradicts Lemma 2 which implies that the second-best welfare changes at a rate $> -\gamma$.

[25]For one-sided rules, we can prove a related, weaker result: making the same assumptions as in Proposition 3, there exists a one-sided linear mechanism-equilibrium pair that yields a higher welfare than the best compulsory voting rule if the participation cost is small; see Appendix B.

Corollary 1. *In any neutral environment, the voluntary majority rule maximizes the ex-ante welfare across all neutral equilibria of neutral mechanisms, voluntary and compulsory.*

Finally, we come to the question whether it is always desirable to lower participation costs when this is feasible. The following result answers this to the negative. If the density of weakly affected types is sufficiently large, then introducing a small participation cost can increase welfare; a similar result was obtained by Chakravarty et al. (2014) for the standard majority rule. The proof is immediate from Lemma 2 and $q \leq n - t^* - 1$.

Corollary 2. *Assume (25). Suppose that F has a continuous density in a neighborhood of 0, and*

$$F'(0) \quad > \quad \frac{2}{-\overline{y}(t^* - 1)}.$$

Consider any deterministic two-sided linear voting rule M that is optimal if participation is compulsory. If participation is made voluntary and the participation cost is sufficiently small, then M has an equilibrium that yields a higher welfare than full participation in the setting without participation cost.

It is interesting to contrast Corollary 2 with Drexl and Kleiner (forthcoming) who show that, in the absence of participation costs, a qualified majority rule maximizes welfare among anonymous deterministic dominant-strategy rules even when transfers are feasible. A participation cost can be interpreted as a transfer scheme. Thus, Corollary 2 identifies cases in which there exists a very simple welfare-enhancing transfer scheme if the dominant-strategy requirement is given up.

6 Conclusion

Our results shed light on the complex relationship between voting rules, voter participation and social welfare. The considerable variation in the design of voting rules in real-world institutions suggests that rules are—at least to some extent—tailored to different environments.[26] As far as the standard majority rule, the qual-

[26]Often, the same decision-making body uses different rules to decide on different issues. An example is the University of Mannheim which generally applies a simple majority rule (with a participation quorum of 50 percent) in its decision making bodies, but requires a two-third majority (in combination with a 60 percent approval quorum) for changes of its constitutions, and requires unanimous approval for some other decisions. Cf. https://www2.uni-mannheim.de/1/universitaet/leitung_organe/grundordnung/grundo_neu_ab_4_2015.pdf and https://www2.uni-mannheim.de/1/universitaet/partner_ehrungen/ehrungen/ehrenordnung /ehrenordnung.pdf).

ified majority rules, and the one-sided linear rules are concerned, this view is consistent with our results. But our main linearity result also has a normative angle: it suggests to reconsider the common practice of using non-linear quorum rules.

We contribute to an explanation of why so many institutions frequently rely on voluntary participation. It is obvious that compulsory participation can depress social welfare when voting costs are large. However, technological progress has significantly reduced the physical cost of participating in many collective decisions. Is compulsory participation welfare-enhancing when voting costs are small? We provide an argument for the answer "No": in a broad class of mechanisms that would be optimal with compulsory participation, voluntary participation leads to a higher welfare.

Two warnings are in place. First, our approach is not detail-free: the designer needs to know the informational and preference environment in order to determine the optimal linear voting rule. More research is needed to evaluate the performance of any particular linear rule in a broad class of environments (featuring, for instance, various types of asymmetry, correlation, or interdependence of values) as a basis for a practical recommendation. Second, our model is not primarily intended to represent large electorates such as in many public referenda, where other approaches may be more appropriate (e.g., Feddersen and Sandroni, 2006).

The technical methods that we have developed in this paper may prove useful in other settings as well. First, whenever a mechanism-design problem shares the feature that different mechanisms lead to the same equilibrium, properties of the optimal mechanism may be proven by maximizing welfare subject to a fixed equilibrium. Second, the perturbation method for analyzing equilibria when the participation cost is small is applicable to voting rules beyond the class of linear mechanisms that we consider.

7 Appendix A

Define the shortcuts

$$G^R = E[\tilde{v}\mathbf{1}_{\tilde{v}>0}] > 0, \quad G^S = E[-\tilde{v}\mathbf{1}_{\tilde{v}<0}] > 0. \tag{31}$$

Proof of Proposition 1. We first treat the cases with $E[\tilde{v}] > 0$ (the cases with $E[\tilde{v}] < 0$ are analogous). Following this, we show that full participation is not optimal. The proof of the linearity result for the unbiased case $E[\tilde{v}] = 0$ uses similar techniques and is relegated to Appendix B.

Consider an optimal mechanism-equilibrium pair $(M^*, \tau^{R*}, \tau^{S*})$. Define corresponding pivot probabilities Δ^{R*} and Δ^{S*} via (9).

Case $E[\tilde{v}] > 0$, $\tau^{R*} > 0$, $\tau^{S*} > 0$, and $\tau^{R*} + \tau^{S*} < 1$.
Here,

$$\Pr_{\tau^{R*},\tau^{S*}}(r,s) > 0 \text{ for all tallies } (r,s) \text{ with } r+s<n. \tag{32}$$

Define the (convex and non-empty) set of mechanisms

$$\begin{aligned}\mathcal{M} &= \{M \mid \Delta^{R*} = d^R(M,\tau^{R*},\tau^{S*}),\ \ \Delta^{S*} = d^S(M,\tau^{R*},\tau^{S*}),\\ &\qquad 0 \le M_{rs} \le 1 \quad \text{for all } (r,s)\}.\end{aligned}$$

By optimality, the mechanism M^* solves the following problem:

$$\text{(lin)} \quad \max_M \rho^{M,\tau^{R*},\tau^{S*}}(A) \quad \text{s.t. } M \in \mathcal{M}.$$

Problem (lin) is linear. Hence, the Kuhn-Tucker conditions are necessary, without any constraint qualification. Thus, there exist Lagrange multiplier μ^R, μ^S, and μ_{rs} [27] for all (r,s) such that

$$\mu_{rs} = \frac{\partial}{\partial M_{rs}}\left(\rho^{M,\tau^{R*},\tau^{S*}}(A) + \mu^R d^R(M,\tau^{R*},\tau^{S*}) + \mu^S d^S(M,\tau^{R*},\tau^{S*})\right), \tag{33}$$

where $\mu_{rs} \le 0$ if $M^*_{rs} < 1$ and $\mu_{rs} \ge 0$ if $M^*_{rs} > 0$ (complementary slackness). Put differently,

$$M^*_{rs} = \begin{cases} 1 & \text{if } \mu_{rs} > 0,\\ 0 & \text{if } \mu_{rs} < 0.\end{cases} \tag{34}$$

Consider a tally (r,s) with $r+s<n$. Rewriting (33) and dropping the index (τ^{R*},τ^{S*}) when writing multinomial probabilities,

$$\begin{aligned}\mu_{rs} &= \Pr(r,s)\left(1-\mu^R+\mu^S\right) + \Pr(r-1,s)\mu^R - \Pr(r,s-1)\mu^S\\ &\overset{(6)}{=} \frac{\Pr(r,s)}{(n-r-s)\tau^{R*}\tau^{S*}}\left(r\xi^R - s\xi^S - n\xi\right),\end{aligned} \tag{35}$$

where

$$\begin{aligned}\xi^R &= \mu^R\tau^{S*}(1-\tau^{R*}-\tau^{S*}) - (1-\mu^R+\mu^S)\tau^{R*}\tau^{S*}, &(36)\\ \xi &= -(1-\mu^R+\mu^S)\tau^{R*}\tau^{S*}, &(37)\\ \xi^S &= \mu^S\tau^{R*}(1-\tau^{R*}-\tau^{S*}) + (1-\mu^R+\mu^S)\tau^{R*}\tau^{S*}. &(38)\end{aligned}$$

[27] We use a single Lagrange multiplies for both constraints $0 \le M_{rs}$ and $M_{rs} \le 1$ because only one constraint can be binding.

Using (32), we conclude that (2) holds; it is straightforward to check that (2) also holds for tallies with $r + s = n$.

To show (1), observe first that $\xi^R \geq 0$ (otherwise (34) implies that $M^*_{r+1,s} \leq M^*_{rs}$ for all (r, s), implying $\Delta^{R*} = 0$ and hence $\tau^{R*} = 0$). Similarly, $\xi^S \geq 0$. Moreover, if $\xi = 0$, then $1 - \mu^R + \mu^S = 0$ by (37), implying $\mu^S \geq 0$ by (38). Thus, $\mu^R = 1 + \mu^S \geq 1$, thus $\xi^R > 0$ by (36). Hence, M^* is a linear mechanism.

Case $E[\tilde{v}] > 0$, $\tau^{R*} > 0$ *and* $\tau^{S*} = 0$. (The case $\tau^{R*} = 0$ and $\tau^{S*} > 0$ is analogous).

Note that $\tau^{R*} \leq F^R < 1$. Define the (convex and non-empty) set of mechanisms

$$\begin{aligned}\mathcal{M}^R \;=\; & \{M \mid \Delta^{R*} = d^R(M, \tau^{R*}, 0), \\ & \qquad M_{rs} = M_{r0},\;\; 0 \leq M_{rs} \leq 1 \quad \text{for all } (r, s)\}.\end{aligned}$$

For any $M \in \mathcal{M}^R$, the pair $(\Delta^{R*}, 0)$ is an equilibrium. Thus, by optimality, the mechanism $\hat{M}$, where $\hat{M}_{rs} = M^*_{r0}$, solves the following problem:

$$(\text{lin})^R \quad \max_M \rho^{M,\tau^{R*},0}(A) \quad \text{s.t.}\;\; M \in \mathcal{M}^R.$$

Problem (lin)^R is linear. Hence, there exist Lagrange multipliers μ^R and μ_r for all r such that

$$\mu_r \;=\; \frac{\partial}{\partial M_{r0}} \left(\rho^{M,\tau^{R*},0}(A) + \mu^R d^R(M, \tau^{R*}, 0) \right), \tag{39}$$

where $\mu_r \leq 0$ if $\hat{M}_{r0} < 1$ and $\mu_r \geq 0$ if $\hat{M}_{r0} > 0$ (complementary slackness). Put differently,

$$\hat{M}_{rs} \;=\; \begin{cases} 1 & \text{if } \mu_r > 0, \\ 0 & \text{if } \mu_r < 0. \end{cases}$$

Rewriting (39) yields

$$\mu_r \;=\; \frac{(n-1)!}{(n-r)!r!} (\tau^{R*})^{r-1} (1 - \tau^{R*})^{n-1-r} \left((n-r)\tau^{R*}(1 - \mu^R) + r(1 - \tau^{R*})\mu^R \right).$$

Thus, (2) holds with

$$\xi^R \;=\; (1 - \tau^{R*})\mu^R - \tau^{R*}(1 - \mu^R), \tag{40}$$
$$\xi \;=\; \tau^{R*}(1 - \mu^R), \tag{41}$$
$$\xi^S \;=\; 0.$$

To see (1), observe first that $\xi^R \geq 0$ (otherwise $\hat{M}_{r+1,s} \leq M^*_{rs}$ for all (r,s), implying $\Delta^{R*} = 0$ and hence $\tau^{R*} = 0$). Moreover, if $\xi = 0$, then $1 - \mu^R = 0$ by (41), implying $\xi^R > 0$ by (40). We conclude that $\hat{M}$ is a linear mechanism.

Finally, we show

$$\tau^{R*} + \tau^{S*} \quad < \quad 1.$$

Suppose otherwise. Then $\tau^{R*} = F^R$ and $\tau^{S*} = F^S$. In such an equilibrium, only the tallies (r,s) with $r+s=n$ occur with positive probability.

Let v_R be the smallest positive type in supp(F) and v_S the largest negative type. By assumption, participation is optimal for the types $v_R > 0$ and $v_S < 0$.

Define the (convex and non-empty) set of mechanisms with the property that full participation is an equilibrium:

$$\begin{aligned}\mathcal{M} \quad = \quad & \{M \mid d^R(M, F^R, F^S) \geq c/v_R,\ \ d^S(M, F^R, F^S) \geq -c/v_S, \\ & \quad 0 \leq M_{rs} \leq 1 \quad \text{for all } (r,s)\}.\end{aligned}$$

By optimality, using (12), the mechanism M^* solves the following problem:

$$\text{(lin)} \quad \max_{M \in \mathcal{M}} E[\tilde{v}]\rho^{M,F^R,F^S}(A) + G^R d^R(M, F^R, F^S) + G^S d^S(M, F^R, F^S).$$

Thus, there exist multipliers $\mu^R \geq 0$, $\mu^S \geq 0$, and

$$\begin{aligned}\mu_{rs} \quad = \quad & \frac{\partial}{\partial M_{rs}} \Big(E[\tilde{v}]\rho^{M,\tau^{R*},\tau^{S*}}(A) + (G^R + \mu^R) d^R(M, \tau^{R*}, \tau^{S*}) \\ & \qquad + (G^S + \mu^S) d^S(M, \tau^{R*}, \tau^{S*}) \Big) \Big|_{M=M^*}, \end{aligned} \tag{42}$$

where $\mu_{rs} \leq 0$ if $M^*_{rs} < 1$ and $\mu_{rs} \geq 0$ if $M^*_{rs} > 0$ (complementary slackness). Moreover, (34) holds. Note that

$$\mu_{rs} = 0 \quad \text{for all tallies } (r,s) \text{ with } r+s \leq n-2 \tag{43}$$

because $\Pr(r,s) = 0$. Consider tallies (r,s) with $r+s = n-1$. Then $\mu_{rs} = \Pr(r,s)(-\mu^R + \mu^S)$. This implies $\mu^R = \mu^S \overset{\text{def}}{=} \mu$ (if, say, "$>$", then $M^*_{rs} = 1$ by (34), implying $d^R(M^*, F^R, F^S) \leq 0$, implying $\Delta^{R*} = 0$ by (7), thus $\tau^{R*} = 0$). Thus,

$$\mu_{rs} = 0 \quad \text{for all tallies } (r,s) \text{ with } r+s = n-1. \tag{44}$$

Now consider tallies (r,s) with $r+s=n$. Then

$$\mu_{rs} \quad = \quad \Pr(r-1,s)(G^R + \mu) - \Pr(r,s-1)(G^S + \mu).$$

Rearranging and using (34) yields that $M^*_{r,s} = 1$ if $r > r^*$ and $= 0$ if $r < r^*$, for some r^*.

Because in the linear problem (lin) the Lagrange conditions are sufficient, all mechanisms in the following non-empty set are optimal:

$$\begin{aligned}\overline{\mathcal{M}} &= \{M \mid d^R(M, F^R, F^S) = \Delta^{R*},\ d^S(M, F^R, F^S) = \Delta^{S*},\\ &\quad \forall (r,s):\ 0 \le M_{rs} \le 1,\ \text{if } r+s=n \text{ then } M_{rs} = M^*_{rs}\}.\end{aligned}$$

For any $M \in \overline{\mathcal{M}}$, let $\overline{r}(M)$ be maximal with $M_{\overline{r},n-1-\overline{r}} < 1$ und $\underline{r}(M)$ be minimal with $M_{\underline{r},n-1-\underline{r}} > 0$. Observe that $\overline{r}(M) \ge r^* - 1$ because otherwise $d^R(M, F^R, F^S) \le 0$. Similarly, $\underline{r}(M) \le r^*$. Choose

$$\hat{M} \in \arg\max_{M \in \overline{\mathcal{M}}} \underline{r}(M) - \overline{r}(M).$$

We claim that

$$\underline{r}(\hat{M}) - \overline{r}(\hat{M}) \ge 0. \tag{45}$$

To prove this, consider $M \in \overline{\mathcal{M}}$ such that $\underline{r}(M) - \overline{r}(M) \le -1$. Then $\overline{r} \ge \underline{r} + 1$. Beginning with $M' = M$, lower $M'_{\underline{r},n-1-\underline{r}}$ and increase $M'_{\overline{r},n-1-\overline{r}}$ while keeping, for all rules M' along the resulting path, the following expression constant:

$$\Pr(\underline{r}, n-1-\underline{r})M'_{\underline{r},n-1-\underline{r}} - \Pr(\overline{r}, n-1-\overline{r})M'_{\overline{r},n-1-\overline{r}}.$$

Thus, $M' \in \overline{\mathcal{M}}$ along the path. The path ends with a rule M' such that $M'_{\underline{r},n-1-\underline{r}} = 0$ or $M'_{\overline{r},n-1-\overline{r}} = 1$. Thus, $\underline{r}(M') - \overline{r}(M') > \underline{r}(M) - \overline{r}(M)$. This proves (45).

From (45), $\hat{M}_{r^*-1,n-r^*} = 0$ or $\hat{M}_{r^*,n-1-r^*} = 1$. Consider the first case (the second is treated analogously). Because $d^S(\hat{M}, F^R, F^S) > 0$,

$$\hat{M}_{r^*,n-1-r^*} > \hat{M}_{r^*,n-r^*}. \tag{46}$$

Define the R-one-sided rule $\hat{M}^R$ via $\hat{M}^R_{rs} = \hat{M}_{r,n-r}$. Then $(F^R, 0)$ is an equilibrium in $\hat{M}^R$ because $d^R(\hat{M}^R, F^R, 0) > d^R(\hat{M}, F^R, F^S)$ by (46). Moreover, $W(\hat{M}^R, F^R, 0) = W(\hat{M}, F^R, F^S) + G(0)c$, showing that $(\hat{M}, F^R, F^S)$ is not optimal. Thus, (42) follows. This completes the proof of Proposition 1.

Proof of Remark 1. By the mean-value theorem, we can represent the relevant conditional expectations below (15) with weighted types close to v^R and v^S, respectively:

$$\begin{aligned}&\exists\, \nu^R \in [v^R - \epsilon, v^R + \epsilon]:\ \eta^R(p^R) = \nu^R g(\nu^R),\\ &\exists\, \nu^S \in [v^S - \epsilon, v^S + \epsilon]:\ \eta^A(p^R, 0) = \nu^S g(\nu^S).\end{aligned}$$

Thus, using (15) we find the limit welfare of implementing R given any election result induced by full one-sided participation:

$$\omega_{t,0}(p^R, 0) \quad \to \quad w(t, n-t) \quad \text{for all } t \text{ as } \epsilon \to 0, \tag{47}$$

where here and below the limit applies uniformly across all ϵ-approximations F.

Choose $\epsilon > 0$ so small that, for all $\nu^R \in [v^R - \epsilon, v^R + \epsilon]$, $\nu^S, \nu^A \in [v^S - \epsilon, v^S + \epsilon]$, $t = 0, \ldots, n$ and $s = 0, \ldots, n-t$, the expression

$$tg(\nu^R)\nu^R + sg(\nu^S)\nu^S + (n-t-s)g(\nu^A)\nu^A \text{ has the same sign as } w(t, n-t). \tag{48}$$

Define $t^* = \min\{r | w(r, n-r) > 0\}$ and $\underline{M}$ via $\underline{M}_{rs} = 1_{r \geq t^*}$.

Assume $c < v^R \, \mathrm{Pr}_{p^R,0}(t^* - 1, 0)$. Thus, for all

$$\epsilon < v^R - \frac{c}{\mathrm{Pr}_{p^R,0}(t^* - 1, 0)},$$

the pair $(\tau^R, \tau^S) = (p^R, 0)$ is an equilibrium in the mechanism $\underline{M}$ because $d^R(\underline{M}, p^R, 0) = \mathrm{Pr}_{p^R,0}(t^* - 1, 0))$. Using (14), the resulting welfare is

$$W(\underline{M}, p^R, 0) \quad = \quad \frac{1}{n} \sum_{t=0}^{n} \binom{n}{t} (p^R)^t (p^S)^{n-t} \max\{0, \omega_{t,0}(p^R, 0)\} - \alpha^R(p^R)c. \tag{49}$$

Thus, using (47),

$$\lim_{c \to 0, \; \epsilon \to 0} W(\underline{M}, p^R, 0) - \max\{0, E[\tilde{v}]\} > 0.$$

Let $c > 0$ and $\epsilon > 0$ be so small that

$$W(\underline{M}, p^R, 0) \quad > \quad \max\{0, E[\tilde{v}]\} + 2\epsilon. \tag{50}$$

Consider any optimal $m^* = (M^*, \tau^{R*}, \tau^{S*})$ with corresponding pivot probabilities $\Delta^{R*} = d^R(m^*)$ and $\Delta^{S*} = d^S(m^*)$.

