Evaluating and Comparing Voting Rules behind the Veil of Ignorance: a Brief and Selective Survey and an Analysis of Two-Parameter Scoring Rules*

PETER POSTL†

January 2017

Abstract

We propose a general framework for the study and evaluation of voting rules behind the 'veil of ignorance'. A selective review of the voting literature shows how many important contributions can be couched as special cases of this general framework. As many studies of voting rules for three or more candidates ignore the issue of strategic voting, and many fully-fledged mechanism design approaches to the design of optimal voting rules focus on the case of elections with two candidates, we then present and discuss special case of the general framework in which the cardinal preferences of voters over three alternatives are private information. In this setting, we study voting rules that are 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. Whilst 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.

*I am indebted to Michel Le Breton for his guidance and support, as well as the feedback of two anonymous referees. Any remaining errors are my own.

†Department of Economics, University of Bath, Bath BA2 7AY, United Kingdom, p.postl@bath.ac.uk.

1 Introduction

An important question in the context of the design of voting rules for collective decision-making is whether a 'good' voting rule 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 in this survey. Early attempts to study or design voting systems that allow voters to express preference intensity can be found in the computer science, operations research, and political science literatures (see Meek, 1975; Nurmi, 1981; Merrill, 1984; Cook and Kress, 1985; Merrill, 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.[1]

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.[2] 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 to 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).[3] 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. The key difference with Weber (1978) is that there are instances of our setting where the game induced by each voting rule 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.[4]

The voting rules that we will focus on in the latter part of this survey are two-parameter scoring rules which include, as special cases, well-known voting procedures such as the plurality rule,

[1] In contrast to the so-called axiomatic voting literature that focuses on whether (or not) a voting rule satisfies a number of desirable properties, the literature to which the present survey mostly refers comes primarily from a mechanism design tradition where given voting rules are compared according to some criterion that is often (but not always) taken to be social welfare. A key feature of a large part of this literature is that the behavioral equilibrium responses of the agents to the voting rule are analyzed explicitly as part of the general exploration. In some cases, we can determine the optimal voting rule. For more information on the voting literature, we refer the reader to Brams and Fishburn (2002) and for a lively presentation of the diversity of opinions and arguments among experts on voting rules to Laslier (2012).

[2] 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.

[3] I am grateful to Michel Le Breton for drawing my attention to this paper. Note that in the social choice tradition, the introduction of probability models regarding voter preferences in order to allow for a comparison and optimization of voting rules dates back to Niemi and Weisberg (1972) - specifically, part 1 of their book is devoted to constitutional design. Alongside Weber, we could also cite, without being exhaustive, early pioneers who have used such a framework to compare voting rules: Bordley (1983), Chamberlin and Cohen (1978), Fishburn and Gehrlein (1976), Merrill (1984), and Merrill (1985).

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

the Borda count, and approval voting, among others.[5] Under a two-parameter scoring rule, each voter submits a vector whose components specify the scores that the voter assigns to the available candidates. More specifically, each voter must assign a score of 1 to one candidate, 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 candidate. After component-wise summation of the score-vectors across all voters, the candidate with the highest score is chosen, and any ties are broken randomly and with equal probability.

For our study of two-parameter scoring rules in the latter part of this survey, we adopt a Bayesian setting: each voter is characterized by a privately observed vector of three Bernoulli utilities, one for each candidate. 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 candidates (see e.g. Schmitz and Tröger, 2012). This latter property implies that, for a given voter, all ordinal rankings over the three candidates are equally likely.[6] At the interim stage at which voting takes place, each voter is fully aware of his 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.[7] 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 and Degan, 2017).

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 candidate. The threshold, given by a weighted average of the Bernoulli utilities associated with the voter's most and least preferred candidates, 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

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

[6]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 (2016).

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

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 computational methods.[8] 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.[9] 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. 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.[10]

The remainder of this survey is structured as follows: In Section 2, we introduce the general model and basic concepts. Section 3 contains our selective survey of the related voting literature, while Section 4 specifies our specific Bayesian voting framework. In Section 5 we then present the characterization of symmetric Bayes Nash equilibria for all two-parameter scoring rules and any number of voters. Section 6 contains our computational results comparing the effectiveness-levels of two-parameter scoring rules. Section 7 offers a brief conclusion.

2 The model

In this section, we describe a general voting model that encompasses much of the literature on the evaluation and comparison of different voting rules. The basic structure of the model is described in the next subsection, before the following subsections introduce the model components in more detail.[11]

[8]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.

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

[10]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.

[11]In the interest of readability, some of the concepts introduced below are presented somewhat informally. E.g. in the case of ordinal direct voting rules mentioned below, we do not describe the domain restrictions assumed in some of the papers in the voting literature. In most papers in the literature, the domain of preferences is unrestricted. I.e. any strict preference order over candidates can arise. However, there are some important restricted domains (for instance the domain of single peaked preferences) which are also widely explored in the voting literature. This has an impact on the comparisons of different voting rules. For instance, the Condorcet efficiency of the Borda rule is not the same on the unrestricted domain and the domain of single peaked preferences. These matters are relevant when comparing voting rules (see Gehrlein and Lepelley, 2011).

2.1 Basic structure and timing

We propose here a general voting model in which voting rules will be evaluated behind a 'veil of ignorance' that stems from the fact that the designer is uncertain about the preferences of individual voters in the electorate. The uncertainty about the electorate's preference profile will be captured by a common prior probability distribution over all possible preference profiles, and the designer will choose a voting rule that maximizes the ex ante expected value of some social objective with respect to the prior probability distribution of voters' preference profiles. The timing inherent in this general model is as follows:

Stage 1. The designer chooses a voting rule from a class of available voting rules.

Stage 2. Nature picks a state of the world - which here shall be a preference profile across the electorate - according to the common prior probability distribution.

Stage 3. Voters receive information (either partial or full) about the state of the world and revise their beliefs according to Bayes rule.

Stage 4. Given their private information, their beliefs, and their preferences, voters play a game whose rules are prescribed by designer's chosen voting rule.

Stage 5. For each conceivable prediction of voters' play, the designer computes the value of a given social objective/criterion.

We now spell out in more detail the aforementioned components of the model.

2.2 Voters, candidates, and preferences

Let $N = \{1, 2, \ldots, n+1\}$ be a finite set of players, which consists of $|N| = n+1$ voters. The voters must collectively make a choice from the finite set K of candidates or alternatives by individually casting a vote according to a given voting rule R. Each voter $i \in N$ has a preference ordering on K that depends on his *type*. This will allow us to model the idea that any choice of voting rule, and the subsequent decision on how to vote, will typically have to be taken without perfect knowledge of voters' preferences. For each voter $i \in N$, there is a set T_i of possible types. Each type $t^i \in T_i$ of voter $i \in N$ gives rise to a von Neumann-Morgenstern (vNM) utility function $u_i : K \times T_i \to \mathbb{R}$, $(k, t^i) \mapsto u_i(k, t^i)$ which, in turn, implies a *preference ordering* $\succeq_{t^i}$ over candidates in K.[12,13] E.g. if $K = \{A, B, C\}$ and voter i's type t^i is such that (s.t.) $u_i(B, t^i) > u_i(A, t^i) > u_i(C, t^i)$, then $B \succ_{t^i} A \succ_{t^i} C$.

2.3 Voting rules

A voting rule is a game form R for the set of candidates K which specifies *action sets* $\Sigma_R^1, \ldots, \Sigma_R^{n+1}$ (one for each voter $i \in N$, with typical action $\sigma^i \in \Sigma_R^i$) and an outcome function $r : \prod_{i \in N} \Sigma_R^i \to$

[12]Note that in making the utility functions u_i depend only on K and T_i we are restricting attention to so called *private-value* environments. I.e. only a voter's own type t^i directly affects his preferences over candidates. We therefore do not consider here the *common-value* voting environments that are studied in the Condorcet jury and information aggregation literature (see e.g. Feddersen and Pesendorfer, 1997 in which each voter's utility depends on the entire profile of voters' types).

[13]We assume furthermore that voters' preferences over the set $\Delta(K)$ of probability measures on K satisfy the vNM axioms, which implies that each voter i's preferences over $\Delta(K)$ are represented by the expected value $E_\mu[u_i(k, t^i)]$ with respect to $\mu \in \Delta(K)$ of the vNM utility function $u_i(k, t^i)$.

