Review of Economic Studies (2011) **78**, 1–16 doi: 10.1093/restud/rdq023

On the Justice of Decision Rules

JOSE APESTEGUIA
ICREA and Universitat Pompeu Fabra

MIGUEL A. BALLESTER
Universitat Autonoma de Barcelona

and

ROSA FERRER
Universitat Pompeu Fabra

First version received September 2008*; final version accepted June* 2010 (*Eds.*)

Which decision rules are the most efficient? Which are the best in terms of maximin or maximax? We study these questions for the case of a group of individuals faced with a collective choice from a set of alternatives. A key message from our results is that the set of optimal decision rules is well defined, particularly simple, and well known: the class of scoring rules. We provide the optimal scoring rules for the three different ideals of justice under consideration: utilitarianism (efficiency), maximin, and maximax. The optimal utilitarian scoring rule depends crucially on the probability distribution of the utilities. The optimal maximin (respectively maximax) scoring rule takes the optimal utilitarian scoring rule and applies a factor that shifts it towards negative voting (respectively plurality voting).

Keywords: Decision rules, Scoring rules, Utilitarianism, Maximin, Maximax

JEL Codes: D00, D63, D71, D72

1. INTRODUCTION

The problem of aggregation of preferences in group decision making has been studied extensively in the literature. In practice, such aggregation is generally based on ordinal information; *i.e.* disregarding the intensities of the individuals' preferences. From small committee decisions to voting in national elections, decision rules typically do not allow individuals to directly express their underlying cardinal utilities. This requirement to use ordinal information appears to be a practical demand, in part, because of the difficulty of expressing preferences in numerical terms, that is, assigning an exact utility intensity to each alternative.[1]

On the other hand, classical ideals of justice, like utilitarianism or maximin, are defined in terms of cardinal utilities. In brief terms, utilitarianism evaluates an alternative in terms of the average individual utility value, while the maximin principle disregards the utility values of the best-off individuals to evaluate an alternative on the basis of the utility value of the worst-off agent.

It is therefore the case that, while ideals of justice make use of interpersonal cardinal utilities, actual decision rules use only ordinal information. The question arises as to how to approach the evaluation of decision rules in cardinal terms, given that they use only limited ordinal

1. For a related discussion, see Austen-Smith and Banks (1999).

information. In other words, this paper attempts to answer questions such as which ordinal decision rules perform best in terms of a given cardinal ideal of justice, which are the most efficient, and which are the best in maximin, or maximax, terms.

In our setting, a group of individuals has to choose an alternative from a set of alternatives. The individuals' valuations over the alternatives are realizations from a random variable. We evaluate decision rules on the basis of their expected value according to a given ideal of justice. We say that a decision rule is *optimal* if it always selects the alternatives that are maximal in expectation with respect to the given ideal of justice. We show that, for every number of individuals and alternatives, and for every probability distribution of the utility values, the optimal decision rule is (1) for utilitarianism a scoring rule that depends crucially on the probability distribution of the utilities, (2) for maximin approximately a scoring rule that applies a factor to the optimal utilitarian scoring rule shifting it towards negative voting, and (3) for maximax approximately a scoring rule that applies a factor to the optimal utilitarian scoring rule shifting it towards plurality voting.[2] [3]

More concretely, we show that for utilitarianism, the expected values of the alternatives, given their ranking, completely characterize the weights of the optimal decision rule. The intuition for the relation between the weights of the optimal utilitarian scoring rule and the probability distribution is as follows. When the probability distribution of the individuals' valuations over the alternatives has an increasing density function, individuals are expected to have a generally high regard for the alternatives. In this case, the optimal utilitarian scoring rule will lean towards negative voting, which discriminates more strongly between the lower-ranked than between the higher-ranked alternatives. This is because the utility values of the higher-ranked alternatives will tend to be concentrated, making it less crucial to discriminate between them. There may, in contrast, be sizeable differences between the values of the lower-ranked alternatives, and thus, it is important to discriminate between them. If, on the other hand, the probability distribution of the individuals' valuations over the alternatives has a decreasing density function, individuals are not expected to be too enthusiastic about the alternatives. Then, the optimal utilitarian scoring rule will lean towards plurality voting, which discriminates more strongly among the higher-ranked alternatives than among the lower-ranked alternatives. Finally, if the valuations made by individuals are expected to be evenly distributed, then the optimal scoring rule is Borda in which there are constant differences between the weights assigned to consecutive alternatives in the ranking.

The maximin optimal scoring rule takes the utilitarian scoring rule and applies a factor that shifts it towards negative voting. That is, under maximin, the optimal rule discriminates more strongly between the lower-ranked alternatives than between the higher-ranked alternatives, thus transmitting precise information on the worst-regarded alternatives. Under the maximax rule, the direction is reversed. The maximax optimal scoring rule again takes the utilitarian rule and applies a factor that shifts it towards plurality. Then, the optimal maximax rule discriminates more strongly between the higher-ranked alternatives than between the lower-ranked alternatives in order to convey precise information on the best-regarded alternatives.

1.1. *Related literature*

We are by no means the first to evaluate decision rules. There is a large and still growing literature examining decision rules on the basis of their capacity to meet certain desirable properties

2. A scoring rule is a vector of fixed weights that individuals assign to the different alternatives. The plurality scoring rule, the negative scoring rule, and Borda's scoring rule are especially salient cases. We give precise definitions in Section 2.2.

3. The notion of "approximately a scoring rule" will be made precise in Section 3.

such as anonymity, strategy proofness, consistency of the social preference ordering, Pareto dominance, path independence, probability of selecting the Condorcet winner, and so forth.[4]

There is, however, very little work on evaluating decision rules on the basis of some theory of justice. Notable exceptions are the early simulation studies of Bordley (1983) and Merrill (1984) and the theoretical work of Weber (1978). Bordley and Merrill use simulations to analyse the efficiency of different voting systems, including plurality and Borda. Consistent with our results, they show that plurality may be outperformed in utilitarian terms by other decision rules. Weber (1978) studies the performance of scoring rules for the case of utilitarianism and for the uniform distribution. He shows that, asymptotically, Borda is the best scoring rule in this case. This, again, is consistent with our results.

In another related strand of literature, there are papers that study how to select a voting rule in a constitutional setting where there are two options, the status quo and a second alternative, and individual preferences are uncertain. A voting rule is characterized by the number of votes needed to accept the second alternative over the status quo. The papers that comprise this literature examine issues such as which voting rules maximize efficiency, which are self-stable, how to weight votes in heterogeneous contexts, self-enforcement voting rules, and so forth. For examples in this vein, see Rae (1969), Barbera and Jackson (2004, 2006) and Maggi and Morelli (2006).

Finally, there is a growing literature addressing the question of the transmission of utility intensities in collective decision problems (see Casella, 2005; Jackson and Sonnenschein, 2007; Hortala-Vallve, 2009, 2010). In particular, they show that voting systems in which individuals are allowed to express intensities may lead to social welfare gains. The innovation of these papers is to consider a decision problem repeated over T times and endow individuals with a maximum number of votes to allocate over the T problems. Individuals are then able to transmit intensities by concentrating their votes on those issues that are most relevant to them.

2. ENVIRONMENT

Consider a society composed of a finite set of individuals N, with cardinality $n \geq 2$, who have preferences over a finite set of alternatives K, with cardinality $k \geq 3$. Typical elements of N are denoted by i and j and of K by l and h. Now, we first present the cardinal environment (utilities and ideals of justice), then the ordinal setting (ordinal preferences and decision rules), and then one that links the two worlds (an index to evaluate the success of a decision rule in terms of a given ideal of justice).

2.1. *Cardinal utilities and ideals of justice*

Individuals' evaluations over the set of alternatives are cardinal, interpersonally comparable i.i.d. utility random variables. U_i^l denotes the random variable representing the utility of individual i for alternative l, distributed according to the distribution function F in the interval $I = [0, \bar{u})$, $\bar{u} \leq \infty$. We often refer to F as the *culture* of the society. We assume that F has an absolutely continuous density function f, with finite first moment. Hence, the probability that U_i^l takes a particular value is 0.

