Equilibrium and Effectiveness of Two-Parameter Scoring Rules*

ADAM GILES†

PETER POSTL‡

Abstract

We study a cardinal model of voting with three alternatives where voters' von Neumann Morgenstern utilities are private information. We consider voting protocols given by two-parameter scoring rules, as introduced by Myerson (2002). For these voting rules, we show that all symmetric Bayes Nash equilibria are sincere, and have a very specific form. These equilibria are unique for a wide range of model parameters, and we can therefore compare the equilibrium performance of different rules. Computational results regarding the effectiveness of different scoring rules (where effectiveness is captured by a modification of the effectiveness measure proposed in Weber, 1978) suggest that those which most effectively represent voters' preferences allow for the expression of preference intensity, in contrast to more commonly used rules such as the plurality rule, and the Borda Count. While approval voting allows for the expression of preference intensity, it does not maximize effectiveness as it fails to unambiguously convey voters' ordinal preference rankings.

Keywords: scoring rule; voting; private information; mechanism design.

JEL classification: C72, D71, D78, D82.

*We are indebted to Tomas Sjöström for discussions that gave rise to this work (preliminary results are contained in Postl (2008) and Giles (2011)). We would also like to thank Siddhartha Bandyopadhyay, Michel Le Breton, Antonio Cabrales, Bhaskar Dutta, Shasi Nandeibam, Matías Núñez, Ronny Razin, Rajiv Sarin, Arunava Sen, participants at the XI Intl. Meeting of the Society of Social Choice and Welfare in New Delhi, seminar participants at the Institute of Economic Theory and Statistics at the Karlsruhe Institute of Technology, as well as two anonymous referees.

†Oxera, Park Central, 40/41 Park End Street, Oxford OX1 1JD, United Kingdom.

‡Department of Economics, University of Bath, Bath BA2 7AY, United Kingdom, p.postl@bath.ac.uk, Phone: +44 1225 38 3014.

"It is often the case in committee that the collective choice for chairman is X, even though a majority prefer Y, simply because a substantial minority strongly object to Y. Any theory of voting which does not allow for intensity of preference is certainly incomplete, and any voting system which does not permit its expression cannot be wholly satisfactory." (Meek, 1975)

1 Introduction

The opening quote raises the question whether a 'good' voting system should allow for, and be responsive to voters' expressions of how much they like the available candidates. It is this question that we aim to shed some light on with this paper. Early attempts to design voting systems that allow voters to express preference intensity can be found in the computer science and operations research literatures (see Meek, 1975, Nurmi, 1981, Cook and Kress, 1985, and Nurmi, 1993). Absent from many of these early contributions is a concern about strategic behavior by voters: they can be expected to overstate their preference intensity if that allows them to bias the collective decision in their favor. Therefore, attempts to answer the question of what constitutes a 'good' voting procedure when individuals have private information about their preferences must explore the extent to which voting rules can be responsive to individuals' expressions of preference intensity.

A general way of addressing this question would be to adopt a mechanism design approach in a setting where voters' preference intensity is captured by their privately observed Bernoulli utilities of the candidates. However, there are formidable technical challenges involved in designing mechanisms for environments where monetary transfers are not available as a tool for eliciting voters' private information.[1] To circumnavigate these problems, we study equilibrium voting behavior in a specific class of voting rules. The equilibria of the different rules are then compared according their effectiveness in representing the overall preferences of the electorate.

By asking which voting systems best represent voters' desires, we reprise a theme that originates with Weber (1978).[2] He addressed this question (albeit asymptotically in a setting with an arbitrarily large electorate) by proposing a measure of how effective a voting system is in representing the overall preferences of the electorate. In this paper, we propose a modification of Weber's effectiveness-measure to compare a wide range of voting rules (including many that allow the expression of preference intensity) in our setting with a finite number of voters who have private information about their preferences over candidates. They key difference with Weber (1978) is that there are instances of our setting where the game induced by each voting system features a unique symmetric voting equilibrium. This means that we can meaningfully compare the effectiveness of any two voting systems without having to worry that there might be other equilibria under which the relative ranking of their effectiveness-levels is reversed.[3]

The voting systems that we study in this paper are two-parameter scoring rules which include, as special cases, well-known voting procedures such as the plurality rule, the Borda count, and approval voting, among others.[4] Under a two-parameter scoring rule, each voter submits a vector whose components specify the scores that the voter assigns to the available alternatives. More specifically, each voter must assign a score of 1 to one alternative, a score of 0 to another, and a score of either x or y (where $0 \leq x \leq y \leq 1$) to the remaining alternative. After component-wise

[1] See e.g. Section 6 of Börgers and Postl (2009), and note the added difficulty that arises in voting environments from the multidimensional nature of voters' private information.

[2] We are grateful to Michel Le Breton for drawing our attention to this paper.

[3] The potential issue of equilibrium multiplicity is not addressed in Weber (1978).

[4] The notion of a two-parameter scoring rules for settings with three alternatives goes back to Myerson (2002).

summation of the score-vectors across all voters, the alternative with the highest score is chosen, and any ties are broken randomly and with equal probability.

For our study of two-parameter scoring rules, we adopt a Bayesian setting: each voter is characterized by a privately observed vector of three Bernoulli utilities, one for each alternative. The state of the world consists of the collection of all voters' utility vectors and is, from an ex ante perspective, modeled as a random variable with commonly known prior probability distribution. As is customary in the mechanism design literature, we assume that this utility distribution is symmetric with respect to voters, and neutral with respect to the alternatives (see e.g. Schmitz and Tröger, 2012). This latter property implies that, for a given voter, all ordinal rankings over the three alternatives are equally likely.[5] At the interim stage at which voting takes place, each voter is fully aware of her own utility vector, but not those of the other voters. The effectiveness of a scoring rule will be calculated at the ex ante stage. Following Weber (1978), effectiveness is defined, loosely speaking, as a ratio of the expected utilitarian welfare generated by the candidate actually elected under the scoring rule, and the expected utilitarian welfare of the socially optimal candidate. Both these expectations are taken with respect to the ex ante unknown state of the world. As there are two-parameter scoring rules that permit the expression of preference intensity (namely those with $x < y$), we can address the opening question whether the most effective voting systems will give voters the opportunity to express their intensity of preference.

To capture voting behavior in our setting, we characterize symmetric Bayes Nash equilibria of the game induced by two-parameter scoring rules. Our first main contribution is to show that the symmetric equilibrium strategy used by voters under each scoring rule involves sincere voting. That is, the score-vector submitted by a voter always reflects his true preference ordering of the candidates. Thus, it follows immediately that all scoring rules where voters do not have the option of expressing their preference intensity have a unique sincere Bayes Nash equilibrium.[6] This finding echoes the main result in Carmona (2012). He finds that all symmetric equilibria of ordinal scoring rules in his setting (which differs from ours) are generically sincere. The fact that all equilibria in our model feature sincere voting contrasts with the strategic voting equilibria found in some of the seminal contributions to the voting literature (such as Myerson and Weber, 1993). This contrast is noteworthy because empirical evidence for strategic voting appears scant (see e.g. Blais, 2002).

Our second contribution is to show that for all two-parameter scoring rules with $x < y$, an equilibrium strategy for any voter implies a threshold criterion for deciding whether to assign the lower score of x, or the higher score of y to his middle-ranked alternative. The threshold, given by a weighted average of the Bernoulli utilities associated with the voter's most and least preferred alternatives, may be degenerate (i.e. equal to the utility of the voter's favorite alternative). In this case, the voter will not use the opportunity to convey his preference intensity. We show in our second main result that in settings with three voters, and those with five or more voters (irrespective of whether this number is odd or even), the symmetric equilibrium voting strategy for *every* scoring rule with $x < y$ involves the expression of preference intensity.

It is important to emphasize that whenever preference intensity is conveyed in equilibrium, the precise value of the threshold that characterizes the equilibrium voting strategy depends on the parameters x and y of the scoring rule, and on the utility distribution. Obtaining an analytical expression for the equilibrium threshold as a function of these model parameters will, in general, be impossible. As a result, the value of the equilibrium threshold will have to be obtained by

[5]In our setting, in fact, symmetry and neutrality follow from the stronger assumption that the three components of a voter's utility vector are independently and identically distributed random variables. This assumption is also made in Kim (2012).

[6]Adopting the terminology in Apesteguia et al, 2011, we refer to scoring rules with $x = y$ as 'ordinal scoring rules'.

computational methods.[7] This is the reason why we report computational results regarding the effectiveness-levels generated by the different scoring rules. To obtain these results, we focus on the case of three voters, as this is the smallest number for which preference intensity is conveyed in equilibrium for all two-parameter scoring rules that permit its expression.

Our computational results indicate that the plurality rule and negative voting are the least effective two-parameter scoring rules. The most effective rules tend to be those that feature a relatively small x-value and a large y-value. Such rules allow voters to convey both their full ordinal ranking, as well as some degree of preference intensity. Whilst approval voting is significantly more effective than the plurality rule and negative voting, it is dominated by the best ordinal rule.[8] This is not surprising, because approval voting, while allowing the expression of preference intensity, does not give voters the opportunity to express unambiguously their respective ordinal rankings. In line with the opening quote by Meek (1975), we find that there is a gain in moving from the most effective ordinal rule to the most effective two-parameter scoring rule, albeit a rather small one. Therefore, and in light of the added complexity for voters in real-world elections, it may not actually be worthwhile introducing voting systems that permit the expression of preference intensity.[9]

Our work in this paper is related to two strands of the literature. The first focuses directly on two-parameter scoring rules in settings with three alternatives. For these rules, Buenrostro et al (2013) study voting behavior under weak informational assumptions, focusing on equilibrium in undominated strategies. We, in contrast, explore voting behavior under stronger informational assumptions which are, however, common in other areas of incentive theory. Myerson (2002) characterizes and compares equilibria under two-parameter scoring rules in a setting where the number of voters participating in the election is unknown. Our setting, in contrast, assumes a fixed and known number of voters, and considers a continuum of possible utility vectors per voter. Finally, Ahn and Oliveros (2010) explore how well two-parameter scoring rules aggregate information in a common value setting. The problem of efficient information aggregation is absent in our setting with private values.

The second branch of related literature focuses on mechanism design approaches to collective decision-making in the absence of transferable utility (see e.g. Börgers and Postl, 2009, Jackson and Sonnenschein, 2007, Kim, 2012, Miralles, 2012, and Schmitz and Tröger, 2012). In a setting similar to the one studied here, Apesteguia et al (2011) show that among deterministic mechanisms which consider only the ordinal aspects of voters' utility vectors, scoring rules are optimal when voters are assumed to report their preferences truthfully. Once strategic behavior and incentive compatibility constraints have to be accounted for, the problem of designing an optimal voting rule becomes significantly harder. Schmitz and Tröger (2012) address this problem by characterizing optimal strategy-proof voting rules for settings with two alternatives. However, analogue results for settings with three alternatives (be that under Bayesian or dominant strategy implementation) are so far absent from literature.[10] In independent work, Kim (2012) provides partial insights into this problem. Using the same information structure as the one employed here, he shows that if the Bernoulli utilities of three or more voters are distributed according to the uniform distribution, there exists a direct revelation mechanism that uses voters' full utility vectors (i.e. not just the or-

[7] While our second main result does not establish in general a unique equilibrium weight strictly between zero and one, we find that the equilibrium weight of every two-parameter scoring rule is unique for each of the utility distributions used in our computational results.

[8] For example, in the case where the utility distribution is uniform, the best ordinal rule is the Borda count.

[9] Note, however, that this finding may be driven by the fact that we have restricted our computational study to just three voters. A study of a larger number of voters would be needed to assess the robustness of this intuition.

[10] To get a sense of the difficulties involved in designing optimal strategy-proof voting rules with three alternatives, see Postl (2011) who analyzes a related, but simpler setting.

dinal rankings embodied therein) and which generates higher ex ante expected welfare than any mechanism that bases the collective choice on the voters' ordinal rankings alone.[11] This result tallies with our computational findings for the uniform distribution. However, our work differs from his in three important ways: first, we provide a characterization of symmetric Bayes Nash voting equilibria of all two-parameter scoring rules, and for any utility distribution; second, we compare, albeit computationally, the effectiveness of all two-parameter scoring rules for distributions other than the uniform; and third, we offer some insights into the connection between optimal two-parameter scoring rules and so called 'second best voting rules' (i.e. rules that maximize ex ante welfare among all incentive compatible direct revelation mechanisms).[12]

The remainder of this paper is structured as follows: In Section 2, we introduce the model and basic definitions. Section 3 contains our characterization of symmetric Bayes Nash equilibria for all two-parameter scoring rules and any number of voters. In Section 4, we present our computational results comparing the effectiveness-levels of two-parameter scoring rules. Section 5 offers a brief conclusion. The Appendix in Section 6 contains all longer proofs.

2 The Model

2.1 Basic set-up

Preferences. There are $n+1$ ($n \in \mathbb{N}$) voters who must choose collectively one alternative from the set $K \equiv \{A, B, C\}$. Each voter $i \in I \equiv \{1, 2, \ldots, n+1\}$ has a von Neumann Morgenstern utility function that represents his preferences over lotteries on K. We denote by $\boldsymbol{u}^i \equiv (u_A^i, u_B^i, u_C^i)$ the vector of Bernoulli utilities that voter i assigns to alternatives A, B and C, resp. We refer to $\boldsymbol{u}^i$ as voter i's *type*, and normalize utilities so that $\boldsymbol{u}^i \in [0,1]^3$ for all $i \in I$.[13] Each voter's type implies a preference ordering of the three alternatives in K, and there are six possible preference orderings here. For example, if voter i's type $\boldsymbol{u}^i$ is such that (s.t.) $u_B^i > u_A^i > u_C^i$, then $B \succ_i A \succ_i C$, where $\succ_i$ denotes i's ordinal preference relation on K. In what follows, it will be convenient to denote by u_1^i voter i's utility from his highest-ranked alternative. Similarly, denote by u_2^i voter i's utility from his middle-ranked alternative, and by u_3^i the utility from his lowest-ranked alternative.

Information structure. We assume that each type $\boldsymbol{u}^i$ is a random variable whose realization is observed only by voter i. The types of the $n+1$ voters are stochastically independent and identically distributed (i.i.d.). We assume furthermore that for each voter i, the three Bernoulli utilities u_A^i, u_B^i and u_C^i are drawn independently from the unit interval $[0,1]$ according to a distribution G with continuous and strictly positive density g on $(0,1)$. As a consequence, the probability that a voter has a particular ordinal ranking is the same for all six possible ordinal rankings. The above features of the von Neumann Morgenstern utility functions, and the joint distribution of types $(\boldsymbol{u}^1, \ldots, \boldsymbol{u}^{n+1})$, are common knowledge among the voters. Observe that the information structure assumed here makes our setting symmetric in the following two ways:

[11] The proof is constructive and builds on the direct mechanism equivalent of the Bayes Nash equilibrium voting rule that we characterize in this paper.

[12] Our tentative exploration of second best voting rules is related to Hortala-Vallve (2009) who, albeit in a different setting, characterizes the class of strategy-proof social choice rules which satisfy an appropriate differentiability condition.

[13] Our normalization of utilities is without loss of generality. This is because our model features state-dependent expected utility functions with a continuum of states of the world that determine the voters' utilities. With state-dependent expected utility, each voter's von Neumann Morgenstern utility function is unique up to a positive affine transformation with state-dependent intercept and state-independent slope (see Ritzberger, 2002 for details). As a result, we can find, for each voter $i \in I$, an appropriate intercept and slope parameter so that the utilities u_A^i, u_B^i and u_C^i are in $[0,1]$ in every state of the world.

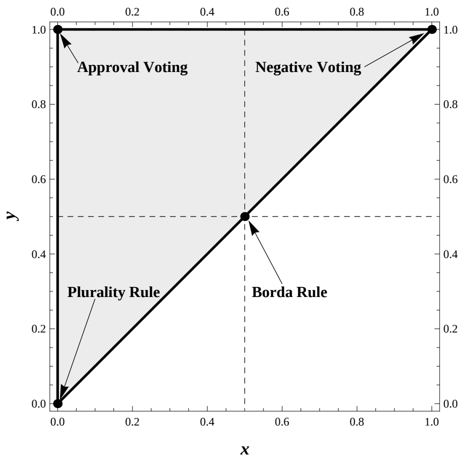


Figure 1: The set of all (x, y)-scoring rules

Symmetry with respect to (w.r.t.) voters, because each type $\boldsymbol{u}^i$ (with $i \in I$) is drawn from the same distribution on $[0,1]^3$.

Symmetry w.r.t. alternatives, because each component u_k^i of $\boldsymbol{u}^i$ (with $k \in K$) is drawn independently from the same distribution G on $[0,1]$.

2.2 Two-parameter scoring rules

In the setup described in Section 2.1, we study and compare voting mechanisms that are scoring rules. Under a scoring rule, each voter is asked to assign a score to every alternative $k \in K$. The scores assigned to each alternative are then added up, and the alternative with the highest aggregate score is chosen. In case of a tie for the highest score, an alternative is chosen randomly from amongst those with the highest score, each with equal probability. As in Myerson (2002), we consider specifically the family of (x, y)-scoring rules, which is characterized by two parameters x and y s.t. $0 \leq x \leq y \leq 1$. Given an (x, y)-scoring rule, each voter must choose a three-vector of scores that is a permutation of either $(1, x, 0)$ or $(1, y, 0)$. That is, each voter must give a score of 1 to one of the three alternatives in K, a score of 0 to one of the other two alternatives in K, and a score of either x or y to the remaining alternative. Many well-known voting rules are special cases of (x, y)-scoring rules. The gray shaded area in Fig. 1 below illustrates the set of all (x, y)-scoring rules, and highlights well-known special cases. For example:

1. **Plurality Rule**: $x = y = 0$. Each voter must choose between three different score-vectors, given by the permutations of $(1, 0, 0)$.

2. **Negative Voting**: $x = y = 1$. Each voter must choose between three different score-vectors, given by the permutations of $(1, 1, 0)$.

3. **Borda Rule**: $x = y = 0.5$. Each voter must choose between six different score-vectors, given by the permutations of $(1, 0.5, 0)$.

4. **Approval Voting**: $x = 0$ and $y = 1$. Each voter must choose between twelve different score-vectors, given by the permutations of $(1, 0, 0)$ and $(1, 1, 0)$.

In order to describe (x,y)-scoring rules more formally, we denote by $\Sigma_{x,y}$ the set of all score-vectors that a voter can choose from under a given (x,y)-scoring rule. A generic score-vector submitted by voter $i \in I$ is denoted by $\boldsymbol{\sigma}^i \equiv (\sigma_A^i, \sigma_B^i, \sigma_C^i)$, where σ_k^i is the score that voter i assigns to alternative k. As (x,y)-scoring rules implement a collective choice on the basis of the highest aggregate score, we define the aggregate score-vector as follows:

Definition 1. *Given any subset $\mathfrak{I} \subseteq I$ of $m \equiv |\mathfrak{I}|$ voters, the **aggregate score-vector $\boldsymbol{s}_m$** across voters in $\mathfrak{I}$ is the sum of their individual score-vectors $\boldsymbol{\sigma}^j$: $\boldsymbol{s}_m \equiv (s_m^A, s_m^B, s_m^C) = (\sum_{j\in\mathfrak{I}} \sigma_A^j, \sum_{j\in\mathfrak{I}} \sigma_B^j, \sum_{j\in\mathfrak{I}} \sigma_C^j)$. We denote by $S_{x,y}^m$ the set of all aggregate score-vectors $\boldsymbol{s}_m$ that can arise for the subset $\mathfrak{I}$ of voters. The set $S_{x,y}^m$ is obtained by adding up, for every profile of m individual score-vectors in $\Sigma_{x,y}^m$, the scores associated with each alternative k.*[14]

Using the notion of aggregate score-vectors, we can introduce formal notation for the probability distribution over outcomes induced by an (x,y)-scoring rule due to the possibility of ties for the highest aggregate score. We use this notation extensively in the remainder of the paper:

Definition 2. *For every subset $\mathfrak{I} \subseteq I$ of $m \equiv |\mathfrak{I}|$ voters, and every aggregate score-vector $\boldsymbol{s}_m \in S_{x,y}^m$, an (x,y)-scoring rule induces a **probability distribution over the set of alternatives** K: $\boldsymbol{\delta}(\boldsymbol{s}_m) \equiv (\delta_A(\boldsymbol{s}_m), \delta_B(\boldsymbol{s}_m), \delta_C(\boldsymbol{s}_m))$. In particular, the probability $\delta_A(\boldsymbol{s}_m)$ associated with alternative A is:*

$$\delta_A(\boldsymbol{s}_m) = \begin{cases} 1 & \text{if } s_m^A > s_m^B \text{ and } s_m^A > s_m^C \\ \frac{1}{2} & \text{if } s_m^A = s_m^B > s_m^C \text{ or } s_m^A = s_m^C > s_m^B \\ \frac{1}{3} & \text{if } s_m^A = s_m^B = s_m^C \\ 0 & \text{if } s_m^A < s_m^B \text{ or } s_m^A < s_m^C \end{cases}$$

The probabilities $\delta_B(\boldsymbol{s}_m)$ and $\delta_C(\boldsymbol{s}_m)$ of alternatives B and C (resp.) are defined analogously, with $\delta_A(\boldsymbol{s}_m) + \delta_B(\boldsymbol{s}_m) + \delta_C(\boldsymbol{s}_m) = 1$ for all $\boldsymbol{s}_m \in S_{x,y}^m$.

3 Equilibrium voting strategies

Every (x,y)-scoring rule gives rise to a Bayesian game with $n+1$ players. In this game, a pure strategy for voter i is a vector-valued function $\boldsymbol{v}^i : [0,1]^3 \to \Sigma_{x,y}$, $\boldsymbol{u}^i \mapsto \boldsymbol{v}^i(\boldsymbol{u}^i)$. That is, $\boldsymbol{v}^i$ specifies for every type $\boldsymbol{u}^i$ the score-vector in $\Sigma_{x,y}$ to be submitted by voter i. We denote by $v_k^i(\boldsymbol{u}^i)$ the score assigned to alternative $k \in K$ under strategy $\boldsymbol{v}^i$ when voter i's type is $\boldsymbol{u}^i$. Given the two types of symmetry inherent in our setting, it is natural to focus on Bayes Nash equilibria $(\boldsymbol{v}^1, \ldots, \boldsymbol{v}^{n+1})$ that are symmetric in the sense of the following two properties:

(S1) Each voter uses the same function $\boldsymbol{v}$ to select a score-vector on the basis of his type: $\boldsymbol{v}^i(\boldsymbol{u}^i) = \boldsymbol{v}(\boldsymbol{u}^i)$ for all $\boldsymbol{u}^i \in [0,1]^3$ and all $i \in I$. I.e. two voters with the same type submit the same score-vector.[15]

(S2) For all voters i, and any $\boldsymbol{u}^i$ and $\tilde{\boldsymbol{u}}^i$, where $\tilde{\boldsymbol{u}}^i$ is obtained from $\boldsymbol{u}^i$ by a permutation of its components, $\boldsymbol{v}$ maps $\tilde{\boldsymbol{u}}^i$ to $\tilde{\boldsymbol{\sigma}}^i$, which is obtained from $\boldsymbol{\sigma}^i = \boldsymbol{v}(\boldsymbol{u}^i)$ by applying the same permutation to the components of $\boldsymbol{\sigma}^i$. For example, let $\boldsymbol{u}^i = (u_A^i, u_B^i, u_C^i)$ and $\tilde{\boldsymbol{u}}^i = (u_B^i, u_A^i, u_C^i)$. If $\boldsymbol{v}(\boldsymbol{u}^i) = (1,x,0)$ then $\boldsymbol{v}(\tilde{\boldsymbol{u}}^i) = (x,1,0)$.