$\Delta(K)$, $(\sigma^1,\sigma^2,\ldots,\sigma^{n+1}) \mapsto r(\sigma^1,\sigma^2,\ldots,\sigma^{n+1})$, where $\Delta(K)$ denotes the set of all probability distributions on K to allow for the possibility of tie between candidates to be broken randomly. The voting rule R (together with the voters' utility functions and beliefs about other voters' types) induce a game between the $n+1$ voters which could be either in normal form or in extensive form, possibly including exogenous uncertainty as captured by a chance move. I.e. this framework is rich enough to encompass 'one shot' voting protocols such as, e.g., the plurality rule or the Borda count, as well as dynamic voting procedures such as plurality with runoff and many others. Given R, every voter i has a set Σ_R^i of (pure) strategies at his disposal, where $\sigma^i \in \Sigma_R^i$ denotes a particular strategy of voter i. Each outcome $r(\sigma)$ of the voting game corresponds to a choice of candidate in K, where we write $\sigma \equiv (\sigma^1,\sigma^2,\ldots,\sigma^{n+1})$ and $\Sigma_R \equiv \prod_{i\in N}\Sigma_R^i$ for ease of notation.

A voting rule is *direct* (denoted by DVR for *direct voting rule*) if it asks each voter to reveal directly to the designer his preferences. I.e. $\Sigma_{DVR}^i = T_i$ for all voters $i \in N$. A DVR is *ordinal* if it is invariant to utility rescaling. I.e. w.l.o.g. we can define an ordinal DVR as one where $\Sigma_{DVR}^i = L_i$, with L_i denoting the set of all possible ordinal rankings of candidates in K induced by T_i. For example, if $K = \{A,B,C\}$, then L_i could consist of up 13 possible orderings, or it could consist of only the six strict orderings if voter i is never indifferent between any two candidates. Henceforth, we denote by L the set of strict preference orderings over K.[14]

2.4 Information structure

We assume that the type-profile $t \equiv (t^1,t^2,\ldots,t^{n+1}) \in T \equiv \prod_{i\in N} T_i$ is unknown ex ante when the designer chooses the voting rule which is subsequently used in making a collective choice from the set of candidates K. The designer's subjective probability distribution features a joint probability distribution λ, where λ^i is its ith marginal probability distribution on T_i. For example, in the second part of this paper we will consider the special case where $K = \{A,B,C\}$, $T_i = [0,1]^3$, $u_i(A,t_i) = t_A^i$, $u_i(B,t_i) = t_B^i$, and $u_i(C,t_i) = t_C^i$ for all voters $i \in N$. In this setting, we will assume furthermore that $\lambda^i(t^i) = \lambda_A^i(t_A^i)\cdot\lambda_B^i(t_B^i)\cdot\lambda_C^i(t_C^i)$ (where $t_k^i \in [0,1]$ for all $k \in K$), and that $\lambda_A^i = \lambda_B^i = \lambda_C^i = g$ for all $i \in N$, where g is a continuous and strictly positive density function on $(0,1)$. These assumptions imply that the types $t^1,t^2,\ldots,t^{n+1}$ of the $n+1$ voters are stochastically independent and identically distributed (i.i.d.), and that the probability that a given voter has a particular ordinal ranking is the same (namely $1/6$) for the six possible ordinal rankings on $K = \{A,B,C\}$. As a result, all profiles of voters' preference orderings are equally likely and occur with probability $(|L|)^{-(n+1)} = 1/6^{n+1}$. This information structure is referred to as '*impartial culture*' (IC) in the statistical social choice literature. See, e.g., Lepelley and Valognes (2003).

An alternative example is the case where $K = \{A,B,C\}$ and each voter's type set T_i corresponds to the set L of all strict orderings of the candidates in K (see Gehrlein and Lepelley, 2001, Gehrlein et al, 2016, and Green-Armytage et al, 2016). In such settings, the term '*impartial anonymous culture*' (IAC) refers to the information structure where the actual type-profiles $t \in T$ of the electorate cannot be observed directly, but summary information about how many of the $n+1$ voters have each of the six orderings in L does become observable. In particular, what becomes observable is a vector $(l_1,l_2,\ldots,l_6)$ with $l_1+l_2+\ldots+l_6 = n+1$, where l_1 is the number of voters who have the first preference ordering in L, l_2 is the number of voters who have the second preference ordering in L, and so on (such a vector $(l_1,l_2,\ldots,l_6)$ is what Lepelley and Valognes (2003) refer to as a *voting situation*). For instance, if there are $n+1 = 3$ voters in N, then there are 56 ways of distributing these three voters across the six strict preference orderings in L. In particular, there are six voting situations where all three voters have the same ordinal ranking, there are 30 voting

[14] If $K = \{A,B,C\}$, we have $L = \{ABC, ACB, BAC, BCA, CAB, CBA\}$, where ABC indicates that A is strictly preferred to B which, in turn, is strictly preferred to C. The other elements of L are interpreted analogously.

situations in which two of the three voters have the same ordinal ranking while the remaining voter has a different ranking, and finally there are 20 voting situations in which each of the three voters has a different ordinal ranking in L. In an IAC setting, all these voting situations are assumed to be equally likely, which means that in the present example, each voting situation arises with probability $1/56$.[15]

At the interim stage (i.e. once the designer has chosen the voting rule and Nature has selected a type-profile $t \in T$, but before voters have cast their votes), we can further distinguish two alternative environments: in the *Bayesian environment*, each voter $i \in N$ learns his type t^i, but not that of the other voters. Given his type t^i, each voter updates his belief about the profile of the other voters' types according to Bayes' rule. This is the assumption used in the specific voting setting presented below in Section 4. An alternative assumption is that all voters learn the entire type-profile t. This is referred to as the *complete information environment* (see, e.g., De Sinopoli et al, 2006 and Buenrostro et al, 2013).

2.5 Voting Behavior

2.5.1 Rational choice in the Bayesian setting

Under this assumption regarding the information structure and voters' behavior, voters play a Bayesian game given a voting rule R, their own types t^i, and their posterior beliefs about the types of the other voters obtained by updating the prior distribution λ. For each voter $i \in N$, a pure strategy is a voting function $v^i : T_i \to \Sigma^i_R$, $t^i \mapsto v^i(t^i)$.[16] That is, v^i specifies for every type t^i a ballot $\sigma^i \in \Sigma_i$ to be submitted by voter i. A *Bayes Nash equilibrium* of this game is a profile $(v^1(\cdot), v^2(\cdot), \ldots, v^{n+1}(\cdot))$ of $n+1$ voting functions (one for each voter), s.t. for every $i \in N$:

$$\mathrm{E}_{t^{-i}}[u_i(r((v^i(t^i), v^{-i}(t^{-i}))), t^i)|t^i] \geq \mathrm{E}_{t^{-i}}[u_i(r((\sigma^i, v^{-i}(t^{-i}))), t^i)|t^i]$$

for all $t^i \in T_i$ and all $\sigma^i \in \Sigma^i_R$, where $v^{-i}(t^{-i}) \equiv (v^1(t^1), v^2(t^2), \ldots, v^{i-1}(t^{i-1}), v^{i+1}(t^{i+1}) \ldots, v^{n+1}(t^{n+1}))$ denotes the profile of voting functions of all voters other than i.

2.5.2 Rational choice in the complete information setting

Under this assumption regarding the information structure and voters' behavior, voters play a complete information game given their knowledge of the entire type-profile t across the electorate (see, e.g., De Sinopoli et al, 2006). For simultaneous move voting games, appropriate equilibrium concepts include Nash equilibrium, undominated Nash equilibrium, trembling hand perfect Nash equilibrium (see Selten, 1975), proper Nash equilibrium (see Myerson, 1978), and Mertens-stable Nash equilibrium (see Mertens, 1989). As pointed out by De Sinopoli et al (2006), the reason for making recourse to refinements of Nash equilibrium is its lack of predictive power in complete

[15]To see the difference between IC and IAC in this example, note that in the case where $T_i = L$ for all $i \in N$, the 56 voting situations are *not* all equally likely under IC, but the underlying type-profiles $t \in T$ are. E.g., under IC there are *three* equally likely type-profiles where two of the three voters have the preference ordering ABC and the remaining voter has the preference ordering ACB. In contrast, there is only one type-profile where all three voters have the preference ordering ABC. Therefore, the voting situation in which two out of three voters have the ranking ABC and the remaining one has ACB is three times more likely than the voting situation where all three voters have the ranking ABC.

[16]We may also wish to allow for voters using mixed strategies in selecting their ballots. In this case, each voting strategy $v^i(t^i)$ induces a lottery or probability distribution $\mu_i(t^i) \in \Delta\left(\Sigma^i_R\right)$ over the set Σ^i_R of pure strategies associated with voting rule R.

information voting games such as the one induced by approval voting. For extensive form voting games, subgame perfect Nash equilibrium and its refinements constitute appropriate concepts.

2.5.3 Other standards of behavior

In voting theory, it is often assumed that voters vote *sincerely*. Depending on the voting rule at hand, it may not always be straightforward to define sincerity. An example is approval voting, for which a strategy (following Brams and Fishburn, 1978) is said to be sincere if a voter who (sincerely) approves of a given candidate $k \in K$ also approves of any candidate that he prefers to k. However, for *direct* voting rules it is straightforward to define sincerity as the truthful communication of a voter's type. While it is clearly restrictive to simply *assume* sincere voting (as opposed to seeing it arise in the equilibrium of a game where voters may vote strategically or misrepresent their types), there are empirical studies which do not always reject the hypothesis that voters do vote sincerely (see, e.g., Blais and Degan, 2017). Thus there is some justification for studying and comparing the properties of specific voting rules (e.g. their tendency to select the Condorcet winner, whenever one exists) under the assumption of sincere voting.