A Social Welfare Function (SWF) is a mapping W from $I^{n \times k}$ to I^k, where $W^l(u^l) \in I$ denotes the social value of alternative l, given the realization of the vector of random variables

4. Early studies include Brams and Fishburn (1978), Caplin and Nalebuff (1988), Demeyer and Plott (1970), and Nurmi (1983). See also Benoit and Kornhauser (2010), Dasgupta and Maskin (2008), Gehrlein (1997), Levin and Nalebuff (1995), Myerson (2002), Ozkal-Sanver and Sanver (2006), and Saari (1999).

$U^l = (U^l_1, \ldots, U^l_n)$.[5] The three SWFs considered herein are utilitarianism, maximin, and maximax. *Utilitarianism* evaluates an alternative by taking the average of all individual utilities. Formally, a SWF is utilitarian if $W = UT$ with $UT^l(u^l) = \sum_{i \in N} u^l_i / n$. The *maximin* principle evaluates an alternative on the basis of the utility value of the worst-off agent, disregarding any other utility value. In other words, a SWF is of the maximin type if $W = MN$ with $MN^l(u^l) = \min_{i \in N} u^l_i$. Consider also the *maximax* rule, which, in contrast to maximin, focuses on the best-off individuals. That is, a SWF is of the maximax type if $W = MX$ with $MX^l(u^l) = \max_{i \in N} u^l_i$. As an ideal of justice maximax may appear a mere formal curiosity. We shall see, however, that because of its close connection to plurality voting, the maximax principle plays a more important role in democratic political institutions than might be expected.

2.2. *Ordinal preferences and decision rules*

We denote by M the matrix representing the *ordinal* preferences of the individuals, given the realizations of the utility random variables $\{U^l_i\}_{i \in N, l \in K}$. M is an $n \times k$ matrix with entries $m^l_i \in \{1, \ldots, k\}$ denoting the position of alternative l in the preferences of individual i, where the higher m^l_i is, the higher alternative l is ranked by individual i.[6] M^l denotes the l-th column of matrix M, representing the ordinal preferences of all individuals with respect to alternative l. The collection of all possible matrices M is denoted by $\mathcal{M}$. We denote by $\mathbf{l}^{(t)}$ the number of individuals that place alternative l exactly above $t-1$ alternatives. That is, $\mathbf{l}^{(t)} = |\{i \in N\colon m^l_i = t\}|$.

A decision rule D is a correspondence from $\mathcal{M}$ to K. We impose no restriction on the possible set of decision rules other than assuming that it uses individuals' actual ordinal preferences. Scoring rules are a particularly interesting class of decision rules. They are typically simple to implement in practice and encompass a number of widely used decision rules. Formally, consider a vector $S \in [0,1]^k$, with $S^j \leq S^t$ whenever $j \leq t$, $S^1 = 0$, and $S^k = 1$, where S^t denotes the value of an individual's vote for whichever alternative she ranks higher than exactly $t-1$ alternatives. Given M, an alternative h is selected by S if and only if $h \in \arg\max_{l \in K} \sum_{t=1}^{k} \mathbf{l}^{(t)} S^t$.

We say that a scoring rule S is convex (concave) if the differences $S^{t+1} - S^t$, $1 \leq t \leq k-1$, are increasing (decreasing). A convex scoring rule aims to discriminate more finely among the higher-valued alternatives than among the lower-valued alternatives. A concave scoring rule pursues exactly the opposite aim, *i.e.* to discriminate more finely among the lower-valued alternatives than among the higher-valued ones. The most salient scoring rules are plurality, Borda, and negative. A scoring rule is *plurality* if $S = S_{\text{Pl}}$ with $S^t_{\text{Pl}} = 0$ for every $t < k$. That is, plurality allows individuals to indicate only their first choice, and hence, it represents an extreme case of a convex scoring rule. It is *negative* if $S = S_{\text{Ng}}$ with $S^t_{\text{Ng}} = 1$ for every $t > 1$. Negative represents the opposite of plurality since it only allows individuals to signify their least preferred alternative and therefore represents an extreme case of a concave scoring rule. A scoring rule is *Borda* if $S = S_{\text{Bd}}$ with $S^t_{\text{Bd}} = \frac{t-1}{k-1}$ for every t. Borda assigns constant differences between the weights assigned to consecutive alternatives in the ranking. Consequently, Borda represents the intersection between convex and concave scoring rules.

2.3. *Cardinal and ordinal preferences: Evaluating decision rules*

The aim of this paper is to find the decision rule D that, for every single possible composition of M, identifies the optimal alternative(s) in terms of a given ideal of justice W. In pursuit

5. Throughout the paper, we use the terms "social welfare function" and "ideal of justice" interchangeably.

6. Note that, given that the culture is continuous, ties have zero probability, and hence, we can assume without loss of generality that $m^l_i \neq m^h_i$ for all $i \in N$ and for all $l, h \in K$, $l \neq h$.

of this aim, it is important to note that for every given M, there is a whole class of utility realizations of the random variables $\{U_i^l\}_{i\in N,l\in K}$ consistent with it. Thus, we judge an alternative l by its *expected utility value* in terms of a given ideal of justice W, in the class of cardinal utility realizations consistent with M. That is, given M, alternative l is evaluated in terms of the $\mathbb{E}[W^l(U^l) \mid M]$, where $W^l(U^l)$ is a random variable that depends on the vector of random utility values $U^l = (U_1^l, \ldots, U_n^l)$ consistent with M. We are now in a position to introduce the notion of optimal decision rules. The W-optimal decision rule D_W selects, for every single possible composition of M, all the alternatives with the largest expected value in terms of W. That is, D_W is the W-optimal decision rule if, for all $M \in \mathcal{M}$:

$$D_W(M) = \arg\max_{h\in K} \mathbb{E}[W^h(U^h) \mid M].$$

Note that, for the case of utilitarianism, $\mathbb{E}[W^h(U^h) \mid M]$ represents the expected average utility value of h within the class of cardinal utility realizations consistent with M, *i.e.* $\mathbb{E}[W^h(U^h) \mid M] = \mathbb{E}[\sum_{i\in N} U_i^h/n \mid M]$. For maximin, $\mathbb{E}[W^h(U^h) \mid M] = \mathbb{E}[\min_{i\in N} U_i^h \mid M]$, and for maximax, $\mathbb{E}[W^h(U^h) \mid M] = \mathbb{E}[\max_{i\in N} U_i^h \mid M]$.

The computation of $\mathbb{E}[W^h(U^h) \mid M]$ requires the use of the theory of order statistics.[7] Given the random variables $U_i^1, U_i^2, \ldots, U_i^k$, the order statistics $U_i^{(1)} \le U_i^{(2)} \le \cdots \le U_i^{(k)}$ are also random variables, defined by sorting the realizations of $U_i^1, U_i^2, \ldots, U_i^k$ in increasing order of magnitude. $U_i^{(t)}$ denotes the t-th order statistic of individual i, representing the utility value for individual i of that alternative having $t-1$ alternatives with lower utility values. Note that, since the utility random variables are i.i.d., $U_i^{(t)}$ and $U_j^{(t)}$ are identical for every pair of individuals i, j. Hence, for every position t, we will often omit the individual subindex and write $U^{(t)}$ with cumulative distribution (respectively, density) function $F^{(t)}$ (respectively, $f^{(t)}$).

3. RESULTS

3.1. *Utilitarianism*

We first show that, for every n, every k, and every culture, the optimal utilitarian decision rule is a scoring rule. This is good news because it implies that if the interest is to maximize the expected value of utilitarianism, it is advisable to implement a scoring rule, which is a relatively simple decision rule. Furthermore, we provide the exact form of the optimal utilitarian scoring rule, conditional on the culture under consideration. More specifically, the culture determines the expected values of the order statistics, which in turn characterize the optimal value of the scoring rule.