[14] The set $\Sigma_{x,y}^m \equiv \Sigma_{x,y} \times \ldots \times \Sigma_{x,y}$ denotes the m-times Cartesian product of $\Sigma_{x,y}$.

[15] The reverse is obviously *not* true: As there is only a finite number of score vectors in $\Sigma_{x,y}$, if two voters submit the same score-vector, they need not have the same type.

We refer to a strategy $\boldsymbol{v}$ that satisfies property S2 as a *symmetric strategy*, and to a symmetric Bayes Nash Equilibrium (cf. S1) in symmetric strategies (cf. S2) as a *fully symmetric Bayes Nash Equilibrium (FSE)*. To show that such an equilibrium exists, we have to consider the decision of any voter i as to which one of the available score-vectors in $\Sigma_{x,y}$ he wishes to submit when all other voters use the same symmetric strategy $\boldsymbol{v}$. For this purpose, we need to quantify voter i's beliefs about the score-vectors submitted by the other voters. Consider a voter $j \neq i$ whose submitted score-vector is $\boldsymbol{v}(\boldsymbol{u}^j)$. As voter i does not know j's type, he views the score-vector $\boldsymbol{v}(\boldsymbol{u}^j)$ as a random variable with sample space $\Sigma_{x,y}$. As both $\boldsymbol{v}$ and our setting are symmetric, voter i believes that every permutation of the score-vector $\boldsymbol{v}(\boldsymbol{u}^j) = (v_A(\boldsymbol{u}^j), v_B(\boldsymbol{u}^j), v_C(\boldsymbol{u}^j))$ arises with the same probability as $\boldsymbol{v}(\boldsymbol{u}^j)$. In particular, if the (x,y)-scoring rule features $x < y$, then all permutations of the score-vector $(1,x,0)$ arise with probability $\Pr[\boldsymbol{v}(\boldsymbol{u}^j) = (1,x,0)] \equiv p \in [0,1/6]$, and all permutations of $(1,y,0)$ arise with probability $\Pr[\boldsymbol{v}(\boldsymbol{u}^j) = (1,y,0)] = (1/6) - p$. Building on this observation, we can establish our first main result:

Proposition 1. *Any FSE of an (x,y)-scoring rule features a sincere voting strategy $\boldsymbol{v}$. That is, $v_1(\boldsymbol{u}^i) = 1$ and $v_3(\boldsymbol{u}^i) = 0$ for all $i \in I$ and all $\boldsymbol{u}^i \in [0,1]^3$, where $v_1(\boldsymbol{u}^i)$ and $v_3(\boldsymbol{u}^i)$, resp., denote the scores that $\boldsymbol{v}$ assigns to i's highest- and lowest-ranked alternatives when his type is $\boldsymbol{u}^i$.*

The proof of Proposition 1 can be found in the Appendix (see Section 6.1). To gain some intuition for this result, note that we can exploit the symmetry of our information structure to show that, from voter i's perspective, alternatives A, B and C are equally likely to be the 'collective choice' of the n other voters who use the same symmetric voting-strategy $\boldsymbol{v}$.[16] Put differently, in expected terms, the voting behavior of the other voters results in a uniform distribution over the set of alternatives K: $\mathrm{E}[\delta_A(\sum_{j \neq i} \boldsymbol{v}(\boldsymbol{u}^j))] = \mathrm{E}[\delta_B(\sum_{j \neq i} \boldsymbol{v}(\boldsymbol{u}^j))] = \mathrm{E}[\delta_C(\sum_{j \neq i} \boldsymbol{v}(\boldsymbol{u}^j))] = 1/3$. By submitting a score-vector $\boldsymbol{\sigma}^i$ that assigns a score of 1 to his favorite alternative, a score of either x or y to his middle-ranked alternative, and a score of 0 to his least preferred alternative, voter i generates an expected probability distribution $\mathrm{E}[\boldsymbol{\delta}(\sum_{j \neq i} \boldsymbol{v}(\boldsymbol{u}^j) + \boldsymbol{\sigma}^i)]$ over K that first-order stochastically dominates any distribution $\mathrm{E}[\boldsymbol{\delta}(\sum_{j \neq i} \boldsymbol{v}(\boldsymbol{u}^j) + \tilde{\boldsymbol{\sigma}}^i)]$ that prevails if he submits any non-sincere score-vector $\tilde{\boldsymbol{\sigma}}^i$.

Proposition 1 characterizes the unique FSE of (x,y)-scoring with $x = y$ (so called 'ordinal' scoring rules, as they give voters no scope to express any intensity of preference). It also echoes the main result in Carmona (2012) for a setting where only ordinal scoring rules are considered, and where voters' preferences depend not only on the alternative chosen from K, but also on the aggregate scores received by all the alternatives. He shows that symmetric BNE are sincere for almost all probability distributions from which voters' types (i.e. their ordinal rankings) are drawn. Proposition 1 implies furthermore that for (x,y)-scoring rules with $x < y$, a voter's only remaining decision is whether to assign the lower score of x, or the higher score of y to his middle-ranked alternative. As a corollary to Proposition 1, it is straightforward to establish the following result:

Corollary 1. *Any FSE of an (x,y)-scoring rule with $x < y$ features a sincere voting strategy $\boldsymbol{v}$ that assigns a score of $v_2(\boldsymbol{u}^i) \in \{x,y\}$ to the middle-ranked alternative of each voter i according to the following threshold criterion: for some $\alpha \in [0,1]$,*

$$v_2(\boldsymbol{u}^i) = \begin{cases} x & \text{if } u_2^i < \alpha u_1^i + (1-\alpha) u_3^i \\ y & \text{if } u_2^i > \alpha u_1^i + (1-\alpha) u_3^i \end{cases} \tag{1}$$

[16] It is important to emphasize here that the symmetric strategy $\boldsymbol{v}$ used by the other voters need not be sincere. To see this, suppose each voter $j \neq i$ assigns a score of 1 to his middle-ranked alternative, and a score of y to his favorite alternative. In this case, all permutations of $(1,x,0)$ arise with $\Pr[\boldsymbol{v}(\boldsymbol{u}^j) = (1,x,0)] = 0$, and all permutations of $(1,y,0)$ arise with $\Pr[\boldsymbol{v}(\boldsymbol{u}^j) = (1,y,0)] = 1/6$.

As a result of Corollary 1, voter i expects any other voter $j \neq i$ to submit the score-vector $(1,x,0)$ (and any permutation of it) with the following probability:

$$\Pr\left[\boldsymbol{v}(\boldsymbol{u}^j) = (1,x,0)\right] = \int_0^1 \int_0^{u_1^j} \int_{u_3^j}^{\alpha u_1^j + (1-\alpha)u_3^j} g(u_2^j)\,\mathrm{d}u_2^j g(u_3^j)\,\mathrm{d}u_3^j g(u_1^j)\,\mathrm{d}u_1^j \equiv p(\alpha) \tag{2}$$

Similarly, voter i expects j to submit each permutation of $(1,y,0)$ with the complementary probability $\Pr\left[\boldsymbol{v}(\boldsymbol{u}^j) = (1,y,0)\right] = (1/6) - p(\alpha)$. Note that $p(\alpha)$ defined in (2) is a monotonically increasing and differentiable function (i.e. $p'(\alpha) > 0$ for all α), with $p(0) = 0$ and $p(1) = 1/6$.

We now present the proof of Corollary 1, because it forms the basis on which we determine the equilibrium weight α in the threshold criterion in (1). The weight α is the final missing piece in the characterization of the FSE of (x,y)-scoring rules with $x < y$.

Proof. Consider voter i and suppose w.l.o.g. that his type $\boldsymbol{u}^i$ is s.t. $u_A^i > u_B^i > u_C^i$. His expected utility from submitting a score-vector $\boldsymbol{\sigma}^i = (1, \sigma_B^i, 0)$ (with $\sigma_B^i \in \{x,y\}$) is:[17]

$$U_i(\boldsymbol{\sigma}^i, \boldsymbol{u}^i) \equiv \sum_{k \in K} \mathrm{E}[\delta_k(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)] u_k^i, \tag{3}$$

where $\mathrm{E}[\delta_k(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)] = \sum_{\boldsymbol{s}_n \in S_{x,y}^n} \Pr[\boldsymbol{s}_n] \delta_k(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)$, and δ_k is as given in Definition 2.

It is optimal for voter i to submit the score-vector $\boldsymbol{\sigma}^i = (1,y,0)$ if $U_i((1,y,0),\boldsymbol{u}^i) > U_i((1,x,0),\boldsymbol{u}^i)$. Using (3), this inequality can be rearranged as follows:

$$u_B^i > \left(\frac{\mathrm{E}[\delta_A(\boldsymbol{s}_n + (1,x,0))] - \mathrm{E}[\delta_A(\boldsymbol{s}_n + (1,y,0))]}{\mathrm{E}[\delta_B(\boldsymbol{s}_n + (1,y,0))] - \mathrm{E}[\delta_B(\boldsymbol{s}_n + (1,x,0))]} \right) u_A^i$$
$$+ \left(1 - \frac{\mathrm{E}[\delta_A(\boldsymbol{s}_n + (1,x,0))] - \mathrm{E}[\delta_A(\boldsymbol{s}_n + (1,y,0))]}{\mathrm{E}[\delta_B(\boldsymbol{s}_n + (1,y,0))] - \mathrm{E}[\delta_B(\boldsymbol{s}_n + (1,x,0))]} \right) u_C^i \tag{4}$$

The condition in (4) states that voter i's best response to the symmetric strategy $\boldsymbol{v}$ used by all other voters is to assign the higher score of y to his middle-ranked alternative B if his utility u_B^i from it exceeds a weighted average of the utilities associated with his favorite and least-preferred alternatives. If the converse holds, a best response involves assigning the lower score of x to his middle-ranked alternative. Observe that the weight used in the average on the right-hand side of (4) is a number in $[0,1]$. This follows immediately from the fact that assigning the higher score of y to the middle-ranked alternative B shifts probability mass from alternative(s) A and/or C to B, relative to a situation where the lower score of x is assigned to B: $\delta_B(\boldsymbol{s}_n + (1,y,0)) - \delta_B(\boldsymbol{s}_n + (1,x,0)) \geq \delta_A(\boldsymbol{s}_n + (1,x,0)) - \delta_A(\boldsymbol{s}_n + (1,y,0)) \geq 0$ for all $\boldsymbol{s}_n \in S_{x,y}^n$. □

In order to pin down the specific value(s) of the weight α for which the symmetric voting strategy in Corollary 1 constitutes an equilibrium, we focus on equation (4) above. In particular, our interest centers on the weight attached to the utility of voter i's favorite alternative, which we refer to as the 'loss-gain-ratio':

[17] To ease notation in what follows, we write $\boldsymbol{s}_n$ for the aggregate score-vector $\sum_{j \neq i} \boldsymbol{v}(\boldsymbol{u}^j)$ across the n voters other than i.

Definition 3. *The **loss-gain-ratio** is given by the following expression:*

$$L(p(\alpha)) \equiv \frac{\mathrm{E}[\delta_A(\boldsymbol{s}_n + (1,x,0))] - \mathrm{E}[\delta_A(\boldsymbol{s}_n + (1,y,0))]}{\mathrm{E}[\delta_B(\boldsymbol{s}_n + (1,y,0))] - \mathrm{E}[\delta_B(\boldsymbol{s}_n + (1,x,0))]}$$

$$= \frac{\sum_{\boldsymbol{s}_n \in S^n_{x,y}} \Pr[\boldsymbol{s}_n] \left(\delta_A(\boldsymbol{s}_n + (1,x,0)) - \delta_A(\boldsymbol{s}_n + (1,y,0))\right)}{\sum_{\boldsymbol{s}_n \in S^n_{x,y}} \Pr[\boldsymbol{s}_n] \left(\delta_B(\boldsymbol{s}_n + (1,y,0)) - \delta_B(\boldsymbol{s}_n + (1,x,0))\right)} \tag{5}$$

Note that L is a differentiable function of α because each individual score-vector of the other voters occurs either with probability $p(\alpha)$ or probability $(1/6) - p(\alpha)$.

The numerator of the loss-gain-ratio captures the expected loss in the probability of voter i's favorite alternative when he assigns the higher score of y to his middle-ranked alternative, as opposed to the lower score of x. The denominator of the loss-gain-ratio, in turn, represents the expected gain in the probability of the middle-ranked alternative when i assigns it a score of y (rather than x). Before we use the loss-gain ratio in Definition 5 to characterize the value of the weight α in a FSE, it is instructive to consider an example of a sincere symmetric voting strategy of the form in Corollary 1 that does *not* constitute a FSE in our setting.

Example 1. *Consider the symmetric voting strategy presented in Weber (1978) (who studies asymptotic equilibria of approval voting (i.e. $(x,y) = (0,1)$) in a setting that includes our information structure as a special case but, unlike ours, features a large number of voters). Under Weber's strategy, each voter i chooses from $\Sigma_{0,1}$ the score-vector $(\sigma^i_A, \sigma^i_B, \sigma^i_C)$ that maximizes $\sigma^i_A(u^i_A - \bar{u}^i) + \sigma^i_B(u^i_B - \bar{u}^i) + \sigma^i_C(u^i_C - \bar{u}^i)$, where $\bar{u}^i$ denotes the average of voter i's Bernoulli utilities. This implies that each voter votes sincerely and assigns the lower score of* 0 *to his middle-ranked alternative if $u^i_2 < (u^i_1 + u^i_3)/2$, and the higher score of* 1 *if $u^i_2 > (u^i_1 + u^i_3)/2$. That is, under Weber's strategy each voter's middle-ranked alternative is awarded a score according to the threshold criterion in Corollary 1 with weight $\alpha = 0.5$.*

To see that Weber's strategy does not constitute a FSE of approval voting in our setting with a finite number of voters, suppose there are two voters i and j. If j uses Weber's strategy, then i expects every permutation of the score-vector $(1,0,0)$ to arise with probability $p(0.5) \in (0, 1/6)$ (where $p(0.5)$ is computed according to (2)), and every permutation of $(1,1,0)$ with probability $(1/6) - p(0.5)$. It is easy to verify that there are types of voter i for whom it is not a best response to vote according to Weber's strategy. In particular, we have $U_i((1,y,0),\boldsymbol{u}^i) < U_i((1,x,0),\boldsymbol{u}^i)$ for all $\boldsymbol{u}^i \in [0,1]^3$ s.t.:

$$\frac{u^i_1 + u^i_3}{2} < u^i_2 < \frac{5 - 6p(0.5)}{7 - 12p(0.5)} u^i_1 + \frac{6p(0.5)}{1 + 12p(0.5)} u^i_3$$

This inequality highlights that the weight $\alpha = 0.5$ associated with Weber's strategy differs from voter i's loss-gain-ratio: $L(p(0.5)) = (5 - 6p(0.5))/(7 - 12p(0.5)) > 0.5$ for any distribution G. It is this discrepancy between α and $L(p(\alpha))$ that disqualifies Weber's strategy as FSE of approval voting.

Example 1 highlights the central role played by the loss-gain-ratio in the characterization of the weights α associated with FSE voting strategies of (x,y)-scoring rules with $x < y$. In fact, in equilibrium the value of the weight α must be equal to the loss-gain-ratio of every voter. More formally, a FSE voting strategy must feature a weight $\alpha^* \in [0,1]$ s.t. $L(p(\alpha^*)) = \alpha^*$. In other words, the equilibrium weight α^* is a fixed point of the loss-gain-ratio. The equilibrium characterization therefore boils down to finding fixed points of the function $L(p(\alpha))$. Of particular interest in this context is the question whether there exists a fixed point α^* in the interior of the interval $[0,1]$. The reason is that only for $\alpha^* \in (0,1)$ do voters express in equilibrium their intensity of preference to the extent possible under an (x,y)-scoring rule with $x<y$.

A *sufficient condition* for the loss-gain-ratio to have interior fixed points for given model parameters x, y, n, and G is that $L(p(0))>0$ and $L(p(1))<1$. This is because whenever both these inequalities hold, we can appeal to the Intermediate Value Theorem to establish the existence of a value $\alpha^* \in (0,1)$ s.t. $L(p(\alpha^*))-\alpha^*=0$. The following result shows that the first inequality holds for all model parameters:

Lemma 1. *The loss-gain-ratio is positive at $\alpha=0$: $L(p(0))>0$ for all $x<y$, all n, and all G.*

To give the reader a sense of how this result is established, we present here the proof for the cases of two, three, and four voters, referring the reader to the Appendix (Section 6.2) for details of the proof with five or more voters.

Proof. Consider voter i and suppose w.l.o.g. that his type $\boldsymbol{u}^i$ is s.t. $u_A^i > u_B^i > u_C^i$. Given the definition of the loss-gain-ratio in (5), it suffices to show that there exists an aggregate score-vector $\boldsymbol{s}_n \in S_{x,y}^n$ which occurs with positive probability at $\alpha=0$ (i.e. $\Pr[\boldsymbol{s}_n]=(\frac{1}{6}-p(0))^n=1/6^n$) s.t. the probability of voter i's favorite alternative being chosen is strictly lower when he submits the individual score-vector $\boldsymbol{\sigma}^i=(1,y,0)$ than when he submits the score-vector $\boldsymbol{\sigma}^i=(1,x,0)$. In this case, the sum in the numerator of the loss-gain-ratio features at least one positive element: $\Pr[\boldsymbol{s}_n]\left(\delta_A(\boldsymbol{s}_n+(1,x,0))-\delta_A(\boldsymbol{s}_n+(1,y,0))\right)>0$ for at least one $\boldsymbol{s}_n \in S_{x,y}^n$.

(i) For two voters ($n=1$), the table below shows the (individual and aggregate) score-vector $\boldsymbol{s}_1=(y,1,0)$ which occurs with probability $\Pr[\boldsymbol{s}_1]=1/6$ at $\alpha=0$:

$\boldsymbol{s}_1$	$\boldsymbol{\sigma}^i$	$\boldsymbol{s}_1+\boldsymbol{\sigma}^i$	$\boldsymbol{\delta}(\boldsymbol{s}_1+\boldsymbol{\sigma}^i)$
$(y,1,0)$	$(1,x,0)$	$(1+y,1+x,0)$	$(1,0,0)$
$(y,1,0)$	$(1,y,0)$	$(1+y,1+y,0)$	$(\frac{1}{2},\frac{1}{2},0)$

Submission of the individual score-vector $\boldsymbol{\sigma}^i=(1,y,0)$ by voter i generates a shift in probability mass from alternative A to B, relative to submission of $\boldsymbol{\sigma}^i=(1,x,0)$: $\delta_A(\boldsymbol{s}_1+(1,x,0))-\delta_A(\boldsymbol{s}_1+(1,y,0))=\frac{1}{2}$.

(ii) For three voters ($n=2$), the table below shows the aggregate score-vector $\boldsymbol{s}_2=(y,1,1+y)$ which occurs with probability $\Pr[\boldsymbol{s}_2]=1/36$ at $\alpha=0$:

$\boldsymbol{s}_2$	$\boldsymbol{\sigma}^i$	$\boldsymbol{s}_2+\boldsymbol{\sigma}^i$	$\boldsymbol{\delta}(\boldsymbol{s}_2+\boldsymbol{\sigma}^i)$
$(y,1,1+y)$	$(1,x,0)$	$(1+y,1+x,1+y)$	$(\frac{1}{2},0,\frac{1}{2})$
$(y,1,1+y)$	$(1,y,0)$	$(1+y,1+y,1+y)$	$(\frac{1}{3},\frac{1}{3},\frac{1}{3})$

Submission of the individual score-vector $\boldsymbol{\sigma}^i=(1,y,0)$ by voter i generates a shift in probability mass from alternative A to B, relative to submission of $\boldsymbol{\sigma}^i=(1,x,0)$: $\delta_A(\boldsymbol{s}_1+(1,x,0))-\delta_A(\boldsymbol{s}_1+(1,y,0))=\frac{1}{6}$.

(iii) For four voters ($n=3$), the table below shows the aggregate score-vector $\boldsymbol{s}_3=(1+2y,2+$

$y,0)$ which occurs with probability $\Pr[\boldsymbol{s}_3] = 1/216$ at $\alpha = 0$:

$\boldsymbol{s}_3$	$\boldsymbol{\sigma}^i$	$\boldsymbol{s}_3 + \boldsymbol{\sigma}^i$	$\boldsymbol{\delta}(\boldsymbol{s}_3 + \boldsymbol{\sigma}^i)$
$(1+2y, 2+y, 0)$	$(1,x,0)$	$(2+2y, 2+x+y, 0)$	$(1,0,0)$
$(1+2y, 2+y, 0)$	$(1,y,0)$	$(2+2y, 2+2y, 0)$	$(\frac{1}{2}, \frac{1}{2}, 0)$

Submission of the individual score-vector $\boldsymbol{\sigma}^i = (1,y,0)$ by voter i generates a shift in probability mass from alternative A to B, relative to submission of $\boldsymbol{\sigma}^i = (1,x,0)$: $\delta_A(\boldsymbol{s}_1 + (1,x,0)) - \delta_A(\boldsymbol{s}_1 + (1,y,0)) = \frac{1}{2}$.

In the Appendix (see Section 6.2) we show that a similar argument establishes the result in Lemma 1 for five or more voters. □

We now show for which model parameters the second part of the sufficient condition for equilibrium existence holds (i.e. $L(p(1)) < 1$). This gives rise to our second main result:

Proposition 2. *For three voters, and for five or more voters (odd or even), any FSE of an (x,y)-scoring rule with $x < y$ features a voting strategy of the form in Corollary 1 with weight $\alpha^* \in (0,1)$ regardless of the distribution G from which voters' utilities are drawn.*

The implication of Proposition 2 is that voters express in equilibrium their intensity of preference to the extent allowed by scoring rules with $x < y$. To give the reader a sense of how this result is established, we state here the proof for three voters, referring the reader to the Appendix (Section 6.3) for details of the proof with five or more voters.

Proof. Consider voter i and suppose w.l.o.g. that his type $\boldsymbol{u}^i$ is s.t. $u^i_A > u^i_B > u^i_C$. We now establish that $L(p(1)) < 1$. Given the definition of the loss-gain-ratio in (5), it suffices to show that there exists an aggregate score-vector $\boldsymbol{s}_2 \in S^2_{x,y}$ which occurs with positive probability at $\alpha = 1$ (i.e. $\Pr[\boldsymbol{s}_n] = (p(1))^2 = 1/(6)^n$) s.t. voter i's middle-ranked alternative gains probability mass from his lowest-ranked alternative (and not just from his favorite alternative) when he submits the individual score-vector $\boldsymbol{\sigma}^i = (1,y,0)$ as opposed to the score-vector $\boldsymbol{\sigma}^i = (1,x,0)$. In this case, the numerator of the loss-gain-ratio is strictly smaller than the denominator: $\delta_B(\boldsymbol{s}_n + (1,y,0)) - \delta_B(\boldsymbol{s}_n + (1,x,0)) > \delta_A(\boldsymbol{s}_n + (1,x,0)) - \delta_A(\boldsymbol{s}_n + (1,y,0))$. Now consider the aggregate score-vector $\boldsymbol{s}_2 \in S^2_{x,y}$ in the table below, which occurs with probability $\Pr[\boldsymbol{s}_2] = 1/36$:

$\boldsymbol{s}_2$	$\boldsymbol{\sigma}^i$	$\boldsymbol{s}_2 + \boldsymbol{\sigma}^i$	$\boldsymbol{\delta}(\boldsymbol{s}_2 + \boldsymbol{\sigma}^i)$
$(x, 1, 1+x)$	$(1,x,0)$	$(1+x, 1+x, 1+x)$	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$
$(x, 1, 1+x)$	$(1,y,0)$	$(1+x, 1+y, 1+x)$	$(0,1,0)$

The table shows that for $\boldsymbol{s}_2 = (x, 1, 1+x)$, submission of the score-vector $\boldsymbol{\sigma}^i = (1,y,0)$ by voter i generates a shift in probability mass from alternatives A and C to B, relative to submission of $\boldsymbol{\sigma}^i = (1,x,0)$:

$$\delta_A(\boldsymbol{s}_2 + (1,x,0)) - \delta_A(\boldsymbol{s}_2 + (1,y,0)) = 1/3$$
$$\delta_B(\boldsymbol{s}_2 + (1,y,0)) - \delta_B(\boldsymbol{s}_2 + (1,x,0)) = 2/3$$

As alternative B's gain (which appears in the denominator of L) exceeds alternative A's loss (which appears in the numerator of L), it follows immediately that $L(p(1)) \in (0,1)$. Together with the fact that $L(p(0)) > 0$ by Lemma 1, we can infer by the Intermediate Value Theorem that there exists at least one equilibrium weight $\alpha^* \in (0,1)$ for every (x,y)-scoring rule with $x < y$ and every distribution G. In the Appendix (see Section 6.3), we show that a similar argument establishes the result in Proposition 2 for five or more voters. □

Limitations of Proposition 2. We first point out that Proposition 2 is essentially an existence result: it shows that there is a value α^* s.t. voters' equilibrium strategy involves the expression of preference intensity. However, it neither establishes uniqueness of α^*, nor does it shed light on how α^* varies with the model parameters x, y, n, and the distribution G. Secondly, Proposition 2 does not apply to settings with two and four voters. To address at least partially these limitations, we provide below a full characterization of the loss-gain-ratio for the cases of two and three voters, as the number of possible aggregate score-vectors in these cases is still manageable. Common to both settings is the fact that the continuum of (x,y)-scoring rules is partitioned in equilibrium into a finite number of equivalence classes, with rules in a given class generating the same loss-gain-ratio, and therefore the same value(s) α^*.[18] However, with two voters, we find that contrary to Proposition 2 not all (x,y)-scoring rules with $x < y$ give rise to a FSE in which preference intensity is expressed, even though the voting rule allows it. This feature also arises in settings with four voters, as we show by means of an example at the end of this section.

