THE LIKELIHOOD OF CHOOSING THE BORDA-WINNER WITH PARTIAL PREFERENCE RANKINGS OF THE ELECTORATE

Ömer Eğecioğlu
Department of Computer Science,
University of California,
Santa Barbara CA 93106
omer@cs.ucsb.edu

Ayça E. Giritligil
Institute for Economic Analysis,
Campus UAB Bellaterra,
Barcelona 08193 Spain
ayca.ebru@uab.es

ABSTRACT

This paper aims to empirically investigate the extent of difficulty in implementing Borda outcomes when n voters report only the first r ranks of their *linear* preference rankings over m alternatives. The information contained in the first r ranks is aggregated through a Bordalike method, namely the r-Borda rule. Monte-Carlo experiments are run to detect changes in the likelihood of the r-Borda winner to coincide with the original Borda winner as a function of m, n and r. The voters' preferences are generated through Impartial Anonymous and Neutral Culture model where both the names of the alternatives and voters are immaterial. For a given r, the likelihood of choosing the Borda winner decreases towards zero independently of n as m increases within the computed range of parameter values, $1 \leq m,n \leq 30$. We show empirically that this probability is inversely proportional to m, and determine the constant of proportionality for two different types of likelihood that we consider.

1. INTRODUCTION

A voting rule solves the collective decision problem, where the voters must jointly choose one among the possible outcomes (alternatives), on the basis of the reported ordinal preferences. The choice of a voting rule has been a major ethical question ever since the political philosophy of the Enlightenment. When only two alternatives are at stake, the ordinary majority voting, whose axiomatic formulation is due to May(1952) is unambiguously regarded as the 'fairest' method. For three and more alternatives, plurality voting, where each voter reports the name of (exactly) one alternative on his ballot and the alternative receiving the most votes wins, has been historically the most popular voting rule. The two celebrated critiques of plurality voting, Borda(1781) and Condorcet(1785) noted that plurality voting may elect a poor candidate, namely, one that would lose in a simple pair-wise majority comparison to every other candidate. Borda and Condorcet individually devised different rules to replace plurality voting as attempts to extend majority voting among pairs. Borda introduced the Borda Rule, which assigns points to each candidate, linearly increasing with the candidate's ranking in a voter's opinion, and elects the alternative with the highest total score over all alternatives. On the other hand, Condorcet, a strong critic of Borda Rule, provided the most analyzed non-positionalist voting principle, the principle that, if a candidate defeats every other candidate on the basis of the simple majority rule, then that candidate should be the winner in the election. These two approaches have generated most of the modern scholarly research which focus on the ways of aggregating individual preferences into social decisions in a manner compatible with the fulfillment of a variety of positive and normative criteria.

As shown by Niemi and Riker(1976), Fishburn(1984), Nurmi(1987), Amy(2000), there is no perfect voting procedure, however some procedures are clearly superior to others in satisfying certain criteria. Borda rule certainly occupies a special place among all voting rules due to its important superiorities. Unlike the great majority of the voting rules in the literature, Borda Rule has theoretical characterizations. Among those, Young(1974), Hansson and Sahlquist(1976), Nitzan and Rubinstein(1981) and Debord(1992) are the best known ones. It is shown by Saari(1987, 1989, 1990, 2001) that Borda rule is less susceptible than all other positional scoring rules to some unsettling possibilities and paradoxes. Some of the theoretical and probabilistic results concerning Borda rule are summarized in Brams and Fishburn(2002) and Pattanaik(2002). However, among its shortcomings, its vulnerability to strategic manipulations and the difficulty of implementing it in real life remain its most criticized properties. There are many studies which theoretically and/or probabilistically consider the former issue. In this paper, we focus on the latter which, to our knowledge, has not been studied in detail.

The implementation criterion is concerned with the complexity of the information that a voting procedure requires from voters concerning their preference over the alternatives. Unlike non-ranked single-stage voting procedures (such as plurality voting, negative plurality voting, and approval voting) and non-ranked multistage voting procedures (such as plurality with a run-off and plurality with successive elimination), Borda rule is a ranked procedure as Condorcet-consistent voting rules since it requires the complete preference rankings of the electorate over the set of available alternatives. This requirement is quite difficult to fulfill due to the associated complications both on the side of the voters and the administrators who are to collect this information.

In this paper, we aim to investigate the extent of difficulty in implementing Borda outcomes when voters are asked to rank only a specified number of alternatives. We consider the situations where n voters are required to report only the first r ($1 \leq r < m$) ranks of their *linear* (i.e. full or total) preferences over m alternatives. We assume that the truncated individual preferences are aggregated through a method a lá Borda, namely *r-Borda rule.* The *r-Borda rule* assigns strictly positive points to each alternative that appears in the first r-ranks of a voter's actual preference, linearly increasing with its rank, and gives zero points to the ones that are not among the top *r*-ranks in the voter's decision. Then, the alternative(s) that receive(s) the highest score over the entire electorate is elected as the *r*-Borda winner(s).

We run Monte-Carlo experiments to ascertain the information content of the first *r* ranks of the electorate's preferences from the perspective of implementing the (original) Borda outcome, i.e. the likelihood of *r*-Borda winner to coincide with the Borda outcome that would emerge if the full preferences could be considered, namely the *m*-Borda winner. It should be noted that, the way that *r*-Borda rule aggregates the voter preferences is different than the aggregations implemented by well-known single- and multi-stage non-ranked procedures which permit truncated ballots. In other words, the present study is not aiming to detect the likelihood of any of these rules to choose the Borda winner(s).

The *à la Borda* aggregation of truncated individual preferences is a popular method for sports and contests in real life. The "Most Valuable Player" contest run by the Associated Press and United Press International in NCAA (National Collegiate Athletic Association) sports, as well as the "Most Valuable Player" contest for the National Basketball Association in the United States, the Eurovision Song Contest and the Formula 1 Car Race are best known examples which adopt Borda-like aggregation methods for the truncated ballots. These contests require the voters to rank a specified number of candidates. Each stated candidate is given a score depending on its rank in the preference ordering of a voter, and the candidate that receives the maximum score gets elected as the winner. The number of candidates to be ranked and the scores to be assigned to the ranks change from one contest to another. The People's Remix Music Competition requires the voters to rank their top three candidates and gives three, two and one points to the first-, second- and third-ranks relatively. This is equal to our rank scoring for $r=3$.

Like many other voting procedures, the Borda rule can choose more than one alternative as the winner. Instead of randomly breaking the ties, we adopt two types of probabilities computed for triples of *m*, *n* and *r* as the likelihood of choosing the Borda winner with partially reported linear preference rankings. The first type refers to the likelihood of choosing *exactly* the same set of Borda winners of the original profile by considering only the reported partial profile, i.e. the probability of the set of *m*-Borda winners and *r*-Borda winners to be identical. The second type of probability considers the likelihood of *r*-Borda winners to be included in the set of *m*-Borda winners. We investigate the changes in these values as a function of *m*, *n* and *r* by considering all possible values of these parameters in an appropriate range.

The voters' preferences are generated through Impartial Anonymous and Neutral Culture (IANC) model. As introduced by Eğecioğlu and Giritligil (2005), IANC model treats the voters' preferences through a class of preference profiles, namely *root profiles*, where the names of both voters and alternatives are immaterial.