Suppose that both types $v^R - \epsilon$ and $v^S + \epsilon$ find it optimal to abstain. Then

$$(v^R - \epsilon)\Delta^{R*} \leq c, \quad (-v^S - \epsilon)\Delta^{S*} \leq c.$$

Thus, using (12) and the shortcuts $\alpha^{R*} = \alpha^R(\tau^{R*})$ and $\alpha^{S*} = \alpha^S(\tau^{S*})$,

$$\begin{aligned} W(m^*) \quad & \leq \quad E[\tilde{v}]\rho^{m^*}(A) + \alpha^{R*}((v^R + \epsilon)\Delta^{R*} - c) + \alpha^{S*}((-v^S + \epsilon)\Delta^{S*} - c) \\ & \leq \quad \max\{0, E[\tilde{v}]\} + 2\epsilon, \end{aligned}$$

which by (50) contradicts the optimality of m^*.

Thus, participation is optimal either for all types in $[v^R - \epsilon, v^R + \epsilon]$ (implying $\tau^{R*} = p^R$) or for all types in $[v^S - \epsilon, v^S + \epsilon]$ (implying $\tau^{S*} = p^S$).

Consider the case $\tau^{R*} = p^R$. Thus $\alpha^{R*} = \alpha^R(p^R)$. We can write the welfare (14) by first summing over the number t of R-agents (all of which will participate), and then summing over the number s of S-voters:

$$\begin{aligned} W(m^*) &= \frac{1}{n}\sum_{t=0}^{n}\binom{n}{t}(p^R)^t(p^S)^{n-t}\sum_{s=0}^{n-t}\binom{n-t}{s}(\frac{\tau^{S*}}{p^S})^s(1-\frac{\tau^{S*}}{p^S})^{n-t-s}M^*_{ts}\omega_{ts}(p^R,\tau^{S*}) \\ &\quad -(\alpha^R(p^R)+\alpha^{S*})c, \end{aligned} \tag{51}$$

where

$$\omega_{ts}(p^R,\tau^{S*}) = t\eta^R(p^R) + s\int_{v^S-\epsilon}^{F^{-1}(\tau^{S*})}\frac{v\mathrm{d}G(v)}{\tau^{S*}} + (n-t-s)\int_{F^{-1}(\tau^{S*})}^{v^S+\epsilon}\frac{v\mathrm{d}G(v)}{p^S-\tau^{S*}}$$

is the welfare from implementing R conditional on the tally (t,s). Using the mean-value theorem, there exist $\nu^R \in [v^R - \epsilon, v^R + \epsilon]$ and $\nu^S, \nu^A \in [v^S - \epsilon, v^S + \epsilon]$ such that

$$\omega_{ts}(p^R,\tau^{S*}) = tg(\nu^R)\nu^R + sg(\nu^S)\nu^S + (n-t-s)g(\nu^A)\nu^A.$$

Thus, using (48),

$$\omega_{ts}(p^R,\tau^{S*}) \begin{cases} > 0 \text{ if } t \geq t^*, \\ < 0 \text{ if } t < t^*. \end{cases}$$

Hence, by optimality of M^*,

$$M^* = \underline{M}. \tag{52}$$

By a basic property of the binomial distribution,

$$\sum_{s=0}^{n-t}\binom{n-t}{s}(\frac{\tau^{S*}}{p^S})^s(1-\frac{\tau^{S*}}{p^S})^{n-t-s}\cdot s = \frac{\tau^{S*}}{p^S}(n-t).$$

Thus, for any $t \geq t^*$,

$$\begin{aligned} &\sum_{s=0}^{n-t}\binom{n-t}{s}(\frac{\tau^{S*}}{p^S})^s(1-\frac{\tau^{S*}}{p^S})^{n-t-s}\omega_{ts}(p^R,\tau^{S*}) \\ &= t\eta^R(p^R) + \frac{\tau^{S*}}{p^S}(n-t)\int_{v^S-\epsilon}^{F^{-1}(\tau^{S*})}\frac{v\mathrm{d}G(v)}{\tau^{S*}} + (1-\frac{\tau^{S*}}{p^S})(n-t)\int_{F^{-1}(\tau^{S*})}^{v^S+\epsilon}\frac{v\mathrm{d}G(v)}{p^S-\tau^{S*}} \\ &= t\eta^R(p^R) - (n-t)\eta^S(p^S) \\ &= \omega_{t,0}(p^R, 0). \end{aligned}$$

Hence, using (51) and (52),

$$\begin{aligned} W(M^*, p^R, \tau^{S*}) &= \frac{1}{n}\sum_{t=t^*}^{n}\binom{n}{t}(p^R)^t(p^S)^{n-t}\left(\sum_{s=0}^{n-t}\binom{n-t}{s}(\frac{\tau^{S*}}{p^S})^s(1-\frac{\tau^{S*}}{p^S})^{n-t-s}\right. \\ &\qquad \left.\cdot\,\omega_{t,s}(p^R,\tau^{S*})\right) - (\alpha^R(p^R)+\alpha^{S*})c \\ &\leq \frac{1}{n}\sum_{t=t^*}^{n}\binom{n}{t}(p^R)^t(p^S)^{n-t}\omega_{t,0}(p^R,0) - \alpha^R(p^R)c \\ &\overset{(49)}{=} W(\underline{M}, p^R, 0). \end{aligned}$$

We conclude that the mechanism-equilibrium pair $(\underline{M}, p^R, 0)$ is the unique optimum with $\tau^{R*} = p^R$. Analogously, the mechanism-equilibrium pair $(\mathbf{1}_{s\leq n-t^*}, 0, p^S)$ is the unique optimum with $\tau^{S*} = p^S$. Comparing the welfare levels achieved with these two candidates, the mechanism-equilibrium pair with smaller g-weighted participation cost is optimal. This completes the proof of Remark 1.

The proof of Remark 2 relies on auxiliary notation and on a lemma. Given any three-point distribution $\hat{F} = \hat{F}^{v^0}$, consider the probability distribution that describes an agent's g-weighted valuation:

$$\begin{pmatrix} \hat{v}^S & \hat{v}^0 & \hat{v}^R \\ p^S & p^0 & p^R \end{pmatrix} \overset{\text{def}}{=} \begin{pmatrix} g(v^S)v^S & g(v^0)v^0 & g(v^R)v^R \\ p^S & p^0 & p^R \end{pmatrix},$$

where we have to distinguish the realizations $\hat{v}^S$ and $\hat{v}^0$ even when they happen to be of equal size. Let $(\hat{v}_1, \ldots, \hat{v}_n)$ denote an i.i.d. vector of random variables with the above distribution.

Denote the set of feasible mechanism-participation-rates combinations by

$$\begin{aligned} \mathcal{F} &= \{(M, \tau^R, \tau^S) | \forall (r,s):\ 0 \leq M_{rs} \leq 1,\ \tau^R + \tau^S \leq 1, \\ &\qquad 0 \leq \tau^R \leq p^R + p^0,\ 0 \leq \tau^S \leq p^S + p^0\}. \end{aligned}$$

(In this definition we allow $\tau^R > p^R$ even when $v^0 < 0$ so that the definition is the same as when $v^0 = 0$.)

Consider any $(M, \tau^R, \tau^S) \in \mathcal{F}$. Let $(\tilde{a}_1, \ldots, \tilde{a}_n)$ denote an i.i.d. random vector that describes each agent i's action R, S, or A such that the participation pair (τ^R, τ^S) arises if agent i's valuation is chosen according to $\hat{v}_i$. That is, $\Pr[\tilde{a}_i = R,\ \hat{v}_i = \hat{v}^R] = \min\{p^R, \tau^R\}$, $\Pr[\tilde{a}_i = S,\ \hat{v}_i = \hat{v}^R] = 0$, $\Pr[\tilde{a}_i = R,\ \hat{v}_i = \hat{v}^0] = \max\{p^0, \tau^R - p^R\}$, and so on.

Let $W_{\hat{F},0}(M,\tau^R,\tau^S)$ denote the welfare, ignoring the participation costs:

$$W_{\hat{F},0}(M,\tau^R,\tau^S) \;=\; \frac{1}{n}E\left[\sum_{i=1}^{n}\hat{v}_i M_{\tilde{r},\tilde{s}}\right], \qquad (53)$$

where

$$\tilde{r}=|\{j|\tilde{a}_j=R\}| \quad \text{and} \quad \tilde{s}=|\{j|\tilde{a}_j=S\}|.$$

By construction, analogously to (14),

$$\begin{aligned} W_{\hat{F},0}(M,\tau^R,\tau^S) \;&=\; \frac{1}{n}\sum_{r+s\leq n}\binom{n}{r\ s}(\tau^R)^r(\tau^S)^s(1-\tau^R-\tau^S)^{n-r-s}M_{rs} \\ &\qquad\cdot\left(r\hat{\eta}^R-s\hat{\eta}^S+(n-r-s)\hat{\eta}^A\right), \end{aligned}$$

where

$$\begin{aligned} \hat{\eta}^R &= E[\hat{v}_i\mathbf{1}_{\tilde{a}_i=R}]/\tau^R && \text{if } \tau^R>0, \\ \hat{\eta}^S &= E[\hat{v}_i\mathbf{1}_{\tilde{a}_i=S}]/\tau^S && \text{if } \tau^S>0, \\ \hat{\eta}^A &= E[\hat{v}_i\mathbf{1}_{\tilde{a}_i=A}]/(1-\tau^R-\tau^S) && \text{if } \tau^R+\tau^S<1. \end{aligned}$$

Observe that $W_{\hat{F},0}$ is continuous in (M,τ^R,τ^S) and in v^0.

Given any $\hat{\delta}>0$, we can choose $\epsilon>0$ so small that

$$\max_{v\in[v^S-\epsilon,v^S+\epsilon]}|vg(v)-\hat{v}^S|<\hat{\delta}, \quad \max_{v\in[v^R-\epsilon,v^R+\epsilon]}|vg(v)-\hat{v}^R|<\hat{\delta},$$

and, for all $v^0\in(v^S,0]$ around which g is defined,

$$\max_{v\in[v^0-\epsilon,v^0+\epsilon]}|vg(v)-\hat{v}^0|<\hat{\delta}.$$

Consider any ϵ-approximation F of $\hat{F}$. By construction, using the functions defined below (15),

$$\begin{aligned} |\eta^R(\tau^R)-\hat{\eta}^R| &\leq \hat{\delta} && \text{if } \tau^R>0, \\ |\eta^S(\tau^S)-\hat{\eta}^S| &\leq \hat{\delta} && \text{if } \tau^S>0, \\ |\eta^A(\tau^R,\tau^S)-\hat{\eta}^A| &\leq \hat{\delta} && \text{if } \tau^R+\tau^S<1. \end{aligned}$$

Thus, making the direct dependence of W on F and c explicit with lower indices,

$$|W_{\hat{F},0}(M,\tau^R,\tau^S)-W_{F,c}(M,\tau^R,\tau^S)+(\alpha^R(\tau^R)+\alpha^S(\tau^S))c| \;\leq\; \hat{\delta}. \qquad (54)$$

In particular, as $(c, \epsilon) \to 0$,

$$\sup_{\substack{(M, \tau^R, \tau^S) \in \mathcal{F}, \text{ all } v^0, \\ F \text{ an } \epsilon\text{-approximation of } \hat{F}^{v^0}}} \left| W_{\hat{F}^{v^0},0}(M, \tau^R, \tau^S) - W_{F,c}(M, \tau^R, \tau^S) \right| \quad \to \quad 0. \tag{55}$$

Define the mechanism $\overline{M}$ that, combined with participation of the types around v^R and v^S and abstention of the types around v^0, implements the optimal alternative conditional on any tally (r, s):

$$\overline{M}_{rs} \quad = \quad \mathbf{1}_{w(r,s)>0} = \mathbf{1}_{r\hat{v}^R + s\hat{v}^S + (n-r-s)\hat{v}^0 > 0}. \tag{56}$$

Using the definition (15),

$$|\omega_{rs}(p^R, p^S) - w(r, s)| \quad \leq \quad \hat{\delta}. \tag{57}$$

The following lemma determines the "first-best". The crucial aspect of this result is that we fully characterize the solution set.

The first condition in (59) excludes extreme cases in which a single agent with type v^S outweighs all others' preferences in the electorate. Without the second condition, it would be unnecessary to distinguish types v^S and v^0 to determine the welfare-maximizing alternative—counting the number of agents with type v^R would suffice. This condition also excludes extreme cases in which a single agent with type v^R outweighs all others' preferences in the electorate. Given (21), conditions (58) and (59) are implied by (19) and (20). Note that Lemma 3 includes cases with $v^0 = 0$.

Lemma 3. *Consider a three-point distribution $\hat{F}$ such that*

$$w(r, s) \neq 0 \quad \textit{for all } (r, s) \neq (0, 0) \tag{58}$$

and

$$\begin{aligned} &w(n-1, 1) > 0, \textit{ and} \\ &\exists \hat{r} \in \{1, \ldots, n-2\} : \; w(\hat{r}, n-\hat{r}) < 0, \quad w(\hat{r}, 0) > 0. \end{aligned} \tag{59}$$

The set of solutions to the problem

$$(*) \qquad \max_{(M,\, \tau^R,\, \tau^S) \in \mathcal{F}} W_{\hat{F},0}(M, \tau^R, \tau^S)$$

is given by the singleton $(\overline{M}, p^R, p^S)$ if $\hat{v}^0 < 0$, and is given by the set $\overline{\mathcal{M}} \times \{p^R\} \times \{p^S\}$ if $\hat{v}^0 = 0$.

Proof of Lemma 3. Let (M, τ^R, τ^S) denote a maximizer of $(*)$. Using (53),

$$W_{\hat{F},0}(M, \tau^R, \tau^S) \leq \frac{1}{n} E\left[\max\{0, \sum_{i=1}^{n} \hat{v}_i\}\right] = W_{\hat{F},0}(\overline{M}, p^R, p^S).$$

Because (M, τ^R, τ^S) is optimal, the "$\leq$" is in fact an "$=$". Hence,[28]

$$\text{if } \sum_{i=1}^{n} \hat{v}_i > 0 \text{ then } M_{\tilde{r},\tilde{s}} = 1, \tag{60}$$

$$\text{if } \sum_{i=1}^{n} \hat{v}_i < 0 \text{ then } M_{\tilde{r},\tilde{s}} = 0. \tag{61}$$

First of all, this implies

$$\tau^R \geq p^R \ \text{ or } \ \tau^S \geq p^S. \tag{62}$$

Suppose not. Then $\Pr[\hat{v}_1 = \cdots = \hat{v}_n = \hat{v}^R,\ \tilde{a}_1 = \cdots = \tilde{a}_n = A] > 0$, implying $M_{0,0} = 1$ by (60). Similarly, $\Pr[\hat{v}_1 = \cdots = \hat{v}_n = \hat{v}^S,\ \tilde{a}_1 = \cdots = \tilde{a}_n = A] > 0$, implying $M_{0,0} = 0$ by (61), a contradiction. Next,

$$\text{if } \tau^R < p^R \text{ then } \tau^S = 1 - p^R. \tag{63}$$

Suppose not. Then, using (62),

$$\Pr[\hat{v}_1 = \hat{v}^S,\ \hat{v}_2 = \cdots = \hat{v}_n = \hat{v}^R,\ \tilde{a}_1 = S,\ \tilde{a}_2 = \cdots = \tilde{a}_n = A] \quad > \quad 0,$$

implying $M_{0,1} = 1$ by (59) and (60). On the other hand, using $\tau^S < 1 - p^R$,

$$\Pr[\hat{v}_1 = \hat{v}^S,\ \hat{v}_2 = \cdots = \hat{v}_n = \hat{v}^0,\ \tilde{a}_1 = S,\ \tilde{a}_2 = \cdots = \tilde{a}_n = A] \quad > \quad 0.$$

implying $M_{0,1} = 0$ by (61), a contradiction.

Next,

$$\tau^R \geq p^R \tag{64}$$

Suppose not. Then, using (63) and $\hat{r}$ from (59),

$$\begin{aligned}\Pr[&\hat{v}_1 = \cdots = \hat{v}_{\hat{r}} = \hat{v}^R,\ \hat{v}_{\hat{r}+1} = \cdots = \hat{v}_n = \hat{v}^S,\\ &\tilde{a}_1 = \cdots = \tilde{a}_{\hat{r}} = A,\ \tilde{a}_{\hat{r}+1} = \cdots = \tilde{a}_n = S] > 0,\end{aligned}$$

implying $M_{0,n-\hat{r}} = 0$ by (61). On the other hand,

$$\begin{aligned}\Pr[&\hat{v}_1 = \cdots = \hat{v}_{\hat{r}} = \hat{v}^R,\ \hat{v}_{\hat{r}+1} = \cdots = \hat{v}_n = \hat{v}^0,\\ &\tilde{a}_1 = \cdots = \tilde{a}_{\hat{r}} = A,\ \tilde{a}_{\hat{r}+1} = \cdots = \tilde{a}_n = S] > 0,\end{aligned}$$

[28] Read as "the event described after 'if' implies the event described after 'then'."

implying $M_{0,n-\hat{r}} = 1$ by (60), a contradiction.

Next,

$$\text{if } \tau^S < p^S \text{ then } \tau^R = 1 - p^S. \tag{65}$$

Suppose not. Then

$$\begin{aligned}\Pr[&\hat{v}_1 = \cdots = \hat{v}_{\hat{r}} = \hat{v}^R,\ \hat{v}_{\hat{r}+1} = \cdots = \hat{v}_n = \hat{v}^S,\\ &\tilde{a}_1 = \cdots = \tilde{a}_{\hat{r}} = R,\ \tilde{a}_{\hat{r}+1} = \cdots = \tilde{a}_n = A] > 0,\end{aligned}$$

implying $M_{\hat{r},0} = 0$ by (61). On the other hand,

$$\begin{aligned}\Pr[&\hat{v}_1 = \cdots = \hat{v}_{\hat{r}} = \hat{v}^R,\ \hat{v}_{\hat{r}+1} = \cdots = \hat{v}_n = \hat{v}^0,\\ &\tilde{a}_1 = \cdots = \tilde{a}_{\hat{r}} = R,\ \tilde{a}_{\hat{r}+1} = \cdots = \tilde{a}_n = A] > 0,\end{aligned}$$

implying $M_{\hat{r},0} = 1$ by (60), a contradiction.

Next

$$\tau^S \geq p^S \tag{66}$$

Suppose not. Then, using (65),

$$\Pr[\hat{v}_1 = \hat{v}^S,\ \hat{v}_2 = \cdots = \hat{v}_n = \hat{v}^0,\ \tilde{a}_1 = A,\ \tilde{a}_2 = \cdots = \tilde{a}_n = R] \quad > \quad 0,$$

implying $M_{n-1,0} = 0$ by (61). On the other hand,

$$\Pr[\hat{v}_1 = \hat{v}^S,\ \hat{v}_2 = \cdots = \hat{v}_n = \hat{v}^R,\ \tilde{a}_1 = A,\ \tilde{a}_2 = \cdots = \tilde{a}_n = R] \quad > \quad 0,$$

implying $M_{n-1,0} = 1$ by (59) and (60), a contradiction.