Proposition 3. *With two voters, the set of (x,y)-scoring rules with $x < y$ is partitioned into three categories depending on whether or not preference intensity is expressed in equilibrium:*

1. ***Unique FSE with preference intensity expressed for all** G. The following (x,y)-scoring rules have a unique FSE with $\alpha^* \in (0,1)$ for every distribution G:*
   - *(a) If $x = 0.5$, α^* is the unique solution of $\alpha = (1+4p(\alpha))/(1+8p(\alpha))$;*
   - *(b) If $x < 0.5$ and $y > 1-x$, α^* is the unique solution of $\alpha = (1+6p(\alpha))/(1+12p(\alpha))$;*
   - *(c) If $x < 0.5$ and $y = 1-x$, α^* is the unique solution of $\alpha = (5-6p(\alpha))/(7-12p(\alpha))$;*
2. ***Unique FSE where no preference intensity is expressed for any** G. For (x,y)-scoring rules with $y < 0.5$, and those with $x > 0.5$, the unique FSE features $\alpha^* = 1$;*
3. ***Possibility of multiple FSE/preference intensity expressed for some** G:*
   - *(a) If $y \in (0.5, 1-x)$, there exists a FSE with $\alpha^* = 1$ for all G. This is the unique FSE iff:*

$$6p(\alpha) > \frac{2-3\alpha}{1-2\alpha} \quad \text{for all } \alpha \in [\frac{2}{3},1); \tag{6}$$

     otherwise there exist additional FSE with $\alpha^{**} \in (0,1)$;
   - *(b) If $y = 0.5$, there exists a FSE with $\alpha^* = 1$ for all distributions G. Furthermore, this is the unique FSE iff G satisfies the following condition:*

$$6p(\alpha) > \frac{4-5\alpha}{1-2\alpha} \quad \text{for all } \alpha \in [\frac{4}{5},1); \tag{7}$$

     otherwise there exist additional FSE with $\alpha^{**} \in (0,1)$.

A sufficient condition for (6) and (7) to hold is that the distribution G is concave.

[18]Our results suggest that this is a general phenomenon for any finite n because K contains only three alternatives.

The proof of Proposition 3 is in the Appendix (see Section 6.4). The left-hand panel of Fig. 2 illustrates the partitions in Proposition 3. The middle panel of Fig. 2 provides an illustration of item 3 of Proposition 3 in the case where G is concave.

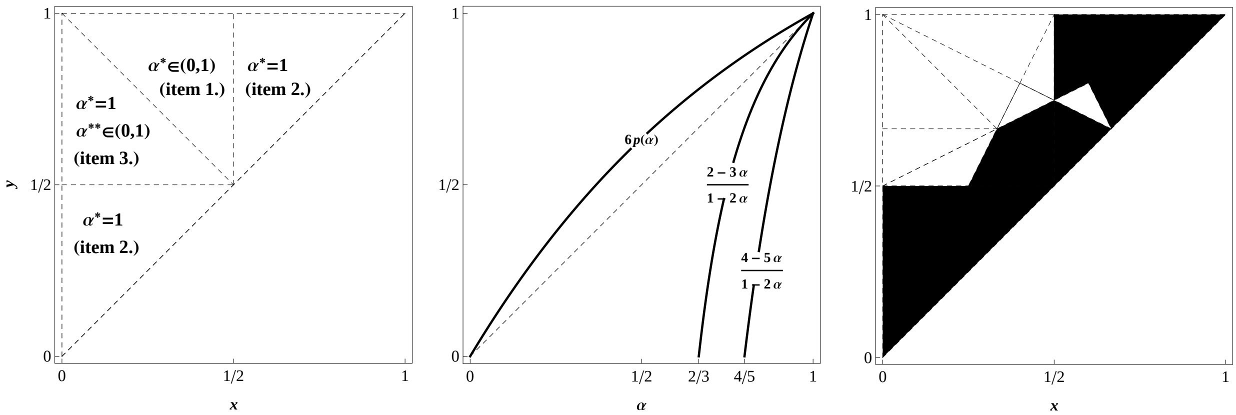


Figure 2: Partitions in Prop. 3 (left); item 3. of Prop. 3 (middle); Prop. 4 (right)

Proposition 4. *With three voters, the FSE voting strategy partitions the set of (x,y)-scoring rules with $x < y$ into 64 equivalence classes regardless of the distribution G from which the voters' Bernoulli utilities are drawn (a detailed list can be found in Table 11 on pages 38 - 41 of the Appendix). For 38 of these equivalence classes - all of which feature a monotone loss-gain ratio - there is a unique equilibrium weight α^* for every G (for an illustration of which rules are contained in these classes, see the black areas in right-hand panel of Fig. 2). For the remaining 26 equivalence classes - for which the loss-gain-ratio is non-monotonic - uniqueness of the equilibrium weight α^* is distribution-dependent.*[19]

The proof of Proposition 4 is in the Appendix (see Section 6.5). There, we list in a table all 64 equivalence classes that arise with three voters, along with the loss-gain-ratio for each equivalence class. For a graphical illustration of these equivalence classes, see the partitions indicated by dashed lines in Fig. 3 below. While our results in Propositions 3 and 4 show that the number of equivalence classes for a given number of voters is not distribution-dependent, it is clear that the value(s) of the equilibrium weight α^* pertaining to each equivalence class does depend on the distribution G through the loss-gain-ratio. Even in the relatively simple cases of two and three voters, it is impossible to obtain an analytical expression for the equilibrium weight α^* as a function of the model parameters x, y, and the distribution G. The reason is that even for the simplest parameterized distributions (such as $G(u) = u^b$, with $b > 0$), the equilibrium condition $L(p(\alpha^*)) = \alpha^*$ generates complicated expressions for which only computational solutions for α^* can be obtained. For three voters and uniform G - a setting in which there is a unique equilibrium weight α^* for every equivalence class because at least one of the sufficient conditions in footnote 19 is satisfied - Fig. 3 shows some of the numerically computed values of α^* as a function of x and y. In order to interpret the figure, note that every number superimposed on a polygon in Fig. 3

[19] Two (not mutually exclusive) sufficient conditions for uniqueness of α^* are: (i) the loss-gain-ratio associated with the equivalence class is a contraction mapping (i.e. $|L'(p(\alpha))p'(\alpha)| \leq k < 1$ for all $\alpha \in (0,1)$); (ii) the loss-gain-ratio associated with the equivalence class is either everywhere strictly convex, or everywhere strictly concave (i.e. $L''(p(\alpha))(p'(\alpha))^2 + L'(p(\alpha))p''(\alpha) \gtrless 0$ for all $\alpha \in (0,1)$).

represents the value α^* pertaining to all two-parameter scoring rules whose values x and y form a point within this polygon (exclusive of the dashed borders).[20]

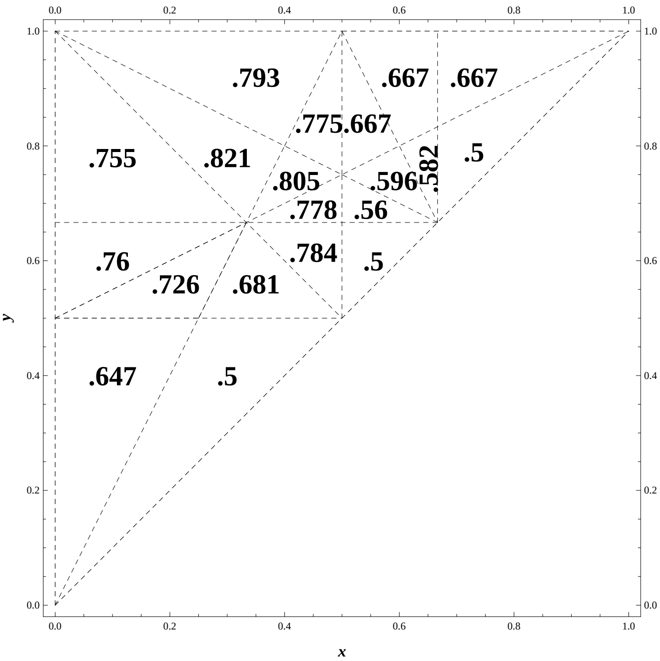


Figure 3: Equilibrium values of α^* for three voters and uniform G

We conclude this section with an example to show that with four voters, not all (x,y)-scoring rules with $x < y$ feature an equilibrium voting strategy in which preference intensity is expressed:

Example 2. *Let $n = 3$, and consider the (x,y)-scoring rule with $x = 0.4$ and $y = 0.45$. With the help of Mathematica, we can compute explicitly the loss-gain-ratio $L(p(\alpha))$ in (5):*

$$
\begin{aligned}
L(p(\alpha)) &= \frac{\frac{9}{2}\left(\left(\frac{1}{6}-p(\alpha)\right)^3+(p(\alpha))^3\right)+12\left(\left(\frac{1}{6}-p(\alpha)\right)^2 p(\alpha)+\left(\frac{1}{6}-p(\alpha)\right)(p(\alpha))^2\right)}{\frac{9}{2}\left(\left(\frac{1}{6}-p(\alpha)\right)^3+(p(\alpha))^3\right)+12\left(\left(\frac{1}{6}-p(\alpha)\right)^2 p(\alpha)+\left(\frac{1}{6}-p(\alpha)\right)(p(\alpha))^2\right)} \\
&= 1 \textit{ for all } \alpha \in [0,1].
\end{aligned}
$$

This loss-gain-ratio implies an equilibrium weight $\alpha^* = 1$, so that no preference intensity is expressed for any distribution G. In fact, in can be verified that with four voters, no (x,y)-scoring rule with $x \in \left(\frac{1}{3},\frac{1}{2}\right)$ and $y \in \left(x,\frac{1+x}{3}\right)$ involves the expression of preference intensity for any G.

4 Effectiveness of two-parameter scoring rules

4.1 Welfare notions and effectiveness

In this section, we build on our equilibrium characterization in Section 3 to compare (x,y)-scoring rules according to their effectiveness in selecting an alternative that is representative of voters'

[20]In order not to overload Fig. 3, we show the values α^* only for the 20 equivalence classes associated with the interiors of the various polygons. The remaining 44 equivalence classes are associated with the dashed boundaries between the polygons, as well as the intersections of the dashed boundaries.

preferences. The notion of effectiveness used here is due to Weber (1978). Expressed in terms from the mechanism design literature, Weber's effectiveness measure is defined as the welfare gain of a given (x,y)-scoring rule over random selection of an alternative from K, divided by the welfare gain of a 'first best' decision rule over random selection. The denominator of this ratio thereby represents the maximal welfare gain that one can hope for in the hypothetical scenario where incentive problems surrounding the revelation of voter-preferences are absent.

When we refer to the welfare of an (x,y)-scoring rule, we mean here expected utilitarian welfare, given by the sum of voters' ex ante expected utilities in the FSE of the scoring rule. Due to the symmetry of our setup in Section 2, where all voters are ex ante identical, it suffices to compute the ex ante expected utility of a representative voter and multiply it by the number of voters (i.e. by $n+1$) in order to obtain expected welfare. Furthermore, as all ordinal rankings of the alternatives in K are equally likely, we can compute the ex ante expected utility of any voter i by fixing a representative ordinal ranking (associated with utilities $u_1^i > u_2^i > u_3^i$), and then multiplying by six voter i's expected utility under the representative ranking.[21] Thus, expected welfare, denoted by $W(x,y)$, is:

$$W(x,y) \equiv 6(n+1)\left(\int_0^1\int_0^{u_1^i}\int_{u_3^i}^{\alpha^* u_1^i+(1-\alpha^*)u_3^i} U_i((1,x,0),\boldsymbol{u}^i)\, g(u_2^i)\,\mathrm{d}u_2^i\, g(u_3^i)\,\mathrm{d}u_3^i\, g(u_1^i)\,\mathrm{d}u_1^i \right.$$
$$\left. + \int_0^1\int_0^{u_1^i}\int_{\alpha^* u_1^i+(1-\alpha^*)u_3^i}^{u_1^i} U_i((1,y,0),\boldsymbol{u}^i)\, g(u_2^i)\,\mathrm{d}u_2^i\, g(u_3^i)\,\mathrm{d}u_3\, g(u_1^i)\,\mathrm{d}u_1^i\right)$$

In order to define formally the effectiveness measure of Weber (1978), we have to quantify the notions of 'first best' and 'random selection'. Both are instances of so called direct revelation mechanisms. Formally, a direct revelation mechanism (DRM) is a function $\boldsymbol{f}:[0,1]^{3(n+1)}\to\Delta(K)$, $\boldsymbol{u}\mapsto(f_A(\boldsymbol{u}),f_B(\boldsymbol{u}),f_C(\boldsymbol{u}))$, where $\Delta(K)$ is the set of probability distributions over K, and $f_k(\boldsymbol{u})$ is the probability that alternative $k\in K$ is chosen by the mechanism when the voters' type-profile is $\boldsymbol{u}\equiv(\boldsymbol{u}^1,\dots,\boldsymbol{u}^{n+1})$. Ex ante expected welfare associated with a DRM is:

$$\text{Welfare of } \boldsymbol{f} = \mathrm{E}[\sum_{k\in K} f_k(\boldsymbol{u})(\sum_{i\in I} u_k^i)] \tag{8}$$

When we speak of 'random selection', we mean here the DRM $\boldsymbol{f}^R$ that selects each alternative in K with equal probability, regardless of the voters' types:

Definition 4. *The DRM $\boldsymbol{f}^R$ implements* ***random selection*** *if $(f_A^R(\boldsymbol{u}),f_B^R(\boldsymbol{u}),f_C^R(\boldsymbol{u}))=(\frac{1}{3},\frac{1}{3},\frac{1}{3})$ for all $\boldsymbol{u}\in[0,1]^{3(n+1)}$.*

A 'first best' decision rule, in contrast, is a DRM that maximizes expected welfare in (8) in the hypothetical scenario where the types $\boldsymbol{u}^i$ of all $n+1$ voters, once realized, become observable before the collective choice has to be made:[22]

Definition 5. *A DRM $\boldsymbol{f}^{FB}$ is a* ***first best decision rule*** *if for every $k\in K$: $\sum_{i\in I}u_k^i > \max_{l\neq k}\{\sum_{i\in I}u_l^i\} \Rightarrow f_k^{FB}(\boldsymbol{u})=1$ and $\sum_{i\in I}u_k^i < \max_{l\neq k}\{\sum_{i\in I}u_l^i\} \Rightarrow f_k^{FB}(\boldsymbol{u})=0$.*

On the basis of Definition 5, ex ante expected welfare of a first best decision rule reduces to the expectation of the order statistic $\max\{w_A,w_B,w_C\}$, where $w_k\equiv\sum_{i\in I}u_k^i$ is a random variable with

[21] See (3) for the definition of $U_i((1,\sigma_B^i,0),\boldsymbol{u}^i)$, where $\sigma_B^i\in\{x,y\}$.

[22] I.e. a first best DRM is an ex ante classically efficient decision rule in the sense of Holmström and Myerson (1983).

support $[0, n+1]$, and any two random variables w_k, w_l ($k, l \in K$, $l \neq k$) are i.i.d.[23] We can now define formally the effectiveness of an (x,y)-scoring rule:

$$\text{Effectiveness} \equiv \frac{W(x,y) - \text{Welfare of } \boldsymbol{f}^R}{\text{Welfare of } \boldsymbol{f}^{FB} - \text{Welfare of } \boldsymbol{f}^R}$$

4.2 Computational results

In this section, we present computational results regarding the effectiveness levels of all (x,y)-scoring rules in a setting with three voters.[24] Our main aim is to gain some insight into which rule is most effective in selecting an alternative that represents voters' preferences. We would also like to know if (x,y)-scoring rules that allow for the expression of preference intensity are more effective than those that only allow voters to convey their ordinal rankings (these latter rules are the ones on the 45-degree line in the unit square, and the Borda rule is one of them).[25]

To obtain our computational results, we have selected specific distributions G from which voters' Bernoulli utilities are drawn, and then used Mathematica to compute the equilibrium weights α^* that characterize FSE voting strategies. To make as broad as possible the scope of our computational results, we used 25 different Beta-distributions for G. The density function $g(u;a,b)$ and the cumulative distribution function $G(u;a,b)$ of the Beta-distribution are parameterized by two shape parameters $a, b > 0$: $g(u;a,b) \equiv u^a(1-u)^b / \int_0^1 s^a(1-s)^b ds$ and $G(u;a,b) \equiv \int_0^u s^a(1-s)^b ds / \int_0^1 s^a(1-s)^b ds$ for $0 \leq u \leq 1$. In order to expedite the computation of the equilibrium weights α^* and their associated welfare levels, we generate our 25 Beta-distributions by varying separately each of the two shape-parameters a and b from 1 to 5 in increments of 1 (as opposed to using smaller increments). E.g. the uniform distribution corresponds to the case where $a = b = 1$.

Note that the effectiveness levels of different (x,y)-scoring rules can be compared meaningfully here because we have chosen our 25 Beta-distributions so that each rule gives rise to a unique equilibrium weight $\alpha^* \in (0,1)$. We have verified uniqueness by checking explicitly the sufficient conditions in footnote 19.[26] Our computational results indicate that the (x,y)-scoring rules which generate the highest average effectiveness across all 25 Beta-distributions are those in the top left corner of Fig. 4 (those yielding 89.06% effectiveness).[27] These rules allow voters to express their

[23]If we denote by H the distribution of the random variable w_k, then the distribution of the order statistic $\max\{w_A, w_B, w_C\}$ is $\bar{H}(w) \equiv \Pr[\max\{w_A, w_B, w_C\}] = (H(w))^3$. Note that analytical expressions for H are difficult to obtain for arbitrary distributions G. In our computational work below, we therefore compute first best welfare using Monte Carlo experiments in Mathematica 8.0 for Windows 7x64.

[24]The reason for focusing on three voters is that this is the smallest number of voters for which the FSE voting strategy in Corollary 1 involves $\alpha^* \in (0,1)$ for all $x < y$.

[25]Kim (2012) proves analytically for the case of the uniform distribution that ordinal rules do not maximize social welfare.

[26]By Proposition 4, we only have to verify uniqueness for (x,y)-scoring rules in the 26 equivalence classes that feature a non-monotonic loss-gain-ratio. For each of these equivalence classes, we have checked sufficient condition (i) in footnote 19 for every one of our 25 Beta-distributions. This showed that 20 of the 26 equivalence generate a loss-gain-ratio that is a contraction mapping. For the remaining six equivalence classes (namely those numbered 7., 15., 16., 28., 29., and 58. in Table 11 at the end of the proof of Proposition 4), condition (i) was violated for some of the 25 Beta-distributions. For these six equivalence classes, we then established uniqueness of α^* across all 25 Beta-distributions by verifying that each loss-gain-ratio is either strictly convex or strictly concave.

[27]As with Fig. 3 above, we show the effectiveness levels only for the 20 equivalence classes associated with the interiors of the various polygons.

intensity of preference because $x < y$. We also find that ordinal rules (i.e. those where $x = y$) do not yield maximal effectiveness.

In Table 1 below we present in the first column the effectiveness levels of well known special cases of (x, y)-scoring rules (the second column will be explained below in Section 4.3). The following pattern emerges: negative voting is the worst-performing two-parameter scoring rule, followed by plurality voting. Approval voting outperforms both of them, but is dominated by the best ordinal rule.[28] This is not surprising, given that approval voting, while allowing for the expression of preference intensity, does not allow voters to convey unambiguously their ordinal rankings. Under the most effective rules, voters can convey both their ordinal ranking *and* express their preference intensity.

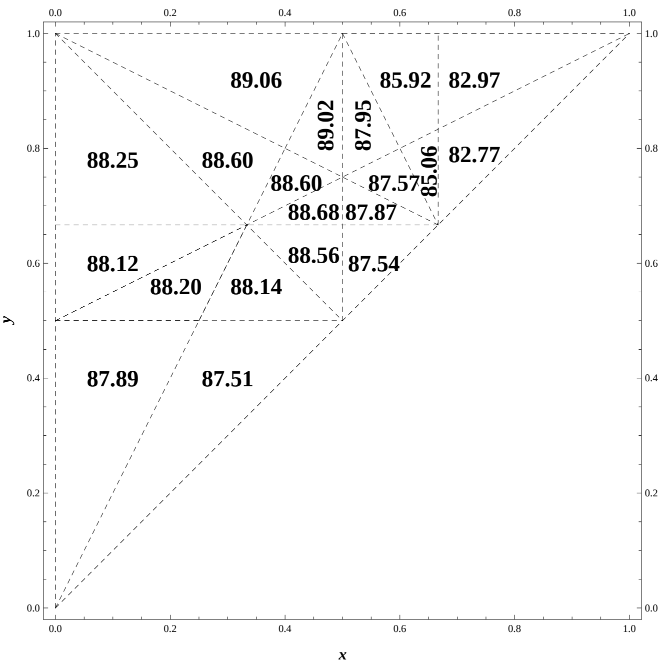


Figure 4: Average effectiveness in % across 25 Beta-distributions

Scoring Rule	(x, y)	Effectiveness w.r.t $\boldsymbol{f}^{FB}$	Effectiveness w.r.t $\boldsymbol{f}^*$
Negative Voting	$(1, 1)$	67.28%	73.83%
Plurality Rule	$(0, 0)$	76.76%	84.25%
Approval Voting	$(0, 1)$	83.71%	91.88%
Best Ordinal rule	$(0.6, 0.6)$	86.41%	94.83%
Most effective rule	$(0.4, 0.9)$	89.06%	97.74%

Table 1: Average effectiveness of well-known voting rules

We conclude our computational section with an attempt to understand better how maximum effectiveness varies with the distribution G. Our results reveal that of the 64 equivalence classes of (x, y)-scoring rules with $x < y$, only five maximize effectiveness (each class for a different subset of our 25 Beta-distributions). The first three classes are indicated by the numbers superimposed onto

[28] Among all rules with $x = y$, the one with $x = y = 0.6$ yields the highest average effectiveness across our 25 Beta-distributions.