Theorem 3.1. *For every n, for every k, and for every culture, D_{UT} is a scoring rule with* $D_{\mathrm{UT}}(M) = \arg\max_{l\in K} \sum_{t=1}^{k} \mathbf{l}^{(t)} S_{\mathrm{UT}}^t$, *where* $S_{\mathrm{UT}}^t = \frac{\mathbb{E}[U^{(t)}]-\mathbb{E}[U^{(1)}]}{\mathbb{E}[U^{(k)}]-\mathbb{E}[U^{(1)}]}$, $1 \le t \le k$.

Theorem 3.1 shows that the optimal utilitarian decision rule is a scoring rule with weights $S_{\mathrm{UT}}^t = \frac{\mathbb{E}[U^{(t)}]-\mathbb{E}[U^{(1)}]}{\mathbb{E}[U^{(k)}]-\mathbb{E}[U^{(1)}]}$. The optimal weight of an alternative ranked at position t is simply the expected value of the t-th order statistic $\mathbb{E}[U^{(t)}]$, normalized to lie in the unit interval. The intuition of the result is as follows. Consider an alternative l and suppose that an individual ranks it at position t. Consequently, that individual's expected utility from this alternative is $\mathbb{E}[U^{(t)}]$. Now, for each position $1 \le t \le k$, there are $\mathbf{l}^{(t)}$ individuals that rank alternative l in

7. See David and Nagaraja (2003) for an introduction to the theory of order statistics.

position t. It follows that the expected utilitarian pay-off for society of alternative l is simply $\sum_{t=1}^{k} \mathbf{l}^{(t)} \mathbb{E}[U^{(t)}]/n$. Therefore, a scoring rule that uses the weights $\mathbb{E}[U^{(t)}]$, normalized to lie in the unit interval, will be optimal and hence the result. The complete proof of Theorem 3.1, and of all the results that follow, is given in the Appendix.

Theorem 3.1 shows that there is a mapping from the culture of the society to the optimal utilitarian scoring rule.[8] We now study the practical implications of this mapping. Theorem 3.2 below relates the shape of the culture to the type of optimal utilitarian scoring rule. The main intuition may be summarized as follows. If the values of the alternatives in some range of the ranking are expected to be close to each other, then the optimal utilitarian scoring rule barely discriminates among these alternatives and consequently assigns similar weights to them. However, when one expects to find relatively high dispersion in the values of the alternatives, then it becomes important to discriminate closely among them.

For the sake of illustration, suppose, *e.g.* that individuals are expected to have a generally high regard for the alternatives. In this case, the optimal utilitarian scoring rule will discriminate more strongly among the lower-ranked alternatives and less strongly among the higher-ranked alternatives. This is because the values of the higher-ranked alternatives will tend to be concentrated, and hence, the need to discriminate among them is less crucial. However, there may be sizeable differences in the values of the lower-end alternatives in which case it will be important to discriminate between them. Consequently, the optimal scoring rule has a shape analogous to negative voting (*i.e.* a concave scoring rule). If, on the other hand, the utility values are expected to be low, then the optimal utilitarian scoring rule will discriminate strongly among the best-ranked alternatives and be less concerned about the lower-ranked alternatives. In this case, therefore, the optimal scoring rule is a version of plurality (a convex scoring rule). Note that the former case arises when the values are drawn from an increasing density function, while the latter when these are drawn from a decreasing density function.

Theorem 3.2. *For every n, for every k, and for every culture with an increasing (decreasing) density function, D_{UT} is a concave (convex) scoring rule.*

Since the uniform distribution has a constant density function, it immediately follows from Theorem 3.2 that Borda is the optimal utilitarian scoring rule in this case.[9] Furthermore, the proof of Theorem 3.2 suggests that for cultures in which the values of the middle-ranked alternatives are concentrated, but there is relatively high dispersion in both the higher- and lower-ranked alternatives, the optimal utilitarian scoring rule represents a combination of the forces discussed above. It discriminates more strongly between the very best and between the very worst alternatives and less strongly between the intermediate alternatives. Consequently, the values of S_{UT}^{t} grow rapidly for the lower-ranked alternatives, then slowly for the middle-ranked alternatives, and then rapidly again for the higher-ranked alternatives. That is, S_{UT} will first have a negative shape and then a plurality shape (*i.e.* first concave and then convex). On the other hand, if one expects agents' evaluations of the alternatives to be polarized, *i.e.* either very highly or very poorly rated, the optimal scoring rule discriminates more strongly between the intermediate alternatives and less closely between the very best and between the very worst alternatives. Consequently, the values of S_{UT}^{t} grow slowly for the lower-ranked alternatives, then rapidly for the middle-ranked alternatives, and then slowly again for the higher-ranked alternatives. That is, it will first be convex and then concave. It is easy to see that the first case is nicely captured

8. Given the results of Theorem 3.1, we can use the terms optimal utilitarian *decision* rule and optimal utilitarian *scoring* rule interchangeably.

9. This result also follows immediately from Theorem 3.1 by using the expected values of the order statistics of the uniform distribution.

by probability distributions such as the normal distribution, or the Cauchy distribution, or for certain parameters of the beta distribution, while the intuition of the latter case is consistent with the U-quadratic distribution or with certain parameters of the beta distribution.

We illustrate the above relations between cultures and optimal utilitarian scoring rules by way of an example. We use the family of beta distributions with different parameters to illustrate the cases of cultures with (1) an increasing density function, (2) a decreasing density function, (3) a symmetric and concave density function, (4) a symmetric and convex density function, and (5) constant density function corresponding to the case of the uniform distribution. Figure 1 reports, for the case of $k = 9$, each of these density functions and the corresponding optimal utilitarian scoring rules, calculated according to Theorem 3.1. For the expected values of the order statistics of beta distributions, see, *e.g.* Thomas and Samuel (2008).

For each of the five distributions, Figure 1 maps the position t, in which an alternative is ordered, with the corresponding weight of the optimal scoring rule S^t_{UT}. Figure 1 neatly represents the previous discussions, showing that the shape of the optimal utilitarian scoring rule depends crucially on the underlying culture of the society.

3.2. *Maximin*

Let us start by first describing an important property that holds for utilitarianism but not for maximin. In utilitarianism, an alternative l is evaluated according to the expected value of the sum of the order statistics associated with the individuals' valuations of alternative l. Since the utility random variables are independent across individuals, the expected value of the sum of the order statistics composing l is simply the sum of the expected value of the order statistics composing l. This property gives tractability to the problem of finding the optimal utilitarian decision rule, allowing us to offer the exact shape of D_{UT} contingent upon the expected values of the order statistics. Thus, D_{UT} provides insights about its relationship with the culture of the society. In the case of maximin, an alternative is evaluated according to the expected value of the minimum of the order statistics associated with individuals' valuations of that alternative. The minimum operator, however, does not preserve the independence of the order statistics across individuals, and hence, the problem of finding the optimal maximin decision rule is less tractable.

In order to address this problem, we first make use of well-known results in reliability theory. This allows us to give the exact shape of the optimal maximin decision rule D_{MN}, for every society and every possible culture. Unfortunately, D_{MN} is somewhat opaque regarding its relationship with the culture of the society. In order to improve on this, in our second step, we use D_{MN} and build on known results in statistical theory to approximate the distributions of the order statistics through the exponential distribution. The exponential approximation makes the problem more tractable and allows us to show that the optimal maximin decision rule is approximately a scoring rule that depends on the expected values of the order statistics. Furthermore, we provide its exact shape and show that it has a particularly interesting relationship with the optimal utilitarian scoring rule. The optimal maximin scoring rule is equal to the optimal utilitarian scoring rule, multiplied by a fraction that makes the maximin scoring rule more concave. Thus, the main message that follows from these results is that, relative to utilitarianism, the optimal maximin scoring rule shifts towards negative voting. Consequently it discriminates more closely between the lower-ranked alternatives, than between the higher-ranked alternatives.