Now we show that

$$\tau^R = p^R \tag{67}$$

Suppose not. Then, using (64) and (66),

$$\Pr[\hat{v}_1 = \hat{v}^S,\ \hat{v}_2 = \cdots = \hat{v}_n = \hat{v}^0,\ \tilde{a}_1 = S,\ \tilde{a}_2 = \cdots = \tilde{a}_n = R] \quad > \quad 0,$$

implying $M_{1,n-1} = 0$ by (61). On the other hand,

$$\Pr[\hat{v}_1 = \hat{v}^S,\ \hat{v}_2 = \cdots = \hat{v}_n = \hat{v}^R,\ \tilde{a}_1 = S,\ \tilde{a}_2 = \cdots = \tilde{a}_n = R] \quad > \quad 0,$$

implying $M_{1,n-1} = 1$ by (59) and (60), a contradiction.

Finally, we show that

$$\tau^S = p^S \tag{68}$$

Suppose not. Then, using (66) and (67),

$$\Pr[\hat{v}_1 = \cdots = \hat{v}_{\hat{r}} = \hat{v}^R,\ \hat{v}_{\hat{r}+1} = \cdots = \hat{v}_n = \hat{v}^S,\\ \tilde{a}_1 = \cdots = \tilde{a}_{\hat{r}} = R,\ \tilde{a}_{\hat{r}+1} = \cdots = \tilde{a}_n = S] > 0,$$

implying $M_{\hat{r},n-\hat{r}} = 0$ by (61). On the other hand,

$$\Pr[\hat{v}_1 = \cdots = \hat{v}_{\hat{r}} = \hat{v}^R,\ \hat{v}_{\hat{r}+1} = \cdots = \hat{v}_n = \hat{v}^0,\\ \tilde{a}_1 = \cdots = \tilde{a}_{\hat{r}} = R,\ \tilde{a}_{\hat{r}+1} = \cdots = \tilde{a}_n = S] > 0,$$

implying $M_{\hat{r},n-\hat{r}} = 1$ by (60), a contradiction.

From (67) and (68), $\tilde{r} = |\{j|\hat{v}_j = \hat{v}^R\}|$ and $\tilde{s} = |\{j|\hat{v}_j = \hat{v}^S\}|$.

Consider any (r,s) and a realization of $(\hat{v}_1,\ldots,\hat{v}_n)$ such that $\tilde{r} = r$ and $\tilde{s} = s$. Thus, $\sum_{i=1}^n \hat{v}_i = r\hat{v}^R + s\hat{v}^S + (n-r-s)\hat{v}^0$. By (58), (i) $\sum_{i=1}^n \hat{v}_i > 0$ or (ii) $\sum_{i=1}^n \hat{v}_i < 0$ if $(r,s) \neq (0,0)$. In the case (i), $M_{rs} = 1$ by (60); in the case (ii), $M_{rs} = 0$ by (61). Moreover, if $\hat{v}^0 < 0$, then $w(0,0) < 0$ so that $M = \overline{M}$. But if $\hat{v}^0 = 0$ then the value of the objective $W_{\hat{F},0}(M,p^R,p^S)$ is independent of $M_{0,0}$. This completes the proof of Lemma 3.

Proof of Remark 2. Assume (21). Moreover, using that (57) holds for any $\hat{\delta}$, we can assume ϵ is so small that, for all $v^0 < 0$ around which g is defined and all (r,s),

$$\omega_{rs}(p^R,p^S) \text{ has the same sign as } w^{v^0}(r,s). \tag{69}$$

Let $W^{**}(v^0)$ denote the maximum value of problem $(*)$ if $\hat{F} = \hat{F}^{v^0}$. Using the definition (22), denote

$$\overline{\Delta}^R = \min_{M\in\overline{\mathcal{M}}} d^R(M,p^R,p^S) > 0 \ \text{ and } \ \overline{\Delta}^S = \min_{M\in\overline{\mathcal{M}}} d^S(M,p^R,p^S) > 0.$$

Let

$$\bar{c} \ = \ \frac{1}{2}\min\{v^R\overline{\Delta}^R, -v^S\overline{\Delta}^S\}.$$

By continuity, there exist $\overline{v}^0(c) > 0$, $\overline{\epsilon}(c,v^0) > 0$ and an open neighborhood $\mathcal{N}$ of the set $\overline{\mathcal{M}} \times \{p^R\} \times \{p^S\}$ such that, for all $c < \bar{c}$, $v^0 < \overline{v}^0(c)$, $\epsilon < \overline{\epsilon}(c,v^0)$ and $(M,\tau^R,\tau^S) \in \mathcal{N} \cap \mathcal{F}$:

$$(v^R - \epsilon)d^R(M,\tau^R,\tau^S) > c, \quad (-v^S - \epsilon)d^S(M,\tau^R,\tau^S) > c,\\ -v^0 + \epsilon < c. \tag{70}$$

In particular, then (p^R, p^S) is the unique equilibrium of M among all participation pairs (τ^R, τ^S) with $(M, \tau^R, \tau^S) \in \mathcal{N} \cap \mathcal{F}$.

Next, we show that there exists $\delta > 0$, $\overline{c}' > 0$, $\overline{v^0}' < 0$, and $\overline{\epsilon}' > 0$ such that (71) holds for all $c < \overline{c}'$, $v^0 > \overline{v^0}'$, $\epsilon < \overline{\epsilon}'$, any ϵ-approximation F of $\hat{F}^{v^0}$, and any $(M, \tau^R, \tau^S) \in \mathcal{F}$:

$$\text{if } W_{F,c}(M, \tau^R, \tau^S) > W^{**}(v^0) - \delta, \text{ then } (M, \tau^R, \tau^S) \in \mathcal{N}. \qquad (71)$$

To see why, suppose (71) fails. Then there exist sequences $\delta_j \to 0$, $v_j^0 \to 0$, $c_j \to 0$, $\epsilon_j \to 0$, a sequence F_j, where F_j is an ϵ_j-approximation of $\hat{F}^{v_j^0}$, and a sequence $(M_j, \tau_j^R, \tau_j^S) \in \mathcal{F}$ such that $W_{F_j,c_j}(M_j, \tau_j^R, \tau_j^S) > W^{**}(v_j^0) - \delta_j$ and $(M_j, \tau_j^R, \tau_j^S) \notin \mathcal{N}$. By Bolzano-Weierstraß, there exists a limit point $(\hat{M}, \hat{\tau}^R, \hat{\tau}^S) \notin \mathcal{N}$ that yields by (55) the limit welfare $W_{\hat{F}^0,0}(\hat{M}, \hat{\tau}^R, \hat{\tau}^S) \geq W^{**}(0)$. Moreover, $(\hat{M}, \hat{\tau}^R, \hat{\tau}^S) \notin \overline{\mathcal{M}}$, contradicting Lemma 3.

Now, w.l.o.g., let $\overline{c}' < \delta/2$ and $\overline{\epsilon}'$ so small that (54) applies with $\hat{\delta} = \delta/2$ if $\epsilon < \overline{\epsilon}'$.

Thus, for all $c < \overline{c}'$, $\epsilon < \overline{\epsilon}'$, $v^0 > \overline{v^0}'$, and any ϵ-approximation F of $\hat{F}^{v^0}$,

$$W_{F,c}(\overline{M}, p^R, p^S) \geq W_{\hat{F}^{v^0},0}(\overline{M}, p^R, p^S) - \hat{\delta} - c > W^{**}(v^0) - \delta. \qquad (72)$$

Now consider any $c < \min\{\overline{c}, \overline{c}'\}$, $v^0 > \max\{\overline{v}^0(c), \overline{v^0}'\}$, $\epsilon < \min\{\overline{\epsilon}(c, v^0), \overline{\epsilon}'\}$ and any ϵ-approximation F of $\hat{F}^{v^0}$. Consider any optimal mechanism-equilibrium pair $(M^*, \tau^{R*}, \tau^{S*})$. Then

$$W_{F,c}(M^*, \tau^{R*}, \tau^{S*}) \geq W_{F,c}(\overline{M}, p^R, p^S).$$

Together with (71) and (72) this implies $(M^*, \tau^{R*}, \tau^{S*}) \in \mathcal{N}$.

Hence, $(\tau^{R*}, \tau^{S*}) = (p^R, p^S)$ by (70). Given these participation rates, the welfare is

$$\begin{aligned} W_{F,c}(M^*, p^R, p^S) &= \frac{1}{n} \sum_{r+s \leq n} \binom{n}{r\ s} (p^R)^r (p^S)^s (p^0)^{n-r-s} \omega_{rs}(p^R, p^S) M_{rs}^* \\ &\quad -(\alpha^R(p^R) + \alpha^S(p^S))c, \end{aligned}$$

implying that the unique best rule is $M^* = \overline{M}$ because of (69). This completes the proof of Remark 2.

Proof of Lemma 2. We will use the shortcut

$$d^* = d^R(M, F^R, F^S) = \Pr_{F^R,F^S}(t^* - 1, n - t^*).$$

For all (r, s) with $r + s \geq t^* + q - 2$,

$$M_{r,s} = \begin{cases} 1 & \text{if } r \geq t^*, \\ 1 & \text{if } r = t^* - 1,\ s = q, \\ 0 & \text{if } r \leq t^* - 1,\ s \geq q + 1. \end{cases}$$

Other tallies (r, s) will not be relevant in the computations below. The remaining free variable is $M_{t^*-2,q}$. At $m^* = (M, F^R, F^S)$,

$$d^R(m^*) = d^* > 0, \quad d^S(m^*) = 0, \quad F^{-1}(F^S) = 0, \quad F^{-1}(1 - F^R) = 0. \tag{73}$$

The function ϕ is continuously differentiable in a neighborhood $\mathcal{N}$ of $(0, m^*)$. For all $(c, m) = (c, M, \tau^R, \tau^S) \in \mathcal{N}$, the Jacobi matrix with respect to the first and third variables is

$$\phi_{\partial c, \partial \tau^R}(m) = \begin{pmatrix} -1 & -(F^{-1})'(1 - \tau^R) d^R(m) + F^{-1}(1 - \tau^R) d^R_{\tau^R}(m) \\ 1 & F^{-1}(\tau^S) d^S_{\tau^R}(m) \end{pmatrix}, \tag{74}$$

where lower indices denote partial derivatives. Using (73),

$$\det \phi_{\partial c, \partial \tau^R}(m^*) = \frac{d^*}{f(0)} > 0,$$

where $f = F'$ denotes the density of F. Also, (29) holds at $(0, m^*)$. Thus, by the implicit-function theorem (IFT), there exists an open neighborhood $\mathcal{U}$ of F^S and an open neighborhood $\mathcal{V}$ of $(0, F^R)$ and a (unique) function $(\check{c}, \check{\tau}^R) : \mathcal{U} \to \mathcal{V}$ such that

$$\{(\check{c}(\tau^S), \check{\tau}^R(\tau^S), \tau^S) \mid \tau^S \in \mathcal{U}\} = \{(c, \tau^R, \tau^S) \mid \tau^S \in \mathcal{U},\ (c, \tau^R) \in \mathcal{V},\ (29)\}. \tag{75}$$

The relevant domain of the implicit function $(\check{c}, \check{\tau}^R)$ is $\{\tau^S \in \mathcal{U} \mid \tau^S < F^S\}$. From (75),

$$\check{c}(F^S) = 0, \quad \check{\tau}^R(F^S) = F^R.$$

In order to compute the first-order derivatives of $(\check{c}, \check{\tau}^R)$, we invert (74),

$$\phi^{-1}_{\partial c, \partial \tau^R} = \frac{1}{\det \phi_{\partial c, \partial \tau^R}} \begin{pmatrix} F^{-1}(\tau^S) d^S_{\tau^R} & (F^{-1})'(1 - \tau^R) d^R - F^{-1}(1 - \tau^R) d^R_{\tau^R} \\ -1 & -1 \end{pmatrix},$$

and consider the derivative

$$\phi_{\partial \tau^S} = \begin{pmatrix} F^{-1}(1 - \tau^R) d^R_{\tau^S} \\ F^{-1}(\tau^S) d^S_{\tau^S} + (F^{-1})'(\tau^S) d^S \end{pmatrix}. \tag{76}$$

From the IFT,

$$\begin{pmatrix} \mathrm{d}\check{c}/\mathrm{d}\tau^S \\ \mathrm{d}\check{\tau}^R/\mathrm{d}\tau^S \end{pmatrix} = -\phi^{-1}_{\partial c, \partial \tau^R} \cdot \phi_{\partial \tau^S}, \tag{77}$$

where $\cdot$ denotes matrix-vector multiplication and where we have dropped the argument $(M, \check{\tau}^R(\tau^S), \tau^S)$ at which both matrix and vector are evaluated.

For any function of τ^S, denote the lth derivative evaluated at F^S by an upper index (l). Let

$$k = n - t^* + 1 - q \geq 2.$$

(In Figure 4, k equals the number of 0s below row $s = q$ in column $r = t^* - 1$.)

Using (77) recursively, we find (for details see Appendix B)

$$\check{c}^{(l)} = 0, \quad (\check{\tau}^R)^{(l)} = 0 \quad \text{for all } 1 \leq l < k \tag{78}$$

and

$$\begin{pmatrix} \check{c}^{(k)} \\ (\check{\tau}^R)^{(k)} \end{pmatrix} = \frac{(n-1)!}{(t^*-1)!q!} (F^R)^{t^*-1} (F^S)^q (-1)^{k-1} \cdot k \cdot \begin{pmatrix} -\frac{1}{f(0)} \\ \frac{1}{d^*} \end{pmatrix}. \tag{79}$$

From (78) and (79)

$$(-1)^l \check{c}^{(l)} \quad \begin{cases} = 0 & \text{if } l < k, \\ > 0 & \text{if } l = k. \end{cases} \tag{80}$$

Thus,[29] there exists $\bar{\epsilon} > 0$ such that $\check{c}'(\tau^S) < 0$ for $\tau^S \in (F^S - \bar{\epsilon}, F^S) \subseteq \mathcal{U}$ and, similarly, $(\check{\tau}^R)'(\tau^S) > 0$. In particular, defining $\bar{c} = \check{c}(F^S - \bar{\epsilon}) > 0$, the function $\check{c}$ has an inverse $\tilde{\tau}^S$ on $[0, \bar{c})$.

Defining $\tilde{\tau}^R(c) = \check{\tau}^R(\tilde{\tau}^S(c))$, the pair $(\tilde{\tau}^R(c), \tilde{\tau}^S(c))$ is an equilibrium at cost $c \in (0, \bar{c})$.

[29] A simple calculus lemma. If a k times continuously differentiable function $h(x)$ has $k-1$ derivatives equal to 0 at $x = 0$ and the kth derivative is strictly positive at 0, then the function is strictly increasing in a right-neighborhood of 0. Denoting derivatives with a lower index, $h_{k-1}(x) = \int_0^x h_k(y)\mathrm{d}y > 0$ by the fundamental theorem of calculus, hence $h_{k-2}(x) = \int_0^x h_{k-1}(y)\mathrm{d}y > 0$, and so on continuing inductively until we find $h_1(x) > 0$ for all x in a right neighborhood of 0, showing that h is strictly increasing. Now consider the function $\hat{h}(x) = h(-x)$ in a left-neighborhood of 0, then the kth derivative $\hat{h}_k(0) = (-1)^k h_k(0)$. So, if $(-1)^k \hat{h}_k(0) > 0$ then $\hat{h}$ is strictly decreasing in a left neighborhood of 0.

In order to prove (30), we begin by writing the welfare as a function

$$\begin{aligned}\check{W}(\tau^S) &= W(\check{c}(\tau^S), M, \check{\tau}^R(\tau^S), \tau^S) \\ &= E[\tilde{v}]\check{\rho}(\tau^S) + \int_{v\geq 0} \max\{v\check{d}^R(\tau^S) - \check{c}(\tau^S), 0\}\mathrm{d}G(v) \\ &\quad + \int_{v\leq 0} \max\{-v\check{d}^S(\tau^S) - \check{c}(\tau^S), 0\}\mathrm{d}G(v),\end{aligned}$$

where we have used the shortcuts

$$\begin{aligned}\check{\rho}(\tau^S) &= \rho^{M,\check{\tau}^R(\tau^S),\tau^S}(A) \\ \check{d}^R(\tau^S) &= d^R(M, \check{\tau}^R(\tau^S), \tau^S), \\ \check{d}^S(\tau^S) &= d^S(M, \check{\tau}^R(\tau^S), \tau^S).\end{aligned}$$

From (29),

$$\check{d}^R(\tau^S) = \frac{\check{c}(\tau^S)}{F^{-1}(1-\check{\tau}^R(\tau^S))}, \quad \check{d}^S(\tau^S) = \frac{-\check{c}(\tau^S)}{F^{-1}(\tau^S)}. \tag{81}$$

Thus,

$$\begin{aligned}\check{W}(\tau^S) &= E[\tilde{v}]\check{\rho}(\tau^S) + \int_{v\geq F^{-1}(1-\check{\tau}^R(\tau^S))} (\check{d}^R(\tau^S)v - \check{c}(\tau^S))\mathrm{d}G(v) \\ &\quad + \int_{v\leq F^{-1}(\tau^S)} (\check{d}^S(\tau^S)(-v) - \check{c}(\tau^S))\mathrm{d}G(v).\end{aligned} \tag{82}$$

Observe that, from (81), the integrands in (82) equal 0 at the variable boundary of the respective integration area. Thus,

$$\begin{aligned}\frac{\mathrm{d}\check{W}}{\mathrm{d}\tau^S} &= \check{\rho}'(\tau^S)E[\tilde{v}] \\ &\quad + \int_{v\geq F^{-1}(1-\check{\tau}^R(\tau^S))} ((\check{d}^R)'(\tau^S)v - \check{c}'(\tau^S))\mathrm{d}G(v) \\ &\qquad + \int_{v\leq F^{-1}(\tau^S)} ((\check{d}^S)'(\tau^S)(-v) - \check{c}'(\tau^S))\mathrm{d}G(v). \\ &= \check{\rho}'(\tau^S)E[\tilde{v}] \\ &\quad + (\check{d}^R)'(\tau^S)\int_{v\geq F^{-1}(1-\check{\tau}^R(\tau^S))} v\mathrm{d}G(v) + (\check{d}^S)'(\tau^S)\int_{v\leq F^{-1}(\tau^S)} (-v)\mathrm{d}G(v). \\ &\quad - \int_{v\geq F^{-1}(1-\check{\tau}^R(\tau^S))} \mathrm{d}G(v)\check{c}'(\tau^S) - \int_{v\leq F^{-1}(\tau^S)} \mathrm{d}G(v)\check{c}'(\tau^S).\end{aligned} \tag{83}$$

Using this, we find (for details see Appendix B)

$$\check{W}^{(l)} \quad = \quad 0 \quad \text{for all } 1 \leq l \leq k-1. \tag{84}$$

and

$$\check{W}^{(k)} \quad = \quad \check{c}^{(k)}\left(-F'(0)\left(1-\frac{1}{n-t^*+1-q}\right)\overline{y}(t^*-1)-\gamma\right). \tag{85}$$

The desired first-order welfare effect can be computed as

$$\begin{aligned}
\lim_{c\to 0}\frac{\mathrm{d}}{\mathrm{d}c}W_{F,c}(M,\tilde{\tau}^R(c),\tilde{\tau}^S(c)) &= \lim_{c\to 0}\frac{\mathrm{d}}{\mathrm{d}c}\check{W}(\tilde{\tau}^S(c)) \\
&= \lim_{c\to 0}\check{W}'(\tilde{\tau}^S(c))\cdot(\tilde{\tau}^S)'(c) \\
&= \lim_{c\to 0}\frac{\check{W}'(\tilde{\tau}^S(c))}{\check{c}'(\tilde{\tau}^S(c))} \\
&= \lim_{\tau^S\to F^S}\frac{\check{W}'(\tau^S)}{\check{c}'(\tau^S)}.
\end{aligned}$$

Due to (78) and (84), we can apply L'Hospital's rule $k-1$ times, so that

$$\lim_{\tau^S\to F^S}\frac{\check{W}'(\tau^S)}{\check{c}'(\tau^S)} \quad = \quad \frac{\check{W}^{(k)}}{\check{c}^{(k)}}. \tag{86}$$

Now (85) implies (30). This completes the proof of Lemma 2.