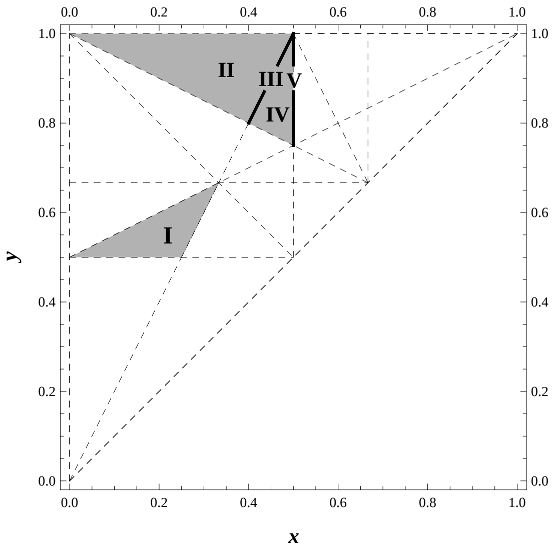


Figure 5: Graphical illustration of classes I-IV

the shaded gray areas in Fig. 5 (classes I, II, and IV are represented by a gray triangle each), while classes III and V are represented by thick black lines.[29] The figure confirms what we expected on the basis of average effectiveness levels, namely that ordinal rules do not generate maximal effectiveness for *any* of the 25 Beta-distributions.

To illustrate the types of distributions associated with each of the maximally effective equivalence classes, we list in Table 2 on page 20 the five classes along with the Beta-distributions for which each class is most effective. We also provide the indicative shape of the density under each Beta-distribution so as to convey a sense of how the maximally effective class of scoring rules varies with the concentration of probability mass under each distribution. The observation that emerges from Table 2 is that for distributions which place much of their probability mass on high utilities, the most effective (x,y)-scoring rules involve a value of x close to 0.5, and a value of $y > 1-0.5x$ (see rules IV and V); for distributions that spread their mass more evenly, the most effective rules involve a wider spread between the values of x and y (see rules II and III) than the spread that emerges under rules IV and V. However, as Fig. 4 indicates, the effectiveness levels of classes II, III, and IV are virtually identical (the effectiveness of class III, which was omitted from Fig. 4, is 89.04%). This suggests that any of these rules are likely to perform very well across a wide range of distributions. Only when probability mass is heavily concentrated on low values (as in $\mathrm{B}(1,5)$) is class I most effective.

Finally, to make it easier for the reader to glean how many of our 25 Beta-distributions are actually associated with each of the maximally effective equivalence classes, we show in Fig. 6 a histogram that conveys for how many distributions, and for which ones, each of the classes I-V in Fig. 5 is most effective (in the histogram, $\mathrm{B}(a,b)$ denotes the Beta-distribution with shape-parameters a and b).

[29] Dashed lines around shaded areas indicate that the boundaries are excluded. In terms of the 64 equivalence classes listed in Table 11 at the end of the proof of Proposition 4, classes II, III, IV, and V in the figure correspond to classes numbered 58., 57., 56, and 55., resp. in Table 11. Class I in the figure corresponds to class 15. in Table 11.

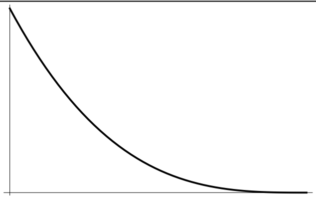
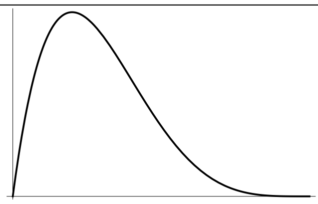
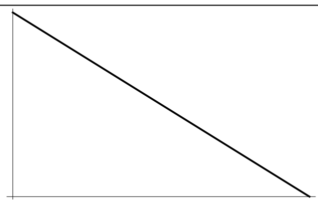
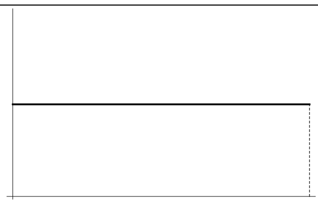
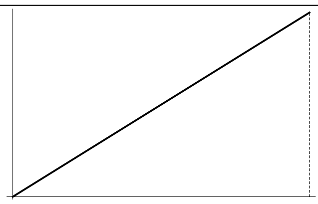
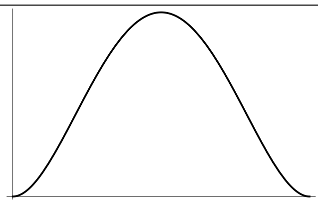
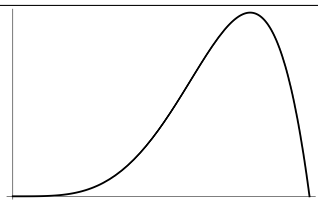
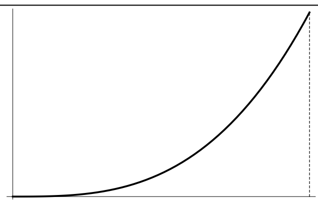
Indicative shape of $g(u;a,b)$	Parameters a and b	Most effective scoring rules
	$a=1, b=5$ $a=1, 2<b<5$	Class I Class II
	$2 \leq a < b \leq 5$	Class II
	$a=1, b=2$	Class II
	$a=1, b=1$	Class II
	$a=2, b=1$	Class II
	$a=b>1$	Class III
	$2 \leq b < a \leq 5$	Class IV
	$b=1, 2<a<5$	Class V

Table 2: Most effective (x,y)-scoring rule associated with each Beta-distribution

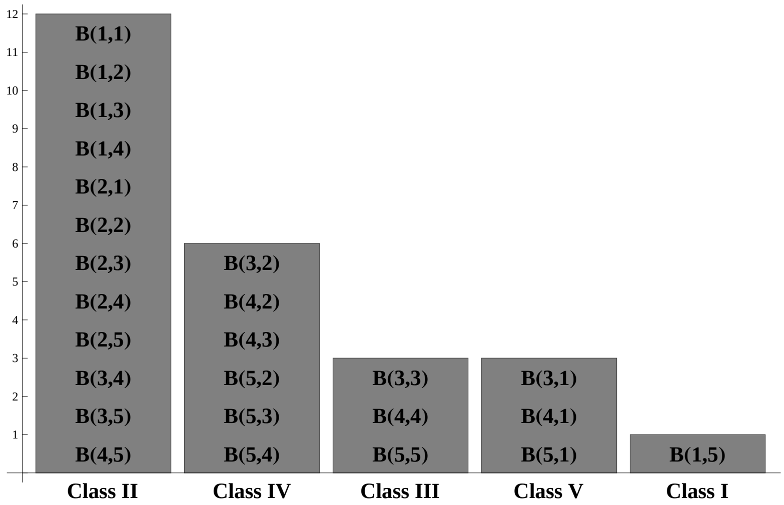


Figure 6: Histogram showing Beta-distributions associated with classes I-V

4.3 A superior measure of effectiveness

It is important to note that evaluating the effectiveness of two-parameter scoring rules against first best decision rules gives a somewhat unfavorable picture. The reason is that first best decision rules do not have to satisfy constraints to ensure that all voters report truthfully their types. Therefore, we would like to have a more realistic benchmark against which to compare two-parameter scoring rules. Ideally, we would like to compare them to so called ‘second best decision rules’ (i.e. rules that maximize ex ante welfare among all incentive compatible DRMs). However, a complete characterization of incentive compatible DRMs is beyond the scope of this paper. In order to make some progress in spite of this, we highlight here an equilibrium property of (x,y)-scoring rules that must also hold in any incentive compatible DRM: Consider any voter i, and pick two proportional types $\hat{\boldsymbol{u}}^i, \tilde{\boldsymbol{u}}^i \in [0,1]^3$ that reflect the same ordinal ranking of the alternatives. By ‘proportional’ we mean that $\tilde{\boldsymbol{u}}^i = c(\hat{u}^i_1, \hat{u}^i_2, \hat{u}^i_3)$ for some real-valued scalar $c > 0$. Observe that under the symmetric voting strategy in Corollary 1, if type $\hat{\boldsymbol{u}}^i$ submits the individual score-vector $\boldsymbol{\sigma}^i = (1, y, 0)$, then so will type $\tilde{\boldsymbol{u}}^i$. This implies that all types $\boldsymbol{u}^i \in [0,1]^3$ with the same ‘normalized type’ $(1, (\hat{u}^i_2 - \hat{u}^i_3)/(\hat{u}^i_1 - \hat{u}^i_3), 0)$ will submit the same score-vector $\boldsymbol{\sigma}^i$ as $\hat{\boldsymbol{u}}^i$.

It is important to emphasize that under state-dependent expected utility assumed in our model, it is not possible w.l.o.g. to replace voters’ original types $\boldsymbol{u}^i$ with their normalized types (only the Bernoulli utility u^i_3 of the lowest-ranked alternative can be normalized to zero w.l.o.g. for every type $\boldsymbol{u}^i \in [0,1]^3$; see also footnote 13). However, the standard characterization of incentive compatible DRMs (see e.g. Rochet, 1987) implies that *if* a DRM gives voters the incentive to reveal truthfully their types, *then* the mechanism can only be responsive to voters’ normalized types. In other words, an incentive compatible DRM must treat the type $\tilde{\boldsymbol{u}}^i$ the same as the type $\hat{\boldsymbol{u}}^i$. This is referred to in Hortala-Vallve (2009) as ‘bunching of proportional types’. If this was not the case, and a voter could expect a more favorable distribution over alternatives simply by exaggerating all components of his type-vector by the same factor, then it would be in the voter’s interest to do so. While incentive compatibility of a DRM may require constraints in addition to the bunching of proportional types, it is clear that no incentive compatible DRM can generate higher ex ante expected welfare than a mechanism that respects *only* the bunching requirement. Therefore, the latter mechanism should serve as a more realistic basis for computing the effectiveness of (x,y)-scoring rules than first best decision rules.

We propose now a particular DRM $\boldsymbol{f}^*$ that treats the same all voters who have the same normalized type. Consider a voter i who reports a type $\hat{\boldsymbol{u}}^i = (\hat{u}^i_1, \hat{u}^i_2, \hat{u}^i_3)$, with associated ordinal ranking $\succ_i$. On the basis of his report, voter i is assigned the following 'secondary type' $\boldsymbol{\tau}^i$:

$$\boldsymbol{\tau}^i \equiv (\tau^i_1, \tau^i_2, \tau^i_3) = \left(\mathrm{E}[u^i_1 | \succ_i], \theta_i \mathrm{E}[u^i_1 | \succ_i] + (1-\theta_i)\mathrm{E}[u^i_3 | \succ_i], \mathrm{E}[u^i_3 | \succ_i]\right)$$

where $\theta_i \equiv (\hat{u}^i_2 - \hat{u}^i_3)/(\hat{u}^i_1 - \hat{u}^i_3)$, and $\mathrm{E}[u^i_k | \succ_i]$ is the expectation of u^i_k conditional on the ordinal ranking associated with voter i's reported type $\hat{\boldsymbol{u}}^i$.[30] Given the secondary types $\boldsymbol{\tau}^i$ of all voters, let $\boldsymbol{f}^*$ select a probability distribution over K that is first best in the sense of Definition 5, but with respect to the profile of secondary types $\boldsymbol{\tau} \equiv (\boldsymbol{\tau}^1, \ldots, \boldsymbol{\tau}^{n+1})$ rather than the reported types. It is straightforward to show that $\boldsymbol{f}^*$ gives voters the incentive to reveal truthfully their ordinal rankings when all other voters do so.[31]

We would like to argue that the DRM $\boldsymbol{f}^*$ is a more realistic benchmark for evaluating the effectiveness of two-parameter scoring rules, because both $\boldsymbol{f}^*$ and the FSE of (x,y)-scoring rules bunch proportional types. We have computed numerically the ex ante expected welfare of $\boldsymbol{f}^*$ across all 25 Beta-distributions, and have used the results to obtain the average effectiveness of (x,y)-scoring rules relative to $\boldsymbol{f}^*$, instead of $\boldsymbol{f}^{FB}$.[32] The results are reported in the second column of Table 1, and they reveal that the most effective rule now comes very close to full effectiveness (Equivalence class II in Fig. 5 is again the most effective; in fact, the ranking of effectiveness-levels is not affected by using the benchmark $\boldsymbol{f}^*$ instead of $\boldsymbol{f}^{FB}$). This suggests that the lower effectiveness of class II relative to first best can be attributed almost entirely to the need to bunch proportional types.[33] Even though class II may not be a second best DRM, its very small loss in effectiveness relative to $\boldsymbol{f}^*$ suggests that it comes very close.

5 Conclusion

In this paper, we have characterized symmetric equilibria of two-parameter scoring rules. With three voters, these symmetric equilibria are unique for the 25 Beta-distributions considered here. We can therefore safely compare the performance of different scoring rules across these distributions. As a measure of performance, we have reprised the notion of effectiveness proposed in Weber (1978). Our results show that the plurality rule and negative voting are the least effective in representing the preferences of the electorate. Whilst approval voting performs much better than either of these rules, it does not perform as well as optimal (x,y)-scoring rules, which feature a relatively small x-value and large y-value. Our computational results suggest that voting rules which allow voters to express their strength of preference perform better than those which do not. In future work, we intend to explore further the characteristics of voting mechanisms that are optimal in the class of incentive compatible mechanisms (i.e. so called 'second best' decision rules). However, any such work will have to address the considerable mathematical difficulties that an analytical mechanism design approach to this question entails.

[30]Observe that both the reported type $\hat{\boldsymbol{u}}^i$ and the corresponding secondary type $\boldsymbol{\tau}^i$ give rise to the same normalized type: $(1, \theta_i, 0)$.

[31]Note, however, that we do not make any claims as to whether $\boldsymbol{f}^*$ is fully incentive compatible. Voters may still have an incentive to misrepresent their types in ways other than simply multiplying them by a constant.

[32]Ex ante expected welfare of $\boldsymbol{f}^*$ has been obtained by means of Monte Carlo experiments.

[33]This observation tallies with the very small welfare losses found in Börgers and Postl (2009), who studied a model where the Bernoulli utility of each agents' favorite alternative was 'normalized' to 1, and that of the least preferred alternative was normalized to 0. As a result, there were no proportional types to be bunched in their model.

6 Appendix

6.1 Proof of Proposition 1

Consider voter $i \in I$. From his perspective, the score-vector $\boldsymbol{v}(\boldsymbol{u}^j)$ submitted by any voter $j \neq i$ is a r.v. with sample space $\Sigma_{x,y}$ and generic realization $\boldsymbol{\sigma}^j = (\sigma_A^j, \sigma_B^j, \sigma_C^j) \in \Sigma_{x,y}$. To obtain the probability distribution of $\boldsymbol{v}(\boldsymbol{u}^j)$, note that the information structure in Section 2.1, together with the fact that the voting strategy $\boldsymbol{v}$ is symmetric in the sense of property S2 implies that any permutation $\hat{\boldsymbol{\sigma}}^j$ of a given score-vector $\boldsymbol{\sigma}^j \in \Sigma_{x,y}$ occurs with the same probability as $\boldsymbol{\sigma}^j$. In formulating his best response to the voting strategy $\boldsymbol{v}$ used by all voters $j \neq i$, voter i must compute the expected probability distribution $\mathrm{E}[\boldsymbol{\delta}(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)]$ over K implied by the (x,y)-scoring rule for the aggregate score-vector $\boldsymbol{s}_n + \boldsymbol{\sigma}^i$. Note that $\boldsymbol{s}_n = \sum_{j \neq i} \boldsymbol{v}(\boldsymbol{u}^j)$ is a r.v. whose distribution is derived from that of the individual score-vectors $\boldsymbol{v}(\boldsymbol{u}^j)$ submitted by voters $j \neq i$. In Lemmas 2 and 3 below, we derive for the r.v. $\boldsymbol{s}_n$ the relevant characteristics of its sample space $S_{x,y}^n$ and the probability distribution on it.

Lemma 2. *The probability that alternative k has the highest aggregate score is the same for all $k \in K$.*

Proof of Lemma 2. To show this, we introduce the following notation: The subset $S_k^n \subset S_{x,y}^n$ contains all aggregate score-vectors s.t. alternative k has the highest aggregate score. To ease notation, we omit the explicit reference to the values x and y associated with the given (x,y)-scoring rule: we write Σ instead of $\Sigma_{x,y}$, and S^n instead of $S_{x,y}^n$. Now pick any aggregate score-vector $\boldsymbol{s}_n \in S_A^n$. That is, the components of $\boldsymbol{s}_n$ are s.t. $\sum_{j \neq i} \sigma_A^j > \max\{\sum_{j \neq i} \sigma_B^j, \sum_{j \neq i} \sigma_C^j\}$. Consequently, the aggregate ranking of alternatives in K across all voters $j \neq i$ is: $A \succ_n B$ and $A \succ_n C$, where $\succ_n$ denotes the aggregate preference relation. Next, generate for each voter $j \neq i$ a new score-vector $\hat{\boldsymbol{\sigma}}^j$ by interchanging in the original score-vectors $\boldsymbol{\sigma}^j$ the scores assigned to alternatives A and B: $\hat{\boldsymbol{\sigma}}^j = (\sigma_B^j, \sigma_A^j, \sigma_C^j)$. Adding up these new score-vectors $\hat{\boldsymbol{\sigma}}^j$, we obtain an aggregate score-vector $\hat{\boldsymbol{s}}_n \equiv (\sum_{j \neq i} \sigma_B^j, \sum_{j \neq i} \sigma_A^j, \sum_{j \neq i} \sigma_C^j)$ under which alternative B has the highest score (i.e. $B \succ_n A$ and $B \succ_n C$). By virtue of the fact that the set Σ contains all permutations of $(1,x,0)$ and $(1,y,0)$, it follows that $\hat{\boldsymbol{s}}_n \in S_B^n \subset S^n$. Furthermore, $\Pr[\boldsymbol{v}(\boldsymbol{u}^j) = \hat{\boldsymbol{\sigma}}^j] = \Pr[\boldsymbol{v}(\boldsymbol{u}^j) = \boldsymbol{\sigma}^j]$ for every permutation $\hat{\boldsymbol{\sigma}}^j$ of $\boldsymbol{\sigma}^j$, it follows that the aggregate score-vectors $\boldsymbol{s}_n$ and $\hat{\boldsymbol{s}}_n$ occur with the same probability. By interchanging in all score-vectors $\boldsymbol{\sigma}^j$ the scores assigned to alternatives A and C, we can obtain an aggregate score-vector $(\sum_{j \neq i} \sigma_C^j, \sum_{j \neq i} \sigma_B^j, \sum_{j \neq i} \sigma_A^j) \in S_C^n \subset S^n$ (i.e. $C \succ_n A$ and $C \succ_n B$) which occurs with the same probability as the aggregate score-vector $\boldsymbol{s}_n$. Thus, for *every* aggregate score-vector in S^n s.t. alternative A has the highest score, there is a score-vector in S^n s.t. alternative B (C, resp.) has the highest score. Therefore, the number $|S_k^n|$ of aggregate score-vectors s.t. alternative k has the highest score is the same for all $k \in K$. Furthermore, the probability that the aggregate score-vector $\boldsymbol{s}_n$ is s.t. alternative k has the highest score (and therefore $\delta_k(\boldsymbol{s}_n) = 1$) is the same for all $k \in K$: $\Pr[S_A^n] = \Pr[S_B^n] = \Pr[S_C^n]$. □

Lemma 3. *The probability that any two alternatives $k, l \in K$ ($k \neq l$) are in a tie for the highest score is the same for all pairs (k,l).*

Proof of Lemma 3. Let $S_{kl}^n \subset S^n$ denote the set containing all aggregate score-vectors s.t. two alternatives k and l are in a tie for the highest score while beating the remaining alternative. Now pick any aggregate score-vector $\bar{\boldsymbol{s}}_n \in S_{AB}^n$. That is, the components of $\bar{\boldsymbol{s}}_n$ are s.t. $\sum_{j \neq i} \bar{\sigma}_A^j = \sum_{j \neq i} \bar{\sigma}_B^j > \sum_{j \neq i} \bar{\sigma}_C^j$ (i.e. $A \sim_n B \succ_n C$). Next, generate for each voter $j \neq i$ a new score-vector $\check{\boldsymbol{\sigma}}^j$ by interchanging in the original score-vectors $\bar{\boldsymbol{\sigma}}^j$ the scores assigned to alternatives A and C:

$\check{\boldsymbol{\sigma}}^j = (\bar{\sigma}_C^j, \bar{\sigma}_B^j, \bar{\sigma}_A^j)$. Adding up these new individual score-vectors $\check{\boldsymbol{\sigma}}^j$, we obtain the aggregate score-vector $\check{\boldsymbol{s}}_n = (\sum_{j \neq i} \bar{\sigma}_C^j, \sum_{j \neq i} \bar{\sigma}_B^j, \sum_{j \neq i} \bar{\sigma}_A^j) \in S_{BC}^n \subset S^n$ (i.e. $B \sim_n C \succ_n A$). Note that aggregate score-vectors $\bar{\boldsymbol{s}}_n$ and $\check{\boldsymbol{s}}_n$ occur with the same probability. This argument shows that for every aggregate score-vector in S^n s.t. alternatives A and B tie while beating C, there exists an aggregate score-vector in S^n s.t. alternatives B and C tie while beating A. Consequently, the number $|S_{AB}^n|$ of aggregate score-vectors in S^n is the same as the number of aggregate score-vectors $|S_{BC}^n|$ in S^n. Furthermore, $\Pr[S_{AB}^n] = \Pr[S_{BC}^n]$. Next, consider an aggregate score-vector $\check{\boldsymbol{s}}_n \in S_{AC}^n$ s.t. alternatives A and C tie while beating alternative B (i.e. $A \sim_n C \succ_n B$). By interchanging the A- and B-scores in the original score-vectors $\check{\boldsymbol{\sigma}}^j$, we obtain an aggregate score-vector s.t. alternatives B and C tie for the highest score (i.e. $B \sim_n C \succ_n A$), and this new aggregate score-vector occurs with the same probability as the original aggregate score-vector $\check{\boldsymbol{s}}_n$. We can therefore conclude that for every aggregate score-vector in S^n s.t. alternatives A and C tie while beating B, there exists an aggregate score-vector in S^n s.t. alternatives B and C tie while beating A. Consequently, the number of aggregate score-vectors $|S_{AC}^n|$ in S^n is the same as the number of aggregate score-vectors $|S_{BC}^n|$ in S_n. In summary: $|S_{AC}^n| = |S_{BC}^n| = |S_{AB}^n|$ and $\Pr[S_{AC}^n] = \Pr[S_{BC}^n] = \Pr[S_{AB}^n]$. □

We are now ready to continue the proof of Proposition 1 in the main text. Our objective is to show that it is optimal for voter i to assign a score of 1 to his favorite alternative, and a score of zero to his lowest ranked alternative when all other voters use the same strategy that is symmetric in the sense of S2. We start with the following observation: Using Lemmas 2 and 3 above, it is easy to see that the expected probability $\mathrm{E}[\delta_k(\boldsymbol{s}_n)]$ assigned by the (x,y)-scoring rule $\boldsymbol{\delta}$ to any alternative k on the basis of the aggregate score-vector $\boldsymbol{s}_n$ is the same for all $k \in K$:[34]

$$\begin{aligned}\mathrm{E}[\delta_A(\boldsymbol{s}_n)] &\equiv \Pr[S_A^n] + \frac{1}{2}\left(\Pr[S_{AB}^n] + \Pr[S_{AC}^n]\right) + \frac{1}{3}\Pr[S_{ABC}^n] \\ &= \frac{1}{3} = \mathrm{E}[\delta_B(\boldsymbol{s}_n)] = \mathrm{E}[\delta_C(\boldsymbol{s}_n)]\end{aligned}$$

Now suppose w.l.o.g. that voter i submits the score-vector $\boldsymbol{\sigma}^i = (1, \sigma_B^i, 0)$, which assigns a score of 1 to alternative A, a score of $\sigma_B^i \in \{x, y\}$ to alternative B, and a score of 0 to alternative C. In what follows, we show that this results in an expected probability distribution $\mathrm{E}[\boldsymbol{\delta}(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)]$ over the set of alternatives K s.t.: $\mathrm{E}[\delta_A(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)] > 1/3 > \mathrm{E}[\delta_C(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)]$ and $\mathrm{E}[\delta_A(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)] > \mathrm{E}[\delta_B(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)] = \mathrm{E}[\delta_C(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)]$ for $\sigma_B^i = 0$; $\mathrm{E}[\delta_A(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)] > \mathrm{E}[\delta_B(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)] > \mathrm{E}[\delta_C(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)]$ for all $\sigma_B^i \in (0,1)$; and $\mathrm{E}[\delta_A(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)] = \mathrm{E}[\delta_B(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)] > \mathrm{E}[\delta_C(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)]$ if $\sigma_B^i = 1$. To show this, we distinguish three cases: Case I, where $\sigma_B^i = 0$; Case II, where $\sigma_B^i \in (0,1)$; and Case III, where $\sigma_B^i = 1$. In each case, we compute for every $k \in K$ the expected change $\mathrm{E}[\Delta\delta_k(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)]$ in the probability of alternative k, where $\Delta\delta_k(\boldsymbol{s}_n + \boldsymbol{\sigma}^i) \equiv \delta_k(\boldsymbol{s}_n + \boldsymbol{\sigma}^i) - \delta_k(\boldsymbol{s}_n)$. This allows us to capture the way in which the score-vector $\boldsymbol{\sigma}^i$ submitted by voter i affects the probability distribution $\mathrm{E}[\boldsymbol{\delta}(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)]$. To facilitate the computation of $\Delta\delta_k(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)$ for every $\boldsymbol{s}_n \in S^n$, we split the set S^n on the basis of the alternatives' aggregate scores, and the aggregate ordinal ranking these scores give rise to. In addition to the subsets S_k^n and S_{kl}^n defined in the proofs of Lemmas 2 and 3 above, and the subset S_{ABC}^n defined in footnote 34, we will also encounter the following subsets of S_C^n: $S_{C,A}^n$ and $S_{C,B}^n$ (which are associated with the aggregate ordinal rankings $C \succ_n A \succ_n B$ and $C \succ_n B \succ_n A$, resp.), as well as $S_{C,AB}^n$ (associated with $C \succ_n A \sim_n B$). With the exception of subsets where alternative A holds or shares with one or more alternatives the highest aggregate score, we may have to consider further subsets of S^n depending on how the aggregate A-score s_n^A relates to the scores s_n^B and s_n^C of the other two alternatives.