2. CONTRIBUTION AND RELATION TO LITERATURE

To our knowledge, this study is the first attempt in the literature to analyze the extent of difficulty in implementing Borda outcomes when voters are asked to rank only a specified number of alternatives, and where the underlying model is as structurally general as is possible. The contribution of the paper and its relation to literature can be discussed based on two main issues: aggregation of truncated preferences and generation of voters' preferences.

Aggregation of truncated preferences

A positional scoring vector for the (finite) alternative set A of m elements is a real vector $s = (s_1,s_2,..,s_m)$ for which $s_1 \geq s_2 \geq \ldots . \geq s_m$ and $s_1 > s_m$. As the most well-known positional scoring procedure, Borda rule is defined by the scoring vector $s(B_m) = (m-1, \quad m-2, \ldots, 0)$. That is, $s_i = m-i$ for all i, and the difference in scores s_i-s_j is proportional to j-i for all i and j. In this paper, we consider the situation where the voters are asked to state only the first r ranks of their linear preferences over the m alternatives and investigate the importance of the information revealed in the first r ranks of the electorate's preferences from the perspective of implementing the Borda outcome. Here, the way that this partial information is aggregated plays the key role. We aggregate this information in a Borda fashion. That is, an alternative in rank i is assigned the score $s_i(B_r) = r+1-i$ if $i \leq r$, and $s_i(B_r) = 0$ otherwise.

Nurmi(1987) considers the situations where a proposal of a lower level body concerning the order of priority among the candidates to an office is to be submitted to the nominating authority. He stresses the importance of determining non-arbitrary ways of coming up with a collective preference order on a subset of candidates whose cardinality is fixed by the information asked of the voter, and proposes a method which guarantees some degree of "proportionality". Although this method is not exactly equal to the r-Borda rule we adopt here, it has similar peculiarities to it.

It should be noted that r-Borda rule aggregates the truncated preferences unlike constant scoring rules. A constant scoring rule asks each voter to indicate a given (and constant) number of alternatives. Each of the indicated alternatives receives one point whereas the all the others get zero, and the alternative with the most votes is elected. Hence, for $m \geq 3$, the scoring vector imposed by r-Borda rule is not the same as the one assigned by a constant scoring procedure unless r=1, i.e. (1,0…,0), which is the vector identifying of the most popular constant scoring rule, namely plurality rule. That is, the outcomes of the plurality rule and 1-Borda procedure coincide for $m \geq 3$. The probability of constant scoring voting rules to choose the Borda outcome has been studied by Gehrlein(1981), Gehrlein and Lepelley(2000) and Vandercruyssen(1999). Gehrlein and Fishburn(1980) and Gehrlein etal.(1982), on the other hand, provide the propensity of pairs of score vectors for a set A of alternatives and a nonempty proper subset of A to yield the same ranking over the subset for an arbitrary profile of linear orders on A.

The method we adopt to aggregate truncated preferences is also different than some other procedures that permit truncated ballots. Among these, approval voting has been considered the most in both theoretical and practical grounds. Under approval voting, each voter indicates those alternatives that s/he approves of. Each approved alternative receives a point, and the winner is the alternative with the highest point total, summed over all voters. Axiomatized by Fishburn (1978) and Sertel(1978), approval voting is extensively analyzed with comparisons to Borda rule, as well as other rules. Brams and Fishburn (2002) provide a summary of the theoretical debate between approval voting and the Borda rule.

Recently, a version of approval voting, fallback voting, was proposed by Brams and Sanver(2005) which is the application of fallback bargaining of Brams and Kilgour(2001) to voting. Fallback voting requires the voters to rank their approved alternatives and successively falls back on lower-ranked approved alternatives if no higher-ranked approved alternatives receive majority approval. For both approval voting and fallback voting, the number of alternatives to be indicated or ranked by the electorate is not given or homogeneous across the voters. Majoritarian compromise, on the other hand, requires the ranking of $\lceil card(A)/2 \rceil$ alternatives, where $\lceil . \rceil$ is the ceiling function). Introduced by Sertel(1987), majoritarian compromise selects the candidate that has the support of the majority in the best degree possible. Clearly, these rules aggregate the truncated preferences in a quite different fashion than the way we consider in this paper. Therefore, the likelihood of the coincidence of the outcomes of these rules and of the Borda rule is a different question than what we ask here.

Let us consider a social planner who believes that Borda Rule is the "fairest" voting rule to be adopted for an election, but is unable to ask the voters to report their full preference rankings over m alternatives because of the practical complications alluded to above. As it is generally practiced, he can ask the electorate to report only their first-best alternatives. Or, he can require the voters to state the first r ($1 \leq r < m$) ranks of their linear preference rankings in the hope of increasing the probability of choosing the Borda winner that would be elected if the full linear preference rankings could be collected. Especially in such a case like this, it seems natural to aggregate the reported partial rankings via the original Borda fashion for the sake of preserving some "consistency" in the aggregation method used for declared preferences.

Generation of voters' preferences

An immense literature has been devoted to analyze the "behaviors" of various social choice rules through the use of computer simulations and of probability models designed to generate voters' preferences. There are two basic and most commonly used probability models in this literature, namely, Impartial Culture (IC) and Impartial Anonymous (IAC) conditions. IC condition has been introduced in social choice literature by Guilbaud(1952). It is the multinomial equiprobable preference profiles model that assumes that each voter selects her preference according to a uniform probability distribution. On the other hand, IAC, which has first been introduced by Fishburn and Gehrlein(1978), relies also on an equiprobability assumption, but without taking the identity of the voters into account. The details about these assumptions and their extensive use in the literature are presented in Berg and Lepelly (1994) and Gehrlein(1997).

In this paper, we generate the voters' preferences through IANC model which is introduced by Eğecioğlu and Giritligil(2005). IANC is also an equiprobability assumption, however, unlike IC and IAC, it is able to neglect the names of both alternatives and voters. IANC regards each root profile from which all the preference profiles can be generated through renaming the alternatives and voters, equally probable. Since the number of root profiles is small relative to the number of profiles that can be generated for m alternatives and n voters, IANC enables the researcher to obtain quite accurate probabilities even for large parameter values. As it is noted in Section 3.2, the probabilities computed through

IAC and IANC models coincide only in the vanishingly small likelihood of $m!$ and n being relatively prime.

We would like to remark that when a voting rule is invariant under a large group of symmetries of the underlying set of preference profiles, structurally the most accurate theory is the one which takes all of the symmetries that leave the outcome invariant into account. Regarding no set of symmetries gives the simplest model IC; taking into account only the symmetry of the voters gives the commonly used model IAC; and taking the symmetries of both the voters and the alternatives into account gives the strongest model IANC. Taking more symmetries into account makes it possible to characterize structural properties of the voting rules more accurately, however the mathematics becomes more complicated with each step, with resulting algorithmic difficulties in generating preference profiles "fairly" (i.e. requiring that each distinct profile should be equally likely to be selected) from the equivalence classes under the group of symmetries assumed. This paper is the first study in the literature which generates the electorate's preferences through IANC model.