We now turn to the formal presentation of the maximin results. In our first result, we use tools from reliability theory to find the exact shape of D_{MN}, for every society and culture. In particular, we adapt the notion of the failure rate function for systems comprising independent

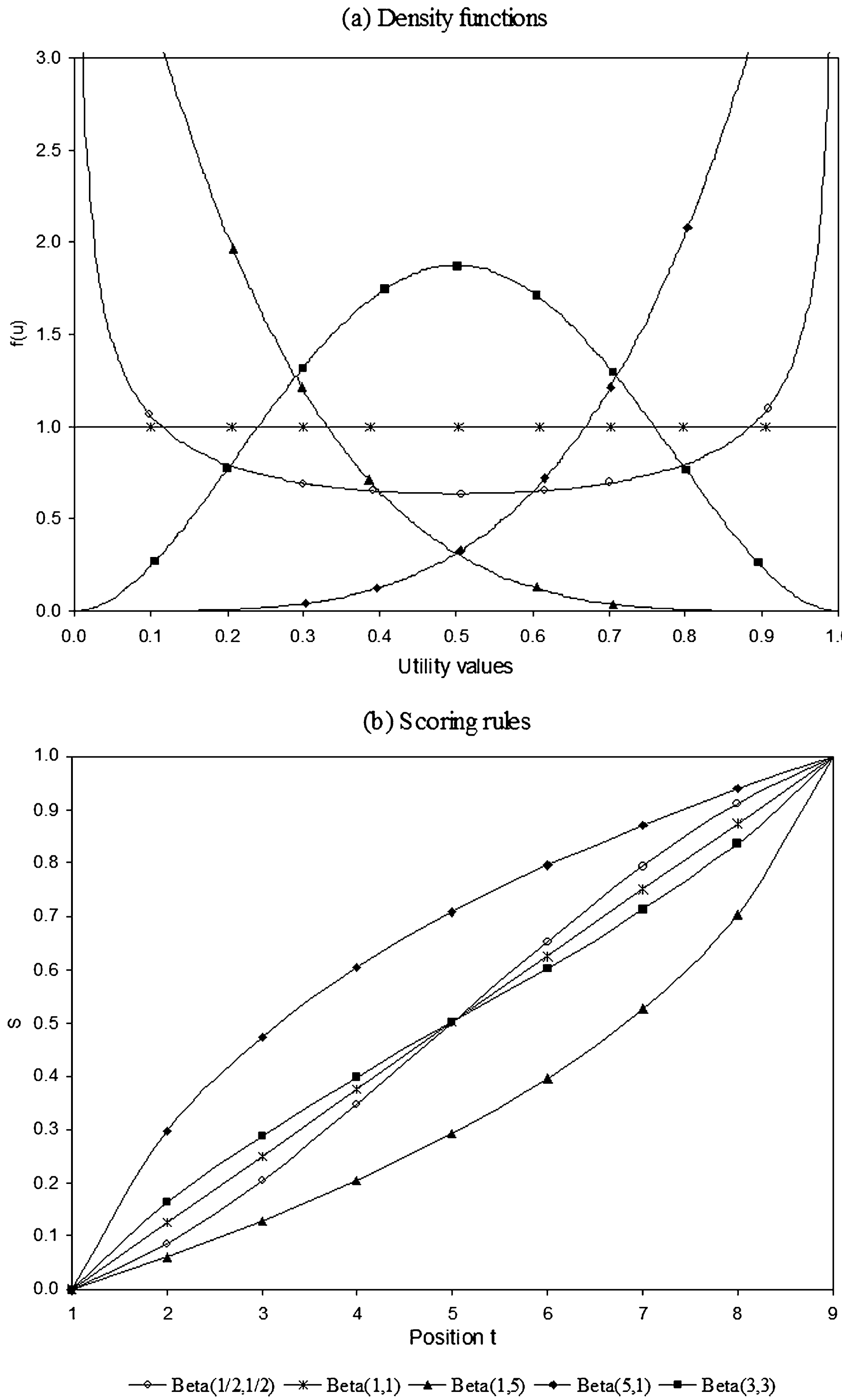

FIGURE 1

Density functions for the five beta distributions and the corresponding optimal utilitarian scoring rules components to our context.[10] We, thus, introduce the notion of the satiation rate of an order statistic, which has a natural interpretation in our setting. The satiation rate of an order statistic is simply the probability that the order statistic satiates at an utility value u, conditional on having reached that value u. More formally, the satiation rate of the t-th order statistic $U^{(t)}$, denoted by $z^{(t)}(u)$, is the probability that $U^{(t)} \in (u, u+\varepsilon)$, knowing that $U^{(t)}$ has reached the value u or simply $z^{(t)}(u) = \frac{f^{(t)}(u)}{1-F^{(t)}(u)}$. This allows us to show that, as in the case of D_{UT}, D_{MN} depends on the order statistics. However, unlike in the case of D_{UT}, D_{MN} does not depend on the expected values of the order statistics, but on their satiation rates.

Theorem 3.3. *For every n, for every k, and for every culture, the optimal maximin decision rule is* $D_{\mathrm{MN}}(M) = \arg\max_{l \in K} \int_I \exp\left(-\int_0^u \sum_{t=1}^{k} \mathbf{I}^{(t)} z^{(t)}(v)\,dv\right) du.$

It emerges that D_{MN} may not generally constitute a scoring rule. To provide more intuition regarding the relationship between the culture of the society and the shape of the optimal maximin decision rule, we now approximate the distribution functions of the order statistics with a single family of distributions: exponential distributions. The foundations for exponential approximations have been studied extensively in the literature providing sharp bounds.[11] Exponential approximations are widely used in statistical theory and its applications (*e.g.* in reliability theory, insurance risk management, etc.), one of the main reasons being that the exponential distribution is more manageable because of its memorylessness, which, in our context, implies that the satiation rate is a constant function. The exponential distribution is also well known as the maximum entropy distribution among all continuous distributions with support on the positive real numbers with a given mean. Maximizing entropy minimizes the amount of prior information built into the distribution, thus giving the exponential distribution the necessary flexibility to approach any possible distribution of order statistics with information only on the mean, as is the purpose here.

Thus, we approximate the distribution function of the t-th order statistic $F^{(t)}$ through the exponential distribution function with parameter $1/\mathbb{E}[U^{(t)}]$. We then say that a decision rule is approximately optimal in terms of the ideal of justice W, and denote it by $\widetilde{D}_W$, if it is optimal whenever we replace the order statistics of the culture with their exponential approximations.

Theorem 3.4. *For every n, for every k, and for every culture,* $\widetilde{D}_{\mathrm{MN}}$ *is a scoring rule with* $\widetilde{D}_{\mathrm{MN}}(M) = \arg\max_{l \in K} \sum_{t=1}^{k} S^t_{\mathrm{MN}}$, *where* $S^t_{\mathrm{MN}} = S^t_{\mathrm{UT}} \frac{\mathbb{E}[U^{(k)}]}{\mathbb{E}[U^{(t)}]}$.

Theorem 3.4 is fundamental for a better understanding of the optimal maximin decision rule. It tells us that the approximate optimal maximin decision rule is a scoring rule, that, as in the case of utilitarianism, relies on the expected values of the order statistics. Moreover, it establishes that S^t_{MN} takes the optimal utilitarian scoring rule S^t_{UT} and applies a factor $\frac{\mathbb{E}[U^{(k)}]}{\mathbb{E}[U^{(t)}]}$ that makes S_{MN} more concave, shifting it in the direction of negative voting. Note that if the utilitarian scoring rule S_{UT} is very convex, then the distance between $\mathbb{E}[U^{(k)}]$ and $\mathbb{E}[U^{(t)}]$ is large, and hence, the factor $\frac{\mathbb{E}[U^{(k)}]}{\mathbb{E}[U^{(t)}]}$ makes S_{MN} very concave. These intuitions can immediately be seen in Figure 2.