[34] The event that an alternative ties with the other two alternatives is denoted by $S_{ABC}^n \subset S^n$.

	Subset of S^n	$\delta_A(\boldsymbol{s}_n)$	$\delta_B(\boldsymbol{s}_n)$	$\Delta\delta_A(\boldsymbol{s}_{n+1})$	$\Delta\delta_B(\boldsymbol{s}_{n+1})$
I.1	S^n_A	1	0	0	0
I.2	S^n_{AB}	1/2	1/2	1/2	−1/2
I.3	S^n_{AC}	1/2	0	1/2	0
I.4	S^n_{ABC}	1/3	1/3	2/3	−1/3
I.5	S^n_{BC}: $s^B_n - s^A_n < 1$	0	1/2	1	−1/2
I.6	S^n_{BC}: $s^B_n - s^A_n = 1$	0	1/2	1/3	−1/6
I.7	S^n_{BC}: $s^B_n - s^A_n > 1$	0	1/2	0	0
I.8	S^n_B: $s^B_n - s^A_n < 1$	0	1	1	−1
I.9	S^n_B: $s^B_n - s^A_n = 1$	0	1	1/2	−1/2
I.10	S^n_B: $s^B_n - s^A_n > 1$	0	1	0	0
I.11	S^n_C: $s^C_n - s^A_n < 1$	0	0	1	0
I.12	S^n_C: $s^C_n - s^A_n = 1$	0	0	1/2	0
I.13	S^n_C: $s^C_n - s^A_n > 1$	0	0	0	0

Table 3: Change in δ_k when $\boldsymbol{\sigma}^i = (1,0,0)$

6.1.1 Case I: $\sigma^i_B = 0$

Table 3 below displays the values of $\delta_A(\boldsymbol{s}_n)$ and $\delta_B(\boldsymbol{s}_n)$, as well as $\Delta\delta_A(\boldsymbol{s}_{n+1})$ and $\Delta\delta_B(\boldsymbol{s}_{n+1})$, for $\boldsymbol{s}_n$ in all relevant subsets of S^n. To ease notation, we write $\boldsymbol{s}_{n+1}$ as shorthand for $\boldsymbol{s}_n + \boldsymbol{\sigma}^i$. In order to reduce the dimensions of the table, we omit $\delta_C(\boldsymbol{s}_n)$, as $\delta_C(\boldsymbol{s}_n) = 1 - \delta_A(\boldsymbol{s}_n) - \delta_B(\boldsymbol{s}_n)$. We also omit $\Delta\delta_C(\boldsymbol{s}_{n+1})$ from the table, as $\Delta\delta_A(\boldsymbol{s}_{n+1}) + \Delta\delta_B(\boldsymbol{s}_{n+1}) + \Delta\delta_C(\boldsymbol{s}_{n+1}) = 0$. As Table 3 reveals, the change $\Delta\delta_A(\boldsymbol{s}_n + (1,0,0))$ is non-negative for all aggregate score-vectors $\boldsymbol{s}_n$, while $\Delta\delta_B(\boldsymbol{s}_n + (1,0,0))$ and $\Delta\delta_C(\boldsymbol{s}_n + (1,0,0))$ are both always non-positive, and sometimes negative. It follows immediately that $\mathrm{E}[\delta_A(\boldsymbol{s}_n + (1,0,0))] > 1/3$, and $\mathrm{E}[\delta_B(\boldsymbol{s}_n + (1,0,0))] = \mathrm{E}[\delta_C(\boldsymbol{s}_n + (1,0,0))] < 1/3$.

Now consider w.l.o.g. a voter i with ordinal ranking $A \succ_i B \succ_i C$ (i.e. Bernoulli utilities $u^i_A > u^i_B > u^i_C$). By the argument in the preceding paragraph, his choice problem of which score-vector to submit essentially boils down to choosing the alternative to which he wishes to assign a coefficient $\pi_H > 1/3$ in his expected utility function, while assigning a coefficient $\pi_L < 1/3$ to the remaining two alternatives.[35] For instance, when choosing the score-vector $(1,0,0)$, voter i's expected utility is $U_i((1,0,0),\boldsymbol{u}^i) = \pi_H u^i_A + \pi_L(u^i_B + u^i_C)$. The expected utilities from the other two score-vectors are obtained analogously. Now observe that $U_i((1,0,0),\boldsymbol{u}^i) - U_i((0,1,0),\boldsymbol{u}^i) = (\pi_H - \pi_L)(u^i_A - u^i_B) > 0$ and $U_i((0,1,0),\boldsymbol{u}^i) - U_i((0,0,1),\boldsymbol{u}^i) = (\pi_H - \pi_L)(u^i_B - u^i_C) > 0$, which shows that it is optimal for voter i to assign the score of 1 to his favorite alternative A.

6.1.2 Case II: $\sigma^i_B \in (0,1)$

Table 4 below displays the values of $\delta_A(\boldsymbol{s}_n)$ and $\delta_B(\boldsymbol{s}_n)$, as well as $\Delta\delta_A(\boldsymbol{s}_{n+1})$ and $\Delta\delta_B(\boldsymbol{s}_{n+1})$, for $\boldsymbol{s}_n$ in all relevant subsets of S^n. As can be inferred from Table 4, the change $\Delta\delta_C(\boldsymbol{s}_n + (1,\sigma^i_B,0))$ is non-positive for all, and negative for some aggregate score-vectors $\boldsymbol{s}_n$. Thus, $\mathrm{E}[\delta_C(\boldsymbol{s}_n + (1,\sigma^i_B,0))] < 1/3$. Therefore, it is optimal for voter i to assign a score of 0 to his least preferred alternative. While the change $\Delta\delta_A(\boldsymbol{s}_n + (1,\sigma^i_B,0))$ is non-negative for all, and positive for some aggregate score-vectors $\boldsymbol{s}_n$, the change $\Delta\delta_B(\boldsymbol{s}_n + (1,\sigma^i_B,0))$ takes both positive values for some, and negative values for other aggregate score-vectors $\boldsymbol{s}_n$.

[35] We use the shorthand π_H to denote the probability of the alternative that receives the score of 1 in voter i's chosen score-vector, while π_L denotes the probability of the alternatives that receive zero scores.

We begin by showing that $\mathrm{E}[\delta_A(\boldsymbol{s}_n+(1,\sigma_B^i,0))]$ exceeds $\mathrm{E}[\delta_B(\boldsymbol{s}_n+(1,\sigma_B^i,0))]$. For this purpose, we compute the difference between these two expected probabilities using Table 4:

$$\begin{aligned}
& \mathrm{E}[\delta_A(\boldsymbol{s}_n+(1,\sigma_B^i,0))]-\mathrm{E}[\delta_B(\boldsymbol{s}_n+(1,\sigma_B^i,0))] \\
= \; & \mathrm{E}[\Delta\delta_A(\boldsymbol{s}_n+(1,\sigma_B^i,0))]-\mathrm{E}[\Delta\delta_B(\boldsymbol{s}_n+(1,\sigma_B^i,0))]+\underbrace{\mathrm{E}[\delta_A(\boldsymbol{s}_n)]}_{=1/3}-\underbrace{\mathrm{E}[\delta_B(\boldsymbol{s}_n)]}_{=1/3} \\
= \; & \Pr[\text{II.2}]+\frac{1}{2}\Pr[\text{II.3}]+\Pr[\text{II.4}]+\frac{3}{2}\Pr[\text{II.5}]+\frac{1}{2}\Pr[\text{II.6}] \\
& -\frac{1}{2}\Pr[\text{II.7}]+2\Pr[\text{II.8}]+\Pr[\text{II.9}]+\Pr[\text{II.11}]+\frac{1}{2}\Pr[\text{II.12}] \\
& +\Pr[\text{II.14}]+\frac{1}{2}\Pr[\text{II.16}]-\frac{1}{2}\Pr[\text{II.18}]-\Pr[\text{II.20}] \qquad (9)
\end{aligned}$$

To show that this difference is positive, we need to argue that any negative terms in (9) are canceled out or exceeded by the positive terms. Intuitively, this means showing that for every aggregate score-vector $\boldsymbol{s}_n$ s.t. alternative B gains probability mass from alternative C while A does not, there exists an aggregate score-vector $\hat{\boldsymbol{s}}_n$ which arises with the same probability as $\boldsymbol{s}_n$ s.t. alternative A gains the same amount of probability mass from C while B does not.[36]

	Subset of S^n	$\delta_A(\boldsymbol{s}_n)$	$\delta_B(\boldsymbol{s}_n)$	$\Delta\delta_A(\boldsymbol{s}_{n+1})$	$\Delta\delta_B(\boldsymbol{s}_{n+1})$
II.1	S^n_A	1	0	0	0
II.2	S^n_{AB}	1/2	1/2	1/2	−1/2
II.3	S^n_{AC}	1/2	0	1/2	0
II.4	S^n_{ABC}	1/3	1/3	2/3	−1/3
II.5	S^n_{BC}: $s^B_n - s^A_n < 1-\sigma^i_B$	0	1/2	1	−1/2
II.6	S^n_{BC}: $s^B_n - s^A_n = 1-\sigma^i_B$	0	1/2	1/2	0
II.7	S^n_{BC}: $s^B_n - s^A_n > 1-\sigma^i_B$	0	1/2	**0**	**1/2**
II.8	S^n_B: $s^B_n - s^A_n < 1-\sigma^i_B$	0	1	1	−1
II.9	S^n_B: $s^B_n - s^A_n = 1-\sigma^i_B$	0	1	1/2	−1/2
II.10	S^n_B: $s^B_n - s^A_n > 1-\sigma^i_B$	0	1	0	0
II.11	$S^n_{C,A} \cup S^n_{C,AB}$: $s^C_n - s^A_n < 1$	0	0	1	0
II.12	$S^n_{C,A} \cup S^n_{C,AB}$: $s^C_n - s^A_n = 1$	0	0	1/2	0
II.13	$S^n_{C,A} \cup S^n_{C,AB}$: $s^C_n - s^A_n > 1$	0	0	0	0
II.14	$S^n_{C,B}$: $s^C_n - s^A_n < 1 \wedge s^B_n - s^A_n < 1-\sigma^i_B$	0	0	1	0
II.15	$S^n_{C,B}$: $s^C_n - s^A_n < 1 \wedge s^B_n - s^A_n = 1-\sigma^i_B$	0	0	1/2	1/2
II.16	$S^n_{C,B}$: $s^C_n - s^A_n = 1 \wedge s^B_n - s^A_n < 1-\sigma^i_B$	0	0	1/2	0
II.17	$S^n_{C,B}$: $s^C_n - s^A_n = 1 \wedge s^B_n - s^A_n = 1-\sigma^i_B$	0	0	1/3	1/3
II.18	$S^n_{C,B}$: $s^C_n - s^A_n > 1 \wedge s^C_n - s^B_n = \sigma^i_B$	0	0	**0**	**1/2**
II.19	$S^n_{C,B}$: $s^C_n - s^A_n > 1 \wedge s^C_n - s^B_n > \sigma^i_B$	0	0	0	0
II.20	$S^n_{C,B}$: $s^B_n - s^A_n > 1-\sigma^i_B \wedge s^C_n - s^B_n < \sigma^i_B$	0	0	**0**	**1**

Table 4: Change in δ_k when $\boldsymbol{\sigma}^i=(1,\sigma_B^i,0)$, where $\sigma_B^i\in(0,1)$

We consider first an aggregate score-vector $\boldsymbol{s}_n$ in subset II.7. Such a score-vector has the property that $B\sim_n C\succ_n A$ with $s_n^B-s_n^A>1-\sigma_B^i$. Now interchange, in each individual score-vector

[36] In Table 4, see events II.7, II.18, and II.20 highlighted in bold, in which B gains while A does not.

$\boldsymbol{\sigma}^j = (\sigma_A^j, \sigma_B^j, \sigma_C^j)$ associated with $\boldsymbol{s}_n$, the scores σ_B^j and σ_A^j. This give rise to new individual score-vectors $\hat{\boldsymbol{\sigma}}^j = (\sigma_B^j, \sigma_A^j, \sigma_C^j)$ for all $j \neq i$, and an associated aggregate score-vector $\hat{\boldsymbol{s}}_n = (\hat{s}_n^A, \hat{s}_n^B, \hat{s}_n^C) = (s_n^B, s_n^A, s_n^C)$, where $A \sim_n C \succ_n B$ and $\hat{s}_n^A - \hat{s}_n^B > 1 - \sigma_B^i$. Note that the new aggregate score-vector $\hat{\boldsymbol{s}}_n$ is an element of subset II.3 (i.e. $\hat{\boldsymbol{s}}_n \in S_{AC}^n$), and that $\Pr[\hat{\boldsymbol{s}}_n] = \Pr[\boldsymbol{s}_n]$ (by virtue of the fact that $\hat{\boldsymbol{s}}_n$ is a permutation of $\boldsymbol{s}_n$, and that the permuted individual score-vectors $\hat{\boldsymbol{\sigma}}^j$ occur with the same probability as the original individual score-vectors $\boldsymbol{\sigma}^j$). This implies that in the expression for the difference in equation (9) above, the term $-(1/2)\Pr[\text{II.7}]$ is counteracted by the positive term $(1/2)\Pr[\text{II.3}]$.[37]

Next, we consider aggregate score-vectors $\boldsymbol{s}_n$ in subsets II.18 and II.20. As these subsets depend on the value of σ_B^i, we distinguish below settings where $\sigma_B^i \leq 1/2$ from settings where $\sigma_B^i > 1/2$:

Case II.1: $\boldsymbol{\sigma_B^i \leq 1/2}$. Consider first an aggregate score-vector $\boldsymbol{s}_n$ in subset II.20. Such a score-vector has the property that $C \succ_n B \succ_n A$, with $s_n^C - s_n^B < \sigma_B^i$ and $s_n^B - s_n^A > 1 - \sigma_B^i$. Now generate a new aggregate score-vector as follows: in the individual score-vectors $\boldsymbol{\sigma}^j = (\sigma_A^j, \sigma_B^j, \sigma_C^j)$ $(j \neq i)$ associated with $\boldsymbol{s}_n$, permute the scores so that $\hat{\sigma}_A^j = \sigma_B^j$, $\hat{\sigma}_B^j = \sigma_C^j$, and $\hat{\sigma}_C^j = \sigma_A^j$. This yields a new aggregate score-vector $\hat{\boldsymbol{s}}_n$ s.t. $B \succ_n A \succ_n C$, with $\hat{s}_n^B - \hat{s}_n^A < \sigma_B^i \leq 1/2 \leq 1 - \sigma_B^i$ and $\hat{s}_n^A - \hat{s}_n^C > 1 - \sigma_B^i$. Observe that $\hat{\boldsymbol{s}}_n$ is an element of subset II.8, and that $\Pr[\hat{\boldsymbol{s}}_n] = \Pr[\boldsymbol{s}_n]$. This implies that in the difference in equation (9) above, the negative term $-\Pr[\text{II.20}]$ is counteracted by the positive term $2\Pr[\text{II.8}]$.

Next, note that we can deal with aggregate score-vectors $\boldsymbol{s}_n$ in subset II.18 in a similar fashion. After appropriate permutation of the individual score-vectors $\boldsymbol{\sigma}^j$, we obtain $\hat{\boldsymbol{s}}_n$ s.t. $B \succ_n A \succ_n C$, with $\hat{s}_n^B - \hat{s}_n^A = \sigma_B^i \leq 1/2 \leq 1 - \sigma_B^i$, which is an element of subset II.8 if $\sigma_B^i < 1/2$, and an element of subset II.9 if $\sigma_B^i = 1/2$.

Case II.2: $\boldsymbol{\sigma_B^i > 1/2}$. Consider first an aggregate score-vector $\boldsymbol{s}_n$ in subset II.20. Such a score-vector has the property that $C \succ_n B \succ_n A$, with $s_n^C - s_n^B < \sigma_B^i$ and $s_n^B - s_n^A > 1 - \sigma_B^i$. In the individual score-vectors $\boldsymbol{\sigma}^j = (\sigma_A^j, \sigma_B^j, \sigma_C^j)$ $(j \neq i)$ associated with $\boldsymbol{s}_n$, permute the scores so that $\hat{\sigma}_A^j = \sigma_B^j$, $\hat{\sigma}_B^j = \sigma_A^j$, and $\hat{\sigma}_C^j = \sigma_C^j$. This yields $\hat{\boldsymbol{s}}_n$ s.t. $C \succ_n A \succ_n B$, with $\hat{s}_n^C - \hat{s}_n^A < \sigma_B^i < 1$ and $\hat{s}_n^A - \hat{s}_n^B > 1 - \sigma_B^i$. Note that $\hat{\boldsymbol{s}}_n$ is an element of subset II.11, and that $\Pr[\hat{\boldsymbol{s}}_n] = \Pr[\boldsymbol{s}_n]$. This implies that in the expression for the difference in equation (9) above, the negative term $-\Pr[\text{II.20}]$ is counteracted by the positive term $\Pr[\text{II.11}]$.

Note that we can deal with aggregate score-vectors $\boldsymbol{s}_n$ in subset II.18 in a similar fashion. After appropriate permutation of the individual score-vectors $\boldsymbol{\sigma}^j$, we obtain $\hat{\boldsymbol{s}}_n$ s.t. $C \succ_n B \succ_n A$, with $\hat{s}_n^C - \hat{s}_n^A = \sigma_B^i < 1$, which is an element of subset II.11. In summary, we can conclude that $\mathrm{E}[\delta_A(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)] > \mathrm{E}[\delta_B(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)]$ for all $\sigma_B^i \in (0,1)$. Having shown this, we now show that $\mathrm{E}[\delta_B(\boldsymbol{s}_n + (1, \sigma_B^i, 0))] > \mathrm{E}[\delta_C(\boldsymbol{s}_n + (1, \sigma_B^i, 0))]$. For this purpose, we compute the difference between these two

[37] In fact, as subset II.3 also contains aggregate score-vectors $\check{\boldsymbol{s}}_n$ where $\check{s}_n^A = \check{s}_n^C > \check{s}_n^B$ and $\check{s}_n^A - \check{s}_n^B \leq 1 - \sigma_B^i$, it follows that $(1/2)\Pr[\text{II.3}] - (1/2)\Pr[\text{II.7}] > 0$.

expected probabilities. Using Table 4, this difference reduces to:

$$
\begin{aligned}
& \mathrm{E}[\Delta\delta_B(\boldsymbol{s}_n+(1,\sigma_B^i,0))]-\mathrm{E}[\Delta\delta_C(\boldsymbol{s}_n+(1,\sigma_B^i,0))] \\
= \; & \Pr[\text{II.7}]+\Pr[\text{II.18}]+\frac{3}{2}\Pr[\text{II.15}]+\Pr[\text{II.17}]+2\Pr[\text{II.20}] \\
& -\frac{1}{2}\Pr[\text{II.2}]-\Pr[\text{II.8}]-\frac{1}{2}\Pr[\text{II.9}] \\
& +\frac{1}{2}\Pr[\text{II.3}]+\frac{1}{2}\Pr[\text{II.6}]+\Pr[\text{II.11}] \\
& +\frac{1}{2}\Pr[\text{II.12}]+\Pr[\text{II.14}]+\frac{1}{2}\Pr[\text{II.16}] \qquad (10)
\end{aligned}
$$

To show that this difference is positive, we need to argue that any negative terms in (10) are canceled out or exceeded by the positive terms. First, it is easy to see that $\Pr[\text{II.2}] = \Pr[\text{II.3}]$ because event II.2 corresponds to the subset S_{AB}^n and event II.3 corresponds to the subset S_{AC}^n. By Lemma 3, we know that $\Pr[S_{AB}^n] = \Pr[S_{AC}^n]$. Next, consider event II.8. This event can be partitioned into two sub-events:

Case II.3: $B \succ_n A \succeq_n C$ with $s_n^B - s_n^A < 1-\sigma_B^i$. Now generate a new aggregate score-vector $\hat{\boldsymbol{s}}_n$ by interchanging the scores of alternatives B and C. This new score-vector is associated with the ordinal ranking $C \succ_n A \succeq_n B$ with $\hat{s}_n^C - \hat{s}_n^A < 1-\sigma_B^i < 1$. It is easy to see that $\hat{\boldsymbol{s}}_n$ is an element of event II.11. In other words, for every aggregate score-vector $\boldsymbol{s}_n$ that falls into the first partition of II.8, there is an aggregate score-vector $\hat{\boldsymbol{s}}_n$ in II.11 that arises with the same probability as $\boldsymbol{s}_n$.