2.6 Social objectives

In order to capture the goal of the designer, we make recourse to a *social objective* or *criterion*. Denote by $W(\sigma,t)$ the function that assigns a value for the social objective to every possible vote-profile $\sigma \in \Sigma_R$ and every type-profile $t \in T$. For a given voting rule R, denote by $v_R(t)$ the associated equilibrium voting strategy-profile across all $n+1$ voters if it is unique, and by $V_R(t)$ the set of equilibrium voting strategies if there are multiple equilibria. In case of multiple voting equilibria (some of which may be mixed), we have to settle on a single value for the social objective which, e.g., could be done by setting:[17]

$$W_R(t) \equiv \inf_{v \in V_R(t)} \int W(\sigma,t) d\mu_R(\sigma) \text{ or } W_R(t) \equiv \sup_{v \in V_R(t)} \int W(\sigma,t) d\mu_R(\sigma)$$

where $\mu_R(\sigma) \equiv \mu_R^1(\sigma^1) \cdot \mu_R^2(\sigma^2) \cdot \ldots \cdot \mu_R^{n+1}(\sigma^{n+1})$ is the probability distribution over strategy-profiles $\sigma \in \Sigma_R$ induced by the voters' mixed strategies.

As the designer chooses a voting rule R at the ex ante stage (i.e. under uncertainty about the actual type-profile), we must compute the expected value $W_R \equiv \mathrm{E}_t[W_R(t)]$ of the social objective under voting rule R with respect to the prior probability distribution λ on the set T of type-profiles. An optimal voting rule R^* is then one that solves $\max_{R \in \mathscr{R}} W_R$, where $\mathscr{R}$ denotes the set of voting rules among which the designer can choose. Below is a list of examples of social objectives or criteria that have been used in the voting literature:

1. **Utilitarian welfare:** for each possible type-profile $t \in T$, there is an additive social welfare function that consists of the sum of the individual voters' vNM utility functions: $W(\sigma,t) = \sum_{i \in N} u_i(r(\sigma),t^i)$.

2. **Rawlsian welfare:** $W(\sigma,t) = \min_{i \in N} u_i(r(\sigma),t^i)$.

3. **Condorcet efficiency:** A Condorcet candidate is one who would be preferred by the majority of voters in a binary contest with any other candidate standing in the election. We can then

[17] The two examples stated here capture the worst and best case scenarios, resp. However, other aggregation devices are also conceivable.

formalize the social objective of Condorcet efficiency as follows:

$$W(\sigma,t)=\begin{cases} 1 & \text{if } r(\sigma) \text{ is a Condorcet candidate for } t \text{ when such a winner exists} \\ 0 & \text{if } r(\sigma) \text{ is not a Condorcet candidate for } t \text{ given that such a winner exists} \end{cases}$$

Note that here $W(\sigma,t)$ is defined exclusively for type-profiles t for which a Condorcet candidate exists at t. Denote the set of type-profiles for which a Condorcet candidate exists by T^C. Therefore, strictly speaking, the designer must choose a voting rule so as to maximize the *conditional* probability $\mathrm{E}_t[W_R(t)|t \in T^C]$ that the Condorcet candidate is elected. Note also that if Condorcet efficiency is the social objective and if the set $\mathscr{R}$ of voting rules among which the designer can choose is unrestricted, then any Condorcet social choice correspondence will solve the designer's objective. However, if $\mathscr{R}$ is restricted (e.g. to a certain sub-class of voting rules), then the designer has to take additional constraints into account in pursuing the social objective.

4. Robustness to manipulation:

$$W(\sigma,t)=\begin{cases} 1 & \text{if } \sigma \text{ is a profile of sincere voting strategies for } t \\ 0 & \text{if } \sigma \text{ is } \textit{not} \text{ a profile of sincere voting strategies for } t \end{cases}$$

5. Effectiveness (Weber (1978)): Weber's effectiveness measure is defined as the welfare gain of a given voting rule over random selection of a candidate 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. We will use this measure of social welfare below in Section 6, where it will be defined formally.

As the above list demonstrates, there is an abundance of social objectives to choose from. Out of the five considered here, three correspond to social objectives that genuinely depend on the social alternative which is selected by the voting rule and the profile of types. The first two criteria are welfare-based, while the third is not directly related to welfare considerations. The fourth criterion depends on the profile of votes. When we deal with the multiplicity of equilibria by looking at the most favorable one (as defined above) then in the case where there is a well-defined unique sincere strategy (like in the case of DVRs), we have $W_R(t) = 1$ if and only if the profile of sincere strategies is an equilibrium profile. In the complete information case and Nash equilibrium, we obtain a measure of the resistance of the voting rule to individual strategic voting behavior. With strong Nash equilibrium, we obtain a measure of resistance of the voting rule to coalitional strategic voting behavior. As we will see in Section 3, these measures have been mostly developed and used in settings with DVRs and ordinal DVRs. We could alternatively define a measure of vulnerability to strategic manipulation as an index where the places of 0 and 1 in the social objective are inverted. Of course, many alternative ways of quantifying a voting rule's susceptibility to strategic manipulation can be thought of. Some of these measures do not fit the formal framework considered here. For instance, Smith (1999) considers four different indices of manipulability. Some of them fit perfectly into our framework while, others do not. We could even ignore the prior probability distribution as done, for instance, in Carroll (2013) and instead adopt a numerical measure of susceptibility to manipulation defined as the largest expected utility that a generic agent can gain by manipulating, where the supremum is taken over all types, all deviations, and all beliefs that the agent may hold about the behavior of other agents under the voting rule. These different measures may result in different relative rankings of voting rules to the ones cited in our survey below.

3 A brief and selective survey

In this section, we describe branches of the voting and mechanism design literature that are related to the Bayesian comparison of voting rules presented in Sections 4-6 of this paper. The first branch is the literature on statistical social choice which consists of both early and very recent papers that evaluate the performance of well-known multi-candidate electoral systems (among them the plurality rule, approval voting, the Borda rule, as well as different multi-round voting schemes in which candidates are eliminated on the basis of how many (or how few) first place votes they have received) under the assumption that voters vote sincerely, i.e. in line with their true preference orderings of the available candidates. This is clearly a heroic assumption, and various contributors to this literature acknowledge the importance of this issue by at least partially addressing the question of how susceptible to strategic manipulation the various voting rules are, or how they could be modified to induce sincere voting (see Merrill (1984) or Pivato, 2016).

3.1 Statistical social choice

Contributions to this literature typically differ in the information structure they assume, meaning the probability distribution over the type-profiles or voting situations that may arise in elections with three (or more) candidates. One information structure that has been studied in this literature is the '*impartial culture*' assumption that has been described in Sec. 2.4 above, and which will be assumed in the Bayesian comparison of voting rules below in Sections 3-5.

3.1.1 Impartial culture

One social objective that has been used in this literature (see Merrill, 1984) - and which will also be used below in Sec. 4 of the present paper - is a voting rule's effectiveness. This social welfare-based criterion is superior to Condorcet efficiency if voters' intensity of preference is to be acknowledged (even though Merrill, 1984 studies only one rule (namely approval voting) under which voters can express their preference intensity). In order to assess the performance of the various voting rules, Merrill (1984) simulates large numbers of elections with 25 voters and a varying number of candidates (from two to five candidates, and with seven and ten candidates, resp.) In order to generate voter types, he draws independently from the same distribution each element of a voter's type-vector (however, unlike in our Bayesian study below, Merrill (1984) restricts himself to the uniform distribution). While our Bayesian comparison of the effectiveness of voting rules below will focus on the case of three voters, we nevertheless compare here the effectiveness of voting rules in Merrill (1984) for three-candidate elections and 25 voters with that found in our computational results for the uniform distribution: we obtain an effectiveness-level of 83.98% for the plurality rule (as opposed to 83% in Merrill, 1984, where voting is assumed to be sincere), effectiveness of 91.91% for approval voting (as opposed to 95.4% in Merrill, 1984), and effectiveness of 94.47% for the Borda rule (as opposed to 94.8% in Merrill, 1984). It is interesting to note that approval voting narrowly outperforms the Borda rule in Merrill (1984), while the Borda rule clearly outperforms approval voting in our setting (and not only for the uniform distribution, but on average across all the 25 Beta-distributions we consider). Our Example 1 in Section 3 below suggests that this discrepancy is due to the fact that in the Bayes Nash equilibrium of approval voting in our setting, voters 'approve' of their middle-ranked alternative less often than in Merrill (1984). Under the IC assumption, there are many contributions to the literature that aim to compare and evaluate ordinal DVRs from the perspective of individual manipulation. For instance, Smith (1999) compares five popular ordinal voting rules, and Maus et al (2007) characterize the class of ordinal DVRs that minimize the susceptibility to individual manipulation among those ordinal