References

Aguiar-Conraria, L., and Magalhaes, P. C. (2010). “How quorum rules distort referendum outcomes: Evidence from a pivotal voter model,” *European Journal of Political Economy* **26**, 541–557.

Azrieli, Y., and S. Kim (2014), “Pareto efficiency and weighted majority rules,” *International Economic Review* **55**, 1067–1088.

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

Bierbrauer, F., and M. Hellwig (2016), "Robustly Coalition-Proof Incentive Mechanisms for Public Good Provision are Voting Mechanisms and Vice Versa," *Review of Economic Studies* **83**, 1440–1464.

Blackwell, C. W. (2003), “Athenian Democracy: a brief overview”, in C. W. Blackwell, ed., Dēmos: Classical Athenian Democracy (A. Mahoney and R. Scaife, edd., The Stoa: a consortium for electronic publication in the humanities [www.stoa.org]) edition of February 28, 2003.

Bognar, K., T. Börgers, and M. Meyer-ter-Vehn (2015), “An optimal Voting System when Voting is costly,” *Journal of Economic Theory* **159**, 1056–1073.

Börgers, T. (2000), “Costly Voting,” *Working Paper*.

Börgers, T. (2004), “Costly Voting,” *American Economic Review* **94**, 57–66.

Campbell, C. M. (1999), “Large electorates and decisive minorities,” *Journal of Political Economy* **107**, 1199–1217.

Celik, G., and O. Yilankaya (2009), “Optimal Auctions with Simultaneous and Costly Participation," *The B.E. Journal of Theoretical Economics*, DOI: https://doi.org/10.2202/1935-1704.1522.

Chakravarty, S., Kaplan, T. R., and Myles, G. D. (2010). “The benefits of costly voting,” Available at SSRN 1570094.

Charletyy, P., M.-C. Fagart, and S. Souam (2015), “Quorum Rules and Shareholder Power”, mimeo, Paris Descartes.

Corte-Real, P. P., and Pereira, P. T. (2004), “The voter who wasn’t there: referenda, representation and abstention,” *Social Choice and Welfare* 22, 349–369.

Drexl, M., and A. Kleiner (forthcoming) “Why Voting? - A Welfare Analysis," *American Economic Journal: Microeconomics*.

Feddersen, T. J., and W. Pesendorfer (1996), “The swing voter’s curse,” *The American Economic Review* **86**, 408–424.

Feddersen, T. J., and W. Pesendorfer (1997), “Voting behavior and information aggregation in elections with private information,” Econometrica **65**, 1029–1058.

Feddersen, T., and A. Sandroni (2006), “A theory of participation in elections,” *The American Economic Review* **96**, 1271–1282.

Gershkov, A., B. Moldovanu, and X. Shi (2017), “Optimal voting rules,” *The Review of Economic Studies* **84**, 688–717.

Ghosal, S., and B. Lockwood (2009), “Costly voting when both information and preferences differ: is turnout too high or too low?,” *Social Choice and Welfare* **33**, 25–50.

Goeree, J. K., and J. Grosser (2007), “Welfare reducing polls,” *Economic Theory* **31**, 51–68.

Herrera, H. and A. Matozzi (2010), “Quorum and Turnout in Referenda,” *Journal of the European Economic Association* **8**, 838–871.

Herrera, H. , M. Morelli, and T. Palfrey (2014), “Turnout and Power Sharing,” *The Economic Journal* **124**, 131–162.

Kartal, M. (2015), “A comparative welfare analysis of electoral systems with endogenous turnout,” *The Economic Journal.* **125**, 1369–1392.

Kos, N. (2012), “Communication and Efficiency in Auctions,” *Games and Economic Behavior* **75**, 233–249.

Krasa, S., and M. Polborn (2009), “Is Mandatory Voting Better than Voluntary Voting?,” *Games and Economic Behavior* **66**, 275–291.

Krishna, V., and J. Morgan (2015), “Majority rule and utilitarian welfare,” *American Economic Journal: Microeconomics* **7**, 339–375.

Laruelle, A., and F. Valenciano (2011), “Majorities with a quorum,” *Journal of Theoretical Politics* **23**, 241–259.

Ledyard, J. O. (1984), “The pure theory of large two-candidate elections,” *Public Choice* **44**, 7–41.

Ledyard, J.O., and T.R. Palfrey (2002), “The approximation of efficient public good mechanisms by simple voting schemes," *Journal of Public Economics* **83**, 153–171.

Maniquet, F., and M. Morelli (2015), “Approval quorums dominate participation quorums," *Social Choice and Welfare* **45**, 1–27.

Myatt, D.P. 2015, “A Theory of Voter Turnout,” *Working paper*, London Business School.

Myerson, R.B. (1998) “Population uncertainty and Poisson games," *International Journal of Game Theory* **27**, 375–392.

Palfrey, T. R., and H. Rosenthal (1983), "A strategic calculus of voting," *Public Choice* **41**, 7-53.

Palfrey, T. R., and H. Rosenthal (1985), "Voter participation and strategic uncertainty," *American Political Science Review* **79**, 62–78.

Qvortrup, M. (2005), *A Comparative Study of Referendums*, Manchester University Press, Manchester UK.

Rae, D. W. (1969), "Decision-rules and individual values in constitutional choice," *American Political Science Review* **63**, 40–56.

Schmitz, P. W., and T. Tröger (2012), "The (sub-) optimality of the majority rule," *Games and Economic Behavior* **74**, 651–665.

Taylor, C. R., and H. Yildirim (2010), "A unified analysis of rational voting with private values and group-specific costs," *Games and Economic Behavior* **70**, 457–471.

Appendix B

Heterogenous participation costs

The game-theoretic literature on costly voting largely relies on (special cases of) the following preference model. Voters have stochastically independent types (v_i, c_i) distributed according to a c.d.f. $H(v_i, c_i)$, where v_i is the Bernoulli utility from outcome R, $v_i - c_i$ from iR, 0 from S and $-c_i$ from iS. The marginal distribution of v_i/c_i is assumed to be atom-free. The objective is to maximize the ex-ante expected utility. We show here that any such objective is captured by our model if the weight density function is chosen appropriately.

Let $(\tilde{v}, \tilde{c})$ denote a random vector with distribution H.

Let $c > 0$ denote a fixed parameter, e.g. $c = 1$.

The distribution H can be alternatively represented by a two-step experiment, where first $c\tilde{v}/\tilde{c}$ is realized and then $\tilde{c}/c$ is realized. Let F denote the c.d.f. for $c\tilde{v}/\tilde{c}$ and J_v denote the c.d.f. for $\tilde{c}/c$, conditional on $c\tilde{v}/\tilde{c} = v$.

The equilibrium conditions (7) and (8) and definitions of τ^R and τ^S apply unchanged. Let m denote a mechanism-equilibrium pair. The ex-ante expected utility is

$$\begin{aligned}
W_H(m) &= \rho^m(A)E_H[\tilde{v}] + \int_{v_i>0} \max\{0, v_i\Delta^R - c_i\}\mathrm{d}H(v_i, c_i) \\
&\quad + \int_{v_i<0} \max\{0, -v_i\Delta^S - c_i\}\mathrm{d}H(v_i, c_i) \\
&= \rho^m(A)\int \frac{cv_i}{c_i}\frac{c_i}{c}\mathrm{d}H(v_i, c_i) + \int_{v>0} \max\{0, \frac{cv_i}{c_i}\Delta^R - c\}\frac{c_i}{c}\mathrm{d}H(v_i, c_i) \\
&\quad + \int_{v<0} \max\{0, -\frac{cv_i}{c_i}\Delta^S - c\}\frac{c_i}{c}\mathrm{d}H(v_i, c_i), \\
&= \rho^m(A)\int vg(v)\mathrm{d}F(v) + \int_{v>0} \max\{0, v\Delta^R - c\}g(v)\mathrm{d}F(v) \\
&\quad + \int_{v<0} \max\{0, -v\Delta^S - c\}g(v)\mathrm{d}F(v),
\end{aligned}$$

where we use the weights $g(v) = \int(c_i/c)\mathrm{d}J_v(c_i)$.

A second, alternative, welfare expression is obtained by aggregating across agents and dividing through the population size n. To obtain this formula, note that

$$
\begin{aligned}
W_H(m) &= \rho^m(A)\int v\mathrm{d}G(v) + \int_{v\Delta^R>c}\left(v\Delta^R - c\right)\mathrm{d}G(v) \\
&\quad + \int_{-v\Delta^S>c}\left(-v\Delta^S - c\right)\mathrm{d}G(v), \\
&= \rho^m(A)\int_{v\Delta^R<c,-v\Delta^S<c} v\mathrm{d}G(v) + \rho^m(R)\int_{v\Delta^R>c} vg(v)\mathrm{d}G(v) \\
&\quad +\rho^m(S)\int_{-v\Delta^S>c} v\mathrm{d}G(v) - \int_{v\Delta^R>c \text{ or } -v\Delta^S>c} c\mathrm{d}G(v) \\
&= \rho^m(A)(1-\tau^R-\tau^S)y^A + \rho^m(R)\tau^R y^R + \rho^m(S)\tau^S(-y^S) - y^c \\
&= \sum_{r+s\le n-1} M_{rs}\overset{n}{\Pr}(r,s)\frac{n-r-s}{n}y^A + \sum_{r+s\le n-1} M_{r+1,s}\overset{n}{\Pr}(r+1,s)\frac{r+1}{n}y^R \\
&\quad + \sum_{r+s\le n-1} M_{r,s+1}\overset{n}{\Pr}(r,s+1)\frac{s+1}{n}(-y^S) - y^c \\
&= \sum_{r+s\le n} M_{rs}\overset{n}{\Pr}(r,s)\frac{n-r-s}{n}y^A + \sum_{r'+s\le n} M_{r',s}\overset{n}{\Pr}(r',s)\frac{r'}{n}y^R \\
&\quad + \sum_{r+s'\le n} M_{r,s'}\overset{n}{\Pr}(r,s')\frac{s'}{n}(-y^S) - y^c \\
&= \frac{1}{n}\sum_{r+s\le n} M_{rs}\overset{n}{\Pr}(r,s)((n-r-s)y^A + ry^R - sy^S) - y^c,
\end{aligned}
$$

where $\Pr^n(r,s) = \begin{pmatrix} n \\ r\ s \end{pmatrix}(\tau^R)^R(\tau^S)^S(1-\tau^R-\tau^S)^{n-r-s}$ denotes the equilibrium probability of tally (r,s), and

$$
\begin{aligned}
y^c &= c\int_{v\Delta^R>c \text{ or } -v\Delta^S>c}\mathrm{d}G(v), \\
y^A &= \frac{1}{1-\tau^R-\tau^S}\int_{v\Delta^R<c,-v\Delta^S<c} v\mathrm{d}G(v), \\
y^R &= \frac{1}{\tau^R}\int_{v\Delta^R>c} v\mathrm{d}G(v), \\
y^S &= \frac{1}{\tau^S}\int_{-v\Delta^S>c}(-v)\mathrm{d}G(v).
\end{aligned}
$$

Proof for the "unbiased case" $E[\tilde{v}] = 0$ in Proposition 1.

We need an auxiliary result, Lemma A. Here, the welfare (11) is independent of M. Consider a relaxed maximization problem in which the equilibrium conditions are replaced by inequalities:

$$
\begin{aligned}
\text{(relax)} \quad \max_{(M,\Delta^R,\Delta^S)} \quad & \int \max\{v\Delta^R - c, 0\}\mathrm{d}G(v) + \int \max\{-v\Delta^S - c, 0\}\mathrm{d}G(v) \\
\text{s.t.} \quad & \Delta^R - d^R(M, l^R(\Delta^R), l^S(\Delta^S)) \le 0, \quad (R) \\
& \Delta^S - d^S(M, l^R(\Delta^R), l^S(\Delta^S)) \le 0, \quad (S) \\
& \Delta^R \ge 0, \quad \Delta^S \ge 0, \\
& 0 \le M_{rs} \le 1 \quad \text{for all } (r, s).
\end{aligned}
$$

Lemma A below justifies our focus on the relaxed problem.

Lemma A. *Problem* $(relax)$ *always has a solution such that both* (R) *and* (S) *are satisfied with equality. In particular, if* $E[\tilde{v}] = 0$*, then any solution to (opt) also solves problem* $(relax)$*.*

Proof. Observe first that (relax) always has a solution. This follows from Weierstrass' maximum-value theorem (note that (Δ^R, Δ^S) belongs to the compact set $[0, 1]^2$ because $d^R \le 1$ and $d^S \le 1$).

Consider any solution (M, Δ^R, Δ^S) to (relax). We will construct from it another solution to (relax) such that the constraints (R) and (S) are satisfied with equality. Let $\tau^S = l^S(\Delta^S)$ and $\tau^R = l^R(\Delta^R)$.

Case 1: $\tau^S > 0$ and $\tau^R > 0$ and $\tau^R + \tau^S < 1$. We claim that at (M, Δ^R, Δ^S) both constraints (R) and (S) are satisfied with equality.

Because $\Delta^R > 0$ (from $\tau^R > 0$), constraint (R) implies that there exists $(\hat{r}, \hat{s})$ such that $M(\hat{r} + 1, \hat{s}) - M(\hat{r}, \hat{s}) > 0$. Thus,

$$M(\hat{r} + 1, \hat{s}) > 0 \text{ and } M(\hat{r}, \hat{s}) < 1. \tag{87}$$

Suppose that constraint (R) is satisfied with strict inequality. Then locally only constraint (S) is relevant. Thus, (87) implies[30]

$$0 \le \Pr_{\tau^R,\tau^S}(\hat{r} + 1, \hat{s}) - \Pr_{\tau^R,\tau^S}(\hat{r} + 1, \hat{s} - 1), \tag{88}$$

$$0 \ge \Pr_{\tau^R,\tau^S}(\hat{r}, \hat{s}) - \Pr_{\tau^R,\tau^S}(\hat{r}, \hat{s} - 1). \tag{89}$$

(If (88) does not hold, then one can slightly decrease $M(\hat{r} + 1, \hat{s})$, thus making the left-hand-side of (S) strictly smaller than 0, followed by an increase of Δ^R (or Δ^S)

[30] We use the convention $\Pr_{\tau^R,\tau^S}(r, s) = 0$ if not $r \ge 0$, $s \ge 0$, and $s + r \le n - 1$.

that is so small that both constraints remain satisfied; the increase of Δ^R increases the welfare. A similar contradiction is obtained by increasing $M(\hat{r}, \hat{s})$ if (89) does not hold).

But the expression

$$\left(\Pr_{\tau^R,\tau^S}(r,s) - \Pr_{\tau^R,\tau^S}(r,s-1)\right) \frac{r!s!(n-r-s)!}{(n-1)!(\tau^R)^r(\tau^S)^{s-1}(1-\tau^R-\tau^S)^{n-1-r-s}}$$
$$= (n-r-s)\tau^S - s(1-\tau^R-\tau^S)$$

is strictly decreasing in r, a contradiction to (88) and (89). The proof that constraint (S) is satisfied with equality is analogous.

Case 2: $\tau^R + \tau^S = 1$. Then $\tau^R = F^R$ and $\tau^S = F^S$. We claim that at (M, Δ^R, Δ^S) both constraints (R) and (S) are satisfied with equality.

The steps leading to (88) and (89) are as in *Case 1*. Here, $\hat{s} = n-1-\hat{r}$ in (87) because only the tallies (r,s) with $r+s = n-1$ occur with positive probability from any agent's point of view. Thus, (88) and (89) simplify to

$$\Pr_{\tau^R,\tau^S}(\hat{r}+1, n-2-\hat{r}) = 0, \quad \Pr_{\tau^R,\tau^S}(\hat{r}, n-1-\hat{r}) = 0.$$

The first equation implies $\hat{r} = n-1$, a contradiction to the second equation.

Case 3: $\tau^S = 0$ and $\tau^R > 0$. (The case $\tau^R = 0$ and $\tau^S > 0$ is analogous). In the objective of (relax), the right-most integral $= 0$. Thus, another solution to (relax) is given by $(\hat{M}, \Delta^R, 0)$ with $\hat{M}_{rs} = M_{r0}$ for all r and s. At $(\hat{M}, \Delta^R, 0)$, constraint (S) is satisfied with equality. Also, constraint (R) is satisfied with equality because otherwise one could increase Δ^R (while (S) remains satisfied with equality).

Case 4: $\tau^S = 0$ and $\tau^R = 0$. Then the objective of (relax) obtains the value 0 so that another solution to (relax) is given by $(\hat{M}, 0, 0)$ with $\hat{M}_{rs} = 0$ for all (r,s). The solution $(\hat{M}, 0, 0)$ satisfies both constraints (R) and (S) with equality. This completes the proof of Lemma A.

By Lemma A, $(M^*, \Delta^{R*}, \Delta^{S*})$ solves problem (relax).

Case $E[\tilde{v}] = 0$, $\tau^{R*} > 0$, $\tau^{S*} > 0$, and $\tau^{R*} + \tau^{S*} < 1$.

It is not possible to change M^* such that both constraints (S) and (R) become strict, because if so then one could increase Δ^{R*} and Δ^{S*} slightly while keeping the constraints satisfied and increasing the objective. In other words, by the separating hyperplane theorem, there exist $\mu^R \geq 0$ and $\mu^S \geq 0$ with

$$\mu^R > 0 \text{ or } \mu^S > 0 \tag{90}$$

such that $M = M^*$ solves the problem

$$\max_M \mu^S d^S(M, \tau^{R*}, \tau^{S*}) + \mu^R d^R(M, \tau^{R*}, \tau^{S*})$$
$$\text{s.t. } 0 \leq M_{rs} \leq 1 \quad \text{for all } (r, s).$$

The objective of this problem is linear in M, that is, there exist coefficients $\mu_{rs} \in \mathbb{R}$ such that the objective can be written

$$\sum_{r,s} \mu_{rs} M_{rs} + \text{const.}$$

By optimality, (34) holds. Rewriting the coefficients yields that (35) holds with

$$\begin{aligned} \xi^R &= \tau^{S*}(\mu^R - \mu^R \tau^{S*} - \mu^S \tau^{R*}), \\ \xi^S &= \tau^{R*}(\mu^S - \mu^S \tau^{R*} - \mu^R \tau^{S*}), \quad (91) \\ \xi &= \tau^{S*}\tau^{R*}(\mu^R - \mu^S). \quad (92) \end{aligned}$$

We conclude that (2) holds for all tallies with $r + s < n$; the same conclusion holds if $r + s = n$.

To show (1), observe first that $\xi^R \geq 0$ (otherwise (34) implies that $M^*_{r+1,s} \leq M^*_{rs}$ for all (r, s), implying $\Delta^{R*} = 0$ and hence $\tau^{R*} = 0$). Similarly, $\xi^S \geq 0$. Moreover, if $\xi = 0$, then $\mu^R = \mu^S$ by (92), implying $\mu^R > 0$ and $\mu^S > 0$ by (90), implying $\xi^R = \tau^{S*}\mu^R(1 - \tau^{R*} - \tau^{S*}) > 0$ by (91). Hence, M^* is a linear mechanism.

Case $E[\tilde{v}] = 0$, $\tau^{R*} > 0$, *and* $\tau^{S*} = 0$. (The case with $\tau^{R*} = 0$ and $\tau^{S*} > 0$ is analogous).