Case II.4: $B \succ_n C \succ_n A$ with $\hat{s}_n^B - \hat{s}_n^A < 1-\sigma_B^i$. Now generate a new aggregate score-vector $\tilde{\boldsymbol{s}}_n$ by interchanging the scores of alternatives B and C. This new score-vector is associated with the ordinal ranking $C \succ_n B \succ_n A$ with $\tilde{s}_n^C - \tilde{s}_n^A < 1-\sigma_B^i$. As $\tilde{s}_n^B < \tilde{s}_n^C$, it follows that $\tilde{s}_n^B - \tilde{s}_n^A < 1-\sigma_B^i$. It is now easy to see that $\tilde{\boldsymbol{s}}_n$ is an element of event II.14. In other words, for every aggregate score-vector $\boldsymbol{s}_n$ that falls into the second partition of II.8, there is an aggregate score-vector $\tilde{\boldsymbol{s}}_n$ in II.14 that arises with the same probability as $\boldsymbol{s}_n$. Together, Cases II.3 and II.4 show that the term $-\Pr[\text{II.8}]$ in (10) is fully canceled out by the positive terms $\Pr[\text{II.11}]$ and $\Pr[\text{II.14}]$.

Now consider event II.9. This event can be partitioned into two sub-events:

Case II.5: $B \succ_n A \succeq_n C$ with $s_n^B - s_n^A = 1-\sigma_B^i$. Now generate a new aggregate score-vector $\hat{\boldsymbol{s}}_n$ by interchanging the scores of alternatives B and C. This new score-vector is associated with the ordinal ranking $C \succ_n A \succeq_n B$ with $\hat{s}_n^C - \hat{s}_n^A = 1-\sigma_B^i < 1$. It is easy to see that $\hat{\boldsymbol{s}}_n$ is an element of event II.11. In other words, for every aggregate score-vector $\boldsymbol{s}_n$ that falls into the first partition of II.9, there is an aggregate score-vector $\hat{\boldsymbol{s}}_n$ in II.11 that arises with the same probability as $\boldsymbol{s}_n$.

Case II.6: $B \succ_n C \succ_n A$ with $s_n^B - s_n^A = 1-\sigma_B^i$. Now generate a new aggregate score-vector $\tilde{\boldsymbol{s}}_n$ by interchanging the scores of alternatives B and C. This new score-vector is associated with the ordinal ranking $C \succ_n B \succ_n A$ with $\tilde{s}_n^C - \tilde{s}_n^A = 1-\sigma_B^i$. As $\tilde{s}_n^B < \tilde{s}_n^C$, it follows that $\tilde{s}_n^B - \tilde{s}_n^A < 1-\sigma_B^i$. It is now easy to see that $\tilde{\boldsymbol{s}}_n$ is an element of event II.14. In other words, for every aggregate score-vector $\boldsymbol{s}_n$ that falls into the second partition of II.9, there is an aggregate score-vector $\tilde{\boldsymbol{s}}_n$ in II.14 that arises with the same probability as $\boldsymbol{s}_n$. Together, Cases II.5 and II.6 show that the term $-\Pr[\text{II.9}]$ in (10) is fully canceled out by the positive terms $\Pr[\text{II.11}]$ and $\Pr[\text{II.14}]$.

In conclusion of Case II, we can state that $\mathrm{E}[\delta_A(\boldsymbol{s}_n+\boldsymbol{\sigma}^i)] > \mathrm{E}[\delta_B(\boldsymbol{s}_n+\boldsymbol{\sigma}^i)] > \mathrm{E}[\delta_C(\boldsymbol{s}_n+\boldsymbol{\sigma}^i)]$ for all $\sigma_B^i \in (0,1)$. Now consider w.l.o.g. a voter i with Bernoulli utilities $u_A^i > u_B^i > u_C^i$. His choice problem of which score-vector to submit essentially boils down to choosing the alternative to which he wants to assign a coefficient π_H, to which he wants to assign a coefficient π_M, and to which he wants to assign π_L, with $\pi_H > \pi_M > \pi_L$.[38] Of the six possible assignments of coefficients to alternatives, only the following three require further consideration (the remaining three are dominated in expected utility terms):

(i) $U_i((1,\sigma_B^i,0),\boldsymbol{u}^i) = \pi_H u_A^i + \pi_M u_B^i + \pi_L u_C^i$

(ii) $U_i((\sigma_B^i,0,1),\boldsymbol{u}^i) = \pi_M u_A^i + \pi_L u_B^i + \pi_H u_C^i$

(iii) $U_i((0,1,\sigma_B^i),\boldsymbol{u}^i) = \pi_L u_A^i + \pi_H u_B^i + \pi_M u_C^i$

Forming the difference between (i) and (ii), and adding and subtracting $\pi_M u_C^i$, it is easy to see that:

$$U_i((1,\sigma_B^i,0),\boldsymbol{u}^i) - U_i((\sigma_B^i,0,1),\boldsymbol{u}^i) = (\pi_H-\pi_M)(u_A^i-u_C^i) + (\pi_M-\pi_L)(u_B^i-u_C^i) > 0.$$

Similarly, forming the difference between (i) and (iii), and adding and subtracting $\pi_L u_B^i$, it is easy to see that:

$$U_i((1,\sigma_B^i,0),\boldsymbol{u}^i) - U_i((0,1,\sigma_B^i),\boldsymbol{u}^i) = (\pi_H-\pi_L)(u_A^i-u_B^i) + (\pi_M-\pi_L)(u_B^i-u_C^i) > 0.$$

We can therefore conclude that the unique utility-maximizing score-vector for agent i is $(1,\sigma_B^i,0)$, which coincides with agent i's true ordinal ranking.

6.1.3 Case III: $\sigma_B^i = 1$

Table 5 below displays the values of $\delta_A(\boldsymbol{s}_n)$ and $\delta_B(\boldsymbol{s}_n)$, as well as $\Delta\delta_A(\boldsymbol{s}_{n+1})$ and $\Delta\delta_B(\boldsymbol{s}_{n+1})$, for $\boldsymbol{s}_n$ in all relevant subsets of S^n. As can be inferred from Table 5, the change $\Delta\delta_C(\boldsymbol{s}_n+(1,1,0))$ is non-positive for all, and negative for some aggregate score-vectors $\boldsymbol{s}_n$. Therefore, $\mathrm{E}[\delta_C(\boldsymbol{s}_n+(1,1,0))] < 1/3$. Consequently, it is optimal for voter i to assign a score of 0 to his least preferred alternative. Note that both $\Delta\delta_A(\boldsymbol{s}_n+(1,1,0))$ and $\Delta\delta_B(\boldsymbol{s}_n+(1,1,0))$ are non-negative for all possible aggregate score-vectors $\boldsymbol{s}_n$.

In order to show that the expected probabilities of alternatives A and B are equal, we compute the difference between $\mathrm{E}[\delta_A(\boldsymbol{s}_n+(1,1,0))]$ and $\mathrm{E}[\delta_B(\boldsymbol{s}_n+(1,1,0))]$:[39]

$$\begin{aligned} & \mathrm{E}[\delta_A(\boldsymbol{s}_n+(1,1,0))] - \mathrm{E}[\delta_B(\boldsymbol{s}_n+(1,1,0))] \\ = \; & \mathrm{E}[\Delta\delta_A(\boldsymbol{s}_n+(1,1,0))] - \mathrm{E}[\Delta\delta_B(\boldsymbol{s}_n+(1,1,0))] \\ = \; & \Pr[\mathrm{III.7}] + \frac{1}{2}\Pr[\mathrm{III.8}] - \frac{1}{2}\Pr[\mathrm{III.14}] - \Pr[\mathrm{III.15}] - \Pr[\mathrm{III.16}] \end{aligned}$$

This difference is zero because for every aggregate score-vector $\boldsymbol{s}_n$ in subsets III.15 and III.16 resp., there exists an aggregate score-vector $\hat{\boldsymbol{s}}_n$ in subset III.7 that occurs with the same probability as $\boldsymbol{s}_n$, and vice versa. Any such score-vector is s.t. $C \succ_n B \succ_n A$, with $s_n^B > s_n^C - 1$. Now swap in the individual score-vectors $\boldsymbol{\sigma}^j$ that give rise to the aggregate score-vector $\boldsymbol{s}_n$ the A-score and the B-score. This gives rise to individual score-vectors $\hat{\boldsymbol{\sigma}}^j = (\sigma_B^j, \sigma_A^j, \sigma_C^j)$, and an aggregate score-vector $\hat{\boldsymbol{s}}_n = (\hat{s}_n^A, \hat{s}_n^B, \hat{s}_n^C) = (s_n^B, s_n^A, s_n^C)$. The new aggregate score-vector $\hat{\boldsymbol{s}}_n$ is s.t. $C \succ_n A \succ_n B$, with $\hat{s}_n^C - \hat{s}_n^A < 1$. This makes $\hat{\boldsymbol{s}}_n$ an element of subset III.7. Furthermore, $\Pr[\hat{\boldsymbol{s}}_n] = \Pr[\boldsymbol{s}_n]$. As a result, the

[38] We use the shorthand π_M to denote the probability of the alternative that receives the score of σ_B^i in voter i's chosen score-vector, while π_H and π_L are as defined in footnote 35.

[39] Note that, by Lemma 3, the terms $(1/2)\Pr[\mathrm{III.3}]$ and $(1/2)\Pr[\mathrm{III.5}]$ cancel out in the expression for $\mathrm{E}[\Delta\delta_A(\boldsymbol{s}_n+(1,1,0))] - \mathrm{E}[\Delta\delta_B(\boldsymbol{s}_n+(1,1,0))]$.

	Subset of S^n	$\delta_A(\boldsymbol{s}_n)$	$\delta_B(\boldsymbol{s}_n)$	$\Delta\delta_A(\boldsymbol{s}_{n+1})$	$\Delta\delta_B(\boldsymbol{s}_{n+1})$
III.1	S_A^n	1	0	0	0
III.2	S_{AB}^n	1/2	1/2	0	0
III.3	S_{AC}^n	1/2	0	1/2	0
III.4	S_{ABC}^n	1/3	1/3	1/6	1/6
III.5	S_{BC}^n	0	1/2	**0**	**1/2**
III.6	S_B^n	0	1	0	0
III.7	$S_{C,A}^n$: $s_n^C - s_n^A < 1$	0	0	1	0
III.8	$S_{C,A}^n$: $s_n^C - s_n^A = 1$	0	0	1/2	0
III.9	$S_{C,A}^n$: $s_n^C - s_n^A > 1$	0	0	0	0
III.10	$S_{C,AB}^n$: $s_n^C - s_n^A < 1$	0	0	1/2	1/2
III.11	$S_{C,AB}^n$: $s_n^C - s_n^A = 1$	0	0	1/3	1/3
III.12	$S_{C,AB}^n$: $s_n^C - s_n^A > 1$	0	0	0	0
III.13	$S_{C,B}^n$: $s_n^A + 1 < s_n^C \wedge s_n^B + 1 < s_n^C$	0	0	0	0
III.14	$S_{C,B}^n$: $s_n^A + 1 < s_n^C \wedge s_n^B + 1 = s_n^C$	0	0	**0**	**1/2**
III.15	$S_{C,B}^n$: $s_n^A + 1 < s_n^C \wedge s_n^B + 1 > s_n^C$	0	0	**0**	**1**
III.16	$S_{C,B}^n$: $s_n^A + 1 \geq s_n^C$	0	0	**0**	**1**

Table 5: Change in δ_k when $\boldsymbol{\sigma}^i = (1,1,0)$

difference Pr[III.7] $-$ Pr[III.15] $-$ Pr[III.16] is zero. A very similar construction allows us to show that the difference $(1/2)\,\text{Pr[III.8]} - (1/2)\,\text{Pr[III.14]}$ is also zero. Consequently, $\text{E}[\delta_A(\boldsymbol{s}_n + (1,1,0))] = \text{E}[\delta_B(\boldsymbol{s}_n + (1,1,0))] > 1/3 > \text{E}[\delta_C(\boldsymbol{s}_n + (1,1,0))]$.

Now consider w.l.o.g. a voter i with Bernoulli utilities $u_A^i > u_B^i > u_C^i$. His choice problem of which one of the score-vectors $(1,1,0)$, $(1,0,1)$, and $(0,1,1)$ to submit essentially boils down to choosing the alternative to which he wishes to assign a coefficient $\pi_L < 1/3$ in his expected utility function, while assigning a coefficient $\pi_H > 1/3$ to the remaining two alternatives. For instance, when choosing the score-vector $(1,1,0)$, voter i's expected utility is $U_i((1,1,0),\boldsymbol{u}^i) = \pi_H(u_A^i + u_B^i) + \pi_L u_C^i$. The expected utilities from the other two score-vectors are obtained analogously. Now observe that $U_i((1,1,0),\boldsymbol{u}^i) - U_i((1,0,1),\boldsymbol{u}^i) = (\pi_H - \pi_L)(u_B^i - u_C^i) > 0$ and $U_i((1,0,1),\boldsymbol{u}^i) - U_i((0,1,1),\boldsymbol{u}^i) = (\pi_H - \pi_L)(u_A^i - u_B^i) > 0$, which shows that it is optimal for voter i to assign the score of 0 to his least preferred alternative C. □

6.2 Proof of Lemma 1 (for five or more voters)

We present the proof by considering, in turn, an odd number of voters greater or equal to five (i.e. $n \geq 4$ and even), followed by an even number of voters greater or equal to six (i.e. $n \geq 5$ and odd).

6.2.1 Five or more voters (odd)

We highlight a particular aggregate score-vector $\boldsymbol{s}_n$ (where $n \geq 4$ and even) with $\Pr[\boldsymbol{s}_n] = (\frac{1}{6} - p(0))^n = \frac{1}{6^n}$ s.t. $\delta_A(\boldsymbol{s}_n + (1,x,0)) - \delta_A(\boldsymbol{s}_n + (1,y,0)) > 0$. We construct this aggregate score-vector $\boldsymbol{s}_n$ by building on the score-vector $\boldsymbol{s}_2 = (y,1,1+y)$ featured in the proof of Lemma 1 for $n = 2$ (see Section 3 in the main body of the paper). Let $n = 2(1+m)$ (where $m \in \mathbb{N}$) and suppose that the aggregate score-vector across these $2(1+m)$ voters is composed as follows:

$$\begin{aligned} \boldsymbol{s}_n &= \boldsymbol{s}_2 + \boldsymbol{s}_{2m} \\ &= (y,1,1+y) + \boldsymbol{s}_{2m} \end{aligned}$$

where $\boldsymbol{s}_{2m} = (m(1+y), m(1+y), 0)$. That is, of the $2m$ voters whose individual score-vectors are aggregated into $\boldsymbol{s}_{2m}$, half have the individual score-vector $(1,y,0)$, while the remaining half have the individual score-vector $(y,1,0)$. On the basis of $\boldsymbol{s}_{2m}$, alternatives A and B are in a two-way tie. Upon adding the score-vectors $\boldsymbol{s}_2$ and $\boldsymbol{s}_{2m}$, we obtain the following aggregate score-vector $\boldsymbol{s}_n$ across all $2(1+m)$ voters other than i:

$$\boldsymbol{s}_n = (m(1+y)+y, m(1+y)+1, 1+y)$$

Table 6 shows that if voter i submits the score-vector $\boldsymbol{\sigma}^i = (1,y,0)$, then alternatives A and B are in a tie for the highest score. If, instead, he submits the score-vector $\boldsymbol{\sigma}^i = (1,x,0)$, then alternative A wins outright. Therefore, submission of $\boldsymbol{\sigma}^i = (1,y,0)$ implies a shift in probability mass from A to B for any $m \in \mathbb{N}$, relative to submission to $\boldsymbol{\sigma}^i = (1,x,0)$.

$\boldsymbol{\sigma}^i$	$\boldsymbol{s}_n + \boldsymbol{\sigma}^i$	$\boldsymbol{\delta}(\boldsymbol{s}_n + \boldsymbol{\sigma}^i)$
(1,x,0)	((1+y)(m+1),m(1+y)+1+x,1+y)	(1,0,0)
(1,y,0)	((1+y)(m+1),(1+y)(m+1),1+y)	($\frac{1}{2}$,$\frac{1}{2}$,0)

Table 6: Score-vector $\boldsymbol{s}_n$ that shows $L(p(0)) > 0$ for $n = 2(1+m)$

As alternative A's loss in probability features in the numerator of the loss-gain-ratio (5), it follows immediately that $L(p(0)) > 0$.

6.2.2 Six or more voters (even)

We highlight a particular aggregate score-vector $\boldsymbol{s}_n$ (where $n \geq 5$ and odd) with $\Pr[\boldsymbol{s}_n] = (\frac{1}{6} - p(0))^n = \frac{1}{6^n}$ s.t. $\delta_A(\boldsymbol{s}_n + (1,x,0)) - \delta_A(\boldsymbol{s}_n + (1,y,0)) > 0$. We construct this aggregate score-vector $\boldsymbol{s}_n$ by building on $\boldsymbol{s}_2 = (y,1,1+y)$ featured in the proof of Lemma 1 for $n = 2$. Let $n = 2(1+m) + 3\hat{m}$, where $\hat{m} \in \mathbb{N}$ and $m \in \mathbb{N} \cup \{0\}$. The idea behind this construction is as follows: Starting from some odd $n \geq 5$, subtract 2 from n. If the resulting odd number $n-2$ is divisible by 3, set $m = 0$ and $\hat{m} = (n-2)/3$. Otherwise, choose the smallest integer-value for m s.t. $n - 2(1+m)$ is divisible by three, and set $\hat{m} = (n - 2(1+m))/3$. For example, if $n = 9$, then $m = 2$ and $\hat{m} = 1$. If, instead, $n = 13$, then $m = 1$ and $\hat{m} = 3$. Now suppose that the aggregate score-vector across these $2(1+m) + 3\hat{m}$ voters is composed as follows:

$$\begin{aligned} \boldsymbol{s}_n &= \boldsymbol{s}_2 + \boldsymbol{s}_{3\hat{m}} + \boldsymbol{s}_{2m} \\ &= (y,1,1+y) + \boldsymbol{s}_{3\hat{m}} + \boldsymbol{s}_{2m} \end{aligned}$$

where $\boldsymbol{s}_{2m} = (m(1+y), m(1+y), 0)$ and $\boldsymbol{s}_{3\hat{m}} = (\hat{m}(1+y), \hat{m}(1+y), \hat{m}(1+y))$. That is, of the $2m$ voters whose individual score-vectors are aggregated into $\boldsymbol{s}_{2m}$, half have the individual score-vector $(1,y,0)$, while the remaining half have the individual score-vector $(y,1,0)$. On the basis of $\boldsymbol{s}_{2m}$, alternatives A and B are in a two-way tie. Of the $3\hat{m}$ voters whose individual score-vectors are aggregated into $\boldsymbol{s}_{3\hat{m}}$, a third have the individual score-vector $(1,y,0)$, another third have $(0,1,y)$, and the final third have $(y,0,1)$. That is, on the basis of $\boldsymbol{s}_{3\hat{m}}$, alternatives A, B, and C are in a three-way tie. Upon adding the score-vectors $\boldsymbol{s}_2$, $\boldsymbol{s}_{3\hat{m}}$, and $\boldsymbol{s}_{2m}$, we obtain the following aggregate score-vector $\boldsymbol{s}_n$ across the $2(1+m)+3\hat{m}$ voters other than i:

$$\boldsymbol{s}_n = (y+(1+y)(\hat{m}+m), 1+(1+y)(\hat{m}+m), (1+y)(\hat{m}+1))$$

Table 7 shows that if voter i submits the score-vector $\boldsymbol{\sigma}^i = (1,y,0)$, then alternatives A and B are in a tie for the highest score. If, instead, he submits the score-vector $\boldsymbol{\sigma}^i = (1,x,0)$, then alternative A wins outright. Therefore, submission of $\boldsymbol{\sigma}^i = (1,y,0)$ implies a shift in probability from A to B for any $\hat{m} \in \mathbb{N}$ and $m \in \mathbb{N} \cup 0$, relative to submission of $\boldsymbol{\sigma}^i = (1,y,0)$:

$\boldsymbol{\sigma}^i$	$\boldsymbol{s}_n+\boldsymbol{\sigma}^i$	$\delta(\boldsymbol{s}_n+\boldsymbol{\sigma}^i)$
$(1,x,0)$	$((1+y)(\hat{m}+m+1),1+x+(1+y)(\hat{m}+m),(1+y)(\hat{m}+1))$	$(1,0,0)$
$(1,y,0)$	$((1+y)(\hat{m}+m+1),(1+y)(\hat{m}+m+1),(1+y)(\hat{m}+1))$	$(\frac{1}{2},\frac{1}{2},0)$

Table 7: Score-vector $\boldsymbol{s}_n$ that shows $L(p(0)) > 0$ if $n = 2(1+m)+3\hat{m}$

As alternative A's loss in probability features in the numerator of the loss-gain-ratio in (5), it therefore immediately that $L(p(0)) > 0$. □

6.3 Proof of Proposition 2 (for five or more voters)

We present the proof by considering, in turn, an odd number of voters greater or equal to five (i.e. $n \geq 4$ and even), followed by an even number of voters greater or equal to six (i.e. $n \geq 5$ and odd).

6.3.1 Five or more voters (odd)

We have to show that $L(p(1)) < 1$. To this end, we highlight an appropriate aggregate score-vector $\boldsymbol{s}_n$ (where $n \geq 4$ and even) s.t. $\delta_B(\boldsymbol{s}_n + (1,y,0)) - \delta_B(\boldsymbol{s}_n + (1,x,0)) > \delta_A(\boldsymbol{s}_n + (1,x,0)) - \delta_A(\boldsymbol{s}_n + (1,y,0))$. More specifically, we construct this aggregate score-vector $\boldsymbol{s}_n$ by building on the score-vector $\boldsymbol{s}_2 = (x,1,1+x)$ featured in the proof of Proposition 2 for $n=2$ (see Section 3 in the main body of the paper). Let $n = 2(1+m)$ (where $m \in \mathbb{N}$) and suppose that the aggregate score-vector across these $2(1+m)$ voters is composed as follows:

$$\begin{aligned} \boldsymbol{s}_n &= \boldsymbol{s}_2 + \boldsymbol{s}_{2m} \\ &= (x,1,1+x) + \boldsymbol{s}_{2m} \end{aligned}$$

where $\boldsymbol{s}_{2m} = (0, m(1+x), m(1+x))$. That is, of the $2m$ voters whose individual score-vectors are aggregated into $\boldsymbol{s}_{2m}$, half have the individual score-vector $(0,1,x)$, while the remaining half have the individual score-vector $(0,x,1)$. On the basis of $\boldsymbol{s}_{2m}$, alternatives B and C are in a two-way tie. Upon adding the score-vectors $\boldsymbol{s}_2$ and $\boldsymbol{s}_{2m}$, we obtain the following aggregate score-vector $\boldsymbol{s}_n$ across the $2(1+m)$ voters other than i:

$$\boldsymbol{s}_n = (x, m(1+x)+1, (1+x)(m+1))$$

Note that $\Pr[\boldsymbol{s}_n] = p(\alpha)^n$, which is positive $\alpha = 1$. Table 8 shows that if voter i submits the score-vector $\boldsymbol{\sigma}^i = (1,y,0)$, then alternatives B wins outright. If, instead, he submits the score-vector $\boldsymbol{\sigma}^i = (1,x,0)$, then alternatives B and C are in a tie for the highest score. Therefore, submission of

$\boldsymbol{\sigma}^i$	$\hat{\boldsymbol{s}}_n+\boldsymbol{\sigma}^i$	$\boldsymbol{\delta}(\hat{\boldsymbol{s}}_n+\boldsymbol{\sigma}^i)$
$(1,x,0)$	$(1+x,(1+x)(m+1),(1+x)(m+1))$	$(0,\frac{1}{2},\frac{1}{2})$
$(1,y,0)$	$(1+x,(1+x)m+1+y,(1+x)(m+1))$	$(0,1,0)$

Table 8: Score-vector $\boldsymbol{s}_n$ that shows $L(p(1))<1$ for $n=2(1+m)$

$\boldsymbol{\sigma}^i=(1,y,0)$ implies a shift in probability mass from C to B for any $m\in\mathbb{N}$, relative to submission of $\boldsymbol{\sigma}^i=(1,x,0)$:
As alternative B's gain in probability from C exceeds A's loss (which is zero in event $\boldsymbol{s}_n$), it follows immediately that $L(p(1))<1$. As $L(p(0))>0$ by Lemma 1, we can appeal to the Intermediate Value Theorem to conclude that for any even $n\geq 4$, all $x<y$, and any distribution G, there exists an equilibrium weight $\alpha^*\in(0,1)$.