Public Choice Society adopted approval voting in its presidential election of 2006. The voters were permitted to vote for as many candidates as they liked. Based on the method of generalized spectral analysis introduced by Lawson etal.(2006), Brams etal.(2006) compare the results of the election with the possible outcomes that would have been obtained if plurality, Condorcet, Borda or single transferable vote has been adopted. Through a different method developed by Falmagne and Regenwetter(1996), Regenwetter and Grofman(1998) analyze ten three-candidate elections conducted under approval voting and construct a distribution of rank orders from subset choice data. It should be noted that both Brams etal.(2006) and Regenwetter and Grofman(1998) start with partial information on the voter preferences and adopt different methods in order to obtain complete preference rankings based on this information. However, in the present study, we take the full orderings over the set of alternatives as given and then consider the first r-ranks of these orderings. It should be noted that given r-ranks of preference orderings, assigning probabilities to each alternative to be the Borda winner, and based on these probabilities, checking whether the "possible" Borda winner(s) coincide(s) with the r-Borda winner is a different approach for detecting probabilities, clearly leading to different results.

3. PRELIMINARIES

3. 1. Preference Profiles and the Borda Rule

By a *preference* on a set A we mean any function $p: A \to 2^A$ which assigns to every $a \in A$ a subset ("lower contour set") $p(A) \subset A$ such that, at all $a, b \in A$ we have

(1) $b \in p(a)$ or $a \in p(b)$ [completeness]

(2) $p(b) \subset p(a)$ whenever $b \in p(a)$ [transitivity]

(3) $b \in p(a)$ and $a \in p(b)$ only if $a = b$ [antisymmetry]

Such a preference clearly corresponds to a linear (or total) order on A.

We denote by $\boldsymbol{p}(A)$ the set of all preferences on a set A. For any positive integer n we write $[n] = \{1,2,..,n\}$, and by a *preference profile* for a society of n voters on a set A we mean any family $P_{m,n} = (p_i)_{i\in[n]} \in \boldsymbol{p}(A)^{[n]}$ of preferences p_i on A indexed by voters $i \in [n]$.

Let $card(p_i(\text{a}))$ be the cardinality of the lower counter set of $a \in A$ for $i\in[n]$. Note that the cardinalities of the top- and bottom- ranked alternatives are m and 1, respectively.

The *Borda score* of $a \in A$ for $i\in[n]$ is defined as,

$$B_i^a = card(p_i(\text{a})),$$

and the set of m-Borda winners at each $P_{m,n} \in \boldsymbol{p}(A)^{[n]}$ is determined by setting

$$\text{B}(P_{m,n}) = \underset{a\in A}{\arg\max} \sum_{i\in[n]} B_i^a .$$

Thus, the Borda Rule chooses the candidates who maximize the total Borda score aggregated over the set of all n voters.

Let $P_{m,n}^r$ denote the portion of a preference profile $P_{m,n}$ where only the first r ranks of the voters' preferences can be observed. When we view the profile $P_{m,n}$ as an $m\times n$ matrix with m rows and n columns, then $P_{m,n}^r$ corresponds to the $r\times n$ submatrix of $P_{m,n}$ consisting of the first r rows. Note that, when $r = m$ or $r = m$-1, the Borda outcome of the entire preference profile is detectable. However when $r < m$-1, the observable preference p_i^r of voter i corresponds to a partial strict ordering on A which is transitive and antisymmetric, however incomplete. Let $A_i^r \subseteq A$ be the set of alternatives that appear at p_i^r .

Let $card(p_i^r(\text{a}))$ be the cardinality of the *observable* lower counter set of $a \in A_i^r$ for $i\in[n]$. Note that $1 \leq card(p_i^r(\text{a})) \leq r$. We now re-define the Borda score of $a \in A$ for $i\in[n]$ as,

$$B_i^a = \begin{cases} card(p_i^r(\text{a})), & \text{if } a \in A_i^r \\ 0, & \text{otherwise,} \end{cases}$$

and the set of r-Borda winners at any $P_{m,n}^r$ is given by

$$\text{B}(P_{m,n}^r) = \underset{a\in A}{\arg\max} \sum_{i\in[n]} B_i^a .$$

In other words, if an alternative is among the first r-ranks of a voter's original preference ranking, then its Borda score is equal to the number of the alternatives it beats at the first r-

ranks of this preference. If it is not one of the top r-ranked alternatives, it receives a Borda score of zero. Then, the "modified" Borda rule chooses the alternatives with the highest Borda scores aggregated over all voters as the set of r-Borda winners.

When the preference profile and the dependence on m and n is clear from the context, we denote B($P^r_{m,n}$) simply by B_r, for $1 \leq r \leq m$. Note that B_m (and also $B_{m-1} = B_m$) is the set of Borda winners of the full profile, and B_1 is the set of plurality winners.

3.2 Root Profiles and IANC

Let $\Omega(m,n)$ denote the set of all preference profiles that can be generated for m alternatives and n voters. As shown in Eğecioğlu and Giritligil (2005), a product permutation group on the names of alternatives and of voters acts on $\Omega(m,n)$, and splits it up into a disjoint union of subsets called *orbits*, i.e.

$$\Omega(m,n) = \Theta_1 + \Theta_2 + \ldots + \Theta_\omega,$$

where each Θ_i is an *anonymous and neutral equivalence class* (ANEC). All the preference profiles within an anonymous and neutral equivalence class can be generated from each other through re-labeling the alternatives and/or the voters. The preference profiles of any such class are equal in the sense that any anonymous and neutral social choice rule, such as Borda Rule, yields the same alternative as the winner (however, under different names) for all of them.

A *root profile* is any preference profile that represents an anonymous and neutral equivalence class. That is, all the other preference profiles within the same equivalence class (orbit) can be generated from this root profile via permuting the names of the m alternatives and simultaneously those of the n voters. We denote by $R = R(m,n)$ the collection of root profiles for m alternatives and n voters. Each element of this set represents an equivalence class of preference profiles from $\Omega(m,n)$.

We define a preference profile as a matrix of size $m \times n$ which shows how each of the n voters linearly orders m alternatives. We assume that the voters correspond to the columns and the alternatives correspond to the rows of the matrix.

As an example, let us consider a case with two alternatives, a and b, and three voters labeled v_1, v_2 and v_3. There are two possible linear preference rankings for two alternatives: a is strictly preferred to b, or b is strictly preferred to a. The total number of preference profiles that can be generated is $2!^3 = 8$.

P_1:

v_1	v_2	v_3
a	a	a
b	b	b

P_2:

v_1	v_2	v_3
a	a	b
b	b	a

P_3:

v_1	v_2	v_3
a	b	a
b	a	b

P_4:

v_1	v_2	v_3
b	a	a
a	b	b

P_5: P_6: P_7: P_8:

v_1	v_2	v_3
b	b	a
a	a	b

v_1	v_2	v_3
b	a	b
a	b	a

v_1	v_2	v_3
a	b	b
b	a	a

v_1	v_2	v_3
b	b	b
a	a	a

Note that there are four AECs for this example: $AEC_1=\{P_1\}$, $AEC_2=\{P_2, P_3, P_4\}$, $AEC_3=\{P_5, P_6, P_7\}$ and $AEC_4=\{P_8\}$. There are two possible permutations for the names of the alternatives: one is the identity permutation which leaves the names of the alternatives intact, and the other is the permutation which relabels *a* as *b* and *b* as *a*. If we apply these permutations to the AECs above, we obtain a further partition of the set $\{AEC_1, AEC_2, AEC_3, AEC_4\}$ of anonymous equivalence classes. Note that this new partition gives us only two ANECs: $ANEC_1=\{AEC_1, AEC_4\}$, and $ANEC_2=\{AEC_2, AEC_3\}$. The root representing $ANEC_1$ shows a preference structure at which all voters have the same preference ranking, and the root representing $ANEC_2$ exhibits a structure where one of the preference rankings is adopted by two voters and the other is adopted by one voter.