Defining $\hat{M}(r, s) = M^*(r, 0)$ for all (r, s), $(\hat{M}, \Delta^{R*}, 0)$ also solves problem (relax). Changing any $\hat{M}_{r0}$ (while keeping $\hat{M}_{rs} = \hat{M}_{r0}$ for all s) cannot make constraint (R) become strict, because otherwise one could increase Δ^{R*} slightly while keeping the constraints satisfied and increasing the objective. Thus, $\hat{M}$ solves the problem

$$\max_M d^R(M, \tau^{R*}, 0) \quad \text{s.t. } 0 \leq M_{rs} \leq 1 \quad \text{for all } (r, s).$$

From this one can conclude that $\hat{M}(r, 0)$ is as in an R-one-sided linear rule.

Model variant with 0-cost agents

Suppose that each agent, with probability z^R (resp., z^S), has 0 participation cost and valuation $v_0^R > 0$ (resp., $v_0^S < 0$); call this type 0^R (resp., 0^S).[31] Moreover,

[31] Introducing different positive-valuation types with 0 participation cost is not necessary because these types would have identical incentives; similar for negative-valuation types.

each agent, with probability $1 - z^R - z^S > 0$, has a type v distributed according to the c.d.f. $F/(1 - z^R - z^S)$. The weights for the new types are denoted y^R and y^S, respectively. The weights of the incumbent types are given by $g/(1 - y^R - y^S)$.

An equilibrium is still described by a pair (Δ^R, Δ^S), with the additional understanding that type 0^R takes action R if $\Delta^R > 0$ and otherwise takes action A; similar for type 0^S.

Equilibrium conditions are as before, with

$$\begin{aligned} l^R(\Delta^R) &= z^R \mathbf{1}_{\Delta^R > 0} + 1 - z^R - z^S - F(\frac{c}{\Delta^R}), \\ l^S(\Delta^S) &= z^S \mathbf{1}_{\Delta^S > 0} + F(-\frac{c}{\Delta^S}). \end{aligned}$$

Any equilibrium can alternatively be expressed in terms of the participation pair (τ^R, τ^S), where $\tau^R = 0$ or $\tau^R \geq z^R$; similar for τ^S.

Given a mechanism-equilibrium pair m, the interim-expected utilities of the new types are

$$U^m(0^R) = v_0^R \rho^m(A) + v_0^R d^R(M, \tau^R, \tau^S), \;\; U^m(0^S) = v_0^S \rho^m(A) - v_0^S d^S(M, \tau^R, \tau^S).$$

Using the shortcut $E^0 = z^R v_0^R + z^S v_0^S + \frac{1 - y^R - y^S}{1 - z^R - z^S} \int v \mathrm{d}G(v)$ for the expected valuation, the *ex-ante* expected welfare of each individual is as in (11) and (12), but with $E[\tilde{v}]$ replaced by E^0, and with the additional terms $y^R v_0^R d^R(m) + y^S(-v_0^S) d^S(m)$ in (11) and $y^R v_0^R \Delta^R + y^S(-v_0^S)\Delta^S$ in (12).

Proposition 4. *Consider the model variant with 0-cost agents. There exists an optimal mechanism-equilibrium pair in which the mechanism is a linear voting rule.*

Any optimal mechanism-equilibrium pair $(M^*, \tau^{R*}, \tau^{S*})$ is such that $\tau^{R*} + \tau^{S*} < 1$. Moreover, if $\tau^{R*} > 0$ and $\tau^{S*} > 0$, then M^* is linear; if $\tau^{R*} > 0$ and $\tau^{S*} = 0$, then M^*_{r0} ($r = 0, \ldots, n$) is as an R-one-sided linear rule; if $\tau^{R*} = 0$ and $\tau^{S*} > 0$, then M^*_{0s} ($s = 0, \ldots, n$) is as an S-one-sided linear rule.

Sketch of proof. All arguments are analogous to the proof of Proposition 1. The main case distinction now is $E^0 > 0$ versus $E^0 = 0$. In case $E^0 = 0$ one considers a variant of problem (relax) with the additional terms $y^R v_0^R \Delta^R + y^S(-v_0^S)\Delta^S$ in the objective.

Model variant with separate populations of R-agents and S-agents

Suppose there is a fixed number n_R (resp. n_S) of R-agents (resp., S-agents) with types v stochastically independently distributed according to a distribution F_R on $\mathbb{R}_+$ (resp., F_S on $\mathbb{R}_-$), as in Palfrey and Rosenthal (1985). Let G_R and G_S denote c.d.f.s for the weights. Let $\tilde{v}_R$ and $\tilde{v}_S$ denote random variables with distribution G_R and G_S, respectively.

A voting rule is given by a mapping $M : \{0, 1, \ldots, n_R\} \times \{0, 1, \ldots, n_S\} \to [0, 1]$.

We focus on equilibria in which all R-agents use the same strategy, and so do all S-agents. These strategies are defined via (Δ^R, Δ^S). The strategy of the R-agents (resp., S-agents) is given by (5), restricted to valuations $v > 0$ (resp., $v < 0$). Let

$$l^R(\Delta^R) = 1 - F_R(\frac{c}{\Delta^R}), \quad l^S(\Delta^S) = F_S(-\frac{c}{\Delta^S}).$$

Let $\mathrm{Pr}_{q,m}(t)$ denote the probability of t successes in a binomial distribution with any parameters (q, m). An agent's anticipated pivot probabilities in a mechanism M are given by

$$\begin{aligned} d^R(M, \tau^R, \tau^S) &= \sum_{r \leq n_R - 1} \sum_{s \leq n_S} \Pr_{\tau^R, n_R - 1}(r) \Pr_{\tau^S, n_S}(s)(M_{r+1,s} - M_{r,s}), \\ d^S(M, \tau^R, \tau^S) &= \sum_{r \leq n_R} \sum_{s \leq n_S - 1} \Pr_{\tau^R, n_R}(r) \Pr_{\tau^S, n_S - 1}(s)(M_{r,s} - M_{r,s+1}). \end{aligned}$$

where $\tau^R \in [0, 1]$ (resp., $\tau^S \in [0, 1]$) is the probability that a given R-agent (resp., S-agent) participates. Equilibrium conditions are (7) for the R-agents and (8) for the S-agents. As in the main model, we can alternatively express any equilibrium via the participation pair (τ^R, τ^S).

A mechanism-equilibrium pair $m = (M, \tau^R, \tau^S)$ yields, for an R-agent of type v, the interim expected utility

$$U^m(v) = v\rho^{m,R}(A) + \max\{vd^R(M, \tau^R, \tau^S) - c, 0\},$$

where

$$\rho^{m,R}(A) = \sum_{r \leq n_R - 1} \sum_{s \leq n_S} \Pr_{\tau^R, n_R - 1}(r) \Pr_{\tau^S, n_S}(s) M_{r,s}.$$

denotes an R-agent's expected utility from abstaining. Similarly, for an S-agent of type v,

$$U^m(v) = v\rho^{m,S}(A) + \max\{-vd^S(M, \tau^R, \tau^S) - c, 0\}.$$

where

$$\rho^{m,S}(A) = \sum_{r \leq n_R} \sum_{s \leq n_S - 1} \Pr_{\tau^R, n_R}(r) \Pr_{\tau^S, n_S - 1}(s) M_{r,s}.$$

denotes an S-agent's expected utility from abstaining.

We consider a social planner who is interested in maximizing a weighted average of an R-agent's utility and an S-agent's utility, with weights $\kappa_R > 0$ and $\kappa_S > 0$. For instance, $\kappa_R = n_R$ and $\kappa_S = n_S$. The resulting objective is

$$\begin{aligned} W(m) &= \kappa_R E[U^m(\tilde{v}_R)] + \kappa_S E[U^m \tilde{v}_S] \\ &= \kappa_R E[\tilde{v}_R]\rho^{m,R}(A) + \kappa_S E[\tilde{v}_S]\rho^{m,S}(A) \\ &\quad + \kappa_R \int \max\{v d^R(M,\tau^R,\tau^S) - c, 0\} \mathrm{d}G_R(v) \qquad (93) \\ &\quad + \kappa_S \int \max\{-v d^S(M,\tau^R,\tau^S) - c, 0\} \mathrm{d}G_S(v). \qquad (94) \end{aligned}$$

Proposition 5. *Consider the model variant with fixed populations of R-agents and S-agents.*

There exists an optimal mechanism-equilibrium pair in which the mechanism is linear. Any optimal mechanism-equilibrium pair $(M^*, \tau^{R*}, \tau^{S*})$ is such that $\tau^{R*} < 1$ and $\tau^{S*} < 1$. Moreover, if $\tau^{R*} > 0$ and $\tau^{S*} > 0$, then M^* is linear; if $\tau^{R*} > 0$ and $\tau^{S*} = 0$, then M^*_{r0} is as an R-one-sided linear rule. If $\tau^{R*} = 0$ and $\tau^{S*} > 0$, then M^*_{0s} is as an S-one-sided linear rule.

Proof. Consider an optimal mechanism-equilibrium pair $m^* = (M^*, \tau^{R*}, \tau^{S*})$. Define corresponding pivot probabilities Δ^{R*} and Δ^{S*} via (9).

To show $\tau^{R*} < 1$, suppose that $\tau^{R*} = 1$ (the proof of $\tau^{S*} < 1$ is analogous). Then only tallies of the form (n_R, s) can occur with positive probability. Then $(0, \tau^{S*})$ is an equilibrium in the mechanism $\hat{M}$ defined via $\hat{M}_{rs} = M_{n_R,s}$. Moreover, $W(\hat{M}, 0, \tau^{S*}) = W(m^*) + \kappa_R c$, contradicting the optimality of m^*.

Case 1: $\tau^{R*} > 0$ and $\tau^{S*} > 0$. Then

$$\Pr_{\tau^{R*},n_R}(r) \Pr_{\tau^{S*},n_S}(s) > 0 \text{ for all } (r,s) \text{ with } r \leq n_R \text{ and } s \leq n_S. \qquad (95)$$

Define the (convex and non-empty) set of mechanisms

$$\begin{aligned} \mathcal{M} &= \{M \mid \Delta^{R*} = d^R(M,\tau^{R*},\tau^{S*}),\ \Delta^{S*} = d^S(M,\tau^{R*},\tau^{S*}), \\ &\qquad 0 \leq M_{rs} \leq 1 \quad \text{for all } (r,s)\}. \end{aligned}$$

By optimality, the mechanism M^* solves the following problem:

$$\text{(lin)} \quad \max_{M \in \mathcal{M}} \underbrace{\kappa_R E[\tilde{v}_R]}_{\stackrel{\text{def}}{=} w_R > 0} \rho^{(M,\tau^{R*},\tau^{S*}),R}(A) + \underbrace{\kappa_S E[\tilde{v}_S]}_{\stackrel{\text{def}}{=} w_S > 0} \rho^{(M,\tau^{R*},\tau^{S*}),S}(A).$$

Problem (lin) is linear. Hence, the Kuhn-Tucker conditions are necessary, without any constraint qualification. Thus, there exist Lagrange multipliers μ^R, μ^S, and μ_{rs} [32] for all (r,s) such that

$$\begin{aligned}\mu_{rs} &= \frac{\partial}{\partial M_{rs}}\Big(w_R\rho^{(M,\tau^{R*},\tau^{S*}),R}(A) + w_S\rho^{(M,\tau^{R*},\tau^{S*}),S}(A) \\ &\qquad + \mu^R d^R(M,\tau^{R*},\tau^{S*}) + \mu^S d^S(M,\tau^{R*},\tau^{S*})\Big), \qquad (96)\end{aligned}$$

where $\mu_{rs} \leq 0$ if $M^*_{rs} < 1$ and $\mu_{rs} \geq 0$ if $M^*_{rs} > 0$ (complementary slackness). Put differently, (34) holds.

Consider a tally (r,s). Rewriting (96) and using upper indices R and S when writing binomial probabilities,

$$\begin{aligned}\mu_{rs} &= \underset{n_R}{\overset{R}{\Pr}}(r)\underset{n_S}{\overset{S}{\Pr}}(s)\left(w_R\frac{n_R - r}{n_R(1-\tau^{R*})} + w_S\frac{n_S - s}{n_S(1-\tau^{S*})}\right. \\ &\qquad \left. + \mu^R\left(\frac{r}{n_R\tau^{R*}} - \frac{n_R - r}{n_R(1-\tau^{R*})}\right) + \mu^S\left(\frac{s}{n_S\tau^{S*}} - \frac{n_S - r}{n_S(1-\tau^{S*})}\right)\right) \\ &= \underset{n_R}{\overset{R}{\Pr}}(r)\underset{n_S}{\overset{S}{\Pr}}(s)\left(r\xi^R - s\xi^S - n\xi\right), \qquad (97)\end{aligned}$$

where

$$\xi = \frac{w_R - \mu^R}{1-\tau^{R*}} + \frac{w_S - \mu^S}{1-\tau^{S*}}, \qquad (98)$$

$$\xi^R = \frac{\mu^R}{n_R\tau^{R*}} - \frac{w_R - \mu^R}{n_R(1-\tau^{R*})}, \qquad (99)$$

$$\xi^S = \frac{\mu^S}{n_S\tau^{S*}} - \frac{w_S - \mu^S}{n_S(1-\tau^{S*})}. \qquad (100)$$

Using (95), we conclude that (2) holds.

To show (1), observe first that $\xi^R \geq 0$ (otherwise (34) implies that $M^*_{r+1,s} \leq M^*_{rs}$ for all (r,s), implying $\Delta^{R*} = 0$ and hence $\tau^{R*} = 0$). Similarly, $\xi^S \geq 0$. Now suppose that $\xi^R = \xi^S = 0$. Then $\mu^R \neq w_R$ by (99), hence (99) yields

$$\frac{\mu^R/w_R}{1-\mu^R/w_R} = \frac{\tau^{R*}}{1-\tau^{R*}},$$

implying $\mu^R/w_R = \tau^{R*}$. Similarly, $\mu^S/w_S = \tau^{S*}$ by (100). Thus, $\xi = w_R + w_S > 0$ by (98). Hence, M^* is a linear mechanism.

[32] We use a single Lagrange multiplier for both constraints $0 \leq M_{rs}$ and $M_{rs} \leq 1$ because only one constraint can be binding.

Case 2: $\tau^{R*} > 0$ and $\tau^{S*} = 0$.

Define the (convex and non-empty) set of mechanisms

$$\begin{aligned} \mathcal{M}^R &= \{M \mid \Delta^{R*} = d^R(M, \tau^{R*}, 0), \\ &\qquad M_{rs} = M_{r0}, \ \ 0 \le M_{rs} \le 1 \quad \text{for all } (r,s)\}. \end{aligned}$$

For any $M \in \mathcal{M}^R$, the pair $(\Delta^{R*}, 0)$ is an equilibrium. Thus, by optimality, the mechanism $\hat{M}$, where $\hat{M}_{rs} = M^*_{r0}$, solves the following problem:

$$(\text{lin})^R \quad \max_{M \in \mathcal{M}^R} w_R \rho^{(M,\tau^{R*},0),R}(A) + w_S \rho^{(M,\tau^{R*},0),S}(A).$$

Problem $(\text{lin})^R$ is linear. Hence, there exist Lagrange multipliers μ^R and μ_r for all r such that

$$\mu_r = \frac{\partial}{\partial M_{r0}} \left(w_R \rho^{(M,\tau^{R*},0),R}(A) + w_S \rho^{(M,\tau^{R*},0),S}(A) + \mu^R d^R(M, \tau^{R*}, 0) \right), \tag{101}$$

where $\mu_r \le 0$ if $\hat{M}_{r0} < 1$ and $\mu_r \ge 0$ if $\hat{M}_{r0} > 0$ (complementary slackness). Put differently,

$$\hat{M}_{rs} = \begin{cases} 1 & \text{if } \mu_r > 0, \\ 0 & \text{if } \mu_r < 0. \end{cases}$$

Rewriting (101) yields

$$\begin{aligned} \mu_r &= \Pr_{n_R}^{R}(r) \left(w_R \frac{n_R - r}{n_R(1-\tau^{R*})} + w_S + \mu^R \left(\frac{r}{n_R \tau^{R*}} - \frac{n_R - r}{n_R(1-\tau^{R*})} \right) \right) \\ &= \Pr_{n_R}^{R}(r) \left(r\xi^R - n\xi \right), \end{aligned}$$

where

$$\xi = \frac{w_R - \mu^R}{1-\tau^{R*}} + w_S, \tag{102}$$

$$\xi^R = \frac{\mu^R}{n_R \tau^{R*}} - \frac{w_R - \mu^R}{n_R(1-\tau^{R*})}. \tag{103}$$

Using that $\Pr_{n_R}^R(r) > 0$ for all r, we conclude that (2) holds.

To see (1), observe first that $\xi^R \ge 0$ (otherwise $\hat{M}_{r+1,s} \le M^*_{rs}$ for all (r,s), implying $\Delta^{R*} = 0$ and hence $\tau^{R*} = 0$). Moreover, using arguments analogous to *Case 1* one can show that $\xi^R = 0$ implies $\xi > 0$. Hence, $\hat{M}$ is a linear mechanism.

Proof of Remark 3

We use the weights $g(v) = 1$ for all v, that is, $G = F$.

In the following we list several properties of the environment considered in Remark 3. Lemma 4 below shows that these properties together are sufficient to prove Remark 3. First, (58) is satisfied. Second, the expectation

$$v^R p^R + v^0 p^0 + v^S p^S \quad = \quad 0, \tag{104}$$

and, third,

$$v^0 \quad < \quad 0. \tag{105}$$

Using the definition (56), define

$$\Delta^{Rg} = d^R(\overline{M}, p^R, p^S) \approx 0.149, \quad \Delta^{Sg} = d^S(\overline{M}, p^R, p^S) \approx 0.148.$$

The fourth property is that

$$-v^S \Delta^{Sg} > c > -v^0 \Delta^{Sg}, \quad c < v^R \Delta^{Rg}. \tag{106}$$

This implies that (p^R, p^S) is an equilibrium in $\overline{M}$ if F is an ϵ-approximation of (23) with sufficiently small ϵ.

We will use the shortcut

$$\overline{W} \quad = \quad W_{\hat{F},0}(\overline{M}, p^R, p^S) - c(p^R + p^S).$$

Define

$$\begin{aligned} W^S \quad = \quad \max\{0, p^S(&\max_{p \in [p^S, p^S + p^0]} \overline{m}_{n-1}(p)(-v^S) - c), \\ &\overline{m}_{n-1}(p^S + p^0)(p^S v^S + p^0 v^0 - c)\}, \end{aligned}$$

where, for any (l, z), $\overline{m}_l(z)$ denotes the probability at the mode of the binomial distribution with parameters (l, z). The final required property of the environment is that

$$\overline{W} > W^S. \tag{107}$$

One can check that this holds for (23) at $c = 1$ and $n = 15$.

The lemma below provides sufficient conditions such that the mechanism-equilibrium pair $(\overline{M}, p^R, p^S)$ is the unique optimum.

Lemma 4. *Consider a three-point distribution $\hat{F}$, a participation cost c, and a population size n such that (58), (104), (105), (106), and (107) hold. If F is an ϵ-approximation of $\hat{F}$ with sufficiently small ϵ, then the mechanism-equilibrium pair $(\overline{M}, p^R, p^S)$ is the unique optimum.*

First we provide a sketch of the proof. Note that by (104), the first term in the welfare (12) vanishes. If either not all types around v^R participated or not all types around v^S participated, then by the equilibrium conditions one of the conditional expectations in (12) would vanish approximately as well. Thus, using Lemma 5 below the welfare would essentially be bounded above by the welfare of a one-sided rule. Lemma 6 states that the welfare of one-sided rules is essentially bounded above by W^S.

By (106), the pair (p^R, p^S) is an equilibrium in $\overline{M}$ if ϵ is small, yielding approximately welfare $\overline{W}$.

Now we can conclude from (107) that in any optimum (M, τ^R, τ^S) all types around v^R as well as all types around v^S will participate. That is, $\tau^S \geq p^S$ and (using (105)) $\tau^R = p^R$.