6.3.2 Six or more voters (even)

Next, we show that $L(p(1))<1$. To this end, we highlight an aggregate score-vector $\boldsymbol{s}_n$ s.t. $\delta_B(\boldsymbol{s}_n+(1,y,0))-\delta_B(\boldsymbol{s}_n+(1,x,0))>\delta_A(\boldsymbol{s}_n+(1,x,0))-\delta_A(\boldsymbol{s}_n+(1,y,0))$. We construct this aggregate score-vector $\boldsymbol{s}_n$ by building on the score-vector $\boldsymbol{s}_2=(x,1,1+x)$ featured in the proof of Proposition 2 for $n=2$. Again, let $n=2(1+m)+3\hat{m}$ (where $\hat{m}\in\mathbb{N}$ and $m\in\mathbb{N}\cup 0$) and suppose that the aggregate score-vector across these $2(1+m)+3\hat{m}$ voters is composed as follows:

$$\begin{aligned}\boldsymbol{s}_n &= \boldsymbol{s}_2+\boldsymbol{s}_{3\hat{m}}+\boldsymbol{s}_{2m}\\ &= (x,1,1+x)+\boldsymbol{s}_{3\hat{m}}+\boldsymbol{s}_{2m}\end{aligned}$$

where $\boldsymbol{s}_{2m}=(0,m(1+x),m(1+x))$ and $\boldsymbol{s}_{3\hat{m}}=(\hat{m}(1+x),\hat{m}(1+x),\hat{m}(1+x))$. That is, of the $2m$ voters whose individual score-vectors are aggregated into $\boldsymbol{s}_{2m}$, half have the individual score-vector $(0,1,x)$, while the remaining half have the individual score-vector $(0,x,1)$. On the basis of $\boldsymbol{s}_{2m}$, alternatives B and C are in a two-way tie. Of the $3\hat{m}$ voters whose individual score-vectors are aggregated into $\boldsymbol{s}_{3\hat{m}}$, a third have the individual score-vector $(1,x,0)$, another third have $(0,1,x)$, and the final third have $(x,0,1)$. That is, on the basis of $\boldsymbol{s}_{3\hat{m}}$, alternatives A, B, and C are in a three-way tie. Upon adding the score-vectors $\boldsymbol{s}_2$, $\boldsymbol{s}_{2m}$, and $\boldsymbol{s}_{3\hat{m}}$, we obtain the following aggregate score-vector $\boldsymbol{s}_n$ across the $2(1+m)+3\hat{m}$ voters other than i:

$$\boldsymbol{s}_n=(x+m(1+x),1+(1+x)(\hat{m}+m),(1+x)(\hat{m}+m+1))$$

Note that $\Pr[\boldsymbol{s}_n]=((1/6)p(\alpha))^n$, which is positive $\alpha=1$. Table 9 shows that if voter i submits the score-vector $\boldsymbol{\sigma}^i=(1,y,0)$, then alternatives B wins outright. If, instead, he submits the score-vector $\boldsymbol{\sigma}^i=(1,x,0)$, then alternatives B and C are in a tie for the highest score. Therefore, submission of $\boldsymbol{\sigma}^i=(1,y,0)$ implies shift in probability from C to B for any $\hat{m}\in\mathbb{N}$ and $m\in\mathbb{N}\cup 0$, relative to submission of $\boldsymbol{\sigma}^i=(1,x,0)$:

$\boldsymbol{\sigma}^i$	$\hat{\boldsymbol{s}}^1_n+\boldsymbol{\sigma}^i$	$\boldsymbol{\delta}(\hat{\boldsymbol{s}}_n+\boldsymbol{\sigma}^i)$
$(1,x,0)$	$((1+x)(m+1),(1+x)(\hat{m}+m+1),(1+x)(\hat{m}+m+1))$	$(0,\frac{1}{2},\frac{1}{2})$
$(1,y,0)$	$((1+x)(m+1),1+y+(1+x)(\hat{m}+m),(1+x)(\hat{m}+m+1))$	$(0,1,0)$

Table 9: Score-vector $\boldsymbol{s}_n$ that shows $L(p(1))<1$ if $n=2(1+m)+3\hat{m}$

As alternative B's gain in probability from C exceeds A's loss (which is zero in event $\boldsymbol{s}_n$), it follows immediately that $L(p(1)) < 1$. As $L(p(0)) > 0$ by Lemma 1, we can appeal to the Intermediate Value Theorem to conclude that for any odd $n \geq 5$, all $x < y$, and any distribution G, there exists an equilibrium weight $\alpha^* \in (0,1)$. □

6.4 Proof of Proposition 3

In the case where voter i interacts with just one other voter (i.e. $n = 1$), it is easy to list the twelve realizations of the aggregate score-vector $\boldsymbol{s}_1$ that can arise when voter $j \neq i$ uses a symmetric strategy in the sense of property S2. In fact, the aggregate score-vector $\boldsymbol{s}_1$ is simply given by voter j's individual score-vector $\boldsymbol{\sigma}^j$. In Table 10 we list all twelve possible realizations of $\boldsymbol{s}_1$.

Row	$\boldsymbol{s}_1$	$\Pr[\boldsymbol{s}_1]$	$\boldsymbol{s}_1+(1,x,0)$	$\boldsymbol{\delta}(\boldsymbol{s}_1+(1,x,0))$	$\boldsymbol{s}_1+(1,y,0)$	$\boldsymbol{\delta}(\boldsymbol{s}_1+(1,y,0))$
1.	$(1,x,0)$	$p(\alpha)$	$(2,2x,0)$	$(1,0,0)$	$(2,x+y,0)$	$(1,0,0)$
2.	$(1,y,0)$	$\frac{1}{6}-p(\alpha)$	$(2,x+y,0)$	$(1,0,0)$	$(2,2y,0)$	$(1,0,0)$
3.	$(1,0,x)$	$p(\alpha)$	$(2,x,x)$	$(1,0,0)$	$(2,y,x)$	$(1,0,0)$
4.	$(1,0,y)$	$\frac{1}{6}-p(\alpha)$	$(2,x,y)$	$(1,0,0)$	$(2,y,y)$	$(1,0,0)$
5.	$(x,1,0)$	$p(\alpha)$	$(1+x,1+x,0)$	$(\frac{1}{2},\frac{1}{2},0)$	$(1+x,1+y,0)$	$(0,1,0)$
6.	$(y,1,0)$	$\frac{1}{6}-p(\alpha)$	$(1+y,1+x,0)$	$(1,0,0)$	$(1+y,1+y,0)$	$(\frac{1}{2},\frac{1}{2},0)$
7.	$(0,1,x)$	$p(\alpha)$	$(1,1+x,x)$	$(0,1,0)$	$(1,1+y,x)$	$(0,1,0)$
8.	$(0,1,y)$	$\frac{1}{6}-p(\alpha)$	$(1,1+x,y)$	$(0,1,0)$	$(1,1+y,y)$	$(0,1,0)$
9.	$(x,0,1)$	$p(\alpha)$	$(1+x,x,1)$	$(1,0,0)$	$(1+x,y,1)$	$(1,0,0)$
10.	$(y,0,1)$	$\frac{1}{6}-p(\alpha)$	$(1+y,x,1)$	$(1,0,0)$	$(1+y,y,1)$	$(1,0,0)$
11.	$(0,x,1)$	$p(\alpha)$	$(1,2x,1)$	$(\delta_A,\delta_B,\delta_C)$	$(1,x+y,1)$	$(\delta_A,\delta_B,\delta_C)$
12.	$(0,y,1)$	$\frac{1}{6}-p(\alpha)$	$(1,x+y,1)$	$(\delta_A,\delta_B,\delta_C)$	$(1,2y,1)$	$(\delta_A,\delta_B,\delta_C)$

Table 10: All realizations of aggregate score-vector $\boldsymbol{s}_1$ in case of two voters

In our proof, we focus on the loss-gain-ratio $L(p(\alpha))$ in order to identify the equilibrium weight α^* as fixed point of L. Note that the loss-gain-ratio (defined in (5)) can be obtained readily from Table 10: The expected loss in the probability of alternative A (which appears in the numerator of the loss-gain-ratio) is given by the probability-weighted sum of the differences between the respective first entries of the probability-vectors in the fifth and seventh columns of the table. Similarly, the expected gain in the probability of alternative B (which appears in the denominator of the loss-gain-ratio) is given by the probability-weighted sum of the differences between the respective second entries of the probability-vectors in the seventh and fifth columns.

Obviously, only those aggregate score-vectors $\boldsymbol{s}_1$ for which voter i is pivotal matter for the loss-gain-ratio. As can be seen from Table 10, the vectors for which voter i is pivotal depend on the values of the model parameters x and y. In particular, in the 11th and 12th rows of the table, the distribution $(\delta_A, \delta_B, \delta_C)$ induced by the voters' individual score-vectors can only be determined conclusively by making assumptions about the values of x and y:

1. For all x and y in items 1.(a) - 1.(c) of Proposition 3, it is easy to verify using Table 10 that $L(p(\alpha))$ in each case is given by the respective fraction in items 1.(a) - 1.(c). Furthermore, in cases 1.(a) and 1.(b) we have $L(p(0)) > L(p(1))$ and $L'(p) < 0$ for all $p \in [0, \frac{1}{6}]$, while in case 1.(c) we have $L(p(0)) < L(p(1))$ and $L'(p) > 0$. It therefore follows immediately that L has a unique interior fixed point $\alpha^* \in (0,1)$ in each of these three cases.

2. For all values of x and y in item 2 of Proposition 3, it is easy to verify using Table 10 that $L(p(\alpha)) = 1$ for all α. Therefore, the unique fixed point of L is $\alpha^* = 1$.

3. For all x and y in item 3.(a), we have $L(p(\alpha)) = (2-6p(\alpha))/(3-12p(\alpha))$, with $L(p(0)) = 2/3$, $L(p(1)) = 1$, and $L'(p) > 0$ for all $p \in [0, \frac{1}{6}]$. Similarly, for all x and y in item 3.(b), the LGR is $L(p(\alpha)) = (4-6p(\alpha))/(5-12p(\alpha))$, with $L(p(0)) = 4/5$, $L(p(1)) = 1$, and $L'(p) > 0$ for all $p \in [0, \frac{1}{6}]$. Therefore, in both cases L has a fixed point at $\alpha^* = 1$. This is the unique fixed point of L iff $L(p(\alpha)) > \alpha$ for all $\alpha \in [0,1)$. By solving this inequality for $6p(\alpha)$ in cases 3.(a) and 3.(b), we obtain the conditions in (6) and (7), resp. To see that concavity of G is sufficient for these two conditions to hold, note that the function $6p(\alpha)$ is increasing and concave if G is concave:

$$p'(\alpha) = \int_0^1 \left(\int_0^{u_1} g(\alpha u_1 + (1-\alpha)u_3)(u_1 - u_3)g(u_3)du_3 \right) g(u_1)du_1 > 0,$$

$$p''(\alpha) = \int_0^1 \left(\int_0^{u_1} g'(\alpha u_1 + (1-\alpha)u_3)(u_1 - u_3)^2 g(u_3)du_3 \right) g(u_1)du_1 \leq 0.$$

As the function $6p(\alpha)$ takes the values $6p(0) = 0$ and $6p(1) = 1$, it is obvious that for concave G it holds that $6p(\alpha) \geq \alpha$ for all α.[40] Furthermore, it is easy to verify that the functions $(2-3\alpha)/(1-2\alpha)$ and $(4-5\alpha)/(1-2\alpha)$ in (6) and (7), resp., are strictly smaller than α for all $\alpha \in [0,1)$.[41] Therefore, it follows immediately that the conditions in (6) and (7) are satisfied. □

6.5 Proof of Proposition 4

We show first that the continuum of (x,y)-scoring rules with $x < y$ is partitioned in equilibrium into 64 equivalence classes. To see this, consider w.l.o.g. a voter i with ordinal ranking $A \succ_i B \succ_i C$. In order to determine the equilibrium weight α^* associated with the FSE voting strategy for given parameters x, y, and G, we need to compute the loss-gain-ration $L(p(\alpha))$ in (5). This, in turn, requires voter i to determine for each aggregate score-vector $\boldsymbol{s}_2 = (s_2^A, s_2^B, s_2^C)$ across the other two voters the distributions over outcomes $\boldsymbol{\delta}(\boldsymbol{s}_2 + (1,x,0))$ and $\boldsymbol{\delta}(\boldsymbol{s}_2 + (1,y,0))$ that arise when he assigns a score of x and y, resp., to his middle-ranked alternative. To obtain these distributions (which are defined in Definition 2) voter i must solve for every $\boldsymbol{s}_2 \in S_{x,y}^2$:

$$\max\{s_2^A + 1, s_2^B + x, s_2^C\} \text{ and } \max\{s_2^A + 1, s_2^B + y, s_2^C\}.$$

Observe that the components s_2^k (where $k \in \{A,B,C\}$) may themselves depend on the parameters x and y. To make explicit this dependence, we can rewrite the aggregate score-vector $\boldsymbol{s}_2$ by introducing the following vector: $\boldsymbol{\kappa} = (a_1, a_x, a_y, b_1, b_x, b_y, c_1, c_x, c_y)$, where a_1 is the number of voters other than i who assign a score of 1 to alternative A, a_x is the number of other voters who assign a score of x to alternative A, and a_y is the number of other voters who assign a score of y to alternative A. The remaining elements of the vector $\boldsymbol{\kappa}$ are defined analogously.[42] We can now write: $s_A^2 = a_1 + a_x x + a_y y$, and the remaining two components of $\boldsymbol{s}_2$ can be expressed analogously.

[40]To see this, note that by definition of concavity: $\alpha 6p(1) + (1-\alpha)6p(0) \leq 6p(\alpha \cdot 1 + (1-\alpha) \cdot 0) \Leftrightarrow \alpha \leq 6p(\alpha)$.

[41]To see this, note that (for $\alpha \in [0.5,1)$): $(2-3\alpha)/(1-2\alpha) < \alpha \Leftrightarrow (\alpha-1)^2 > 0$ and $(4-5\alpha)/(1-2\alpha) < \alpha \Leftrightarrow (\alpha-1)(\alpha-2) > 0$.

[42]Note that with three voters (i.e. $n = 2$) each component of $\boldsymbol{\kappa}$ is in $\{0,1,2\}$. Furthermore, $c_1 = 2 - a_1 - b_1$, and $c_x + c_y = 2 - a_x - a_y - b_x - b_y$.

With this notation, the distributions over outcomes $\boldsymbol{\delta}(\boldsymbol{s}_2+(1,x,0))$ and $\boldsymbol{\delta}(\boldsymbol{s}_2+(1,y,0))$, resp., are obtained by solving the following problems for every possible realization of the vector $\boldsymbol{\kappa}$:

$$\max\{a_1+1+a_x x+a_y y, b_1+(b_x+1)x+b_y y, c_1+c_x x+c_y y\} \tag{11}$$

and

$$\max\{a_1+1+a_x x+a_y y, b_1+b_x x+(b_y+1)y, c_1+c_x x+c_y y\}. \tag{12}$$

This shows that for every given $\boldsymbol{\kappa}$, the winning alternative in 11 and 12, resp., is determined by up to three linear inequalities involving the model parameters x and y. These inequalities partition the set $\{(x,y)\in[0,1]^2|0\le x<y\le 1\}$ of possible (x,y)-scoring rules. The intersection of all these partitions across all possible vectors $\boldsymbol{\kappa}$ gives rise to the different equivalence classes, each of which is associated with a distinct loss-gain-ratio. With three voters ($n=2$), there are a total of $12^n=144$ constellations of individual score-vectors across the two voters other than i, and these give rise to 75 distinct aggregate score-vectors $\boldsymbol{s}_2$.[43] For each $\boldsymbol{s}_2$ (and the associated vector $\boldsymbol{\kappa}$), we have computed the distributions over outcomes $\boldsymbol{\delta}(\boldsymbol{s}_2+(1,x,0))$ and $\boldsymbol{\delta}(\boldsymbol{s}_2+(1,y,0))$ as a function of the parameters x and y by solving (11) and (12). In addition to partitioning the set of (x,y)-scoring rules into its interior and its boundaries (as well as its corner points $(0,0)$, $(0,1)$, and $(1,1)$), the solutions to (11) and (12) give rise to the interior partitions of the set $\{(x,y)\in[0,1]^2|0\le x<y\le 1\}$ listed in the following table. The (in-)equalities in the first column of the table give rise to the dashed lines in Fig. 3 that demarcate the 64 equivalence classes.

Partition	Aggregate score-vector $\boldsymbol{s}_2$ giving rise to partition
$y \lesseqqgtr 1-x$	$(2,1+x+y,x)$, $(2,1+x+y,y)$, $(1+x+y,y,2)$, $(1+x+y,x,2)$
$y \lesseqqgtr 2-2x$	$(2,2x+y,1)$, $(1,2x+y,2)$
$y \lesseqqgtr \frac{1}{2}$	$(2,1+2y,x)$, $(2,1+2y,y)$, $(1+2y,y,2)$, $(1+2y,x,2)$
$y \lesseqqgtr 2x$	$(1+2x,1+y,1)$, $(1,1+2x,1+y)$
$y \lesseqqgtr 1-\frac{1}{2}x$	$(1,x+2y,2)$, $(2,x+2y,1)$
$x \lesseqqgtr \frac{1}{2}$	$(1+2x,y,2)$, $(1+2x,x,2)$, $(2,1+2x,y)$, $(2,1+2x,x)$
$y \lesseqqgtr \frac{2}{3}$	$(2,3y,1)$, $(1,3y,2)$
$x \lesseqqgtr \frac{2}{3}$	$(2,3x,1)$, $(1,3x,2)$
$y \lesseqqgtr \frac{1}{2}+\frac{1}{2}x$	$(2+x,1+2y,0)$, $(1+2y,2+x,0)$

Having established that the continuum of (x,y)-scoring rules is partitioned into a finite number of distinct equivalence classes regardless of the distribution G, we now compute the loss-gain-ratio for each equivalence class. Observe that for $n=2$, the loss-gain-ratio in (5) can be expressed equivalently in the following form:

$$L(p(\alpha))=\frac{n_1\left(\frac{1}{6}-p(\alpha)\right)^2+n_2\left(\frac{1}{6}-p(\alpha)\right)p(\alpha)+n_3\left(p(\alpha)\right)^2}{d_1\left(\frac{1}{6}-p(\alpha)\right)^2+d_2\left(\frac{1}{6}-p(\alpha)\right)p(\alpha)+d_3\left(p(\alpha)\right)^2} \tag{13}$$

[43] For the sake of brevity, we do not list explicitly here these 75 aggregate score-vectors. However, it is easy to do so (to facilitate this process, we have used Mathematica).

The values of n_1, n_2, and n_3 (d_1, d_2, and d_3, resp.) depend on the equivalence class being considered. For example, n_1 (d_1, resp.) represents the sum $\sum_{\boldsymbol{s}_2}(\delta_A(\boldsymbol{s}_2+(1,x,0))-\delta_A(\boldsymbol{s}_2+(1,y,0)))$ $(\sum_{\boldsymbol{s}_2}(\delta_B(\boldsymbol{s}_2+(1,y,0))-\delta_B(\boldsymbol{s}_2+(1,x,0)))$, resp.) over all aggregate score-vectors $\boldsymbol{s}_2$ that arise when both voters other than i assign the higher score of y to their respective middle-ranked alternatives. The remaining coefficients of the loss-gain-ratio in (13) are defined analogously.

Table 11 at the end of this proof lists for each equivalence class the associated values of the coefficients n_1, n_2, n_3, d_1, d_2, and d_3 in the loss-gain-ratio in (13). Furthermore, the table lists the values of the loss-gain-ratio $L(p(\alpha))$ at $\alpha = 0$ (where $p(0) = 0$) and $\alpha = 1$ (where $p(1) = 1/6$), as well providing information about the monotonicity of the loss-gain-ratio. In the column titled 'Mono' we indicate by '$+$' that the loss-gain-ratio is strictly increasing, by '$-$' the fact that the loss-gain-ratio is strictly decreasing, and by 'const.' the fact that the loss-gain-ratio is constant. As the values of $L(0)$ and $L(\frac{1}{6})$ are strictly between 0 and 1 for every equivalence class, it follows immediately that there is a unique interior fixed point (i.e. a point α^* s.t. $L(p(\alpha^*)) = \alpha^*$) for every equivalence class that generates a monotone loss-gain-ratio.[44]

Observe, however, that Table 11 also contains equivalence classes for which the loss-gain-ratio is non-monotonic. In the column titled 'Mono', we denote by '$\mp$' a loss-gain-ratio that is decreasing up to some turning point $\bar{\alpha}$, and which is increasing thereafter. Similarly, the symbol '$\pm$' indicates a loss-gain-ratio that is increasing up to some turning point $\bar{\alpha}$, and which is decreasing thereafter. For equivalence classes with a non-monotonic loss-gain-ratio, the column titled 'TP' provides an implicit statement of the associated turning point $\bar{\alpha}$ by listing the value of $p(\bar{\alpha})$. As $p(\bar{\alpha}) \in (0, 1/6)$ for each non-monotonic loss-gain ratio, it follows that the corresponding value $\bar{\alpha}$ is strictly between 0 and 1.[45]

In order to establish uniqueness of the equilibrium weight α^* for equivalence classes with a non-monotonic loss-gain-ratio, we can draw upon the two sufficient conditions stated in Proposition 4: (i) if the loss-gain-ratio $L(p(\alpha))$ is a contraction of $[0,1]$, then $L(p(\alpha))$ has a unique fixed point. As the loss-gain-ratio is a differentiable function of α, the following condition is necessary and sufficient for $L(p(\alpha))$ to be a contraction: $|L'(p(\alpha))p'(\alpha)| \leq k < 1$ for all $\alpha \in (0,1)$. Unfortunately, it is impossible to verify this condition in general and without specifying the distribution G that underlies the function $p(\alpha)$; (ii) if the loss-gain-ratio is either everywhere strictly convex, or everywhere strictly concave, it follows that the function $L(p(\alpha))$ has a unique fixed point because it takes values $L(p(0)), L(p(1)) \in (0,1)$ at the boundaries of the interval $[0,1]$. This implies immediately that the function $L(p(\alpha))$ can intersect the 45-degree line only once. Unfortunately, it is impossible to verify convexity or concavity of the loss-gain-ratio in general, as the sign of its second derivative $L''(p(\alpha))(p'(\alpha))^2 + L'(p(\alpha))p''(\alpha)$ depends on the magnitudes of the derivatives $p'(\alpha)$ and $p''(\alpha)$, and therefore on the distribution G that underlies the function $p(\alpha)$. □

[44]Note that the first derivative of the loss-gain-ratio is given by $L'(p(\alpha))p'(\alpha)$. As $p'(\alpha) > 0$ by definition, the sign of the derivative depends on the sign of the function $L'(p)$. As long as $L'(p) \lesseqgtr 0$ for all $p \in [0, \frac{1}{6}]$, it follows that the loss-gain-ratio is monotone.

[45]For equivalence classes with non-monotonic loss-gain-ratio, our results show that either $L'(p) < 0$ for $0 < p < p(\bar{\alpha})$ and $L'(p) > 0$ for $p(\bar{\alpha}) < p < 1/6$, or vice versa. In the former case, the first derivative of the loss-gain-ratio $L'(p(\alpha))p'(\alpha)$ has the following sign: $L'(p(\alpha))p'(\alpha) < 0$ for $0 < \alpha < \bar{\alpha}$ and $L'(p(\alpha))p'(\alpha) > 0$ for $\bar{\alpha} < \alpha < 1$. In the latter case, the reverse holds.