IANC uses root profiles to represent voters' preferences through an application of the Dixon Wilf algorithm which enables the roots to be generated from the uniform distribution for *m* alternatives and *n* voters. That is, each root profile is generated uniformly with probability 1/*card*(*R(m,n)*). This allows for Monte-Carlo algorithms for the empirical analysis of the behaviors and applications of anonymous and neutral social choice rules even for large electorate size and high number of alternatives as we describe below in the analysis of selecting the Borda winner, when only partial preference rankings of the voters are available.

Although an immediate way to compute the number of roots (ANECs) for a given *m* and *n* seems to be dividing the number of AECs by *m*!, there are cases for which this does not work. Let us demonstrate the problem via the following example: For *m*=2 and *n*=2, there are four preference profiles:

P_1:

v_1	v_2
a	a
b	b

P_2:

v_1	v_2
a	b
b	a

P_3:

v_1	v_2
b	a
a	b

P_4:

v_1	v_2
b	b
a	a

There are three AECs for this case: $AEC_1=\{P_1\}$, $AEC_2=\{P_2, P_3\}$, $AEC_3=\{P_4\}$. Here 3/2! is not an integer and there are two ANECs: $ANEC_1=\{AEC_1, AEC_3\}$, and $ANEC_2=\{AEC_2\}$.

For $ANEC_2$, no permutations of alternatives lead to anything other than $ANEC_2$ **i**tself, so it has only one AEC. Hence, an ANEC might contain less than *m*! AECs. A special case of a formula for the calculation of the number of ANECs in Eğecioğlu(2005) generalizing a result of Giritligil and Doğan(2004), as well as results presented in

Eğecioğlu and Giritligil (2005) explain the effect of the numerical properties of *n* and *m* in determining the number of roots. In particular, given a pair of integers *m* and *n*, each ANEC root represents *m*! AECs if and only if *m*! and *n* are relatively prime. However, the probability of *m*! and *n* to be relatively prime is vanishingly small, a proof of this fact can be found in the Appendix.

4. LIKELIHOOD MEASURES

4.1 Types of Likelihood:

Two types of probabilities are considered for measure of the likelihood of implementing the Borda outcome with partially observed linear preference rankings of the voters.

- $\boldsymbol{Pr_1} = \boldsymbol{Pr_1}(m,n,r)$ refers to the likelihood of choosing exactly the set of Borda winners themselves by considering only the first *r* rows of a preference profile, $P^r_{m,n}$. In other words $\boldsymbol{Pr_1}$ is the probability that $B_r = B_m$. For a given preference profile $P_{m,n}$ we consider the random variable

$$f_1(P^r_{m,n}) = \begin{cases} 1, & \text{if } B(P^r_{m,n}) = B(P_{m,n}) \\ 0, & \text{otherwise.} \end{cases}$$

Then,

$$Pr_1 = \frac{1}{card\ (R(m,n))} \sum_{P_{m,n} \in R(m,n)} f_1(P^r_{m,n})$$

- $\boldsymbol{Pr_2} = \boldsymbol{Pr_2}(m,n,r)$ is the likelihood that an *r*-Borda winner (chosen at random from among all *r*-Borda winners) is one of the Borda winners of the full profile. Thus it is the likelihood that an element of B_r is actually an element of B_m. For a given $P_{m,n}$, consider the random variable

$$f_2(P^r_{m,n}) = \frac{card\ (B(P_{m,n}) \cap B(P^r_{m,n}))}{card\ (B(P^r_{m,n}))} \qquad (1)$$

Then,

$$Pr_2 = \frac{1}{card\ (R(m,n))} \sum_{P_{m,n} \in R(m,n)} f_2(P^r_{m,n})$$

Note that the above definition of $\boldsymbol{f_2}(P^r_{m,n})$ is intuitive: If $B(P^r_{m,n}) \subseteq B(P_{m,n})$, then $\boldsymbol{f_2}(P^r_{m,n}) = 1$, since any *r*-Borda winner selected on the basis of the first *r* ranks is automatically a Borda winner; if $B(P^r_{m,n})$ and $B(P_{m,n})$ are disjoint, then $\boldsymbol{f_2}(P^r_{m,n}) = 0$, as in this case no r-Borda winner can possibly be a Borda winner; and if B($P^r_{m,n}$) contains, say, two elements of B($P_{m,n}$), then $f_2(P^r_{m,n}) = \frac{2}{card\ (B(P^r_{m,n}))}$, i.e. only two elements out of a total of B($P^r_{m,n}$) would be Borda winners for the complete profile.

Given the distribution of profiles to be generated for a given m and n, and a given r, the approximate probability (of both types) is computed through Monte-Carlo integration by making use of the law of large numbers. The law of large numbers implies that the average of a random sample from a large population is likely to be close to the mean of the whole population. We apply this to the averages that define $\boldsymbol{Pr_1}$ and $\boldsymbol{Pr_2}$ through the random variables $\boldsymbol{f_1}$ and $\boldsymbol{f_2}$.

The tools provided by Eğecioğlu and Giritligil (2005) allow for the generation of roots profiles from *R(m,n)* with probability 1/*card*(*R(m,n)*). Then we can approximate the actual probability as follows: we generate a large number of root profiles from *R(m,n)* with uniform probability 1/*card*(*R(m,n)*), where each selection is done independently of the others. Let *S(m,n)* denote the set of these generated profiles. Then the law of large numbers implies that

$$\boldsymbol{Pr_1} = \frac{1}{card\ (R(m,n))} \sum_{P_{m,n} \in R(m,n)} f_1(P^r_{m,n}) \approx \frac{1}{card\ (S(m,n))} \sum_{P_{m,n} \in S(m,n)} f_1(P^r_{m,n}) \qquad (2)$$

Hence, when only the first r-ranks of the n voters' preferences over m available alternatives are reported, the probability of implementing the Borda outcome by using only these partial preferences equals to the computed probabilities aggregated over all generated profiles in *S(m,n)*, divided by the total number generated. Note that for the formulation (2) to result in a valid Monte-Carlo algorithm for the computation of $\boldsymbol{Pr_1}$, it is essential that each $P_{m,n}$ in *S(m,n)* be drawn from the uniform probability on *R(m,n)*.

The computation of $\boldsymbol{Pr_2}$ by a Monte-Carlo method is similar. We generate a set *S(m,n)* of root profiles of m alternatives and n voters, where each profile in *S(m,n)* is generated uniformly and independently. Then

$$\boldsymbol{Pr_2} \approx \frac{1}{card\ (S(m,n))} \sum_{P_{m,n} \in S(m,n)} f_2(P^r_{m,n}) \qquad (3)$$

Again, for the formulation (3) to result in a valid Monte-Carlo algorithm for the computation of $\boldsymbol{Pr_2}$, each $P_{m,n}$ in *S(m,n)* must be drawn from the uniform probability on *R(m,n)*.