The remainder of the proof uses techniques similar to the proof of Remark 2. We consider the limit setting in which types are distributed according to the three-point distribution $\hat{F}$ (but, in contrast to the proof of Remark 2, the participation cost is fixed). We consider problem (*'): maximize the welfare across all mechanisms and all participation pairs (τ^R, τ^S) with $\tau^R = p^R$ and $\tau^S \geq p^S$, ignoring equilibrium conditions. Extending the ideas of Lemma 3, the unique solution to (*') is $(\overline{M}, p^R, p^S)$. To see this, note that at $(\overline{M}, p^R, p^S)$ the alternative R is implemented if and only if the sum of the agents' valuations is positive, and moreover the total cost spent on participation is minimal across the feasible set.

Uniqueness in (*') implies that any maximizer of any variant of (*') with slightly different objective must be close to $(\overline{M}, p^R, p^S)$.

Now move away from the limit setting and suppose that ϵ is small. Then (p^R, p^S) is an equilibrium in $\overline{M}$ and, moreover, the welfare objective is only slightly different from the objective in (*'). Thus, any optimal mechanism-equilibrium pair is close to $(\overline{M}, p^R, p^S)$.

Using the gaps in the support of F, the equilibrium conditions imply that in any mechanism sufficiently close to $\overline{M}$, any equilibrium sufficiently close to (p^R, p^S) is in fact equal to (p^R, p^S). Thus, in any optimal mechanism-equilibrium pair the equilibrium (p^R, p^S) is played. Given these participation rates, the welfare conditional on any election result (r, s) equals $w(r, s)$ so that by (58) no mechanism other than $\overline{M}$ can be optimal.

We will present the formal proof of Lemma 4 after stating and proving Lemma 5 and Lemma 6.

Lemma 5. *Consider any F, n, and c. Consider a linear rule M and an equilibrium (τ^R, τ^S). Then there exists an R-one sided rule $\hat{M}$ and an equilibrium $(\hat{\tau}^R, 0)$ such that $d^R(\hat{M}, \hat{\tau}^R, 0) \geq d^R(M, \tau^R, \tau^S)$. Similarly, there exists an S-one sided rule $\hat{M}^S$ and an equilibrium $(0, \hat{\tau}^S)$ such that $d^S(\hat{M}^S, 0, \hat{\tau}^S) \geq d^S(M, \tau^R, \tau^S)$.*

Sketch of proof. We distinguish two cases. If M is monotonic, then we form a one-sided rule by taking the average along the desired dimension, say the S-dimension so that an R-one-sided rule $\hat{M}$ is obtained. Using the monotonicity of M, we show that, at the original participation rate τ^R, the R-pivotality is higher in $\hat{M}$ than in M. Starting from τ^R, one can imagine to further increase the R-participation in $\hat{M}$ (and, thus, increase the R-pivotality) until an equilibrium $(\hat{\tau}^R, 0)$ is reached. If M is not monotonic, then, by linearity, it is a "sandwich". In an "R-sandwich", there is a threshold number of R-votes $\hat{r}$ below which S is implemented and above which R is implement (an "S-sandwich" is defined analogously). We consider an R-one-sided rule $\hat{M}$ that requires $\hat{r}$ or $\hat{r}+1$ R-votes to implement R and again show that at the original participation rate the R-pivotality is higher in $\hat{M}$ than in M.

Proof of Lemma 5. We prove the R-part. First suppose that M is not an R-sandwich. Then M is monotonically decreasing in s, that is, $M_{r,s+1} \leq M_{r,s}$ for all (r, s).

Given r votes for R, the conditional distribution of the number of S-votes is denoted p_r on $\{0, 1, \ldots, n-r\}$:

$$p_r(s') = \binom{n-r}{s'} \left(\frac{\tau^S}{1-\tau^R}\right)^{s'} \left(\frac{1-\tau^R-\tau^S}{1-\tau^R}\right)^{n-r-s'}. \qquad (108)$$

For all (r, s), define the mean

$$\hat{M}_{rs} = \sum_{s'=0}^{n-r} p_r(s') M_{rs'}.$$

Observe that $\hat{M}$ is R-one-sided. Moreover,

$$\begin{aligned} d^R(M, \tau^R, \tau^S) &= \sum_{r=0}^{n-1} \sum_{s'=0}^{n-1-r} \Pr_{\tau^R}(r) p_{r+1}(s')(M_{r+1,s'} - M_{r,s'}) \\ &= \sum_{r=0}^{n-1} \Pr_{\tau^R}(r)(\hat{M}_{r+1,0} - \sum_{s'=0}^{n-1-r} p_{r+1}(s') M_{r,s'}), \qquad (109) \end{aligned}$$

where $\Pr_{\tau^R}(r)$ denotes the probability of r successes in a binomial distribution with parameters $(n-1, \tau^R)$.

The distribution p_r dominates the distribution p_{r+1} in terms of the monotone-likelihood-ratio property, that is, the ratio

$$\frac{p_r(s')}{p_{r+1}(s')} = \frac{n-r}{n-r-s'} \frac{1-\tau^R-\tau^S}{1-\tau^R}$$

is increasing in s'. Thus, using the monotonicity of $M_{rs'}$ in s',

$$\sum_{s'=0}^{n-1-r} p_{r+1}(s') M_{r,s'} \geq \sum_{s'=0}^{n-r} p_r(s') M_{r,s'} = \hat{M}_{r,0}.$$

Combining this with (109) we find

$$d^R(M, \tau^R, \tau^S) \leq \sum_{r=0}^{n-1} \Pr_{\tau^R}(r)(\hat{M}_{r+1,0} - \hat{M}_{r,0}) = d^R(\hat{M}, \tau^R, 0).$$

Using the shortcut $\Delta^R = d^R(M, \tau^R, \tau^S)$, it follows that

$$\Delta^R \leq d^R(\hat{M}, l^R(\Delta^R), 0),$$

Hence, using the intermediate value theorem, there exists $\hat{\Delta}^R \geq \Delta^R$ such that

$$d^R(\hat{M}, l^R(\hat{\Delta}^R), 0) = \hat{\Delta}^R. \tag{110}$$

Thus, $(\hat{\tau}^R, 0) = (l^R(\hat{\Delta}^R), 0)$ is an equilibrium in $\hat{M}$, and $\hat{\tau}^R \geq \tau^R$. This completes the proof in the non-sandwich case.

Now suppose that M is an R-sandwich. That is, there exists $\hat{r}$ such that, for all (r, s), $M_{rs} = 1$ if $r > \hat{r}$ and $M_{rs} = 0$ if $r < \hat{r}$. Thus,

$$d^R(M, \tau^R, \tau^S) = \Pr_{\tau^R}(\hat{r}-1) \sum_{s=0}^{n-\hat{r}} p_{\hat{r}}(s) M_{\hat{r},s} + \Pr_{\tau^R}(\hat{r}) \sum_{s=0}^{n-\hat{r}-1} p_{\hat{r}+1}(s)(1 - M_{\hat{r},s}) \tag{111}$$

and

$$d^S(M, \tau^R, \tau^S) = \Pr_{\tau^R}(\hat{r}) \sum_{s=0}^{n-1-\hat{r}} p_{\hat{r}+1}(s)(M_{\hat{r},s} - M_{\hat{r},s+1}).$$

If $\hat{r} = n$, then there is nothing to prove because $d^S(M, \tau^R, \tau^S) = 0$ and M is an R-one-sided rule. In the following, assume $\hat{r} < n$.

In equilibrium, $d^S(M, \tau^R, \tau^S) \geq 0$. Thus,

$$\sum_{s=0}^{n-1-\hat{r}} p_{\hat{r}+1}(s) M_{\hat{r},s} \geq \sum_{s=0}^{n-1-\hat{r}} p_{\hat{r}+1}(s) M_{\hat{r},s+1}. \tag{112}$$

Below we will prove that condition (112) implies the condition

$$\sum_{s=0}^{n-\hat{r}} p_{\hat{r}}(s) M_{\hat{r},s} \leq \sum_{s=0}^{n-\hat{r}-1} p_{\hat{r}+1}(s) M_{\hat{r},s}. \tag{113}$$

Define $\hat{M}$ via $\hat{M}_{rs} = \mathbf{1}_{r \geq r'}$, where $r' \in \arg\max_{r \in \{\hat{r}, \hat{r}+1\}} \Pr_{\tau^R}(r-1)$. Then (113) together with (111) implies

$$d^R(M, \tau^R, \tau^S) \leq \Pr_{\tau^R}(r'-1) = d^R(\hat{M}, \tau^R, 0).$$

By the same argument as in the non-sandwich case, there exists $\hat{\Delta}^R \geq d^R(M, \tau^R, \tau^S)$ such that (110) holds and we are done.

It remains to prove condition (113). Using the new variable $s' = s+1$ on the r.h.s. of (112), rewrite (112) as

$$\begin{aligned} &\sum_{s=0}^{n-1-\hat{r}} \binom{n-1-\hat{r}}{s} \left(\frac{\tau^S}{1-\tau^R}\right)^s \left(\frac{1-\tau^R-\tau^S}{1-\tau^R}\right)^{n-1-\hat{r}-s} M_{\hat{r},s} \\ &\geq \sum_{s'=1}^{n-\hat{r}} \binom{n-1-\hat{r}}{s'-1} \left(\frac{\tau^S}{1-\tau^R}\right)^{s'-1} \left(\frac{1-\tau^R-\tau^S}{1-\tau^R}\right)^{n-\hat{r}-s'} M_{\hat{r},s'}. \end{aligned}$$

Equivalently, after plugging in (108),

$$\sum_{s=0}^{n-\hat{r}-1} \frac{n-\hat{r}-s}{n-\hat{r}} \frac{1-\tau^R}{1-\tau^R-\tau^S} p_{\hat{r}}(s) M_{\hat{r},s} \quad \geq \quad \sum_{s'=1}^{n-\hat{r}} \frac{s'}{n-\hat{r}} \frac{1-\tau^R}{\tau^S} p_{\hat{r}}(s') M_{\hat{r},s'}.$$

After cancelling factors and multiplying with denominators,

$$\sum_{s=0}^{n-\hat{r}-1} (n-\hat{r}-s)\tau^S p_{\hat{r}}(s) M_{\hat{r},s} \quad \geq \quad \sum_{s=0}^{n-\hat{r}} s(1-\tau^R-\tau^S) p_{\hat{r}}(s) M_{\hat{r},s}.$$

After cancelling terms $s\tau^S p_{\hat{r}}(s) M_{\hat{r},s}$,

$$\sum_{s=0}^{n-\hat{r}-1} (n-\hat{r})\tau^S p_{\hat{r}}(s) M_{\hat{r},s} \quad \geq \quad \sum_{s=0}^{n-\hat{r}} s(1-\tau^R) p_{\hat{r}}(s) M_{\hat{r},s} - (n-\hat{r})\tau^S p_{\hat{r}}(n-\hat{r}) M_{\hat{r},n-\hat{r}}.$$

Moving the rightmost term to the left-hand side and dividing by $(n-\hat{r})(1-\tau^R)$,

$$\sum_{s=0}^{n-\hat{r}} \frac{\tau^S}{1-\tau^R} p_{\hat{r}}(s) M_{\hat{r},s} \quad \geq \quad \sum_{s=0}^{n-\hat{r}} \frac{s}{n-\hat{r}} p_{\hat{r}}(s) M_{\hat{r},s}.$$

Moving terms across sides,

$$-\sum_{s=0}^{n-\hat{r}-1}\frac{s}{n-\hat{r}}p_{\hat{r}}(s)M_{\hat{r},s} \quad\geq\quad p_{\hat{r}}(n-\hat{r})M_{\hat{r},n-\hat{r}}-\sum_{s=0}^{n-\hat{r}}\frac{\tau^S}{1-\tau^R}p_{\hat{r}}(s)M_{\hat{r},s}.$$

Adding $\sum_{s=0}^{n-\hat{r}} p_{\hat{r}}(s)M_{\hat{r},s}$ on both sides,

$$\sum_{s=0}^{n-\hat{r}-1}\frac{n-\hat{r}-s}{n-\hat{r}}p_{\hat{r}}(s)M_{\hat{r},s} \quad\geq\quad p_{\hat{r}}(n-\hat{r})M_{\hat{r},n-\hat{r}}+\sum_{s=0}^{n-\hat{r}}\frac{1-\tau^R-\tau^S}{1-\tau^R}p_{\hat{r}}(s)M_{\hat{r},s}.$$

Thus,

$$\sum_{s=0}^{n-\hat{r}-1}\frac{n-\hat{r}-s}{n-\hat{r}}p_{\hat{r}}(s)M_{\hat{r},s} \quad\geq\quad \sum_{s=0}^{n-\hat{r}}\frac{1-\tau^R-\tau^S}{1-\tau^R}p_{\hat{r}}(s)M_{\hat{r},s}.$$

Now (113) follows because $\frac{n-\hat{r}-s}{n-\hat{r}}p_{\hat{r}}(s) = \frac{1-\tau^R-\tau^S}{1-\tau^R}p_{\hat{r}+1}(s)$. This completes the proof of Lemma 5.

Define

$$W^R \quad=\quad \max\{0,\overline{m}_{n-1}(p^R)(p^Rv^R-c)\}.$$

Lemma 6. *Consider a three-point distribution $\hat{F}$, c, and n such that (104) and (105) hold. Suppose that F is an ϵ-approximation of $\hat{F}$ for some $\epsilon > 0$. Then the welfare that can be achieved in any equilibrium of an R-(resp., S-)one-sided mechanism is bounded above by $W^R + 3\epsilon$ (resp., $W^S + 3\epsilon$).*

Proof of Lemma 6. First consider any point $(M,\tau^R,0)$ with $\tau^R \geq 0$ and M (linear and) R-one-sided. If $\tau^R < p^R$, then type $d^R(M,\tau^R,0)(p^R-\epsilon) < c$ by the equilibrium condition; the welfare

$$\begin{aligned} W_{F,c}(M,\tau^R,0) \quad&=\quad E[\tilde{v}]\rho^{M,\tau^R,0}(A)+\int_{F^{-1}(1-\tau^R)}^{v^R+\epsilon}(d^R(M,\tau^R,0)v-c)\mathrm{d}v \\ &\overset{(104)}{\leq}\quad \epsilon+p^R(d^R(M,\tau^R,0)(p^R+\epsilon)-c) < 3\epsilon \leq W^R+3\epsilon. \end{aligned}$$

The remaining case is $\tau^R = p^R$; then

$$W_{F,c}(M,\tau^R,0) = d^R(M,p^R,0)(p^Rv^R-c) \leq W^R.$$

Now consider any point $(M,0,\tau^S)$ with $\tau^S \geq 0$ and M (linear and) S-one-sided.

If $\tau^S < p^S$, then $W_{F,c}(M, 0, \tau^S) < W^S + 3\epsilon$, arguing as in the R-one-sided case.

Suppose that $p^S \leq \tau^S < p^S + p^0$. Then type $v^0 + \epsilon$ abstains, so that $d^S(M, 0, \tau^S)(-v^0 - \epsilon) < c$ by the equilibrium condition; the welfare

$$\begin{aligned}
W_{F,c}(M, 0, \tau^S) &= E[\tilde{v}]\rho^{M,0,\tau^S}(A) + \\
&\quad p^S(-d^S(M, 0, \tau^S)v^S - c) + \int_{v^0-\epsilon}^{F^{-1}(\tau^S)} (d^S(M, 0, \tau^S)(-v) - c)\mathrm{d}v \\
&\leq \epsilon + p^S(-d^S(M, 0, \tau^S)v^S - c) + p^0 2\epsilon \\
&\leq \epsilon + p^S(\max_{p \in [p^S, p^S+p^0]} \overline{m}_{n-1}(p)(-v^S) - c) + 2\epsilon \\
&\leq W^S + 3\epsilon.
\end{aligned}$$

In the remaining case $\tau^S = p^S + p^0$ we argue again as in the R-one-sided case. This completes the proof of Lemma 6.

Proof of Lemma 4. Using that (57) holds for any $\hat{\delta}$, we can assume ϵ is so small that, for all (r, s),

$$\omega_{rs}(p^R, p^S) \text{ has the same sign as } w(r, s). \tag{114}$$

Next, note that, because $\overline{m}_{n-1}(p^R) = \overline{m}_{n-1}(1 - p^R)$ by basic properties of binomial distributions, and using $p^R v^R = p^S v^S - p^0 v^0$ from (104):

$$W^R \leq W^S. \tag{115}$$

Let

$$\epsilon \quad < \quad \overline{\epsilon} = \frac{1}{8}\left(\overline{W} - W^S\right). \tag{116}$$

Denote

$$\overline{\Delta}^R = d^R(\overline{M}, p^R, p^S) > 0 \text{ and } \overline{\Delta}^S = d^S(\overline{M}, p^R, p^S) > 0.$$

By (106), there exists $\overline{\epsilon}'$ and an open neighborhood $\mathcal{N}$ of $(\overline{M}, p^R, p^S)$ such that, for all $\epsilon < \overline{\epsilon}'$ and $(M, \tau^S, \tau^S) \in \mathcal{N}$,

$$(-v^S - \epsilon)d^S(M, \tau^S, \tau^S) > c > (-v^0 + \epsilon)d^S(M, \tau^S, \tau^S), \quad c < (v^R - \epsilon)d^R(M, \tau^S, \tau^S). \tag{117}$$

Using (54), where $\hat{\delta} = \epsilon$ because $g(v) = 1$ for all v,

$$W_{F,c}(\overline{M}, p^R, p^S) \quad \geq \quad \overline{W} - \epsilon. \tag{118}$$

Next we show, for all $\epsilon < \min\{\overline{\epsilon}, \overline{\epsilon}'\}$,

$$\text{if } (M, \tau^R, \tau^S) \text{ is optimal, then } \tau^R = p^R \text{ and } \tau^S \geq p^S. \tag{119}$$

To see this, define $m^* = (M, \tau^R, \tau^S)$, $\Delta^{R*} = d^R(m^*)$ and $\Delta^{S*} = d^S(m^*)$.

Suppose that $\tau^R < p^R$. Then type $v^R - \epsilon$ abstains so that

$$(v^R - \epsilon)\Delta^{R*} \leq c.$$

Thus,

$$\begin{aligned} W_{F,c}(m^*) &\overset{(104)}{\leq} \epsilon + p^R((v^R + \epsilon)\Delta^{R*} - c) + \int_{v^S-\epsilon}^{0} \max\{0, v\Delta^{S*} - c\}\mathrm{d}F(v) \\ &\overset{\text{Lemma 5}}{\leq} \epsilon + 2\epsilon + W_{F,c}(\hat{M}^S, 0, \hat{\tau}^S) + \epsilon \end{aligned}$$

for some S-one-sided mechanism $\hat{M}^S$ with equilibrium $(0, \hat{\tau}^S)$. Using Lemma 6 we conclude that

$$W_{F,c}(m^*) \leq 7\epsilon + W^S \overset{(116)}{<} \overline{W} - \frac{1}{8}\left(\overline{W} - W^S\right) \overset{(116),\,(118)}{<} W_{F,c}(\overline{M}, p^R, p^S).$$

contradicting the optimality of m^* because (p^R, p^S) is an equilibrium in $\overline{M}$ by (117). We conclude that $\tau^R \geq p^R$. An analogous argument, using (115), shows that $\tau^S \geq p^S$. Actually, $\tau^R = p^R$ from (105). This completes the proof of (119).