DVRs that are anonymous, surjective and tops-only when the number of voters $|N|$ is larger than the number of candidates $|K|$. We refer the reader to references therein for further results on this topic.[18]

A second social objective considered under this information structure is *Condorcet efficiency*. For the case of $|K| = 3$ candidates, Merrill (1985) finds that approval voting (with a level of Condorcet efficiency of 76%) is dominated by the plurality rule (with Condorcet efficiency of 79.1%) which, in turn, is dominated by the Borda rule (with Condorcet efficiency of 90.8%). The best among the rules studied by Merrill (1985), however, is Hare with a Condorcet efficiency of 96.2%.[19]

Two final social objectives studied under the 'impartial culture' assumption are utilitarian efficiency and resistance to manipulability, as well as a combination of these two objectives. Green-Armytage et al (2016) define manipulability resistance to manipulability as robustness to manipulation (as defined in Section 2.6) with strong Nash as the equilibrium concept. They find that in a setting with $|N| = 99$ voters, the Hare rule and the Condorcet-Hare rule (which selects the Condorcet candidate if one exists, and otherwise selects the winner according to the Hare rule) both outperform all other 54 voting rules considered on the measure of resistance to manipulability (with more than 80% of elections resulting in a non-manipulable outcome). Note that Green-Armytage et al (2016) do not actually impose incentive compatibility on voting behavior as a constraint, because they simply assume that voting is sincere. Consequently, their results are silent on whether insincere voting actually occurs, and also cannot quantify the welfare loss associated with any instance of strategic voting.

On the measure of utilitarian efficiency, the Borda rule performed best (with utilitarian welfare being maximized in approx. 74% of elections), which is not surprising given that each voter ballot conveys voters' complete ordinal ranking of the three candidates. When the two social objectives of utilitarian efficiency and resistance to manipulability are imposed jointly, the optimal trade-off between the two is achieved by Hare and Condorcet-Hare, both reaching near-identical scores of approx. 71% and 85% for efficiency and resistance to manipulability, resp.

Finally, on the social objective of utilitarian efficiency, Apesteguia et al (2011) show that among direct ordinal voting rules, scoring rules (i.e. rules where each voter assigns a *score* to each candidate $k \in K$ (subject to some constraints) and the candidate with the highest aggregate score across the $|N|$ voters wins) maximize ex ante expected utilitarian efficiency when voters are assumed to report their preferences truthfully. Furthermore, they characterize *which* particular scoring rule is optimal depending on the prior distribution λ over voters' type-profiles. In a related setting with a more general information structure (allowing for 'impartial culture', 'impartial anonymous culture' and others as a special cases), Pivato (2016) studies a version of our general voting model in Sec. 2 with a large electorate (i.e. $|N| \to \infty$). His results show that simple scoring rules, such as the Borda rule or approval voting, select with near certainty a candidate who maximizes (or almost maximizes) utilitarian efficiency.

[18] An ordinal DVR R is *surjective* if for all $k \in K$, there exists a profile of linear orders s.t. $r(\sigma) = k$. It is *tops-only* if for any pair of profiles $\sigma, \hat{\sigma}$ we have $r(\sigma) = r(\hat{\sigma})$ whenever $Top(\sigma^i) = Top(\hat{\sigma}^i)$ for all $i \in N$, where for all strict preferences P over K, $Top(P)$ denotes the highest-ranked candidate in K according to P.

[19] The Hare rule is a multi-stage voting rule where the plurality loser is eliminated in each round. With just three candidates, this implies that the candidate with the least first-round votes is eliminated. In the second round, the winner is determined on the basis of a simple majority vote. Among one-stage scoring rules, the Borda rule maximizes Condorcet efficiency when the number of voters tends to ∞ (see Gehrlein and Lepelley, 2011).

3.1.2 Impartial anonymous culture

The second important information structure that has been studied extensively in the statistical social choice literature is the IAC assumption that has been described in Sec. 2.4 above. On the social objective of *Condorcet efficiency*, Gehrlein and Lepelley (2001) compare a number of important voting rules in three-candidate elections: plurality rule, negative plurality rule (which asks each voter to say which one of the three available candidates he wishes to eliminate), the Borda rule, and the two-stage procedures of plurality elimination (which, for $|K| = 3$, corresponds to the Hare rule described above) and negative plurality elimination (in round one, eliminate the most nominated candidate; then in round two choose the majority candidate among the remaining two). Regardless of the number of voters $|N|$, the following ranking of voting rules according to Condorcet efficiency emerges consistently: negative plurality elimination dominates the Borda rule (by approx. 6 percentage points) which, in turn, dominates the plurality rule by approx. 3 percentage points.[20]

In a recent paper, Gehrlein et al (2016) refine the comparison of three 'one shot' voting rules (namely plurality, negative plurality, and Borda) under IAC in terms of Condorcet efficiency by considering only those voting situations in which the three voting rules do not all yield the same election winner (in voting situations where all three yield the same winner, they will also generate the same Condorcet efficiency score).[21] This modified information structure is labeled 'modified impartial anonymous culture' (MIAC). While MIAC generates lower Condorcet efficiency scores across the board than IAC, the ranking of the three voting rules according to their Condorcet efficiency remains unchanged: Borda (82.4% under MIAC, 90.7% under IAC) dominates the plurality rule (75.9% under MIAC, 87.6% under IAC) which, in turn, dominates the negative plurality rule (21.6% under MIAC, 62% under IAC).

The aforementioned paper by Green-Armytage et al (2016) also considers 54 voting rules under IAC in a setting with $|K| = 3$ candidates and $|N| = 99$ voters. As with IC, the Hare rule and the Condorcet-Hare rule both outperform all other 54 voting rules considered on the measure of resistance to manipulability, while the Borda rule performed best in terms of utilitarian welfare. When the two social objectives of utilitarian efficiency and resistance to manipulability are imposed jointly, the optimal trade-off between the two is again achieved by Hare and Condorcet-Hare, both reaching near-identical scores of approx. 85% and 90% for efficiency and resistance to manipulability, resp.

In summary, while the unrealistic assumption of sincere voting significantly limits the appeal of the insights from the statistical social choice literature, it would nevertheless be ill-advised to dismiss these insights, especially as this literature has studied many important voting rules that more sophisticated voting models have ignored thus far, especially the multi-stage voting rules such as plurality elimination and others.

A second important strand of the voting literature is concerned with the equilibria of voting games (Nash equilibrium and its refinements) under complete information about voters' types.

3.2 Rational voting under complete information

If the type-profile t of voters is commonly known at the start of Stage 2 of our general voting framework, then the game induced by the designer's voting rule is one of complete information.

[20] In contrast to IC, under IAC the voting rule that maximizes Condorcet efficiency among one-stage scoring rules when the number of voters tends to ∞ is no longer the Borda rule. Instead, it is a scoring rule where sincere voters attach a weight of 0.37228 to their second ranked alternative. I.e. this rule is slightly biased towards plurality (see Cervone et al, 2005 and Gehrlein and Lepelley, 2011).

[21] This comparison is based on $|K| = 3$ candidates and $|N| = 101$ voters.

De Sinopoli et al (2006) study the complete information game induced by the approval voting rule in which voters can cast a vote for as many candidates in K as they want.[22] Taking as their starting point the result of Fishburn and Brams (1981) that if the type-profile t is s.t. a Condorcet candidate exists, then the game has a Nash equilibrium in undominated strategies that selects the Condorcet candidate as winner of the election. However, as there are other Nash equilibria, it is natural to ask if the Nash equilibrium selecting the Condorcet candidate survives if refinements of the Nash equilibrium concept are applied. De Sinopoli et al (2006) show through examples that this is *not* the case. First, there is a type-profile for $|N| = |K| = 3$ for which a Condorcet candidate exists s.t. in the unique equilibrium obtained by iterative deletion of dominated strategies there is a tie in the number of votes that each of the three candidates receives. Given the assumption that ties are broken by an equi-probable lottery on the set of tied candidates, the Condorcet candidate as well as all the other Candidates are therefore elected with the same probability. Another example with $|N| = 3$ voters and $|K| = 4$ candidates shows that there exist type-profiles s.t. in the unique equilibrium obtained by iterative deletion of dominated strategies the Condorcet candidate receives no votes at all. And finally, an example is constructed with $|K| = 4$ candidates and any odd number of voters $|N| \geq 3$ in which the *unique* best response of one voter to other voters' strategies (two of which are playing mixed strategies while the remaining ones play their dominant strategies) involves an *insincere* ballot. In contrast, if in this example the plurality rule is used instead of approval voting, then the stable equilibrium involves sincere voting by all voters. This shows that the case for using approval voting is more nuanced than Fishburn and Brams (1981) suggest.