4.2. Examples:

Below we present a preference profile as a matrix of size $m \times n$ where the voters correspond to the columns and the alternatives correspond to the rows.

Example 1: Let us consider the following profile:

a	a	b	c	d
b	b	d	d	a
c	c	c	b	b
d	d	a	a	c

The original Borda winner for the above profile is 'b', i.e. B($P_{4,5}$) = {b}. If only the first ranked alternatives are reported, then 'a' is chosen as the 1-Borda (or Plurality) winner, i.e. B($P^1_{4,5}$) = {a}. Since {a} ≠ {b}, f_1($P^1_{4,5}$) = 0. On the other hand, f_2($P^1_{4,5}$) = 0 because {a}∩ {b}= ∅. If the first two ranks of the voters' preferences are reported, i.e. if we pick r = 2, then B($P^2_{4,5}$) = {a}.[1] Since B($P_{4,5}$) = {b}, we again have f_1($P^2_{4,5}$) = 0 and f_2($P^2_{4,5}$) = 0.

Example 2: Now let us consider the profile below:

a	a	d	d	b
b	d	c	c	c
c	c	b	a	a
d	b	a	b	d

The set of Borda winners is B($P_{4,5}$) ={a,c,d}and the set of 1-Borda winners is B($P^1_{4,5}$) = {a,d}. Then, for this profile f_1($P^1_{4,5}$) = 0 and f_2($P^1_{4,5}$) = 1. For r=2, we find that B($P^2_{m,n}$) = {d}, and therefore the profile yields f_1($P^2_{4,5}$) = 0 and f_2($P^2_{4,5}$) = 1.

Example 3:

a	a	b	b	c
b	c	c	c	a
c	b	a	d	b
d	d	d	a	d

For the above profile, the set of Borda winners is B($P_{4,5}$) = {b,c}, and the set of 1-Borda (Plurality) winners is B($P^1_{4,5}$) = {a,b}. Then, for this profile f_1($P^1_{4,5}$) = 0 and f_2($P^1_{4,5}$) = 1/2. We find that B($P^2_{m,n}$) = {a,b,c}, and hence the profile yields f_1($P^2_{4,5}$) = 0 and f_2($P^2_{4,5}$) = 2/3.

We observe that when f_1($P^r_{m,n}$) = 1, then B($P^r_{m,n}$) = B($P_{m,n}$), and consequently

[1] The set of *(m-1)*-Borda winners is necessarily equal to the set of original Borda winners (i.e. the set of *m*-Borda winners).

$$f_2(P^r_{m,n}) = \frac{card\ (B(P_{m,n}) \cap B(P^r_{m,n}))}{card\ (B(P^r_{m,n}))} = \frac{card\ (B(P^r_{m,n}))}{card\ (B(P^r_{m,n}))} = 1.$$

On the other hand when $f_1(P^r_{m,n}) = 0$, then $B(P^r_{m,n}) \neq B(P_{m,n})$, but it is still possible that these two sets are not disjoint. In other words, when $f_1(P^r_{m,n}) = 0$, it may happen that $f_2(P^r_{m,n}) \geq 0$. It follows that $f_1(P^r_{m,n}) \leq f_2(P^r_{m,n})$. As it follows from the definition of Pr_1 and Pr_2, for any selection of *m, n, r*, we always have $Pr_1 \leq Pr_2$.

5. MONTE-CARLO EXPERIMENTS

At the heart of the Monte-Carlo experiments given here is the Mathematica program `GenerateRoot[m,n]` (the Mathematica notebook containing this function can be accessed online for experimentation: see Eğecioğlu (2004)). The program `GenerateRoot[m,n]` takes two integers *m* and *n* as input parameters and generates an IANC preference profile as an $m \times n$ matrix as its output. The preference profile returned each time by `GenerateRoot[m,n]` is guaranteed to be distributed over the *R(m,n)* roots uniformly. As we remarked before, to be able to estimate the probabilities Pr_1 and Pr_2 through the formulations (2) and (3) by using the law of large numbers, we need the preference profiles generated be uniform over the set of roots *R(m,n)*. `GenerateRoot[m,n]` does exactly that.

The design of the Monte-Carlo experiments was as follows. We have generated 1000 root profiles for each value of the parameters *m, n* under consideration. Thus we took *card(S(m,n))* = 1000. Taking larger sample sizes is a matter of computational resources available, as this particular algorithm is of the type referred to as "embarrassingly parallel". The ranges of the parameters were taken to be $1 \leq m,\ n \leq 30$ for most Monte-Carlo experiments carried out. Below are the basic steps followed for the computation of both Pr_1 and Pr_2 type probabilities in the symbolic algebra package Mathematica:

Step 1: We have generated the values of *m* and *n* themselves, $1 \leq m, n \leq 30$ iteratively by means of two nested loops.

Step 2: For the given values of *m* and *n,* we have invoked the function `GenerateRoot[m,n]`, which returned a preference profile $P_{m,n}$ from the uniform distribution on the set of root profiles *R(m,n)*.

Step 3: We have computed the set of Borda winners $B(P_{m,n})$ for the complete profile $P_{m,n}$ returned.

Step 4: In addition, for every value of *r* in the range $1 \leq r < m$, at this point we also computed the set of Borda winners $B(P^r_{m,n})$ by considering only the first *r* rows of the profile $P_{m,n}$.

Step 5: Using the sets $B(P_{m,n})$ and $B(P^r_{m,n})$ available, we evaluated the random variables f_1 and f_2 corresponding to the triple *m, n, r*.

Step 2 through Step 5 were executed *card(S(m,n))* times. The approximations to $\boldsymbol{Pr_1}$ and $\boldsymbol{Pr_2}$ for given *m*, *n* and *r* were calculated afterwards by dividing the sum of the computed values of $\boldsymbol{f_1}$ and $\boldsymbol{f_2}$ in Step 5 by *card(S(m,n))*.

5.1 Experimental results on $\boldsymbol{Pr_1}$ type probabilities

We present the computed $\boldsymbol{Pr_1}$ type probabilities of the set of *r*-Borda winners to be exactly equal to the set of *m*-Borda winners in Table 1. To be able to present these results in a comprehensible fashion, we have used a fixed value $n = 30$ for the tabulation given. The range of *m* is $1 \leq m \leq 30$. For *m* alternatives, the value of *r* changes in the range $1 \leq r \leq m$. The rows of Table 1 are indexed by *m*, and the columns are indexed by *r*. For the description of the values in this table, we consider a special case: suppose $m = 5$. We generated 1000 IANC profiles from the uniform distribution, and computed the fraction of the cases for which B_r and B were identical, where B_r is the set of Borda winners obtained by scoring the first *r* rows of the profile, and $B = B_m$ is the set of Borda winners for the profile. These are the computed probabilities listed in the table. Let B_1, B_2, B_3, B_4, B_5 denote the *r*-Borda winners obtained by considering the first 1, 2, 3, 4, and 5 rows respectively. Thus B_1 is the set of plurality winners, $B = B_5$ the set of Borda winners. Then we can read from the fifth row of Table 1 that

$\Pr[B_1=B] = 0.51$,

$\Pr[B_2=B] = 0.671$,

$\Pr[B_3=B] = 0.818$,

$\Pr[B_4=B] = \Pr[B_5=B] = 1$.