Class	(x,y)	n_1	n_2	n_3	d_1	d_2	d_3	$L(0)$	$L(\frac{1}{6})$	Mono	TP
1.	$x \in (0, \frac{1}{2})$ $y \in (x, \min\{2x, \frac{1}{2}\})$	$\frac{1}{3}$	1	$\frac{2}{3}$	$\frac{2}{3}$	2	$\frac{4}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	const.	N/A
2.	$x \in (0, \frac{1}{4})$ $y = 2x$	$\frac{1}{3}$	2	$\frac{5}{3}$	$\frac{2}{3}$	4	$\frac{7}{3}$	$\frac{1}{2}$	$\frac{5}{7}$	$+$	N/A
3.	$x \in (0, \frac{1}{4})$ $y \in (2x, \frac{1}{2})$	$\frac{1}{3}$	3	$\frac{8}{3}$	$\frac{2}{3}$	6	$\frac{10}{3}$	$\frac{1}{2}$	$\frac{4}{5}$	$+$	N/A
4.	$x = 0$ $y \in (0, \frac{1}{2})$	$\frac{4}{3}$	2	$\frac{8}{3}$	$\frac{8}{3}$	6	$\frac{16}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	$\mp$	$\frac{\sqrt{2}-1}{6}$
5.	$x \in (\frac{1}{4}, \frac{1}{2})$ $y = \frac{1}{2}$	$\frac{4}{3}$	2	$\frac{2}{3}$	$\frac{5}{3}$	3	$\frac{4}{3}$	$\frac{4}{5}$	$\frac{1}{2}$	$-$	N/A
6.	$x = \frac{1}{4}$ $y = \frac{1}{2}$	$\frac{1}{3}$	3	$\frac{5}{3}$	$\frac{5}{3}$	5	$\frac{7}{3}$	$\frac{1}{5}$	$\frac{5}{7}$	$+$	N/A
7.	$x \in (0, \frac{1}{4})$ $y = \frac{1}{2}$	$\frac{4}{3}$	4	$\frac{8}{3}$	$\frac{5}{3}$	7	$\frac{10}{3}$	$\frac{4}{5}$	$\frac{4}{5}$	$\mp$	$\frac{\sqrt{2}-1}{6}$
8.	$x = 0$ $y = \frac{1}{2}$	$\frac{17}{6}$	6	$\frac{8}{3}$	$\frac{25}{6}$	8	$\frac{16}{3}$	$\frac{17}{25}$	$\frac{1}{2}$	$\pm$	$\frac{\sqrt{368}-19}{6}$
9.	$x \in (\frac{1}{2}, \frac{2}{3})$ $y \in (x, \frac{2}{3})$	$\frac{1}{3}$	1	$\frac{2}{3}$	$\frac{2}{3}$	2	$\frac{4}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	const.	N/A
10.	$x = \frac{1}{2}$ $y \in (\frac{1}{2}, \frac{2}{3})$	$\frac{1}{3}$	2	$\frac{5}{3}$	$\frac{2}{3}$	3	$\frac{7}{3}$	$\frac{1}{2}$	$\frac{5}{7}$	$+$	N/A
11.	$x \in (\frac{1}{3}, \frac{1}{2})$ $y \in (1-x, \frac{2}{3})$	$\frac{1}{3}$	3	$\frac{8}{3}$	$\frac{2}{3}$	4	$\frac{10}{3}$	$\frac{1}{2}$	$\frac{4}{5}$	$+$	N/A
12.	$x \in (\frac{1}{3}, \frac{1}{2})$ $y = 1-x$	$\frac{4}{3}$	3	$\frac{5}{3}$	$\frac{5}{3}$	4	$\frac{7}{3}$	$\frac{4}{5}$	$\frac{5}{7}$	$+$	N/A
13.	$x \in (\frac{1}{4}, \frac{1}{2})$ $y \in (\frac{1}{2}, \min\{2x, 1-x\})$	$\frac{7}{3}$	3	$\frac{2}{3}$	$\frac{8}{3}$	4	$\frac{4}{3}$	$\frac{7}{8}$	$\frac{1}{2}$	$-$	N/A
14.	$x \in (\frac{1}{4}, \frac{1}{3})$ $y = 2x$	$\frac{7}{3}$	4	$\frac{5}{3}$	$\frac{8}{3}$	6	$\frac{7}{3}$	$\frac{7}{8}$	$\frac{5}{7}$	$\mp$	$\frac{7-\sqrt{29}}{12}$
15.	$x \in (0, \frac{1}{3})$ $y \in (\max\{\frac{1}{2}, 2x\}, \frac{1+x}{2})$	$\frac{7}{3}$	5	$\frac{8}{3}$	$\frac{8}{3}$	8	$\frac{10}{3}$	$\frac{7}{8}$	$\frac{4}{5}$	$\mp$	$\frac{\sqrt{57}-7}{6}$
16.	$x \in (0, \frac{1}{3})$ $y = \frac{1+x}{2}$	$\frac{17}{6}$	6	$\frac{8}{3}$	$\frac{19}{6}$	9	$\frac{10}{3}$	$\frac{17}{19}$	$\frac{4}{5}$	$\mp$	$\frac{11}{6} - \sqrt{3}$
17.	$x \in (0, \frac{1}{3})$ $y \in (\frac{1+x}{2}, \frac{2}{3})$	$\frac{10}{3}$	7	$\frac{8}{3}$	$\frac{11}{3}$	10	$\frac{10}{3}$	$\frac{10}{11}$	$\frac{4}{5}$	$\mp$	$\frac{19-\sqrt{246}}{30}$
18.	$x = 0$ $y \in (\frac{1}{2}, \frac{2}{3})$	$\frac{13}{3}$	8	$\frac{8}{3}$	$\frac{17}{3}$	10	$\frac{16}{3}$	$\frac{13}{17}$	$\frac{1}{2}$	$\pm$	$\frac{5-\sqrt{24}}{6}$

Table 11 continues overleaf...

Class	(x,y)	n_1	n_2	n_3	d_1	d_2	d_3	$L(0)$	$L(\frac{1}{6})$	Mono	TP
19.	$x \in (\frac{1}{2}, \frac{2}{3})$ $y = \frac{2}{3}$	$\frac{4}{3}$	1	$\frac{2}{3}$	$\frac{13}{6}$	2	$\frac{4}{3}$	$\frac{8}{13}$	$\frac{1}{2}$	$-$	N/A
20.	$x = \frac{1}{2}$ $y = \frac{2}{3}$	$\frac{4}{3}$	2	$\frac{5}{3}$	$\frac{13}{6}$	3	$\frac{7}{3}$	$\frac{8}{13}$	$\frac{5}{7}$	$+$	N/A
21.	$x \in (\frac{1}{3}, \frac{1}{2})$ $y = \frac{2}{3}$	$\frac{4}{3}$	3	$\frac{8}{3}$	$\frac{13}{6}$	4	$\frac{10}{3}$	$\frac{8}{13}$	$\frac{4}{5}$	$+$	N/A
22.	$x = \frac{1}{3}$ $y = \frac{2}{3}$	$\frac{17}{6}$	5	$\frac{8}{3}$	$\frac{11}{3}$	7	$\frac{10}{3}$	$\frac{17}{22}$	$\frac{4}{5}$	$\mp$	$\frac{11-\sqrt{112}}{6}$
23.	$x \in (0, \frac{1}{3})$ $y = \frac{2}{3}$	$\frac{13}{3}$	7	$\frac{8}{3}$	$\frac{31}{6}$	10	$\frac{10}{3}$	$\frac{26}{31}$	$\frac{4}{5}$	$\mp$	$\frac{39-\sqrt{876}}{90}$
24.	$x = 0$ $y = \frac{2}{3}$	$\frac{16}{3}$	8	$\frac{8}{3}$	$\frac{43}{6}$	10	$\frac{16}{3}$	$\frac{32}{43}$	$\frac{1}{2}$	$\pm$	$\frac{10-\sqrt{85}}{30}$
25.	$x \in (\frac{2}{3}, 1)$ $y \in (x, \frac{1+x}{2})$	$\frac{1}{3}$	1	$\frac{2}{3}$	$\frac{2}{3}$	2	$\frac{4}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	const.	N/A
26.	$x = \frac{2}{3}$ $y \in (\frac{2}{3}, \frac{5}{6})$	$\frac{1}{3}$	1	$\frac{5}{3}$	$\frac{2}{3}$	2	$\frac{17}{6}$	$\frac{1}{2}$	$\frac{10}{17}$	$+$	N/A
27.	$x \in (\frac{6}{10}, \frac{2}{3})$ $y \in (2(1-x), \frac{1+x}{2})$	$\frac{1}{3}$	1	$\frac{8}{3}$	$\frac{2}{3}$	2	$\frac{13}{3}$	$\frac{1}{2}$	$\frac{8}{13}$	$+$	N/A
28.	$x \in (\frac{6}{10}, \frac{2}{3})$ $y = 2(1-x)$	$\frac{1}{3}$	3	$\frac{5}{3}$	$\frac{2}{3}$	5	$\frac{17}{6}$	$\frac{1}{2}$	$\frac{10}{17}$	$\pm$	$\frac{\sqrt{3}-1}{6}$
29.	$x \in (\frac{1}{2}, \frac{2}{3})$ $y \in (\frac{2-x}{2}, \min\{\frac{1+x}{2}, 2(1-x)\})$	$\frac{1}{3}$	5	$\frac{2}{3}$	$\frac{2}{3}$	8	$\frac{4}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	$\pm$	$\frac{\sqrt{2}-1}{6}$
30.	$x \in (\frac{1}{2}, \frac{2}{3})$ $y = \frac{2-x}{2}$	$\frac{4}{3}$	3	$\frac{2}{3}$	$\frac{13}{6}$	5	$\frac{4}{3}$	$\frac{8}{13}$	$\frac{1}{2}$	$-$	N/A
31.	$x \in (\frac{1}{2}, \frac{2}{3})$ $y \in (\frac{2}{3}, \frac{2-x}{2})$	$\frac{7}{3}$	1	$\frac{2}{3}$	$\frac{11}{6}$	2	$\frac{4}{3}$	$\frac{7}{11}$	$\frac{1}{2}$	$-$	N/A
32.	$x = \frac{1}{2}$ $y \in (\frac{2}{3}, \frac{3}{4})$	$\frac{7}{3}$	2	$\frac{5}{3}$	$\frac{11}{3}$	3	$\frac{7}{3}$	$\frac{7}{11}$	$\frac{5}{7}$	$+$	N/A
33.	$x \in (\frac{1}{3}, \frac{1}{2})$ $y \in (\frac{2}{3}, \frac{1+x}{2})$	$\frac{7}{3}$	3	$\frac{8}{3}$	$\frac{11}{3}$	4	$\frac{10}{3}$	$\frac{7}{11}$	$\frac{4}{5}$	$+$	N/A
34.	$x \in (\frac{1}{3}, \frac{1}{2})$ $y = \frac{1+x}{2})$	$\frac{17}{6}$	4	$\frac{8}{3}$	$\frac{25}{6}$	5	$\frac{10}{3}$	$\frac{17}{25}$	$\frac{4}{5}$	$+$	N/A
35.	$x \in (\frac{1}{2}, \frac{2}{3})$ $y \in (\frac{1+x}{2}, \min\{2x, \frac{2-x}{2}\})$	$\frac{10}{3}$	5	$\frac{8}{3}$	$\frac{14}{3}$	6	$\frac{10}{3}$	$\frac{5}{7}$	$\frac{4}{5}$	$\pm$	$\frac{5}{36}$

Table 11 continues on next page...

Class	(x,y)	n_1	n_2	n_3	d_1	d_2	d_3	$L(0)$	$L(\frac{1}{6})$	Mono	TP
36.	$x \in (\frac{1}{3}, \frac{4}{10})$ $y = 2x$	$\frac{10}{3}$	6	$\frac{11}{3}$	$\frac{14}{3}$	8	$\frac{13}{3}$	$\frac{5}{7}$	$\frac{11}{13}$	$+$	N/A
37.	$x \in (0, \frac{4}{10})$ $y \in (\max\{1-x, 2x\}, \frac{2-x}{2})$	$\frac{10}{3}$	7	$\frac{14}{3}$	$\frac{14}{3}$	10	$\frac{16}{3}$	$\frac{5}{7}$	$\frac{7}{8}$	$\mp$	$\frac{\sqrt{50}-7}{6}$
38.	$x \in (0, \frac{1}{3})$ $y = 1-x$	$\frac{13}{3}$	7	$\frac{11}{3}$	$\frac{17}{3}$	10	$\frac{13}{3}$	$\frac{13}{17}$	$\frac{11}{13}$	$\mp$	$\frac{17-\sqrt{245}}{24}$
39.	$x \in (0, \frac{1}{3})$ $y \in (\frac{2}{3}, 1-x)$	$\frac{16}{3}$	7	$\frac{8}{3}$	$\frac{20}{3}$	10	$\frac{10}{3}$	$\frac{4}{5}$	$\frac{4}{5}$	$\mp$	$\frac{2-\sqrt{2}}{6}$
40.	$x = 0$ $y \in (\frac{2}{3}, 1)$	$\frac{19}{3}$	8	$\frac{8}{3}$	$\frac{26}{3}$	10	$\frac{16}{3}$	$\frac{19}{26}$	$\frac{1}{2}$	$\pm$	$\frac{25-\sqrt{472}}{102}$
41.	$x = \frac{2}{3}$ $y = \frac{5}{6}$	$\frac{5}{6}$	2	$\frac{5}{3}$	$\frac{7}{6}$	3	$\frac{17}{6}$	$\frac{5}{7}$	$\frac{10}{17}$	$-$	N/A
42.	$x = \frac{6}{10}$ $y = \frac{8}{10}$	$\frac{5}{6}$	4	$\frac{5}{3}$	$\frac{7}{6}$	6	$\frac{17}{6}$	$\frac{5}{7}$	$\frac{10}{17}$	$-$	N/A
43.	$x \in (\frac{6}{10}, \frac{2}{3})$ $y = \frac{1+x}{2}$	$\frac{5}{6}$	2	$\frac{8}{3}$	$\frac{7}{6}$	3	$\frac{13}{3}$	$\frac{5}{7}$	$\frac{8}{13}$	$-$	N/A
44.	$x \in (\frac{1}{2}, \frac{6}{10})$ $y = \frac{1+x}{2}$	$\frac{5}{6}$	6	$\frac{2}{3}$	$\frac{7}{6}$	9	$\frac{4}{3}$	$\frac{5}{7}$	$\frac{1}{2}$	$-$	N/A
45.	$x = \frac{1}{2}$ $y = \frac{3}{4}$	$\frac{11}{6}$	5	$\frac{5}{3}$	$\frac{8}{3}$	7	$\frac{7}{3}$	$\frac{11}{16}$	$\frac{5}{7}$	$+$	N/A
46.	$x \in (\frac{4}{10}, \frac{1}{2})$ $y = \frac{2-x}{2}$	$\frac{7}{3}$	7	$\frac{8}{3}$	$\frac{19}{6}$	9	$\frac{10}{3}$	$\frac{14}{19}$	$\frac{4}{5}$	$+$	N/A
47.	$x = \frac{4}{10}$ $y = \frac{8}{10}$	$\frac{7}{3}$	8	$\frac{11}{3}$	$\frac{19}{6}$	11	$\frac{13}{3}$	$\frac{14}{19}$	$\frac{11}{13}$	$\mp$	$\frac{\sqrt{149}-11}{84}$
48.	$x \in (0, \frac{4}{10})$ $y = \frac{2-x}{2}$	$\frac{7}{3}$	9	$\frac{14}{3}$	$\frac{19}{6}$	13	$\frac{16}{3}$	$\frac{14}{19}$	$\frac{7}{8}$	$\mp$	$\frac{\sqrt{1032}-25}{222}$
49.	$x \in (\frac{2}{3}, 1)$ $y = \frac{1+x}{2}$	$\frac{5}{6}$	2	$\frac{2}{3}$	$\frac{7}{6}$	3	$\frac{4}{3}$	$\frac{5}{7}$	$\frac{1}{2}$	$-$	N/A
50.	$x \in (\frac{2}{3}, 1)$ $y \in (\frac{1+x}{2}, 1)$	$\frac{4}{3}$	3	$\frac{2}{3}$	$\frac{5}{3}$	4	$\frac{4}{3}$	$\frac{4}{5}$	$\frac{1}{2}$	$-$	N/A
51.	$x = \frac{2}{3}$ $y \in (\frac{5}{6}, 1)$	$\frac{4}{3}$	3	$\frac{5}{3}$	$\frac{5}{3}$	4	$\frac{17}{6}$	$\frac{4}{5}$	$\frac{10}{17}$	$-$	N/A
52.	$x \in (\frac{1}{2}, \frac{2}{3})$ $y \in (\max\{2(1-x), \frac{1+x}{2}\}, 1)$	$\frac{4}{3}$	3	$\frac{8}{3}$	$\frac{5}{3}$	4	$\frac{13}{3}$	$\frac{4}{5}$	$\frac{8}{13}$	$-$	N/A

Table 11 continues overleaf...

Class	(x,y)	n_1	n_2	n_3	d_1	d_2	d_3	$L(0)$	$L(\frac{1}{6})$	Mono	TP
53.	$x \in (\frac{1}{2}, \frac{6}{10})$ $y = 2(1-x)$	$\frac{4}{3}$	5	$\frac{5}{3}$	$\frac{5}{3}$	7	$\frac{17}{6}$	$\frac{4}{5}$	$\frac{10}{17}$	$-$	N/A
54.	$x \in (\frac{1}{2}, \frac{6}{10})$ $y \in (\frac{1+x}{2}, 2(1-x))$	$\frac{4}{3}$	7	$\frac{2}{3}$	$\frac{5}{3}$	10	$\frac{4}{3}$	$\frac{4}{5}$	$\frac{1}{2}$	$-$	N/A
55.	$x = \frac{1}{2}$ $y \in (\frac{3}{4}, 1)$	$\frac{4}{3}$	8	$\frac{5}{3}$	$\frac{5}{3}$	11	$\frac{7}{3}$	$\frac{4}{5}$	$\frac{5}{7}$	$-$	N/A
56.	$x \in (\frac{4}{10}, \frac{1}{2})$ $y \in (\frac{2-x}{2}, 2x)$	$\frac{4}{3}$	9	$\frac{8}{3}$	$\frac{5}{3}$	12	$\frac{10}{3}$	$\frac{4}{5}$	$\frac{4}{5}$	$\mp$	$\frac{\sqrt{2}-1}{6}$
57.	$x \in (\frac{4}{10}, \frac{1}{2})$ $y = 2x$	$\frac{4}{3}$	10	$\frac{11}{3}$	$\frac{5}{3}$	14	$\frac{13}{3}$	$\frac{4}{5}$	$\frac{11}{13}$	$\mp$	$\frac{\sqrt{145}-7}{96}$
58.	$x \in (0, \frac{1}{2})$ $y \in (\max\{\frac{2-x}{2}, 2x\}, 1)$	$\frac{4}{3}$	11	$\frac{14}{3}$	$\frac{5}{3}$	16	$\frac{16}{3}$	$\frac{4}{5}$	$\frac{7}{8}$	$\mp$	$\frac{\sqrt{436}-11}{210}$
59.	$x \in (\frac{2}{3}, 1)$ $y = 1$	$\frac{10}{3}$	6	$\frac{8}{3}$	$\frac{14}{3}$	8	$\frac{10}{3}$	$\frac{5}{7}$	$\frac{4}{5}$	$+$	N/A
60.	$x = \frac{2}{3}$ $y = 1$	$\frac{10}{3}$	6	$\frac{11}{3}$	$\frac{14}{3}$	8	$\frac{29}{6}$	$\frac{5}{7}$	$\frac{22}{29}$	$+$	N/A
61.	$x \in (\frac{1}{2}, \frac{2}{3})$ $y = 1$	$\frac{10}{3}$	6	$\frac{14}{3}$	$\frac{14}{3}$	8	$\frac{19}{3}$	$\frac{5}{7}$	$\frac{14}{19}$	$\pm$	$\frac{\sqrt{3}-1}{6}$
62.	$x = \frac{1}{2}$ $y = 1$	$\frac{10}{3}$	10	$\frac{17}{3}$	$\frac{14}{3}$	14	$\frac{41}{6}$	$\frac{5}{7}$	$\frac{31}{41}$	$+$	N/A
63.	$x \in (0, \frac{1}{2})$ $y = 1$	$\frac{10}{3}$	14	$\frac{20}{3}$	$\frac{14}{3}$	20	$\frac{22}{3}$	$\frac{5}{7}$	$\frac{10}{11}$	$\mp$	$\frac{\sqrt{12}-3}{36}$
64.	$x = 0$ $y = 1$	$\frac{22}{3}$	16	$\frac{20}{3}$	$\frac{32}{3}$	20	$\frac{28}{3}$	$\frac{11}{16}$	$\frac{5}{7}$	$\pm$	$\frac{8-\sqrt{55}}{6}$

Table 11: loss-gain-ratios of all equivalence classes in settings with three voters

References

Ahn DS, Oliveros S (2010) Approval Voting and Scoring Rules with Common Values. University of California Berkeley mimeo.

Apesteguia J, Ballester MA, Ferrer R (2011) On the Justice of Decision Rules. Review of Economic Studies 78:1–16

Blais A (2002) Why Is there So Little Strategic Voting in Canadian Plurality Rule Elections? Political Studies 50:445–454

Börgers T, Postl P (2009) Efficient compromising. Journal of Economic Theory 144:2057–2076

Buenrostro L, Dhillon A, Vida P (2013) Scoring rule voting games and dominance solvability. Social Choice and Welfare 40:329–352

Carmona G (2012) A Voting Model Generically Yielding Sincere Voting in Large Elections. University of Cambridge mimeo.

Cook WD, Kress M (1985) Ordinal Ranking with Intensity of Preference. Management Science 31:26–32

Giles A (2011) Extended Essay: A Comparison of (A,B)-Scoring Rules. London School of Economics mimeo.

Holmström B, Myerson RB (1983) Efficient and durable decision rules with incomplete information. Econometrica 51:1799–1819

Hortala-Vallve R (2009) Inefficiencies on linking decisions. Social Choice and Welfare 34:471–486

Jackson MO, Sonnenschein HF (2007) Overcoming incentive constraints by linking decisions. Econometrica 75:241–257

Kim S (2012) Ordinal versus Cardinal Voting Rules: A Mechanism Design Approach. Ohio State University mimeo.

Meek BL (1975) A transferable voting system including intensity of preference. Mathématique et sciences humaines 50:23–29

Miralles A (2012) Cardinal Bayesian allocation mechanisms without transfers. Journal of Economic Theory 147:179–206

Myerson RB (2002) Comparison of Scoring Rules in Poisson Voting Games. Journal of Economic Theory 103:219–251

Myerson RB, Weber RJ (1993) A Theory of Voting Equilibria. American Political Science Review 87:102–114

Nurmi H (1981) Approaches to collective decision making with fuzzy preference relations. Fuzzy Sets and Systems 6:249–259

Nurmi H (1993) Fuzzy social choice: a selective retrospect. Soft Computing 12:281–288

Postl P (2008) Symmetric Bayes Nash equilibria of three popular voting rules-some simple examples. University of Birmingham mimeo.

Postl P (2011) Strategy-Proof Compromises. The BE Journal of Theoretical Economics 11:Article 20

Ritzberger K (2002) Foundations of Non-Cooperative Game Theory. Oxford University Press, Oxford, UK

Rochet JC (1987) A necessary and sufficient condition for rationalizability in a quasi-linear context. Journal of Mathematical Economics 16:191–200

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

Weber RJ (1978) Comparison of public choice systems. Cowles Foundation Discussion Paper No. 498