10. See Rausand and Hoyland (2004) for an introduction to system reliability theory.

11. See Daley (1988) for some early results on quantifying departure from exponentiality and Cheng and He (1989) for applied results in the context of reliability theory. See Reiss (1989) for a textbook treatment in the context of order statistics.

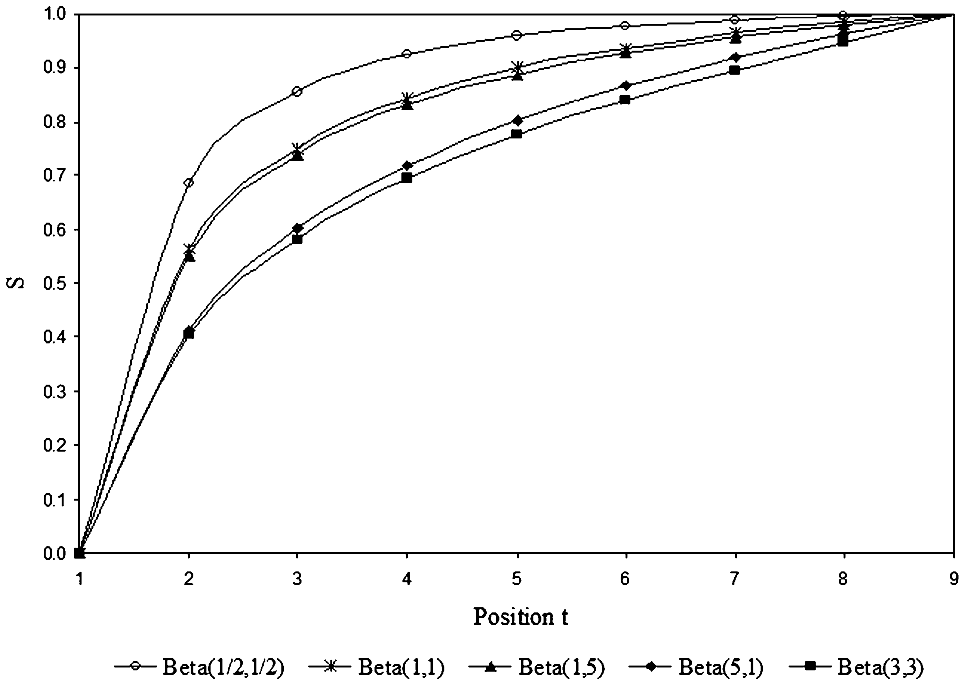

FIGURE 2
Approximate optimal maximin scoring rules for the corresponding density functions of Figure 1(a)

Figure 2 takes the same five distribution functions used for utilitarianism in Figure 1, to represent the approximate optimal maximin scoring rules S_{MN}, as characterized in Theorem 3.4. The figure makes it apparent that the approximate optimal scoring rules S_{MN} are truncated towards negative for the five distributions. For example, the optimal scoring rule for the uniform distribution assigns weights above 0.5 to all alternatives except the worst.

3.3. *Maximax*

Maximax shares the difficulty of maximin in that the maximum operator also fails to preserve the independence of the order statistics. We therefore face a similar tractability problem to the one we met in the previous section. In fact, our analysis of maximax parallels the previous analysis of maximin. To this extent, we first offer the exact optimal maximax decision rule, and see that it is somewhat opaque regarding its relationship with the underlying culture of the society. We then approximate the distributions of the order statistics through exponential distributions and obtain that the optimal maximax decision rule is approximately a scoring rule, characterized by the expected values of the order statistics. Furthermore, the maximax scoring rule simply takes the optimal utilitarian scoring rule and applies a factor that convexifies it, shifting it in the direction of plurality.

We use the same tools as in Theorem 3.3 to obtain the exact shape of the optimal maximax decision rule D_{MX}. To do so, we assume that $\overline{u} < \infty$ and define the inverse random variable $\widehat{U}_i^l = \overline{u} - U_i^l$, with distribution function $\widehat{F}$, where $\widehat{F}(u) = 1 - F(\overline{u} - u)$. Accordingly, $\widehat{z}^{(t)}(\widehat{u})$ denotes the satiation rate of the t-th order statistic $\widehat{U}^{(t)}$. Theorem 3.5 offers the exact shape of the optimal maximax decision rule, characterized by the satiation rates of the inverse order statistics $\widehat{U}_i^l$.

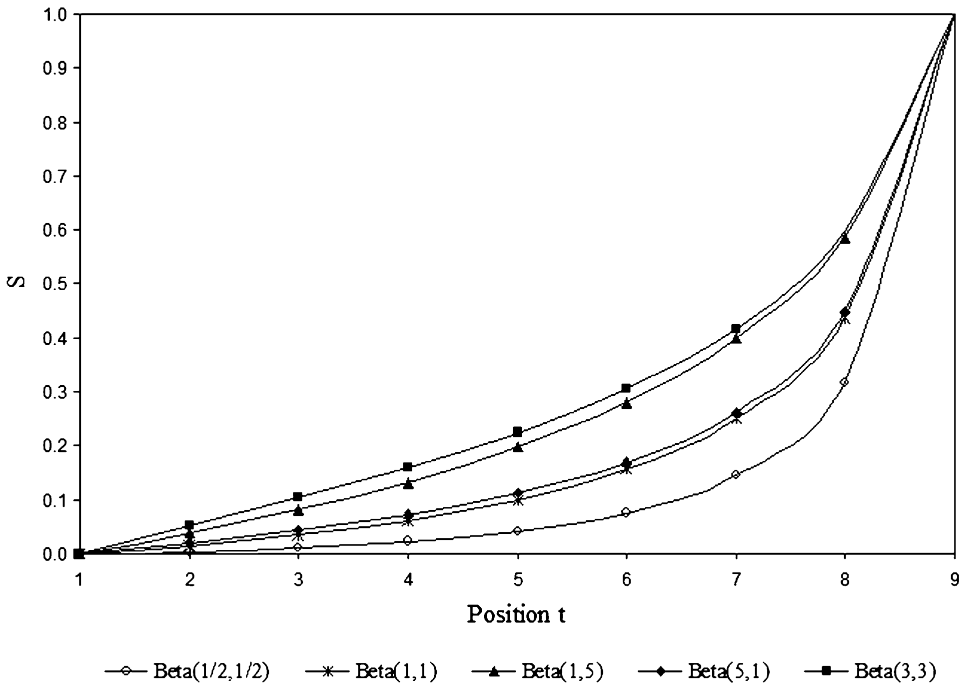

FIGURE 3

Approximate optimal maximax scoring rules for the corresponding density functions of Figure 1(a)

Theorem 3.5. *For every n, for every k, and for every culture, the optimal maximax decision rule is* $D_{\rm MX}(M) = \arg\min_{l \in K} \int_I \exp\left(-\int_0^{\widehat{u}} \sum_{t=1}^{k} \mathbf{l}^{(t)} \widehat{z}^{(k-t+1)}(v)\,dv\right) d\widehat{u}$.

We now approximate the distribution functions of the order statistic $\widehat{F}^{(t)}$ by the exponential distribution with parameter $1/\mathbb{E}[\widehat{U}^{(t)}]$. This enables us to state the following result.

Theorem 3.6. *For every n, for every k, and for every culture,* $\widetilde{D}_{\rm MX}$ *is a scoring rule with* $\widetilde{D}_{\rm MX}(M) = \arg\max_{l \in K} \sum_{t=1}^{k} S^t_{\rm MX}$, *where* $S^t_{\rm MX} = S^t_{\rm UT} \frac{\overline{u} - \mathbb{E}[U^{(k)}]}{\overline{u} - \mathbb{E}[U^{(t)}]}$.