Figure 1 is a three-dimensional plot of the computed $\boldsymbol{Pr_1}$ type probabilities. It is interesting that for the values of the parameters shown, $\boldsymbol{Pr_1}$ appears to be independent of *n* especially as the value of *n* increases. A close observation shows that for *n* fixed at 30, $\boldsymbol{Pr_1} \rightarrow 0$ as *m* gets large for fixed *r*, and the behavior is roughly as $(1+r)/m$. We have made a least squares fit of the sections of this surface at the values of *m*, $1 \leq m \leq 30$, by considering the family of functions of the form $f(m,r) = c(1+r)/m$ (c constant). The best approximating function in the least-squares sense was found to be

$$\boldsymbol{Pr_1} \sim f(m,r) = r/m + 1.4/m \qquad (4)$$

Figure 2 is a three-dimensional plot of the values of this approximating function $f(m,r) = r/m + 1.4/m$. Comparing with the plot of the actual probabilities given in Figure 1, we see that the probability that the set of *r*-Borda winners to be exactly equal to the set of Borda winners very well approximated by the expression for $\boldsymbol{Pr_1}$ in (4).

To summarize, for large values of *n*, the likelihood of choosing exactly the set of Borda winners themselves by considering only the first *r* ($1<r<m-1$) rows of a preference profile is independent of *n*, and increases as the ratio *r/m* increases. However, it is impossible to guarantee the exact Borda outcome unless *r* is set to be equal to *m-1* or *m*.

5.2 Experimental results on Pr_2 type probabilities

In Table 2, we present the computed $\boldsymbol{Pr_2}$ type probabilities of the set of *r*-Borda winners to contain an element which is also an *m*-Borda winner. We present these results for the fixed value $n = 30$. The range of *m* is $1 \leq m \leq 30$. For *m* alternatives, the value of *r* changes in the range $1 \leq r \leq m$. The rows of Table 2 are indexed by *m*, and the columns are indexed by *r*. If we take $m = 5$ and let B_1, B_2, B_3, B_4, B_5 denote the *r*-Borda winners obtained by considering the first 1, 2, 3, 4, and 5 rows respectively, then B_1 is the set of plurality winners, $B = B_5$ the set of Borda winners. We can read off the $\boldsymbol{Pr_2}$ type probabilities for this case from the fifth row of Table 2 as

$\Pr[B_1=B] = 0.607$,

$\Pr[B_2=B] = 0.759$,

$\Pr[B_3=B] = 0.895$,

$\Pr[B_4=B] = \Pr[B_5=B] = 1$.

Figure 3 is a three-dimensional plot of the computed $\boldsymbol{Pr_2}$ type probabilities. We observe that as in the case of the $\boldsymbol{Pr_1}$ type probabilities, and for *n* fixed at 30, $\boldsymbol{Pr_2} \to 0$ as *m* gets large for fixed *r*, and the behavior is roughly as $(1+r)/m$. We can make use of the properties of the analytic approximations for the $\boldsymbol{Pr_2}$ type probabilities and surmise that $\boldsymbol{Pr_2}$ probabilities should be well approximated by

$$\boldsymbol{Pr_2} \sim f(m,r) = r/m + 2.1/m \qquad (5)$$

Indeed, the plot of this function in Figure 4 is a good approximation to the actual values in Figure 3.

Hence, for large values of *n*, the likelihood that an *r*-Borda winner (chosen at random from among all *r*-Borda winners) is one of the Borda winners of the full profile is independent of *n* and increases as the ratio *r/m* increases. Our results show that, for $r = m-2$, $\boldsymbol{Pr_2}$ approaches one as *m* increases. From our data it can also be conjectured that for any fixed *k*, $\boldsymbol{Pr_2}$ approaches 1 for $r=m-k$, but the rate of convergence decreases for larger *k*.

6. CONCLUDING REMARKS

The Borda rule is one of the most considered voting procedures in the theoretical literature of social choice. However, despite its well-known superiorities concerning the fulfillment of important positive and normative criteria, it is not a widely used procedure in the real world. There is no doubt that its requirement from the voters to report their full rankings over the possible alternatives can be blamed for this hesitation.

In this study, we attempt to empirically investigate the extent of difficulty in implementing Borda outcomes when voters are asked to rank only a specified number of alternatives. We consider the situations where *n* voters are required to report only the first *r* ($1 \leq r < m$) ranks of their *linear* preferences over *m* alternatives. We assume that the truncated individual preferences are aggregated through a Bordalike method which we call *r-Borda rule.* We generate the voters' preferences through IANC model which is an equiprobability assumption neglecting the names of both alternatives and voters.

The results of the Monte-Carlo experiments we run indicate that, for large values of n, the likelihood of choosing *exactly* the same set of original Borda winners by considering only the first r ranks of the voters' preference orderings is independent of n, and it approaches to zero as m gets large for fixed r. Through least square fit method, we show that, for any m, it is impossible to guarantee the exact Borda outcome with only partial rankings over the alternatives. We observe that the likelihood that an r-Borda winner is one of the Borda winners of the full profile is also independent of n and it approaches to zero as m gets large for fixed r. Our results show that for $r = m-k$, for k fixed, this probability approaches one as m increases.

There are some immediate directions for further research on the topic. First of all, although the r-Borda rule, as an equal-distance scoring method, is a quite "intuitive" way of aggregating the truncated preferences, empirical studies can be designed to compare the success of assigning different score vectors to the reported ranks for given triples of m, n and r from the perspective of implementing the Borda outcome.

Second, the extent of implementation problems associated with other well-known ranked rules can be empirically investigated. However, especially in the case of pair-wise majority rules (such as Condorcet rule), it should be noted that the methods to be used for aggregating truncated preferences is not as straightforward or intuitive as it is in the case of scoring methods since the lower-counter set cardinalities of the alternatives do not provide the sufficient information for determining the outcomes of such rules. Hence, there is a need for both theoretical and empirical research concerning the matter.

The empirical research on the implementability of ranked voting rules with partial preference orderings can surely contribute to the literature for comparing these rule with the ones that permit truncated ballots.

REFERENCES

Amy, D.C. (2000). *Behind the Ballot Box: A Citizen's Guide to Voting Systems*, Praeger, Westport, CT.

Berg, S. and D. Lepelley (1994) "On Probability Models in Voting Theory", *Statistica Neerlandica* **48**:133-146.

Borda, J.C. de. (1781). Mémoire sur les élections au scrutin. *Histoire de l'Académie Royale des Sciences*, Paris.

Brams, S. and D.M. Kilgour (2001). Fallback bargaining, *Group Decision and Negotiation* **10** : 287-316.

Brams, S. and P.C. Fisburn (2002). Voting Procedures. In K. J. Arrow, A.K. Sen and K. Suzumura (eds.), *Handbook of Social Choice and Welfare* Vol.1, Elsevier Science, Amsterdam.

Brams, S. and M.R. Sanver (2005). Voting Systems that combine approval and preference, mimeo.

Brams, S., M.W. Hansen and M.E. Orrison (2006). Dead heat : The 2006 Public Choice Society election, *Public Choice* **128** : 361-366.

Condorcet, M. de. (1785). Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix, Paris.