As in the proof of Lemma 3 one shows that the tuple $(\overline{M}, p^R, p^S)$ is the unique solution to the problem

$$(*') \qquad \max_{M,\,\tau^R,\,\tau^S} W_{\hat{F},0}(M, \tau^R, \tau^S) - c(\tau^R + \tau^S)$$

$$\begin{aligned} \text{s.t.} \quad & 0 \leq M_{rs} \leq 1 \quad \text{for all } (r, s), \\ & \tau^R = p^R, \\ & p^S \leq \tau^S \leq p^S + p^0. \end{aligned}$$

In particular, the solution value of $(*')$ equals $\overline{W}$. There exists $\delta > 0$ and $\overline{\epsilon}'' > 0$ such that (120) holds for all $\epsilon < \overline{\epsilon}''$, any ϵ-approximation F, and any (M, τ^R, τ^S) in the feasible set of problem $(*')$:

$$\text{if } W_{F,c}(M, \tau^R, \tau^S) > \overline{W} - \delta, \text{ then } (M, \tau^R, \tau^S) \in \mathcal{N}. \tag{120}$$

(Suppose (120) fails. Then there exists a sequence $\delta_j \to 0$, a sequence $\epsilon_j \to 0$, a sequence (F_j), where F_j is an ϵ_j-approximation, and a sequence $(M_j, \tau_j^R, \tau_j^S)$

such that $W_{F_j,c}(M_j, \tau_j^R, \tau_j^S) > \overline{W} - \delta_j$ and $(M_j, \tau_j^R, \tau_j^S) \notin \mathcal{N}$. By (119), we can assume that $\tau_j^R = p^R$ and $\tau_j^S \geq p^S$ for all j. Thus, there exists a limit point $(\hat{M}, \hat{\tau}^R, \hat{\tau}^S) \notin \mathcal{N}$ with $\hat{\tau}^R = p^R$ and $\hat{\tau}^S \geq p^S$. Hence, $W_{\hat{F},0}(\hat{M}, \hat{\tau}^R, \hat{\tau}^S) - c(\tau^R + \tau^S) \geq \overline{W}$, contradicting that $(\overline{M}, p^R, p^S)$ is the unique solution to the problem (*').)

W.l.o.g., $\bar{\epsilon}'' < \delta$.

Using (118), one sees that, for all $\epsilon < \bar{\epsilon}''$,

$$W_{F,c}(\overline{M}, p^R, p^S) > \overline{W} - \delta. \tag{121}$$

Now consider any $\epsilon < \min\{\bar{\epsilon}, \bar{\epsilon}', \bar{\epsilon}''\}$ and any optimal mechanism-equilibrium pair $(M^*, \tau^{R*}, \tau^{S*})$. Then

$$W_{F,c}(M^*, \tau^{R*}, \tau^{S*}) \geq W_{F,c}(\overline{M}, p^R, p^S).$$

Together with (120) and (121) this implies $(M^*, \tau^{R*}, \tau^{S*}) \in \mathcal{N}$.

Hence, $(\tau^{R*}, \tau^{S*}) = (p^R, p^S)$ by (117). Given these participation rates, the welfare is

$$\begin{aligned} W_{F,c}(M^*, p^R, p^S) &= \frac{1}{n} \sum_{r+s \leq n} \binom{n}{r\ s} (p^R)^r (p^S)^s (1 - p^R - p^S)^{n-r-s} \omega_{r,s}(p^R, p^S) M^*_{rs} \\ &\quad - (p^R + p^S)c, \end{aligned}$$

implying by (114) that the unique best rule is $M^* = \overline{M}$.

Proof of (78) and (79).

As an auxiliary step, we establish a result on the multinomial probabilities (6). We will have to deal with higher order partial derivatives of functions of τ^R and τ^S. We will use the lower index $(l)\tau^S$ for the lth partial derivative with respect to τ^S, evaluated at $(\tau^R, \tau^S) = (F^R, F^S)$. Similar so for partial derivatives with respect to τ^R.

Lemma 7. *Let $l = 0, 1, \ldots$. Consider any tally (r, s) with $r + s \leq n - 1$. If $r + s \leq n - 2 - l$, then $\Pr_{(l)\tau^R}(r, s) = 0$ and $\Pr_{(l)\tau^S}(r, s) = 0$. Suppose that $r + s \geq n - 1 - l$. Then*

$$\Pr_{(l)\tau^R}(r, s) = \frac{(n-1)!}{(n-1-s-l)!s!} (F^R)^{n-1-s-l} (F^S)^s \binom{l}{n-1-r-s} (-1)^{n-1-r-s}$$

and

$$\Pr_{(l)\tau^S}(r, s) = \frac{(n-1)!}{r!(n-1-r-l)!} (F^R)^r (F^S)^{n-1-r-l} \binom{l}{n-1-r-s} (-1)^{n-1-r-s}.$$

Proof of Lemma 7. This is a straightforward computation. The only non-vanishing derivative of $(1-\tau^R-\tau^S)^{n-1-r-s}$ is the $(n-1-r-s)$th derivative. Thus both derivatives vanish if $r+s \le n-2-l$. In case $r+s \ge n-1-l$, using the Leibniz rule we find

$$\begin{aligned}&\Pr_{(l)\tau^R}(r,s)\\&=\frac{(n-1)!(-1)^{n-1-r-s}}{r!s!(n-1-r-s)!}\frac{r!(F^R)^{r-l+n-1-r-s}(F^S)^s}{(r-l+n-1-r-s)!}(n-1-r-s)!\binom{l}{n-1-r-s}.\end{aligned}$$

Similarly,

$$\begin{aligned}&\Pr_{(l)\tau^S}(r,s)\\&=\frac{(n-1)!(-1)^{n-1-r-s}}{r!s!(n-1-r-s)!}\frac{s!(F^R)^{r}(F^S)^{s-l+n-1-r-s}}{(s-l+n-1-r-s)!}(n-1-r-s)!\binom{l}{n-1-r-s}.\end{aligned}$$

Cancelling terms yields the expressions in the lemma.

We will use the shortcut

$$\begin{aligned}x &= \Pr_{(k-1)\tau^R}(t^*-1,q)=\Pr_{(k-1)\tau^S}(t^*-1,q)\\&=\frac{(n-1)!}{(t^*-1)!q!}(F^R)^{t^*-1}(F^S)^q(-1)^{k-1}.\end{aligned}$$

In order to compute the relevant higher-order derivatives of $(\check{c},\check{\tau}^R)$, we derive expressions for some higher-order derivatives of d^R, d^S, and $\rho(A)$ with respect to τ^R and/or τ^S, evaluated at $(\tau^R,\tau^S)=(F^R,F^S)$. Note that

$$\begin{aligned}d^S &= \sum_{r+s\le n-k-2}(M_{r,s}-M_{r,s+1})\Pr(r,s)\\&\qquad+\Pr(t^*-1,q)+\mathbf{1}_{t^*>1}\cdot M_{t^*-2,q}\Pr(t^*-2,q).\end{aligned}\tag{122}$$

From Lemma 7, all terms in $(\sum\dots)$ vanish if we take the lth derivative ($l\le k$) w.r.t. τ^S or τ^R at (F^R,F^S). Moreover, using the definitions of k and x,

$$\begin{aligned}\Pr_{(k)\tau^S}(t^*-1,q-1) &= -x\frac{q}{F^S},\\ \Pr_{(k)\tau^S}(t^*-2,q) &= -x\frac{t^*-1}{F^R},\\ \Pr_{(k)\tau^S}(t^*-1,q) &= x\frac{q}{F^S}k.\end{aligned}$$

Thus, for all $l \geq 0$,

$$d^S_{(l)\tau^S} = x \cdot \begin{cases} 0 & \text{if } 1 \leq l \leq k-2, \\ 1 & \text{if } l = k-1, \\ \frac{q}{F^S}k - M_{t^*-2,q}\frac{t^*-1}{F^R} & \text{if } l = k. \end{cases} \tag{123}$$

Next

$$d^S_{(1)\tau^R} = \mathbf{1}_{k=2} \Pr_{(1)\tau^R}(t^*-1, q) = \mathbf{1}_{k=2}x \tag{124}$$

The following formulas (125), (126), (127), and (128) are proved below.

For all $l \geq 0$,

$$d^R_{(l)\tau^S} = x \cdot \begin{cases} 0 & \text{if } 1 \leq l \leq k-2, \\ -1, & \text{if } l = k-1, \\ -(1 - M_{t^*-2,q})\frac{t^*-1}{F^R} - \frac{q}{F^S}(k-1) & \text{if } l = k, \end{cases} \tag{125}$$

Moreover,

$$d^R_{(1)\tau^R} = d^*\frac{t^*-1}{F^R} - \mathbf{1}_{k>2} \cdot d^*\frac{n-t^*}{F^S}. \tag{126}$$

For all $l \geq 0$,

$$\rho(A)_{(l)\tau^S} = x \cdot \begin{cases} 0, & \text{if } 1 \leq l \leq k-2, \\ 1, & \text{if } l = k-1, \\ -M_{t^*-2,q} \cdot \frac{t^*-1}{F^R} + \frac{q(k-1)}{F^S} & \text{if } l = k. \end{cases} \tag{127}$$

Finally,

$$\rho(A)_{(1)\tau^R} = \mathbf{1}_{k>2} \cdot d^*\frac{n-t^*}{F^S}. \tag{128}$$

Proof of (125). Note that

$$\begin{aligned} d^R = & \sum_{r+s \leq n-k-2} (M_{r+1,s} - M_{r,s}) \Pr(r,s) \\ & + \mathbf{1}_{t^*>1} \cdot (1 - M_{t^*-2,q}) \Pr(t^*-2, q) \\ & + \sum_{\hat{s}=0}^{k-2} \Pr(t^*-1, n-t^*-\hat{s}), \end{aligned} \tag{129}$$

where we have used the parameter $\hat{s} = n - t^* - s$ instead of s to write the last sum.[33]

[33] Observe that $k - 1 \leq n - t^*$ by construction of k.

Taking the lth ($l \leq k$) derivative of (129) and evaluating at $1-\tau^R-\tau^S=0$, all terms in the first row vanish because $(1-\tau^R-\tau^S)$ occurs in all terms with an exponent $> k$. Similarly, the second row vanishes unless $l = k$.

We begin by showing (125) for $1 \leq l \leq k-2$. Consider the lth derivative of the third row in (129), evaluated at $1-\tau^R-\tau^S=0$. Within the lth derivative expression as represented according to the general Leibniz product rule, only the term resulting from taking the $\hat{s}$th derivative of $(1-\tau^R-\tau^S)^{\hat{s}}$ (and taking the $(l-\hat{s})$th derivative of $(\tau^S)^{n-t^*-\hat{s}}$) does not vanish. Thus,

$$\begin{aligned}
& d^R_{(l)\tau^S} \\
&= \sum_{\hat{s}=0}^{l} \begin{pmatrix} n-1 \\ t^*-1 \ \hat{s} \end{pmatrix} (F^R)^{t^*-1} \frac{(n-t^*-\hat{s})!}{(n-t^*-\hat{s}-(l-\hat{s}))!} (F^S)^{n-t^*-\hat{s}-(l-\hat{s})} \\
&\qquad \cdot \hat{s}!(-1)^{\hat{s}} \begin{pmatrix} l \\ \hat{s} \end{pmatrix} \\
&= (F^R)^{t^*-1}(F^S)^{n-t^*-l} \sum_{\check{s}=0}^{l} \frac{(n-1)!}{(t^*-1)!(n-t^*-\check{s})!\check{s}!} \frac{(n-t^*-\check{s})!}{(n-t^*-l)!} \\
&\qquad \cdot (\check{s})!(-1)^{\check{s}} \begin{pmatrix} l \\ \check{s} \end{pmatrix} \\
&= (F^R)^{t^*-1}(F^S)^{n-t^*-l} \sum_{\check{s}=0}^{l} \frac{(n-1)!}{(t^*-1)!} \frac{1}{(n-t^*-l)!} \cdot (-1)^{\check{s}} \begin{pmatrix} l \\ \check{s} \end{pmatrix} \\
&= (F^R)^{t^*-1}(F^S)^{n-t^*-l} \frac{(n-1)!}{(t^*-1)!} \frac{1}{(n-t^*-l)!} \sum_{\check{s}=0}^{l} (-1)^{\check{s}} \begin{pmatrix} l \\ \check{s} \end{pmatrix} \\
&= 0. \qquad (130)
\end{aligned}$$

To show (125) for $l = k-1$, we use (129) and the general Leibniz product rule,

$$\begin{aligned}
d^R_{(k-1)\tau^S} &= \sum_{\hat{s}=0}^{k-2} \begin{pmatrix} n-1 \\ t^*-1 \ \hat{s} \end{pmatrix} (F^R)^{t^*-1} (F^S)^{\overbrace{n-t^*-\hat{s}-(k-1-\hat{s})}^{=n-t^*-(k-1)=q}} \\
&\qquad \cdot \frac{(n-t^*-\hat{s})!}{(n-t^*-(k-1))!} \hat{s}!(-1)^{\hat{s}} \begin{pmatrix} k-1 \\ \hat{s} \end{pmatrix} \\
&= \begin{pmatrix} n-1 \\ t^*-1 \ q \end{pmatrix} (k-1)!(F^R)^{t^*-1}(F^S)^q \cdot \sum_{\hat{s}=0}^{k-2} (-1)^{\hat{s}} \begin{pmatrix} k-1 \\ \hat{s} \end{pmatrix} \\
&= -x,
\end{aligned}$$

where we have used the identity $\sum_{\hat{s}=0}^{k-1} (-1)^{\hat{s}} \begin{pmatrix} k-1 \\ \hat{s} \end{pmatrix} = 0$.

To show (125) for $l = k$, note that

$$
\begin{aligned}
d^R_{(k)\tau^S} &= \mathbf{1}_{t^*>1,\ M_{t^*-2,q}=0} \begin{pmatrix} n-1 \\ t^*-2 \ \ q \end{pmatrix} (F^R)^{t^*-2}(F^S)^q k!(-1)^k \\
&\quad + \mathbf{1}_{q>0} \sum_{\hat{s}=0}^{k-2} \begin{pmatrix} n-1 \\ t^*-1\ \hat{s} \end{pmatrix} (F^R)^{t^*-1}(F^S)^{\overbrace{n-t^*-\hat{s}-(k-\hat{s})}^{=n-t^*-k=q-1}} \\
&\quad \cdot \frac{(n-t^*-\hat{s})!}{(n-t^*-k)!} \hat{s}!(-1)^{\hat{s}} \begin{pmatrix} k \\ \hat{s} \end{pmatrix} \\
&= -\mathbf{1}_{t^*>1,\ M_{t^*-2,q}=0} \cdot x \frac{t^*-1}{F^R} \\
&\quad + \mathbf{1}_{q>0} \begin{pmatrix} n-1 \\ t^*-1\ q-1 \end{pmatrix} k!(F^R)^{t^*-1}(F^S)^{q-1} \cdot \sum_{\hat{s}=0}^{k-2} (-1)^{\hat{s}} \begin{pmatrix} k \\ \hat{s} \end{pmatrix} \\
&= -\mathbf{1}_{t^*>1,\ M_{t^*-2,q}=0} \cdot x \frac{t^*-1}{F^R} - \mathbf{1}_{q>0} \cdot x \frac{q}{F^S}(k-1),
\end{aligned}
$$

where we have used the identity

$$
\sum_{\hat{s}=0}^{k-2} (-1)^{\hat{s}} \begin{pmatrix} k \\ \hat{s} \end{pmatrix} = -\sum_{\hat{s}=k-1}^{k} (-1)^{\hat{s}} \begin{pmatrix} k \\ \hat{s} \end{pmatrix} = -(-1)^{k-1}k - (-1)^k = (-1)^k(k-1).
$$

This completes the proof of (125).

Proof of (126). Using (129),

$$
\begin{aligned}
d^R_{(1)\tau^R} &= \mathbf{1}_{t^*>1} \begin{pmatrix} n-1 \\ t^*-1\ 0 \end{pmatrix} (t^*-1)(F^R)^{t^*-2}(F^S)^{n-t^*} \\
&\quad - \mathbf{1}_{k>2} \begin{pmatrix} n-1 \\ t^*-1\ 1 \end{pmatrix} (F^R)^{t^*-1}(F^S)^{n-t^*-1} \\
&= d^* \frac{t^*-1}{F^R} - \mathbf{1}_{k>2} \cdot d^* \frac{n-t^*}{F^S}.
\end{aligned}
$$

Proof of (127). Note that

$$
\begin{aligned}
\rho(A) &= \sum_{r+s\leq n-1} \binom{n-1}{r\ s} (\tau^R)^r (\tau^S)^s (1-\tau^R-\tau^S)^{n-1-r-s} M_{rs} \\
&= \left(\sum_{r+s\leq n-2-k} \cdots \right) \\
&\quad +\mathbf{1}_{q>0} \binom{n-1}{t^*-1\ q-1} (\tau^R)^{t^*-1} (\tau^S)^{q-1} (1-\tau^R-\tau^S)^k \\
&\quad +\mathbf{1}_{t^*>1,\ M_{t^*-2,q}=1} \binom{n-1}{t^*-2\ q} (\tau^R)^{t^*-2} (\tau^S)^{q} (1-\tau^R-\tau^S)^k \\
&\quad + \binom{n-1}{t^*-1\ q} (\tau^R)^{t^*-1} (\tau^S)^{q} (1-\tau^R-\tau^S)^{k-1} \\
&\quad + \sum_{r+s\geq n-1-k,\ r\geq t^*} \binom{n-1}{r\ s} (\tau^R)^r (\tau^S)^s (1-\tau^R-\tau^S)^{\overbrace{n-1-r-s}^{\leq k}}.
\end{aligned}
\tag{131}
$$

Taking the lth derivative ($1 \leq l \leq k-2$), only terms in the last sum can be non-vanishing because $l < k-1$. In the last sum, any term with $n-1-r-s > l$ vanishes, and any term with $s+(n-1-r-s) < l$ vanishes. Thus, using the general Leibniz product rule,

$$
\begin{aligned}
\rho(A)_{(l)\tau^S} &= \sum_{n-1-l\leq r+s\leq n-1,\ r\geq t^*,\ n-1-r\geq l} \binom{n-1}{r\ s} (F^R)^r \binom{l}{n-1-r-s} \\
&\quad \cdot \frac{s!}{(n-1-r-l)!} (F^S)^{\overbrace{s-(l-(n-1-r-s))}^{=n-1-r-l}} (n-1-r-s)!(-1)^{n-1-r-s} \\
&= \sum_{r=t^*}^{n-1-l} (F^R)^r \frac{(n-1)!}{r!(n-1-r-l)!} (F^S)^{n-1-r-l} \\
&\qquad \cdot \sum_{n-1-l-r\leq s\leq n-1-r} \binom{l}{n-1-r-s} (-1)^{n-1-r-s}.
\end{aligned}
$$

The last sum equals 0, as can be seen by using the variable $\check{s} = n-1-r-s$ instead of s. This shows (127) for $1 \leq l \leq k-2$.

The above computation also works if $l = k-1$ or $l = k$, showing that the fifth row on the right-hand-side of (131) can be ignored.

Consider $l = k - 1$. The $(k-1)$th derivative of the fourth row on the right-hand-side of (131) equals x, while the $(k-1)$th derivatives of the second and third rows vanish.