The paper by Buenrostro et al (2013) applies the concept of Nash equilibrium in undominated strategies to a wider class of voting rules (namely scoring rules), which contains approval voting as a special case. They focus on settings with $|K| = 3$ alternatives and $|N| \geq 4$ voters and show that whenever the complete information game induced by a particular scoring rule is dominance solvable (by iterated deletion of dominated actions), the candidates elected in equilibrium coincides with the set of Condorcet candidates. Furthermore, they show that approval voting is dominance solvable whenever any other scoring rule is dominance solvable (thus making it the 'most' dominance solvable rule within the class of scoring rules the authors consider). This suggests that approval is an attractive rule when there are only three candidates. Note, however, that in a setting with three candidates located on a one-dimensional ideology space and voters' preferences being determined by the distance of their individual ideology from that of the election winner, there are instances where approval voting fails to select the Condorcet candidate, while the plurality rule does so for sure (see De Sinopoli et al, 2014).

Finally, the third strand of the related literature deals with incomplete information settings and strategic voter behavior. It is in these settings that incentive compatibility problems and welfare losses due to informational asymmetry can be quantified.

3.3 Bayesian and mechanism design approaches to voting

Myerson (2002) characterizes and compares equilibria of scoring rules in a setting where the number of voters participating in the election is unknown - this setting is not a special case of the one we have described in Section 2 of the present paper, which assumes a fixed and known number of voters, and allows a continuum of possible type vectors per voter. Also related but not a special case of our model in Section 2 is the paper by Ahn and Oliveros (2016) who explore how well scoring rules aggregate information in a common value setting - we focus here explicitly on private-value settings where the problem of efficient information aggregation is absent. The following related

[22] De Sinopoli (2001) studies Nash equilibria and their refinements in the complete information game induced by the plurality rule.

papers focus 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, 2016, Miralles, 2012, and Schmitz and Tröger, 2012), explicitly accounting for strategic voting behavior and incentive compatibility constraints. This makes the problem of designing a utilitarian welfare-maximizing voting rule significantly harder than when sincere voting is assumed. Schmitz and Tröger (2012) address this problem by characterizing welfare-maximizing strategy-proof voting rules for settings with two candidates under an 'impartial culture' information structure, while Barberà and Jackson (2006) characterize utilitarian welfare-maximizing weighted majority rules over two candidates under an asymmetric information structure (i.e. the marginals λ^i need not be the same across all voters $i \in N$).[23] However, analogue results for settings with three candidates (be that under Bayesian or dominant strategy implementation) are so far absent from the literature.[24] In independent work, Kim (2016) provides partial insights into this problem. Using the same information structure as the one employed here, he shows that if the types 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 ordinal 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.[25] 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).[26]

4 A detailed Bayesian comparison of two-parameter scoring rules under the 'impartial culture' assumption

4.1 Basic set-up

We study here a version of the general framework described in Section 2 in which there are three candidates: $K = \{A, B, C\}$ and $n+1$ voters. For every voter $i \in N$, the type set is $T_i = [0,1]^3$. Regarding the information structure, we assume that each type $t^i \equiv (t^i_A, t^i_B, t^i_C) \in T_i$ 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 elements of his type-vector t^i, namely t^i_A, t^i_B and t^i_C 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 func-

[23] See also the survey paper by Le Breton et al (2017) in this special issue, which is dedicated to the mechanism design problem in the two candidate case when the class of voting rules available to the designer is restricted.

[24] 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.

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

[26] 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.

tions, and the joint distribution of types $(t^1, \ldots, t^{n+1})$, are common knowledge among the voters. Observe that the information structure assumed here makes our setting symmetric in the following two ways:

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

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

4.2 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 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 $i \in N$ is denoted by $\sigma^i \equiv (\sigma^i_A, \sigma^i_B, \sigma^i_C)$, where σ^i_k 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 $\mathfrak{N} \subseteq N$ of $m \equiv |\mathfrak{N}|$ voters, the* ***aggregate score-vector*** *s^m across voters in $\mathfrak{N}$ is the sum of their individual score-vectors σ^j: $s^m \equiv (s^m_A, s^m_B, s^m_C) = (\sum_{j\in\mathfrak{N}} \sigma^j_A, \sum_{j\in\mathfrak{N}} \sigma^j_B, \sum_{j\in\mathfrak{N}} \sigma^j_C)$. We denote by $S^m_{x,y}$ the set of all aggregate score-vectors s^m that can arise for the subset $\mathfrak{N}$ of voters. The set $S^m_{x,y}$ is obtained by adding up, for every profile of m individual score-vectors in $\Sigma^m_{x,y}$, the scores associated with each alternative k.*[27]

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

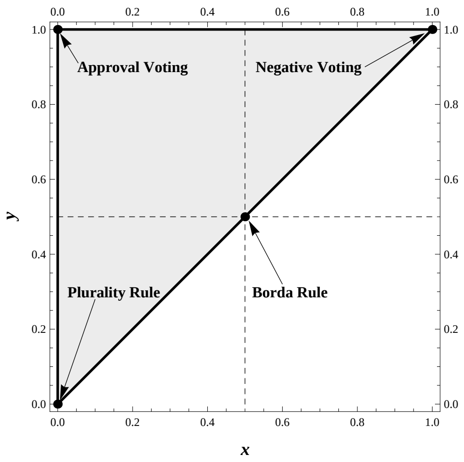


Figure 1: The set of all (x,y)-scoring rules

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

$$\delta_A(s^m) = \begin{cases} 1 & \text{if } s^m_A > s^m_B \text{ and } s^m_A > s^m_C \\ 1/2 & \text{if } s^m_A = s^m_B > s^m_C \text{ or } s^m_A = s^m_C > s^m_B \\ 1/3 & \text{if } s^m_A = s^m_B = s^m_C \\ 0 & \text{if } s^m_A < s^m_B \text{ or } s^m_A < s^m_C \end{cases}$$

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

5 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 $v^i : [0,1]^3 \to \Sigma_{x,y}$, $t^i \mapsto v^i(t^i)$. That is, v^i specifies for every type t^i the score-vector in $\Sigma_{x,y}$ to be submitted by voter i. We denote by $v^i_k(t^i)$ the score assigned to alternative $k \in K$ under strategy v^i when voter i's type is t^i. Given the two types of symmetry inherent in our setting, it is natural to focus on Bayes Nash equilibria $(v^1, \ldots, v^{n+1})$ that are symmetric in the sense of the following two properties:

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

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

We refer to a strategy 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 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 $v(t^j)$. As voter i does not know j's type, he views the score-vector $v(t^j)$ as a random variable with sample space $\Sigma_{x,y}$. As both v and our setting are symmetric, voter i believes that every permutation of the score-vector $v(t^j) = (v_A(t^j), v_B(t^j), v_C(t^j))$ arises with the same probability as $v(t^j)$. In particular, if the (x,y)-scoring rule features $x < y$, then all permutations of the score-vector $(1,x,0)$ arise with probability $\Pr[v(t^j) = (1,x,0)] \equiv p \in [0,1/6]$, and all permutations of $(1,y,0)$ arise with probability $\Pr[v(t^j) = (1,y,0)] = (1/6) - p$. 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 v. That is, $v_1(t^i) = 1$ and $v_3(t^i) = 0$ for all $i \in N$ and all $t^i \in [0,1]^3$, where $v_1(t^i)$ and $v_3(t^i)$, resp., denote the scores that v assigns to i's highest- and lowest-ranked alternatives when his type is t^i.*

The proof of Proposition 1 can be found in the appendix of Giles and Postl (2014) (see Section A.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 v.[28] Put differently, in expected terms, the voting behavior of the other voters results in a uniform distribution over the set of alternatives K: $\mathrm{E}[\delta_A(\sum_{j\neq i} v(t^j))] = \mathrm{E}[\delta_B(\sum_{j\neq i} v(t^j))] = \mathrm{E}[\delta_C(\sum_{j\neq i} v(t^j))] = 1/3$. By submitting a score-vector σ^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 $\mathrm{E}[\delta(\sum_{j\neq i} v(t^j) + \sigma^i)]$ over K that first-order stochastically dominates any distribution $\mathrm{E}[\delta(\sum_{j\neq i} v(t^j) + \tilde{\sigma}^i)]$ that prevails if he submits any non-sincere score-vector $\tilde{\sigma}^i$.