Debord, B. (1992). An axiomatic characterization of Borda's *k*-choice function, *Social Choice and Welfare* **9**: 337-343

Eğecioğlu Ö. (2004). `GenerateRoot[m,n]`, Mathematica function: available online as a Mathematica notebook http://www.cs.ucsb.edu/~omer/randomvote.nb.

Eğecioğlu Ö. (2005). Monte-Carlo Algorithms for Uniform Generation of Anonymous and Neutral Preference Structures, *Technical Report TR2005-25,* Department of Computer Science, University of California, Santa Barbara. http://www.cs.ucsb.edu/~omer/mcma_Jan9_2009.pdf, *Monte Carlo Methods and Applications* (in review, 2009).

Eğecioğlu Ö. and A. E. Giritligil (2005). Public preference structures with Impartial Anonymous and Neutral Culture Model, preprint, http://www.cs.ucsb.edu/~omer/SCpaper_JME_Nov11_2008.pdf

Falmagne, J. and M. Regenwetter (1996). A random utility model for approval voting, *Journal of Mathematical Psychology* **40** : 152-159.

Fishburn, P.C. (1978). Axioms for approval voting: direct proof, *Journal of Economic Theory* **19**: 180-185.

Fishburn, P.C. and W. Gehrlein (1978) "Condorcet Paradox and Anonymous Preference Profiles", *Public Choice* **26**: 1-18.

Fishburn, P.C. (1984). Discrete mathematics in voting and group choice", *SIAM Journal of Algebraic and Discrete Methods* **5**: 263-275.

Gehrlein, W.V. and P.C. Fishburn (1980). Robustness of positional scoring over subsets of alternatives, *Applied Mathematics and Optimization* **6**: 241-255.

Gehrlein, W.V. (1981). Single stage election procedures for large electorates, *Journal of Mathematical Economics* **8**: 263-275.

Gehrlein, W.V., B. Gopinath, J.C. Lagarias and P.C. Fishburn (1982). Optimal pairs of score vectors for positional scoring rules, *Applied Mathematics and Optimization* **8**: 309-324.

Gehrlein, W.V. (1997) “Condorcet’s Paradox, the Condorcet Efficiency of the Voting Rules”, *Mathematica Japonica* **45**:173-199.

Gehrlein, W.V. and D. Lepelley (2000). The probability that all weighted scoring rules elect the same winner, *Economic Letters* **66**, 191-197.

Giritligil, A.E. and O. Doğan (2005). A possibility results for the existence of anonymous, neutral, pareto-optimal and monotonic social choice rules, mimeo.

Guilbaud, G.T. (1952) "Les théories de l'intéret genéral et le problémelogique de l'agrégation", *Economie Appliquée* **5**: 501-584.

Hansson, B. and H. Sahlquist (1976). A proof technique for social choice with variable electorate, *Journal of Economic Theory* **13**: 193-200.

Lawson, B.L., M. Orrison and D.T. Uminsky (2006). Spectral analysis of supreme court, *Mathematics Magazine*, forthcoming.

May, K. (1952). A set of independent necessary and sufficient conditions for simple majority decision, *Econometrica* **20**:680-684.

Niemi, R.G., and W.H. Riker (1976). The choice of voting systems, *Scientific American* **234**: 21-27.

Nitzan, S. and A. Rubinstein (1981). A further characterization of Borda ranking method, *Public Choice* **36**: 153-158.

Nurmi, H. (1987). *Comparing Voting Systems*, Reidel, Dordrecht.

Pattanaik, P.K. (2002). Positional rules of collective decision making, In K. J. Arrow, A.K. Sen and K. Suzumura (eds.), *Handbook of Social Choice and Welfare* Vol.1, Elsevier Science, Amsterdam.

Regenwetter, M. and B. Grofman (1998). Approval voting, Borda winners, and Condorcet winners: evidence from seven elections, *Management Science* **44**:520-533.

Saari, D.G. (1987). The source of some paradoxes from social choice and probability, *Journal of Economic Theory* **41**: 1-22.

Saari, D.G. (1989). A dictionary for voting paradoxes, *Journal of Economic Theory* **48**: 443-475.

Saari, D.G. (1990). Susceptibility to manipulation, *Public Choice* **64**: 21-41.

Saari, D.G.. (2001). *Chaotic Elections! A Mathematician Looks at Voting*, American Mathematical Society, Providence.

Sertel, M.R. (1978). Characterizing approval voting, *Journal of Economic Theory* **45**: 207-211.

Sertel, M.R. (1987) A Non-Dictatorial Compromise. *Social Choice and Welfare* **4**: 1-11.

Vandercruyssen, D. (1999). Analysis of voting procedures in one-seat elections: Condorcet efficiency and Borda efficiency, Discussion Paper Series DPS 99.11, Center for Economic Studies, Kathlieke Universiteit Leuven, Leuven.

Young, H.P. (1974). An axiomatization of Borda's rule, *Journal of Economic Theory* **9**: 43- 52

1.																													
1.	1.																												
0.799	1.	1.																											
0.595	0.8	1.	1.																										
0.51	0.671	0.818	1.	1.																									
0.444	0.582	0.704	0.832	1.	1.																								
0.363	0.521	0.622	0.738	0.843	1.	1.																							
0.344	0.478	0.579	0.691	0.775	0.873	1.	1.																						
0.274	0.404	0.512	0.603	0.7	0.79	0.878	1.	1.																					
0.259	0.38	0.471	0.576	0.664	0.741	0.823	0.909	1.	1.																				
0.226	0.362	0.459	0.548	0.629	0.695	0.774	0.838	0.91	1.	1.																			
0.202	0.308	0.386	0.489	0.558	0.629	0.701	0.772	0.849	0.904	1.	1.																		
0.198	0.312	0.387	0.461	0.531	0.598	0.661	0.721	0.786	0.845	0.925	1.	1.																	
0.208	0.296	0.358	0.44	0.512	0.581	0.632	0.689	0.769	0.824	0.885	0.93	1.	1.																
0.184	0.269	0.352	0.413	0.476	0.546	0.606	0.671	0.734	0.782	0.821	0.867	0.922	1.	1.															
0.16	0.245	0.338	0.387	0.437	0.498	0.572	0.626	0.701	0.737	0.794	0.861	0.903	0.954	1.	1.														
0.175	0.259	0.33	0.387	0.443	0.498	0.551	0.599	0.643	0.691	0.752	0.805	0.851	0.893	0.946	1.	1.													
0.131	0.224	0.286	0.334	0.394	0.449	0.504	0.559	0.617	0.671	0.716	0.762	0.822	0.866	0.901	0.943	1.	1.												
0.133	0.209	0.271	0.332	0.389	0.449	0.51	0.551	0.607	0.655	0.709	0.754	0.791	0.837	0.878	0.914	0.949	1.	1.											
0.117	0.204	0.259	0.326	0.388	0.436	0.475	0.513	0.556	0.613	0.648	0.7	0.748	0.785	0.832	0.875	0.911	0.956	1.	1.										
0.124	0.198	0.259	0.309	0.365	0.416	0.456	0.5	0.543	0.596	0.631	0.662	0.714	0.754	0.801	0.838	0.873	0.914	0.955	1.	1.									
0.116	0.188	0.247	0.298	0.332	0.378	0.435	0.471	0.508	0.562	0.604	0.655	0.691	0.73	0.77	0.809	0.85	0.891	0.922	0.955	1.	1.								
0.089	0.17	0.222	0.269	0.319	0.361	0.39	0.448	0.495	0.532	0.573	0.611	0.657	0.709	0.753	0.779	0.821	0.872	0.904	0.93	0.959	1.	1.							
0.096	0.181	0.211	0.251	0.302	0.341	0.387	0.437	0.48	0.517	0.554	0.601	0.644	0.688	0.733	0.77	0.804	0.85	0.883	0.912	0.934	0.969	1.	1.						
0.107	0.17	0.216	0.263	0.302	0.329	0.38	0.422	0.459	0.506	0.549	0.596	0.635	0.669	0.704	0.738	0.771	0.811	0.835	0.874	0.908	0.94	0.969	1.	1.					
0.079	0.138	0.183	0.22	0.271	0.305	0.346	0.38	0.423	0.454	0.497	0.543	0.573	0.602	0.635	0.676	0.715	0.756	0.796	0.83	0.866	0.892	0.925	0.962	1.	1.				
0.086	0.159	0.203	0.248	0.283	0.312	0.353	0.383	0.419	0.455	0.497	0.534	0.566	0.598	0.63	0.663	0.703	0.746	0.781	0.804	0.853	0.886	0.911	0.936	0.964	1.	1.			
0.078	0.141	0.186	0.232	0.274	0.299	0.337	0.369	0.402	0.438	0.465	0.509	0.543	0.574	0.609	0.644	0.675	0.706	0.733	0.76	0.799	0.838	0.88	0.91	0.947	0.97	1.	1.		
0.076	0.145	0.182	0.229	0.27	0.303	0.339	0.383	0.416	0.444	0.468	0.507	0.54	0.583	0.603	0.633	0.672	0.694	0.726	0.758	0.795	0.818	0.851	0.881	0.909	0.943	0.969	1.	1.	
0.07	0.124	0.173	0.214	0.246	0.291	0.319	0.34	0.374	0.418	0.451	0.491	0.518	0.547	0.572	0.619	0.648	0.69	0.725	0.765	0.799	0.83	0.849	0.878	0.903	0.917	0.941	0.971	1.	1.