Theorem 3.6 shows that the approximate maximax optimal decision rule is a scoring rule that, as in the case of maximin, draws upon the optimal utilitarian scoring rule. For every position t, $S^t_{\rm MX}$ takes the value $S^t_{\rm UT}$ and modifies it by applying a factor that depends on the expected values of the t-th and k-th order statistics $\mathbb{E}[U^{(t)}]$ and $\mathbb{E}[U^{(k)}]$. This factor convexifies $S_{\rm MX}$, with regard to $S_{\rm UT}$, shifting it towards plurality voting. Figure 3 illustrates these considerations.

The figure takes the same five distribution functions as in the case of utilitarianism (and maximin) and computes the maximax scoring rules according to Theorem 3.6. It can be readily seen that the five scoring rules are pushed downwards in the direction of plurality. For example, the optimal scoring rule for the uniform distribution assigns weights below 0.5 to all alternatives except the best.

4. CONCLUSIONS

This paper explores the relationship between ideals of justice and decision rules. Whereas ideals of justice are typically presented in cardinal terms, decision rules are primarily constructed on the basis of ordinal information. We study the cardinal consequences of using ordinal-based decision rules.

We have shown that the optimal choice of decision rule depends on the criterion of justice that one wishes to follow. Among our specific findings, we emphasize that our results identify a particularly prominent set of decision rules as optimal: the set of scoring rules. Interestingly, the optimal scoring rules of the three ideals of justice under consideration are intimately linked. The optimal maximin and maximax scoring rules take the optimal utilitarian scoring rules and apply a factor that shifts them upwards and downwards, respectively. It emerges that maximax is best approached by scoring rules with a plurality shape, maximin by scoring rules with a negative voting shape, and, for the uniform distribution, utilitarianism is best approached by Borda.

APPENDIX A. PROOFS

Proof of Theorem 3.1. Given that for any $l,h \in K$ and $i \neq j$ the random variable U_i^l is independent of the random variable U_j^h, we can write

$$\mathbb{E}\left[\frac{\sum_{i\in N} U_i^l}{n} \;\middle|\; M\right] = \frac{\sum_{i\in N}\mathbb{E}[U_i^l \mid M_i]}{n}.$$

In addition, since the random variables $\{U_i^l\}_{l\in K}$ are i.i.d., we can write

$$\frac{\sum_{i\in N}\mathbb{E}[U_i^l \mid M_i]}{n} = \frac{\sum_{i\in N}\mathbb{E}[U_i^l \mid m_i^l]}{n}.$$

Recall that $U^{(m_i^l)}$ denotes the m_i^l–th order statistic determined by how individual i ranks alternative l. Thus, by definition,

$$\frac{\sum_{i\in N}\mathbb{E}[U_i^l \mid m_i^l]}{n} = \frac{\sum_{i\in N}\mathbb{E}[U^{(m_i^l)}]}{n}.$$

We can write the last expression in terms of the number of individuals who rank alternative l in the same position t. That is,

$$\frac{\sum_{i\in N}\mathbb{E}[U^{(m_i^l)}]}{n} = \sum_{t=1}^{k}\mathbf{l}^{(t)}\frac{\mathbb{E}[U^{(t)}]}{n}.$$

Hence,

$$D_{\mathrm{UT}}(M) = \arg\max_{l\in K}\mathbb{E}\left[\frac{\sum_{i\in N} U_i^l}{n} \;\middle|\; M\right] = \arg\max_{l\in K}\sum_{t=1}^{k}\mathbf{l}^{(t)}\frac{\mathbb{E}[U^{(t)}]}{n}.$$

Finally, we only have to normalize the weights in the last expression to show that D_{UT} is indeed a scoring rule. It is immediate that whenever $A \geq 0$ and $B > 0$,

$$\arg\max_{l\in K}\sum_{t=1}^{k}\mathbf{l}^{(t)}\frac{\mathbb{E}[U^{(t)}]}{n} = \arg\max_{l\in K}\sum_{t=1}^{k}\mathbf{l}^{(t)}\frac{\mathbb{E}[U^{(t)}]-A}{B}.$$

In particular, we can use the values $A = \mathbb{E}[U^{(1)}]$ and $B = \mathbb{E}[U^{(k)}] - \mathbb{E}[U^{(1)}]$. Clearly, it is the case that $A \geq 0$, and the continuity of the density function f guarantees that B is strictly larger than zero. Thus, we obtain

$$D_{\mathrm{UT}}(M) = \arg\max_{l\in K}\sum_{t=1}^{k}\mathbf{l}^{(t)}\frac{\mathbb{E}[U^{(t)}]-\mathbb{E}[U^{(1)}]}{\mathbb{E}[U^{(k)}]-\mathbb{E}[U^{(1)}]}. \quad \|$$

Proof of Theorem 3.2. Denote by $F^{t+1,t}$ the distribution function of the random variable $U^{(t+1)} - U^{(t)}$, $1 \leq t \leq k-1$. Theorem 5.1 in Boland *et al.* (2002, p. 616) shows that if the density function f is increasing, the random variable

$U^{(t)} - U^{(t-1)}$ stochastically dominates the random variable $U^{(t+1)} - U^{(t)}$. That is, $1 - F^{t,t-1}(u) \geq 1 - F^{t+1,t}(u)$ for all $u \in I$. Then, if f is increasing, it follows that $\mathbb{E}[U^{(t)} - U^{(t-1)}] = \int_I (1 - F^{t,t-1}(u))du \geq \int_I (1 - F^{t+1,t}(u))du = \mathbb{E}[U^{(t+1)} - U^{(t)}]$, $2 \leq t \leq k-1$. Now, it is easy to see that,

$$\begin{aligned}\mathbb{E}[U^{(t)} - U^{(t-1)}] \geq \mathbb{E}[U^{(t+1)} - U^{(t)}] &\Rightarrow \mathbb{E}[U^{(t)}] - \mathbb{E}[U^{(t-1)}] \geq \mathbb{E}[U^{(t+1)}] - \mathbb{E}[U^{(t)}]\\ &\Rightarrow \frac{\mathbb{E}[U^{(t)}] - \mathbb{E}[U^{(t-1)}]}{\mathbb{E}[U^{(k)}] - \mathbb{E}[U^{(1)}]} \geq \frac{\mathbb{E}[U^{(t+1)}] - \mathbb{E}[U^{(t)}]}{\mathbb{E}[U^{(k)}] - \mathbb{E}[U^{(1)}]}\\ &\Rightarrow S^t_{\text{UT}} - S^{t-1}_{\text{UT}} \geq S^{t+1}_{\text{UT}} - S^t_{\text{UT}}.\end{aligned}$$

This proves that, when f is increasing, D_{UT} is a concave scoring rule. The claim that, when f is decreasing, D_{UT} is a convex scoring rule is analogous, and hence, it is omitted. ‖

The following lemma will be useful in Theorems 3.3 and 3.5.

Lemma A.1 *For any k, and for every $1 \leq t \leq k$, $(1 - F^{(t)}(u)) = \exp\left(-\int_0^u z^{(t)}(v)dv\right)$.*

Proof of Lemma A.1. Recall that $z^{(t)}(u) = \frac{f^{(t)}(u)}{1-F^{(t)}(u)}$. Therefore, we can write

$$\int_0^u z^{(t)}(v)dv = \int_0^u \frac{f^{(t)}(v)}{1 - F^{(t)}(v)}dv = \int_0^u -\frac{d(ln(1 - F^{(t)}(v)))}{dv}dv.$$

Given that $1 - F^{(t)}(0) = 1$, it follows that $ln(1 - F^{(t)}(0)) = 0$ and therefore,

$$\int_0^u -\frac{d(ln(1 - F^{(t)}(v)))}{dv}dv = -ln(1 - F^{(t)}(u)).$$

Consequently, $(1 - F^{(t)}(u)) = \exp\left(-\int_0^u z^{(t)}(v)dv\right)$, and the claim follows. ‖

Proof of Theorem 3.3. By definition,

$$\mathbb{E}[MN^l(U^l) \mid M] = \mathbb{E}[\min_{i\in N} U_i^l \mid M] = \int_I P(\min_{i\in N} U_i^l > u \mid M)du.$$