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 v that assigns a score of $v_2(t^i) \in \{x,y\}$ to the middle-ranked alternative of each voter i according to the*

[28] It is important to emphasize here that the symmetric strategy 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 $\Pr[v(t^j) = (1,x,0)] = 0$, and all permutations of $(1,y,0)$ arise with $\Pr[v(t^j) = (1,y,0)] = 1/6$.

following threshold criterion: for some $\alpha \in [0,1]$,

$$v_2(t^i) = \begin{cases} x & \text{if } t_2^i < \alpha t_1^i + (1-\alpha)t_3^i \\ y & \text{if } t_2^i > \alpha t_1^i + (1-\alpha)t_3^i \end{cases} \tag{1}$$

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:

$$\Pr\left[v(t^j) = (1,x,0)\right] = \int_0^1 \int_0^{t_1^j} \int_{t_3^j}^{\alpha t_1^j + (1-\alpha)t_3^j} g(t_2^j)\,\mathrm{d}t_2^j g(t_3^j)\,\mathrm{d}t_3^j g(t_1^j)\,\mathrm{d}t_1^j \equiv p(\alpha) \tag{2}$$

Similarly, voter i expects j to submit each permutation of $(1,y,0)$ with the complementary probability $\Pr\left[v(t^j) = (1,y,0)\right] = (1/6) - p(\alpha)$. 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 t^i is s.t. $t_A^i > t_B^i > t_C^i$. His expected utility from submitting a score-vector $\sigma^i = (1, \sigma_B^i, 0)$ (with $\sigma_B^i \in \{x,y\}$) is:[29]

$$U_i(\sigma^i, t^i) \equiv \sum_{k \in K} \mathrm{E}[\delta_k(s^n + \sigma^i)]t_k^i, \tag{3}$$

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

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

$$\begin{aligned} t_B^i > {} & \left(\frac{\mathrm{E}[\delta_A(s^n + (1,x,0))] - \mathrm{E}[\delta_A(s^n + (1,y,0))]}{\mathrm{E}[\delta_B(s^n + (1,y,0))] - \mathrm{E}[\delta_B(s^n + (1,x,0))]}\right) t_A^i \\ & + \left(1 - \frac{\mathrm{E}[\delta_A(s^n + (1,x,0))] - \mathrm{E}[\delta_A(s^n + (1,y,0))]}{\mathrm{E}[\delta_B(s^n + (1,y,0))] - \mathrm{E}[\delta_B(s^n + (1,x,0))]}\right) t_C^i \end{aligned} \tag{4}$$

The condition in (4) states that voter i's best response to the symmetric strategy v used by all other voters is to assign the higher score of y to his middle-ranked alternative B if his utility t_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: $\delta_B(s^n + (1,y,0)) - \delta_B(s^n + (1,x,0)) \geq \delta_A(s^n + (1,x,0)) - \delta_A(s^n + (1,y,0)) \geq 0$ for all $s^n \in S_{x,y}^n$. $\square$

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,

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

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':

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

$$\begin{aligned} L(p(\alpha)) &\equiv \frac{\mathrm{E}[\delta_A(s^n+(1,x,0))]-\mathrm{E}[\delta_A(s^n+(1,y,0))]}{\mathrm{E}[\delta_B(\boldsymbol{s}_n+(1,y,0))]-\mathrm{E}[\delta_B(\boldsymbol{s}_n+(1,x,0))]} \\ &= \frac{\sum_{s^n\in S^n_{x,y}}\Pr[s^n]\,(\delta_A(s^n+(1,x,0))-\delta_A(s^n+(1,y,0)))}{\sum_{s^n\in S^n_{x,y}}\Pr[s^n]\,(\delta_B(s^n+(1,y,0))-\delta_B(s^n+(1,x,0)))} \end{aligned} \tag{5}$$

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 (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 $\sigma^i_A(t^i_A-\bar{t}^i)+\sigma^i_B(t^i_B-\bar{t}^i)+\sigma^i_C(t^i_C-\bar{t}^i)$, where $\bar{t}^i$ 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 $t^i_2<\bar{t}^i\equiv(t^i_1+t^i_3)/2$, and the higher score of 1 if $t^i_2>\bar{t}^i$. I.e. 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 $U_i((1,y,0),t^i)<U_i((1,x,0),t^i)$ for all $t^i\in[0,1]^3$ s.t.:

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

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. We refer the reader to Giles and Postl (2014) for details of the proof with five or more voters.

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

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

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

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

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

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

Submission of the individual score-vector $\sigma^i=(1,y,0)$ by voter i generates a shift in probability mass from alternative A to B, relative to submission of $\sigma^i=(1,x,0)$: $\delta_A(s^2+(1,x,0))-\delta_A(s^2+(1,y,0))=\frac{1}{6}$.

(iii) For four voters ($n=3$), the table below shows the aggregate score-vector $s^3=(1+2y,2+y,0)$ which occurs with probability $\Pr[s^3]=1/216$ at $\alpha=0$:

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

Submission of the individual score-vector $\sigma^i=(1,y,0)$ by voter i generates a shift in probability mass from alternative A to B, relative to submission of $\sigma^i=(1,x,0)$: $\delta_A(s^3+(1,x,0))-\delta_A(s^3+(1,y,0))=\frac{1}{2}$. □

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' type-components 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 Appendix A.3 of Giles and Postl (2014) for details of the proof with five or more voters.

Proof. Consider voter i and suppose w.l.o.g. that his type t^i is s.t. $t^i_A>t^i_B>t^i_C$. 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 $s^2\in S^2_{x,y}$ which occurs with positive probability at $\alpha=1$ (i.e. $\Pr[s^n]=(p(1))^2=1/(6)^n$) 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 $\sigma^i=(1,y,0)$ as opposed to the score-vector $\sigma^i=(1,x,0)$. In this case, the numerator of the loss-gain-ratio is strictly smaller than the denominator: $\delta_B(s^n+(1,y,0))-\delta_B(s^n+(1,x,0))>\delta_A(s^n+(1,x,0))-\delta_A(s^n+(1,y,0))$. Now consider the aggregate score-vector $s^2\in S^2_{x,y}$ in the table below, which occurs with probability $\Pr[s^2]=1/36$:

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

The table shows that for $s^2=(x,1,1+x)$, submission of the score-vector $\sigma^i=(1,y,0)$ by voter i generates a shift in probability mass from alternatives A and C to B, relative to submission of $\sigma^i=(1,x,0)$:

$$\delta_A(s^2+(1,x,0))-\delta_A(s^2+(1,y,0))=1/3$$
$$\delta_B(s^2+(1,y,0))-\delta_B(s^2+(1,x,0))=2/3$$

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. $\square$

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. To address at least partially these limitations, we provide below a full characterization of the loss-gain-ratio for the case of three voters, as the number of possible aggregate score-vectors is still manageable. For a treatment of the setting with two voters, and some insight into the setting with four voters, see Giles and Postl (2014), where it is shown that 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.

Proposition 3. *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. 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 the remaining 26 equivalence classes - for which the loss-gain-ratio is non-monotonic - uniqueness of the equilibrium weight α^* is distribution-dependent.*[30]

For the proof of Proposition 3, we refer the reader to Appendix A.5 of Giles and Postl (2014). There, the interested reader will find a table containing 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. 2 below. While our results in Propositions 3 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 case of 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 $G(\xi) = \xi^b$, 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 30 is satisfied - Fig. 2 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. 2 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).[31]

[30]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 $\alpha \in (0,1)$).

[31]In order not to overload Fig. 2, 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.

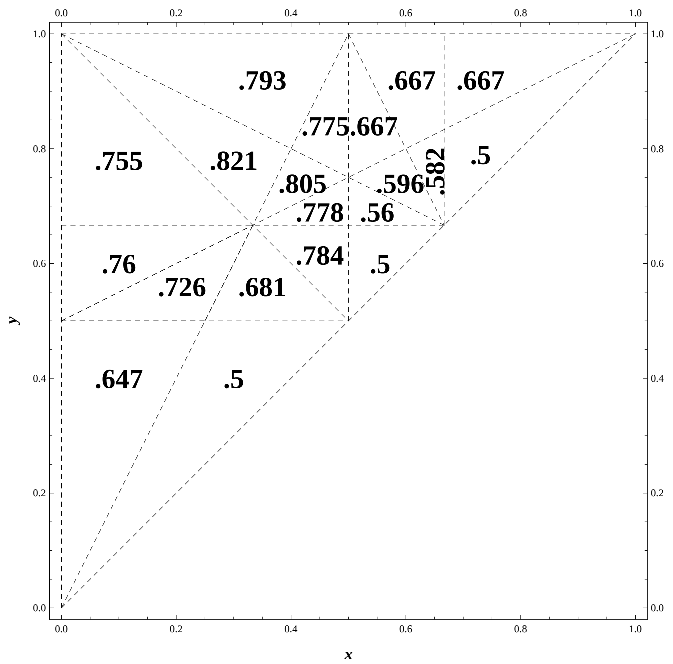


Figure 2: Equilibrium values of α^* for three voters and uniform G

6 Effectiveness of two-parameter scoring rules

6.1 Welfare notions and effectiveness

In this section, we build on our equilibrium characterization in Section 5 to compare (x, y)-scoring rules according to their effectiveness in selecting an alternative that is representative of voters' 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 4, 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 $t_1^i > t_2^i > t_3^i$), and then multiplying by six voter i's expected utility under the representative ranking.[32] Thus, expected welfare, denoted by