Consider $l = k$. The kth derivative of the second and third rows on the right-hand-side of (131) are obtained by taking the kth derivative of $(1 - \tau^R - \tau^S)^k$, yielding the terms

$$\begin{aligned}
& \mathbf{1}_{q>0} \begin{pmatrix} n-1 \\ t^* - 1 \; q-1 \end{pmatrix} (F^R)^{t^*-1} (F^S)^{q-1} k! (-1)^k \\
= \; & \mathbf{1}_{q>0} \frac{(n-1)!}{(t^*-1)!(q-1)!} (F^R)^{t^*-1} (F^S)^{q-1} (-1)^k \\
= \; & -\mathbf{1}_{q>0} \cdot x \frac{q}{F^S}.
\end{aligned} \tag{132}$$

and

$$\begin{aligned}
& \mathbf{1}_{t^*>1,\; M_{t^*-2,q}=1} \begin{pmatrix} n-1 \\ t^* - 2 \; q \end{pmatrix} (F^R)^{t^*-2} (F^S)^{q} k! (-1)^k \\
= \; & \mathbf{1}_{t^*>1,\; M_{t^*-2,q}=1} \cdot \frac{(n-1)!}{(t^*-2)!q!} (F^R)^{t^*-2} (F^S)^{q} (-1)^k \\
= \; & -\mathbf{1}_{t^*>1,\; M_{t^*-2,q}=1} \cdot x \frac{t^*-1}{F^R}.
\end{aligned}$$

The last remaining term is obtained by taking the kth derivative of the fourth row on the right-hand-side of (131). Using the Leibniz product rule, we take the $(k-1)$th derivative of $(1 - \tau^R - \tau^S)^{k-1}$ and the 1st derivative of $(\tau^S)^q$ and multiply with $\begin{pmatrix} k \\ 1 \end{pmatrix} = k$, yielding the term

$$\begin{aligned}
& \mathbf{1}_{q>0} \begin{pmatrix} n-1 \\ t^* - 1 \; q \end{pmatrix} (F^R)^{t^*-1} q (F^S)^{q-1} (k-1)! (-1)^{k-1} k \\
= \; & \mathbf{1}_{q>0} \frac{(n-1)!}{(t^*-1)!(q-1)!} (F^R)^{t^*-1} (F^S)^{q-1} (-1)^{k-1} k \\
= \; & \mathbf{1}_{q>0} \cdot x \frac{q}{F^S} k.
\end{aligned}$$

Summarizing this with (132), we obtain the last term in (127). This completes the proof of (127).

Proof of (128).

$$
\begin{aligned}
\rho(A)_{(1)\tau^R} &= \mathbf{1}_{k=2}\begin{pmatrix} n-1 \\ t^*-1\; q \end{pmatrix}(F^R)^{t^*-1}(F^S)^q(-1) \\
&\quad + \sum_{r+s=n-2,\; r\geq t^*}\begin{pmatrix} n-1 \\ r\; s \end{pmatrix}(F^R)^r(F^S)^s(-1) \\
&\quad + \underbrace{\sum_{r+s=n-1,\; r\geq t^*,\; r\geq 1}\begin{pmatrix} n-1 \\ r\; s \end{pmatrix} r(F^R)^{r-1}(F^S)^s}_{=\sum_{\hat{r}+s=n-2,\; \hat{r}\geq t^*-1}\begin{pmatrix} n-1 \\ \hat{r}\; s \end{pmatrix}(F^R)^{\hat{r}}(F^S)^s,\text{ where } \hat{r}=r-1} \\
&= -\mathbf{1}_{k=2}\begin{pmatrix} n-1 \\ t^*-1\; q \end{pmatrix}(F^R)^{t^*-1}(F^S)^q \\
&\quad + \begin{pmatrix} n-1 \\ t^*-1\; n-1-t^* \end{pmatrix}(F^R)^{t^*-1}(F^S)^{n-1-t^*} \\
&= -\mathbf{1}_{k=2}\cdot d^*\frac{n-t^*}{F^S} + d^*\frac{n-t^*}{F^S} \\
&= \mathbf{1}_{k>2}\cdot d^*\frac{n-t^*}{F^S}.
\end{aligned}
$$

Using (76) and (73), $\phi_{(1)\tau^S} = (0,0)^T$. Thus, (77) implies (78) for $l = 1$.

We proceed by induction over l to show (78). Suppose the formula in (78) holds for some l and we want to show it for $l+1$, where $l+1 < k$. Applying the chain rule and general Leibniz product rule to (77), it is sufficient to show $\phi_{(l')\tau^S} = (0,0)^T$ for all $l' \leq l+1$. Consider the first factor, $F^{-1}(1-\check{\tau}^R(\tau^S))$, of the first component of $\phi_{\partial\tau^S}$. By the chain rule and the induction hypothesis, the first l derivatives of this factor vanish at $\tau^S = F^S$. Hence, the first l derivatives of the first component of $\phi_{\partial\tau^S}$ vanish at m^*. Of the second component of $\phi_{\partial\tau^S}$, the term $F^{-1}(\tau^S)$ vanishes at $\tau^S = F^S$, and, because $l < k-1$, by (123), the first l derivatives of $d^S(M, \check{\tau}^R(\tau^S), \tau^S)$ also vanish at $\tau^S = F^S$. Hence, the first l derivatives of the second component of $\phi_{\partial\tau^S}$ vanish at m^*. This completes the induction.

From (77),

$$
\begin{pmatrix} \mathrm{d}^k\check{c}/\mathrm{d}(\tau^S)^k \\ \mathrm{d}^k\check{\tau}^R/\mathrm{d}(\tau^S)^k \end{pmatrix} = -\frac{\mathrm{d}^{k-1}}{\mathrm{d}(\tau^S)^{k-1}}\left(\phi^{-1}_{\partial c,\partial\tau^R}\cdot\phi_{\partial\tau^S}\right).
$$

Because $\phi_{(l')\tau^S}(m^*) = (0,0)^T$ for all $l' \leq k-1$ from the induction above,

$$
\begin{aligned}
\begin{pmatrix} \left.\frac{\mathrm{d}^k \check{c}}{\mathrm{d}(\tau^S)^k}\right|_{\tau^S=F^S} \\ \left.\frac{\mathrm{d}^k \check{\tau}^R}{\mathrm{d}(\tau^S)^k}\right|_{\tau^S=F^S} \end{pmatrix} &= -\left.\phi^{-1}_{\partial c, \partial \tau^R}\right|_{m^*} \cdot \phi_{(k)\tau^S} \\
&= -\frac{f(0)}{d^*} \begin{pmatrix} 0 & \frac{d^*}{f(0)} \\ -1 & -1 \end{pmatrix} \circ \frac{\mathrm{d}^{k-1}}{\mathrm{d}(\tau^S)^{k-1}} \left.\begin{pmatrix} F^{-1}(1-\check{\tau}^R(\tau^S)) d^R_{\tau^S} \\ F^{-1}(\tau^S) d^S_{\tau^S} + (F^{-1})'(\tau^S) d^S \end{pmatrix}\right|_{\tau^S=F^S}.
\end{aligned}
\tag{133}
$$

By (78) and the chain rule,

$$\frac{\mathrm{d}^{k-1}}{\mathrm{d}(\tau^S)^{k-1}} \left.\left(F^{-1}(1-\check{\tau}^R(\tau^S)) d^R_{\tau^S}\right)\right|_{\tau^S=F^S} = 0.$$

Moreover, using the general Leibniz product rule and (123),

$$\frac{\mathrm{d}^{k-1}}{\mathrm{d}(\tau^S)^{k-1}} \left.\left(F^{-1}(\tau^S) d^S_{\tau^S}\right)\right|_{\tau^S=F^S} = (k-1)\frac{1}{f(0)} d^S_{(k-1)\tau^S} \overset{(123)}{=} \frac{k-1}{f(0)} x.$$

Similarly,

$$\frac{\mathrm{d}^{k-1}}{\mathrm{d}(\tau^S)^{k-1}} \left.\left((F^{-1})'(\tau^S) d^S\right)\right|_{\tau^S=F^S} = \frac{1}{f(0)} x.$$

Thus, (133) implies

$$\begin{pmatrix} \left.\frac{\mathrm{d}^k \check{c}}{\mathrm{d}(\tau^S)^k}\right|_{\tau^S=F^S} \\ \left.\frac{\mathrm{d}^k \check{\tau}^R}{\mathrm{d}(\tau^S)^k}\right|_{\tau^S=F^S} \end{pmatrix} = -\frac{f(0)}{d^*} \begin{pmatrix} 0 & \frac{d^*}{f(0)} \\ -1 & -1 \end{pmatrix} \circ \begin{pmatrix} 0 \\ \frac{k}{f(0)} x \end{pmatrix},$$

yielding (79).

Proof of (84) and (85)

Consider

$$\check{\rho}'(\tau^S) = \frac{\partial \rho(A)}{\partial \tau^R} \cdot (\check{\tau}^R)'(\tau^S) + \frac{\partial \rho(A)}{\partial \tau^S}, \tag{134}$$

$$(\check{d}^R)'(\tau^S) = \frac{\partial d^R}{\partial \tau^R} \cdot (\check{\tau}^R)'(\tau^S) + \frac{\partial d^R}{\partial \tau^S}, \tag{135}$$

$$(\check{d}^S)'(\tau^S) = \frac{\partial d^S}{\partial \tau^R} \cdot (\check{\tau}^R)'(\tau^S) + \frac{\partial d^S}{\partial \tau^S}. \tag{136}$$

Using (134), (127) and (78),

$$(\check{\rho})^{(l)} \quad = \quad 0 \quad \text{for all } 1 \leq l \leq k-2. \tag{137}$$

Using (135), (125) and (78),

$$(\check{d}^R)^{(l)} \quad = \quad 0 \quad \text{for all } 1 \leq l \leq k-2. \tag{138}$$

Similarly, (136), (123), and (78),

$$(\check{d}^S)^{(l)} \quad = \quad 0 \quad \text{for all } 1 \leq l \leq k-2. \tag{139}$$

Applying the general Leibniz product rule to the second term on the right-hand side in equation (83), noting that the first derivative of the second factor in this term vanishes, and using (138), the only non-vanishing term in the $l-1$th derivative comes from taking the $l-1$th derivative of the first factor. Analogous reasoning applies to the third term on the right-hand side in (83). Applying the general Leibniz rule to the fourth and fifth terms and using (78), the $l-1$th derivative of the sum of these terms converges to the lth derivative of c. In summary,

$$\begin{aligned}
&\check{W}^{(l)} := \left.\frac{\mathrm{d}^l \check{W}}{\mathrm{d}(\tau^S)^l}\right|_{\tau^S=F^S} \\
&= (\check{\rho})^{(l)} E[\tilde{v}] + (\check{d}^R)^{(l)} \int_{v\geq 0} v \mathrm{d}G(v) + (\check{d}^S)^{(l)} \int_{v\leq 0} (-v) \mathrm{d}G(v) - \check{c}^{(l)}\gamma. \\
&\quad \text{for all } l = 1, \ldots, k.
\end{aligned} \tag{140}$$

Using this together with (137), (138) and (139), we find (84) for all $l < k-1$.

Using (134), (135), (136), and (78),

$$\begin{aligned}
(\check{\rho})^{(k-1)} &= \rho(A)_{(k-1)\tau^S} \overset{(127)}{=} d^S_{(k-1)\tau^S}, \\
(\check{d}^R)^{(k-1)} &= d^R_{(k-1)\tau^S} \overset{(125)}{=} -d^S_{(k-1)\tau^S}, \\
(\check{d}^S)^{(k-1)} &= d^S_{(k-1)\tau^S}.
\end{aligned}$$

Thus, (140) yields (84) for $l = k-1$.

In order to find $\check{W}^{(k)}$, we have to evaluate the right-hand side of (140) at $l = k$. Note that

$$\begin{aligned}
(\check{\rho})^{(k)} &= (\check{\tau}^R)^{(k)} \cdot \rho(A)_{(1)\tau^R} + \rho(A)_{(k)\tau^S} \\
&= x\mathbf{1}_{k>2} \frac{(n-t^*)k}{F^S} - xM_{t^*-2,q} \cdot \frac{t^*-1}{F^R} + x\frac{q(k-1)}{F^S},
\end{aligned} \tag{141}$$

where the first equation follows from the chain rule, the general Leibniz rule, and (78), and the second equation follows from (127), (128), and (79). Similarly,

$$
\begin{aligned}
(\check{d}^S)^{(k)} &= (\check{\tau}^R)^{(k)} \cdot d^S_{(1)\tau^R} + d^S_{(k)\tau^S}, \\
&= -x\mathbf{1}_{k=2}\frac{(n-t^*)k}{F^S} - xM_{t^*-2,q} \cdot \frac{t^*-1}{F^R} + x\frac{kq}{F^S}, \qquad (142)
\end{aligned}
$$

where the derivatives that occur on the right-hand side of (142) have been computed in (123), (124), and (79). Similarly,

$$
\begin{aligned}
(\check{d}^R)^{(k)} &= (\check{\tau}^R)^{(k)} \cdot d^R_{(1)\tau^R} + d^R_{(k)\tau^S}, \\
&= x \cdot \frac{(t^*-1)k}{F^R} - x\mathbf{1}_{k>2} \cdot \frac{(n-t^*)k}{F^S} - x(1-M_{t^*-2,q}) \cdot \frac{t^*-1}{F^R} - x\frac{q(k-1)}{F^S}, \qquad (143)
\end{aligned}
$$

where the derivatives that occur on the right-hand side of (143) have been computed in (125), (126), and (79).

Plugging (141), (142), (143), and (79) into (140) at $l = k$, the variable x cancels out and we find

$$
\begin{aligned}
\frac{\check{W}^{(k)}}{\check{c}^{(k)}} = & -\frac{f(0)}{k}\left(\mathbf{1}_{k>2}\frac{(n-t^*)k}{F^S} - M_{t^*-2,q} \cdot \frac{t^*-1}{F^R} + \frac{q(k-1)}{F^S}\right) E[\tilde{v}] \\
& -\frac{f(0)}{k}\left(\frac{(t^*-1)k}{F^R} - \mathbf{1}_{k>2} \cdot \frac{(n-t^*)k}{F^S}\right. \\
& \qquad \left. -(1-M_{t^*-2,q})\frac{t^*-1}{F^R} - \frac{q(k-1)}{F^S}\right) G^R \\
& -\frac{f(0)}{k}\left(-\mathbf{1}_{k=2}\frac{(n-t^*)k}{F^S} - M_{t^*-2,q} \cdot \frac{t^*-1}{F^R} + \frac{kq}{F^S}\right) G^S \\
& -\gamma.
\end{aligned}
$$

The terms with $\mathbf{1}_{k>2}$ and $\mathbf{1}_{k=2}$ can be summarized into a single term. Thus

$$
\begin{aligned}
\frac{\check{W}^{(k)}}{\check{c}^{(k)}} = & -\frac{f(0)}{k}\left(-M_{t^*-2,q} \cdot \frac{t^*-1}{F^R} + \frac{q(k-1)}{F^S}\right) E[\tilde{v}] \\
& -\frac{f(0)}{k}\left(\frac{(t^*-1)k}{F^R} - (1-M_{t^*-2,q})\frac{t^*-1}{F^R} - \frac{q(k-1)}{F^S}\right) G^R \\
& -\frac{f(0)}{k}\left(-\frac{(n-t^*)k}{F^S} - M_{t^*-2,q} \cdot \frac{t^*-1}{F^R} + \frac{kq}{F^S}\right) G^S \\
& -\gamma.
\end{aligned}
$$

Similarly, the terms with $1 - M_{t^*-2,q}$ and with $M_{t^*-2,q}$ can be summarized into a single term. Thus,

$$\begin{aligned}
\frac{\check{W}^{(k)}}{\check{c}^{(k)}} &= -\frac{f(0)}{k}\frac{q(k-1)}{F^S}E[\tilde{v}] \\
&\quad -\frac{f(0)}{k}\left(\frac{(t^*-1)k}{F^R} - \frac{t^*-1}{F^R} - \frac{q(k-1)}{F^S}\right)G^R \\
&\quad -\frac{f(0)}{k}\left(-\frac{(n-t^*)k}{F^S} + \frac{kq}{F^S}\right)G^S \\
&\quad -\gamma \\
&= -\frac{f(0)}{k}\left(\frac{(t^*-1)k}{F^R} - \frac{t^*-1}{F^R}\right)G^R - \frac{f(0)}{k}\left(-\frac{(n-t^*)k}{F^S} + \frac{q}{F^S}\right)G^S \\
&\quad -\gamma \\
&= f(0)\left(-\frac{t^*-1}{F^R} + \frac{t^*-1}{kF^R}\right)G^R + f(0)\left(\frac{n-t^*}{F^S} - \frac{n-t^*-k+1)}{kF^S}\right)G^S \\
&\quad -\gamma \\
&= f(0)\left(-\frac{t^*-1}{F^R} + \frac{t^*-1}{kF^R}\right)G^R + f(0)\left(\frac{n-t^*+1}{F^S} - \frac{n-t^*+1}{kF^S}\right)G^S \\
&\quad -\gamma \\
&= -f(0)\left(1 - \frac{1}{k}\right)\left((t^*-1)\frac{G^R}{F^R} - (n-t^*+1)\frac{G^S}{F^S}\right) - \gamma.
\end{aligned}$$

Thus, (85) holds.

Sketch of proof of claim in Footnote 25.

Outline: Considering an appropriate R-one sided rule and an appropriate S-one-sided rule, the terms due to types around 0 abstaining in the first-order welfare effects (145) and (146) are exactly opposite to each other. Thus, one of these terms is non-negative, so that the respective one-sided rule outperforms the best compulsory rule.

A pair $(\tau^R, 0)$ with $\tau^R > 0$ is an equilibrium in an R-one-sided mechanism M if and only if the type $F^{-1}(1-\tau^R) > 0$ is indifferent between participating, that is,

$$F^{-1}(1-\tau^R)d^R(M, \tau^R, 0) - c \quad = \quad 0. \tag{144}$$

In such equilibria, only tallies of the form $(r, 0)$ occur with positive probability. S-one-sided mechanisms are treated analogously.

Lemma 8 establishes existence and properties of equilibria with almost full one-sided participation in the R-one-sided linear rule $\mathbf{1}_{r\geq t^*}$. Lemma 9 is analogous for the S-one-sided case. We use the shortcuts introduced in (31).

Lemma 8. *For all c sufficiently close to 0, there exists an equilibrium $(\tilde{\tau}^R(c), 0)$ ($\to (F^R, 0)$ as $c \to 0$) of the mechanism $M^{R*} = \mathbf{1}_{r\geq t^*}$ that yields the welfare*

$$\lim_{c\to 0} W_{F,c}(M^{R*}, \tilde{\tau}^R(c), 0) = W^*.$$

Moreover,

$$\frac{d}{dc} W_{F,c}(M^{R*}, \tilde{\tau}^R(c), 0)\bigg|_{c=0} \quad = \quad f(0)\left(\frac{G^R}{F^R} - \overline{y}(t^*)\right) - (\gamma - G(0)). \tag{145}$$

To prove Lemma 8, one applies the implicit-function theorem to the equilibrium condition (144) in order to describe the equilibrium $(\tau^R, 0)$ as a function of c. The details are omitted.

To prove the following analogous result for the S-one sided rule $\mathbf{1}_{s<n-t^*}$, one replaces $F(v) \to 1 - F(-v)$, $g(v) \to g(-v)$ and $t^* \to n - t^* + 1$.

Lemma 9. *For all c sufficiently close to 0, there exists an equilibrium $(0, \tilde{\tau}^S(c))$ ($\to (0, F^S)$ as $c \to 0$) of $M^{S*} = \mathbf{1}_{s<n-t^*}$ that yields the welfare*

$$\lim_{c\to 0} W_{F,c}(M^{S*}, 0, \tilde{\tau}^S(c)) = W^*.$$

Moreover,

$$\frac{d}{dc} W_{F,c}(M^{S*}, 0, \tilde{\tau}^S(c))\bigg|_{c=0} \quad = \quad f(0)\left(\frac{G^S}{F^S} + \overline{y}(t^* - 1)\right) - G(0). \tag{146}$$

As $c \to 0$, each rule considered in Lemma 8 and Lemma 9 yields the same welfare W^* as the best compulsory rule. For the best compulsory rule, the first-order welfare effect of introducing a cost is $-\gamma$. A simple computation shows

$$\frac{G^R}{F^R} - \overline{y}(t^*) \quad = \quad -\left(\frac{G^S}{F^S} + \overline{y}(t^* - 1)\right).$$

Thus, either the left-hand side is ≥ 0 or the right-hand side is ≥ 0. Thus, either the left-hand side of (145) is $\geq -F^R > -\gamma$ or the left-hand side of (146) is $\geq -F^S > -\gamma$.