Given that the random variables $\{U_i^l\}_{i\in N, l\in K}$ are i.i.d., the latter is equal to

$$\begin{aligned}\int_I P(\min_{i\in N} U_i^l > u \mid M)du &= \int_I P(\min_{i\in N} U_i^l > u \mid M^l)du\\ &= \int_I \prod_{i\in N} P(U_i^l > u \mid m_i^l)du.\end{aligned}$$

The latter expression can be formulated in terms of the distribution functions of the respective order statistics. Hence,

$$\begin{aligned}\int_I \prod_{i\in N} P(U_i^l > u \mid m_i^l)du &= \int_I \prod_{i\in N} (1 - F^{(m_i^l)}(u))du\\ &= \int_I \prod_{t=1}^k (1 - F^{(t)}(u))^{\mathbf{l}^{(t)}}du.\end{aligned}$$

By Lemma A.1, we know that the distribution function of an order statistic can be expressed in terms of the satiation rate of that order statistic. Thus, we can write the last product as a product dependent on the satiation rates:

$$\begin{aligned}\int_I \prod_{t=1}^k (1 - F^{(t)}(u))^{\mathbf{l}^{(t)}}du &= \int_I \prod_{t=1}^k \exp\left(-\int_0^u z^{(t)}(v)dv\right)^{\mathbf{l}^{(t)}}du\\ &= \int_I \exp\left(-\int_0^u \sum_{t=1}^k \mathbf{l}^{(t)} z^{(t)}(v)dv\right)du.\end{aligned}$$

Therefore,

$$D_{\text{MN}}(M) = \arg\max_{l\in K} \int_I \exp\left(-\int_0^u \sum_{t=1}^k \mathbf{l}^{(t)} z^{(t)}(v)dv\right)du. \quad \|$$

Proof of Theorem 3.4. Consider the approximation of $f^{(t)}(u)$ through the exponential density function $\frac{1}{\mathbb{E}[U^{(t)}]}\exp\left(-\frac{1}{\mathbb{E}[U^{(t)}]}u\right)$. It is straightforward to see that, for every position t, the satiation rate of such an exponential function is constant:

$$\frac{\frac{1}{\mathbb{E}[U^{(t)}]}\exp\left(-\frac{1}{\mathbb{E}[U^{(t)}]}u\right)}{1-\left(1-\exp\left(-\frac{1}{\mathbb{E}[U^{(t)}]}\right)\right)} = -\frac{1}{\mathbb{E}[U^{(t)}]}.$$

Then, we can rewrite the result of Theorem 3.3 as

$$\begin{aligned}\mathbb{E}[MN^l(U^l) \mid M] &= \int_I \exp\left(-\int_0^u \sum_{t=1}^k \mathbf{I}^{(t)} z^{(t)}(v)dv\right)du\\ &\simeq \int_I \exp\left(\int_0^u \sum_{t=1}^k \mathbf{I}^{(t)} \frac{1}{\mathbb{E}[U^{(t)}]}dv\right)du\\ &= \int_I \exp\left(\sum_{t=1}^k \mathbf{I}^{(t)} \frac{1}{\mathbb{E}[U^{(t)}]}u\right)du = \frac{1}{\sum_{t=1}^k \mathbf{I}^{(t)}\frac{1}{\mathbb{E}[U^{(t)}]}}.\end{aligned}$$

Therefore, by the definition of the approximate maximin optimal decision rule $\widetilde{D}_{\mathrm{MN}}$, we have

$$\begin{aligned}\widetilde{D}_{\mathrm{MN}}(M) &= \arg\max_{l\in K} \frac{1}{\sum_{t=1}^k \mathbf{I}^{(t)}\frac{1}{\mathbb{E}[U^{(t)}]}} = \arg\min_{l\in K}\sum_{t=1}^k \mathbf{I}^{(t)}\frac{1}{\mathbb{E}[U^{(t)}]}\\ &= \arg\max_{l\in K}\frac{A-\sum_{t=1}^k \mathbf{I}^{(t)}\frac{1}{\mathbb{E}[U^{(t)}]}}{B},\end{aligned}$$

where the last equality holds whenever $A \geq 0$ and $B > 0$. In particular, we can use the values $A = \frac{1}{\mathbb{E}[U^{(1)}]}$ and $B = \frac{1}{\mathbb{E}[U^{(1)}]} - \frac{1}{\mathbb{E}[U^{(k)}]}$. It immediately follows that $A \geq 0$, and the continuity of f guarantees that B is strictly larger than zero. Thus, we obtain

$$\widetilde{D}_{\mathrm{MN}}(M) = \arg\max_{l\in K}\sum_{t=1}^k \mathbf{I}^{(t)}\frac{\frac{1}{\mathbb{E}[U^{(1)}]}-\frac{1}{\mathbb{E}[U^{(t)}]}}{\frac{1}{\mathbb{E}[U^{(1)}]}-\frac{1}{\mathbb{E}[U^{(k)}]}}.$$

Then, $\widetilde{D}_{\mathrm{MN}}$ is a scoring rule with weights $S^t_{\mathrm{MN}} = \frac{\frac{1}{\mathbb{E}[U^{(1)}]}-\frac{1}{\mathbb{E}[U^{(t)}]}}{\frac{1}{\mathbb{E}[U^{(1)}]}-\frac{1}{\mathbb{E}[U^{(k)}]}} = S^t_{UT}\frac{\mathbb{E}[U^{(k)}]}{\mathbb{E}[U^{(t)}]}$. ‖

Proof of Theorem 3.5. Following the same logical steps as in the proof of Theorem 3.3, we know that

$$\begin{aligned}\mathbb{E}[MX^l(U^l) \mid M] &= \int_I \left(1-\prod_{i\in N} P(U_i^l \leq u \mid m_i^l)\right)du\\ &= \bar{u} - \int_I \prod_{i\in N} F^{(m_i^l)}(u)du.\end{aligned}$$

By the change of variable, we can write

$$\bar{u} - \int_I \prod_{i\in N} F^{(m_i^l)}(u)du = \bar{u} - \int_I \prod_{i\in N}(1-\widehat{F}^{(k-m_i^l+1)}(\widehat{u}))d\widehat{u},$$

and therefore,

$$\bar{u} - \int_I \prod_{i\in N}(1-\widehat{F}^{(k-m_i^l+1)}(\widehat{u}))d\widehat{u} = \bar{u} - \int_I \prod_{t=1}^k (1-\widehat{F}^{(k-t+1)}(\widehat{u}))^{\mathbf{I}^{(t)}}d\widehat{u}.$$

Lemma A.1 guarantees that

$$\begin{aligned}\bar{u} - \int_I \prod_{t=1}^k (1-\widehat{F}^{(k-t+1)}(\widehat{u}))^{\mathbf{I}^{(t)}}d\widehat{u} &= \bar{u} - \int_I \prod_{t=1}^k \exp\left(-\int_0^{\widehat{u}} \widehat{z}^{(k-t+1)}(v)dv\right)^{\mathbf{I}^{(t)}}d\widehat{u}\\ &= \bar{u} - \int_I \exp\left(-\int_0^{\widehat{u}}\sum_{t=1}^k \mathbf{I}^{(t)}\widehat{z}^{(k-t+1)}(v)dv\right)d\widehat{u}\end{aligned}$$

and hence

$$D_{\mathrm{MX}}(M) = \arg\max_{l\in K}\left(\bar{u} - \int_I \exp\left(-\int_0^{\widehat{u}}\sum_{t=1}^{k}\mathbf{l}^{(t)}\widehat{z}^{(k-t+1)}(v)dv\right)d\widehat{u}\right)$$
$$= \arg\min_{l\in K}\int_I \exp\left(-\int_0^{\widehat{u}}\sum_{t=1}^{k}\mathbf{l}^{(t)}\widehat{z}^{(k-t+1)}(v)dv\right)d\widehat{u}. \quad \|$$