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

$W_{x,y}$, is:

$$W_{x,y} \equiv 6(n+1)\left(\int_0^1 \int_0^{t_1^i} \int_{t_3^i}^{\alpha^* t_1^i + (1-\alpha^*)t_3^i} U_i((1,x,0),t^i)\, g(t_2^i)\, \mathrm{d}t_2^i\, g(t_3^i)\, \mathrm{d}t_3^i\, g(t_1^i)\, \mathrm{d}t_1^i \right.$$
$$\left. + \int_0^1 \int_0^{t_1^i} \int_{\alpha^* t_1^i + (1-\alpha^*)t_3^i}^{t_1^i} U_i((1,y,0),t^i)\, g(t_2^i)\, \mathrm{d}t_2^i\, g(t_3^i)\, \mathrm{d}t_3^i\, g(t_1^i)\, \mathrm{d}t_1^i \right)$$

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 voting rules. Formally, a direct voting rule (DVR) in the present model is a function $f : [0,1]^{3(n+1)} \to \Delta(K)$, $t \mapsto (f_A(t), f_B(t), f_C(t))$, where $\Delta(K)$ is the set of probability distributions over K, and $f_k(t)$ is the probability that alternative $k \in K$ is chosen by the mechanism when the voters' type-profile is $t \equiv (t^1, \dots, t^{n+1})$. Ex ante expected welfare associated with a DVR is:

$$\text{Welfare of } f = \mathrm{E}[\sum_{k\in K} f_k(t)(\sum_{i\in N} t_k^i)] \tag{6}$$

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

Definition 4. *The DVR $\bar{f}$ implements* ***random selection*** *if $(\bar{f}_A(t), \bar{f}_B(t), \bar{f}_C(t)) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ for all $t \in [0,1]^{3(n+1)}$.*

A 'first best' DVR, in contrast, is a DVR that maximizes expected welfare in (6) in the hypothetical scenario where the types t^i of all $n+1$ voters, once realized, become observable before the collective choice has to be made:[33]

Definition 5. *A DVR f^{FB} is* ***first best*** *if for every $k \in K$: $\sum_{i\in N} t_k^i > \max_{l\neq k}\{\sum_{i\in N} t_l^i\} \Rightarrow f_k^{FB}(t) = 1$ and $\sum_{i\in N} t_k^i < \max_{l\neq k}\{\sum_{i\in N} t_l^i\} \Rightarrow f_k^{FB}(t) = 0$.*

On the basis of Definition 5, ex ante expected welfare of a first best DVR reduces to the expectation of the order statistic $\max\{w_A, w_B, w_C\}$, where $w_k \equiv \sum_{i\in N} t_k^i$ is a random variable with support $[0, n+1]$, and any two random variables w_k, w_l ($k,l \in K$, $l \neq k$) are i.i.d.[34] We can now define formally the effectiveness of an (x,y)-scoring rule:

$$\text{Effectiveness} \equiv \frac{W_{x,y} - \text{Welfare of } \bar{f}}{\text{Welfare of } f^{FB} - \text{Welfare of } \bar{f}}$$

6.2 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.[35] Our main aim is to gain some insight into which

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

[34] 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.

[35] 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 $x < y$.

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).[36]

To obtain our computational results, we have selected specific distributions G from which voters' type-components t_k^i 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 $g(\xi;a,b)$ and the cumulative distribution function $G(\xi;a,b)$ of the Beta-distribution are parameterized by two shape parameters $a,b>0$: $g(\xi;a,b) \equiv \xi^a(1-\xi)^b / \int_0^1 s^a(1-s)^b ds$ and $G(\xi;a,b) \equiv \int_0^\xi s^a(1-s)^b ds / \int_0^1 s^a(1-s)^b ds$ for $0 \le \xi \le 1$. 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 $\alpha^* \in (0,1)$. 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. 3 (those yielding 89.06% effectiveness).[37] These rules allow voters to express their 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 6.3). 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.[38] 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.

Scoring Rule	(x,y)	Effectiveness w.r.t f^{FB}	Effectiveness w.r.t 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

[36] Kim (2016) proves analytically for the case of the uniform distribution that ordinal rules do not maximize social welfare.

[37] As with Fig. 2 above, we show the effectiveness levels only for the 20 equivalence classes associated with the interiors of the various polygons.

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

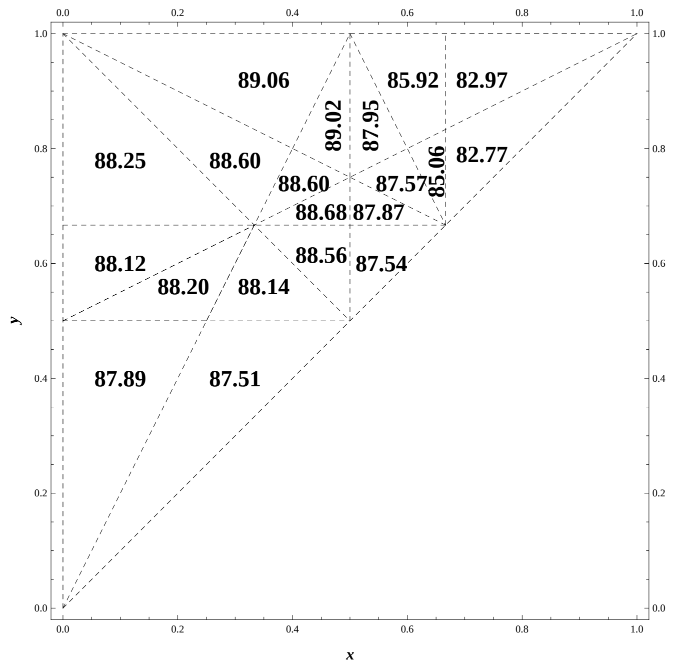


Figure 3: Average effectiveness in % across 25 Beta-distributions

(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 the shaded gray areas in Fig. 4 (classes I, II, and IV are represented by a gray triangle each), while classes III and V are represented by thick black lines.[39] 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 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. 3 indicates, the effectiveness levels of classes II, III, and IV are virtually identical (the effectiveness of class III, which was omitted from Fig. 3, 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.

[39]Dashed lines around shaded areas indicate that the boundaries are excluded. For details of which subset of (x,y)-scoring rules corresponds to classes II, III, IV, and V in the figure, see Table 11 in Giles and Postl (2014).

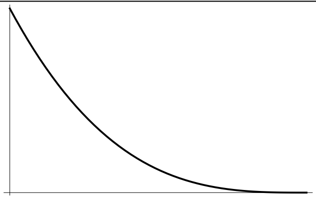
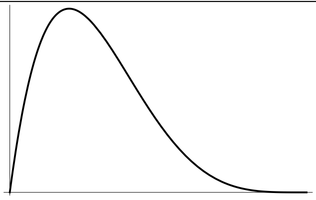
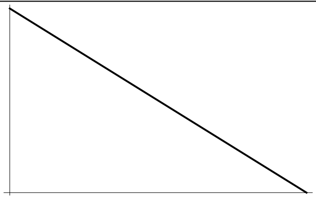
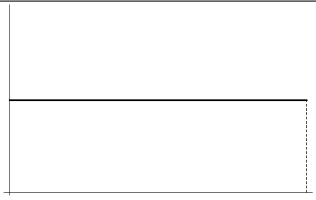
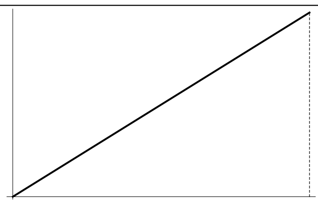
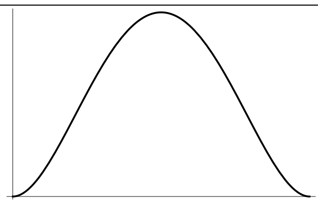
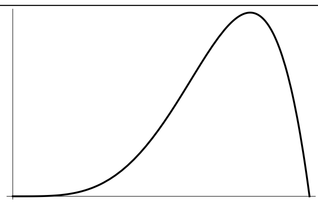
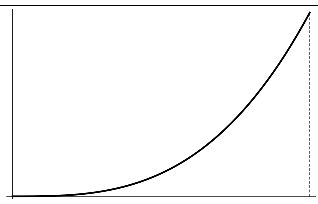
Indicative shape of $g(\xi;a,b)$	Parameters a and b	Most effective scoring rules
	$a = 1, b = 5$ $a = 1, 2 < b < 5$	Class I Class II
	$2 \leq a < b \leq 5$	Class II
	$a = 1, b = 2$	Class II
	$a = 1, b = 1$	Class II
	$a = 2, b = 1$	Class II
	$a = b > 1$	Class III
	$2 \leq b < a \leq 5$	Class IV
	$b =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

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

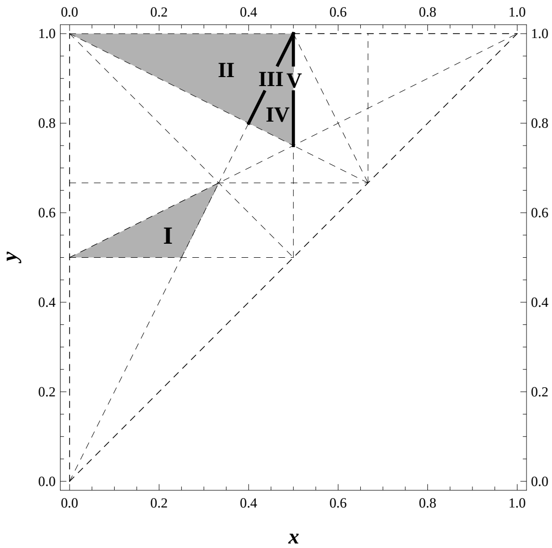


Figure 4: Graphical illustration of classes I-IV

6.3 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 DVRs). However, a complete characterization of incentive compatible DVRs 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 DVR: Consider any voter i, and pick two proportional types $\hat{t}^i, \tilde{t}^i \in [0,1]^3$ that reflect the same ordinal ranking of the alternatives. By 'proportional' we mean that $\tilde{t}^i = c(\hat{t}^i_1, \hat{t}^i_2, \hat{t}^i_3)$ for some real-valued scalar $c > 0$. Observe that under the symmetric voting strategy in Corollary 1, if type $\hat{t}^i$ submits the individual score-vector $\sigma^i = (1, y, 0)$, then so will type $\tilde{t}^i$. This implies that all types $t^i \in [0,1]^3$ with the same 'normalized type' $(1, (\hat{t}^i_2 - \hat{t}^i_3)/(\hat{t}^i_1 - \hat{t}^i_3), 0)$ will submit the same score-vector σ^i as $\hat{t}^i$.

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 t^i with their normalized types (only the Bernoulli utility t^i_3 of the lowest-ranked alternative can be normalized to zero w.l.o.g. for every type $t^i \in [0,1]^3$). However, the standard characterization of incentive compatible DVRs (see e.g. Rochet, 1987) implies that *if* a DVR 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 DVR must treat the type $\tilde{t}^i$ the same as the type $\hat{t}^i$. 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 DVR may require constraints in addition to the bunching of proportional types, it is clear that no incentive compatible DVR 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 DVR f^* that treats the same all voters who have the same normalized type. Consider a voter i who reports a type $\hat{t}^i = (\hat{t}^i_1, \hat{t}^i_2, \hat{t}^i_3)$, with associated ordinal ranking $\succ_i$. On the basis of his report, voter i is assigned the following 'secondary type' τ^i:

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

where $\theta_i \equiv (\hat{t}^i_2 - \hat{t}^i_3)/(\hat{t}^i_1 - \hat{t}^i_3)$, and $\mathrm{E}[t^i_k | \succ_i]$ is the expectation of u^i_k conditional on the ordinal ranking associated with voter i's reported type $\hat{t}^i$.[40] Given the secondary types τ^i of all voters, let 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 $\tau \equiv (\tau^1, \ldots, \tau^{n+1})$ rather than the reported types. It is straightforward to show that f^* gives voters the incentive to reveal truthfully their ordinal rankings when all other voters do so.[41]

We would like to argue that the DVR f^* is a more realistic benchmark for evaluating the effectiveness of two-parameter scoring rules, because both f^* and the FSE of (x,y)-scoring rules bunch proportional types. We have computed numerically the ex ante expected welfare of f^* across all 25 Beta-distributions, and have used the results to obtain the average effectiveness of (x,y)-scoring rules relative to f^*, instead of f^{FB}.[42] 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. 4 is again the most effective; in fact, the ranking of effectiveness-levels is not affected by using the benchmark f^* instead of 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.[43] Even though class II may not be a second best DVR, its very small loss in effectiveness relative to f^* suggests that it comes very close.

7 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.

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

[41] Note, however, that we do not make any claims as to whether 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.

[42] Ex ante expected welfare of f^* has been obtained by means of Monte Carlo experiments.

[43] 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.

References

Ahn DS, Oliveros S (2016) Approval Voting and Scoring Rules with Common Values. Journal of Economic Theory 166:304–310

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

Barberà S, Jackson MO (2006) On the weights of nations: Assigning voting weights in a heterogeneous union. Journal of Political Economy 114:317–339

Blais A, Degan A (2017) L'étude du vote stratégique. In this special issue.

Bordley RF (1983) A pragmatic method for evaluating election schemes through simulation. American Political Science Review 77:123–141

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

Brams SJ, Fishburn PC (1978) Approval voting. American Political Science Review 72:831–847

Brams SJ, Fishburn PC (2002) Voting procedures. In: Arrow KJ, Sen AK, Suzumura K (eds) Handbook of Same Choice and Welfare, Vol. I, Elsevier North Holland, Amsterdam, chap 4, pp 173–236

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.

Carroll G (2013) A Quantitative Approach to Incentives: Application to Voting Stanford U mimeo.

Cervone DP, Gehrlein WV, Zwicker WS (2005) Which Scoring Rule Maximizes Condorcet Efficiency Under Iac? Theory and Decision 58:145–185

Chamberlin JR, Cohen MD (1978) Toward Applicable Social Choice Theory: A Comparison of Social Choice Functions under Spatial Model Assumptions. American Political Science Review 72:1341–1356

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

De Sinopoli F (2001) Sophisticated voting and equilibrium refinements under plurality rule. Social Choice and Welfare 17:655–672

De Sinopoli F, Dutta B, Laslier JF (2006) Approval voting: Three examples. International Journal of Game Theory 35:27–38

De Sinopoli F, Iannantuoni G, Pimienta C (2014) Counterexamples on the superiority of approval versus plurality. Journal of Public Economic Theory 16:824–834

Feddersen T, Pesendorfer W (1997) Voting Behavior and Information Aggregation in Elections With Private Information. Econometrica 65:1029–1058

Fishburn PC, Brams SJ (1981) Approval voting, condorcets principle, and runoff elections. Public Choice 36:89–114

Fishburn PC, Gehrlein WV (1976) An analysis of simple two-stage voting systems. Behavioral Science 21:1–12

Gehrlein WV, Lepelley D (2001) The Condorcet efficiency of Borda Rule with anonymous voters. Mathematical Social Sciences 41:39–50

Gehrlein WV, Lepelley D (2011) Voting Paradoxes and Group Coherence: The Condorcet Efficiency of Voting Rules,. Springer Verlag, Heidelberg

Gehrlein WV, Lepelley D, Plassmann F (2016) Should voters be required to rank candidates in an election? Mathematical Social Sciences 46:707–747

Giles A, Postl P (2014) Equilibrium and effectiveness of two-parameter scoring rules. Mathematical Social Sciences 68:31–52

Green-Armytage J, Tideman TN, Cosman R (2016) Statistical evaluation of voting rules. Mathematical Social Sciences 46:183–212

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 (2016) Ordinal versus Cardinal Voting Rules: A Mechanism Design Approach. National Taiwan University mimeo.

Laslier JF (2012) And the loser is...plurality voting. In: Felsenthal DS, Machover M (eds) Electoral Systems, Springer Verlag, Heidelberg, chap 13, pp 327–351

Le Breton M, Lepelley D, Macé A, Merlin V (2017) Le mécanisme optimal de vote au sein du conseil des représentants dans un système fédéral. In this special issue.

Lepelley D, Valognes F (2003) Voting rules, manipulability and social homogeneity. Public Choice 116:165–184

Maus S, Peters H, Storcken T (2007) Anonymous voting and minimal manipulability. Journal of Economic Theory 135:533–544

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

Merrill S (1984) A Comparison of Efficiency of Multicandidate Electoral Systems. American Journal of Political Science 28:23–48

Merrill S (1985) A statistical model for Condorcet efficiency based on simulation under spatial model assumptions. Public Choice 47:389–403

Mertens JF (1989) Stable equilibria -A reformulation. Part I: Definition and basic properties. Mathematics of Operations Research 14:575–625

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

Myerson RB (1978) Refinements of the Nash equilibrium concept. International Journal of Game Theory 15:133–154

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

Niemi RG, Weisberg HF (eds) (1972) Probability models of collective decision making. C.E. Merrill Publishing Company, Michigan,USA

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

Pivato M (2016) Asymptotic utilitarianism in scoring rules. THEMA, Université de Cergy-Pontoise mimeo.

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

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

Selten R (1975) A reexamination of the perfectness concept for equilibrium points in extensive games. International Journal of Game Theory 4:25–55

Smith DA (1999) Manipulability measures of common social choice functions. Social Choice and Welfare 16:639–661

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