Proof of Theorem 3.6. The same reasoning as applied in the proof of Theorem 3.4 tells us that

$$\widetilde{D}_{\mathrm{MX}}(M) = \arg\min_{l\in K}\frac{1}{\sum_{t=1}^{k}\mathbf{l}^{(t)}\frac{1}{\mathbb{E}[\widehat{U}^{(k-t+1)}]}} = \arg\max_{l\in K}\sum_{t=1}^{k}\mathbf{l}^{(t)}\frac{1}{\mathbb{E}[\widehat{U}^{(k-t+1)}]}.$$

Normalizing the scores to lie in the unit interval, we obtain

$$\widetilde{D}_{\mathrm{MX}}(M) = \arg\max_{l\in K}\sum_{t=1}^{k}\mathbf{l}^{(t)}\frac{\frac{1}{\mathbb{E}[\widehat{U}^{(k-t+1)}]}-\frac{1}{\mathbb{E}[\widehat{U}^{(k)}]}}{\frac{1}{\mathbb{E}[\widehat{U}^{(1)}]}-\frac{1}{\mathbb{E}[\widehat{U}^{(k)}]}}.$$

Clearly, $\widetilde{D}_{\mathrm{MX}}$ is a scoring rule with weights

$$S^t_{\mathrm{MX}} = \frac{\frac{1}{\mathbb{E}[\widehat{U}^{(k-t+1)}]}-\frac{1}{\mathbb{E}[\widehat{U}^{(k)}]}}{\frac{1}{\mathbb{E}[\widehat{U}^{(1)}]}-\frac{1}{\mathbb{E}[\widehat{U}^{(k)}]}} = \frac{\mathbb{E}[\widehat{U}^{(k)}]-\mathbb{E}[\widehat{U}^{(k-t+1)}]}{\mathbb{E}[\widehat{U}^{(k)}]-\mathbb{E}[\widehat{U}^{(1)}]}\frac{\mathbb{E}[\widehat{U}^{(1)}]}{\mathbb{E}[\widehat{U}^{(k-t+1)}]}.$$

Given that, by construction, for every $1\le t\le k$, $\mathbb{E}[U^{(t)}] = \bar{u} - \mathbb{E}[\widehat{U}^{(k-t+1)}]$, we can write

$$S^t_{\mathrm{MX}} = S^t_{\mathrm{UT}}\frac{\bar{u}-\mathbb{E}[U^{(k)}]}{\bar{u}-\mathbb{E}[U^{(t)}]}. \quad \|$$

Acknowledgment. We thank Andrew Daughety, Toni Calvo, Gabor Lugosi, Andrea Moro, Larry Samuelson, and Rann Smorodinsky for helpful comments. We are particularly grateful to the editor Andrea Prat and four anonymous referees for many valuable comments, suggestions, and criticisms that have led to a substantial improvement in the manuscript. Financial support by the Spanish Commission of Science and Technology (ECO2009-12836, ECO2008-04756, EC02008-01768), NSF grant SES08-14312, the Barcelona GSE research network, and the Government of Catalonia is gratefully acknowledged.

REFERENCES

AUSTEN-SMITH, D. and BANKS, J. S. (1999), *Positive Political Theory I: Collective Preference*, Michigan Studies in Political Analysis (Ann Arbor: University of Michigan Press).

BARBERA, S. and JACKSON, M. (2004), "Choosing How to Choose: Self-Stable Majority Rules and Constitutions", *Quarterly Journal of Economics*, **119**, 1011–1048.

BARBERA, S. and JACKSON, M. (2006), "On the Weights of Nations: Assigning Voting Weights in a Heterogeneous Union", *Journal of Political Economy*, **114**, 317–339.

BENOIT, J. P. and KORNHAUSER, L. A. (2010), "Only a Dictatorship is Efficient", *Games and Economics Behavior*, **70**, 261–270.

BOLAND, P. J., HU, T., SHAKED, M. and SHANTHIKUMAR, G. (2002), "Stochastic Ordering of Order Statistics II", in Dror, M., L'Ecuyer, P. and Szidarovszky, F. (eds) *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*(Dordrecht: Kluwer Academic Publishers).

BORDLEY, R. F. (1983), "A Pragmatic Method for Evaluating Election Schemes through Simulation", *American Political Science Review*, **77**, 123–141.

BRAMS, S. J. and FISHBURN, P. C. (1978), "Approval Voting", *American Political Science Review*, **72**, 831–847.

CAPLIN, A. and NALEBUFF, B. (1988), "On 64%-Majority Rule", *Econometrica*, **56**, 787–814.

CASELLA, A. (2005), "Storable Votes", *Games and Economic Behavior*, **51**, 391–419.

CHENG, K. and He, Z. (1989), "On Proximity between Exponential and DMRL Distributions", *Statistics & Probability Letters*, **8**, 55–57.

DALEY, D. J. (1988), "Tight Bounds on the Exponential Approximation of some Aging Distributions", *Annals of Probability*, **16**, 414–423.

DASGUPTA, P. and MASKIN, E. (2008), "On the Robustness of Majority Rule", *Journal of the European Economic Association*, **6**, 949–973.

DAVID, H. A. and NAGARAJA, H. N. (2003), *Order Statistics*, 3rd edn (New Jersey: John Wiley & Sons Inc).

DEMEYER, F. and PLOTT, C.R. (1970), "The Probability of a Cyclical Majority", *Econometrica*, **38**, 345–354.

GEHRLEIN, W. V. (1997), "Condorcet's Paradox and the Condorcet Efficiency of Voting Rules", *Mathematica Japonica*, **45**, 173–199.

HORTALA-VALLVE, R. (2009), "Qualitative Voting" (Economics Series Working Paper No. 320, University of Oxford).

HORTALA-VALLVE, R. (2010), "Inefficiencies on Linking Decisions", *Social Choice and Welfare*, **34**, 471–486.

JACKSON, M. and SONNENSCHEIN, H. F. (2007), "Overcoming Incentive Constraints by Linking Decisions", *Econometrica*, **75**, 241–257.

LEVIN, J. and NALEBUFF, B. (1995), "An Introduction to Vote-Counting Schemes", *Journal of Economic Perspectives*, **9**, 3–26.

MAGGI, G. and MORELLI, M. (2006), "Self-Enforcing Voting in International Organizations", *American Economic Review*, **96**, 1137–1158.

MERRILL, S. (1984), "A Comparison of Efficiency of Multicandidate Electoral Systems", *American Journal of Political Science*, **28**, 23–48.

MYERSON, R. (2002),"Comparison of Scoring Rules in Poisson Voting Games", *Journal of Economic Theory*, **103**, 219–251.

NURMI, H. (1983), "Voting Procedures: A Summary Analysis", *British Journal of Political Science*, **13**, 181–208.

OZKAL-SANVER, I. and SANVER, M. R. (2006), "Ensuring Pareto Optimality by Referendum Voting", *Social Choice and Welfare*, **27**, 211–219.

RAE, D. (1969), "Decision Rules and Individual Values in Constitutional Choice", *American Political Science Review*, **63**, 40–56.

RAUSAND, M. and HOYLAND, A. (2004), *System Reliability Theory: Models, Statistical Methods, and Applications*, 2nd edn (New Jersey: John Wiley & Sons Inc).

REISS, R. D. (1989), *Approximate Distributions of Order Statistics: with Applications to Nonparametric Statistics* (New York: Springer Verlag).

SAARI, D. G. (1999), "Explaining All Three-Alternative Voting Outcomes", *Journal of Economic Theory*, **87**, 313–355.

THOMAS, P. Y. and SAMUEL, P. (2008), "Recurrence Relations for the Moments of Order Statistics from a Beta Distribution", *Statistical Papers*, **49**, 139–146.

WEBER, R. J. (1978), "Comparison of Voting Systems" (Mimeo, Cowles Foundation).