Table 1: For $1 \leq m \leq 30$, $n = 30$ and $1 \leq r \leq m$, the behavior of the probability that the set of r-Borda winners is equal to the set of actual Borda winners. The rows are indexed by m, and the columns are indexed by r.

1.																													
1.	1.																												
0.844	1.	1.																											
0.685	0.87	1.	1.																										
0.607	0.759	0.895	1.	1.																									
0.537	0.663	0.788	0.907	1.	1.																								
0.474	0.618	0.706	0.827	0.919	1.	1.																							
0.429	0.545	0.643	0.754	0.837	0.928	1.	1.																						
0.371	0.475	0.582	0.666	0.761	0.849	0.936	1.	1.																					
0.352	0.444	0.533	0.63	0.715	0.796	0.877	0.954	1.	1.																				
0.313	0.428	0.519	0.602	0.681	0.754	0.827	0.893	0.96	1.	1.																			
0.285	0.362	0.44	0.54	0.604	0.678	0.749	0.82	0.893	0.954	1.	1.																		
0.278	0.372	0.445	0.512	0.579	0.647	0.712	0.774	0.84	0.906	0.971	1.	1.																	
0.278	0.36	0.422	0.489	0.563	0.626	0.678	0.743	0.812	0.871	0.932	0.971	1.	1.																
0.252	0.318	0.395	0.462	0.522	0.592	0.655	0.716	0.773	0.82	0.863	0.913	0.967	1.	1.															
0.225	0.301	0.386	0.437	0.483	0.541	0.616	0.676	0.74	0.784	0.839	0.904	0.944	0.982	1.	1.														
0.239	0.301	0.37	0.418	0.476	0.529	0.582	0.631	0.681	0.73	0.785	0.842	0.891	0.931	0.976	1.	1.													
0.187	0.264	0.328	0.371	0.428	0.483	0.535	0.586	0.644	0.698	0.745	0.794	0.854	0.898	0.928	0.97	1.	1.												
0.196	0.257	0.309	0.365	0.423	0.48	0.542	0.584	0.635	0.684	0.736	0.779	0.822	0.875	0.909	0.946	0.979	1.	1.											
0.18	0.247	0.299	0.36	0.42	0.469	0.506	0.547	0.588	0.639	0.679	0.728	0.776	0.819	0.862	0.908	0.947	0.983	1.	1.										
0.192	0.25	0.303	0.355	0.406	0.446	0.485	0.529	0.576	0.627	0.669	0.699	0.75	0.797	0.839	0.88	0.914	0.954	0.984	1.	1.									
0.173	0.231	0.281	0.323	0.36	0.409	0.465	0.503	0.541	0.591	0.634	0.683	0.724	0.766	0.807	0.844	0.883	0.928	0.956	0.984	1.	1.								
0.151	0.211	0.259	0.302	0.347	0.385	0.423	0.475	0.522	0.558	0.601	0.645	0.687	0.745	0.782	0.813	0.857	0.908	0.934	0.965	0.987	1.	1.							
0.154	0.222	0.252	0.291	0.334	0.371	0.424	0.47	0.51	0.551	0.593	0.635	0.681	0.719	0.765	0.802	0.838	0.883	0.915	0.942	0.964	0.99	1.	1.						
0.154	0.212	0.247	0.293	0.325	0.354	0.4	0.446	0.48	0.531	0.573	0.617	0.66	0.692	0.728	0.76	0.797	0.837	0.869	0.902	0.938	0.966	0.989	1.	1.					
0.127	0.172	0.213	0.256	0.301	0.333	0.377	0.409	0.447	0.486	0.53	0.576	0.604	0.64	0.673	0.714	0.751	0.793	0.837	0.87	0.902	0.93	0.962	0.988	1.	1.				
0.135	0.195	0.233	0.271	0.305	0.338	0.38	0.408	0.45	0.484	0.523	0.557	0.593	0.623	0.656	0.692	0.735	0.774	0.81	0.838	0.886	0.917	0.94	0.967	0.989	1.	1.			
0.129	0.179	0.22	0.261	0.295	0.32	0.362	0.393	0.423	0.459	0.49	0.534	0.57	0.602	0.631	0.667	0.704	0.73	0.758	0.792	0.829	0.868	0.911	0.94	0.97	0.99	1.	1.		
0.132	0.184	0.219	0.259	0.298	0.335	0.371	0.411	0.443	0.472	0.499	0.539	0.579	0.616	0.638	0.668	0.7	0.727	0.765	0.796	0.832	0.858	0.89	0.918	0.946	0.972	0.99	1.	1.	
0.127	0.168	0.207	0.251	0.279	0.322	0.349	0.368	0.406	0.448	0.481	0.52	0.549	0.576	0.603	0.648	0.678	0.724	0.758	0.795	0.832	0.861	0.882	0.914	0.936	0.952	0.972	0.992	1.	1.

Table 2: For $1 \leq m \leq 30$, $n = 30$ and $1 \leq r \leq m$, the behavior of the probability of the set of r-Borda winners to be a subset of the set of m-Borda winners. The rows are indexed by m, and the columns are indexed by r.

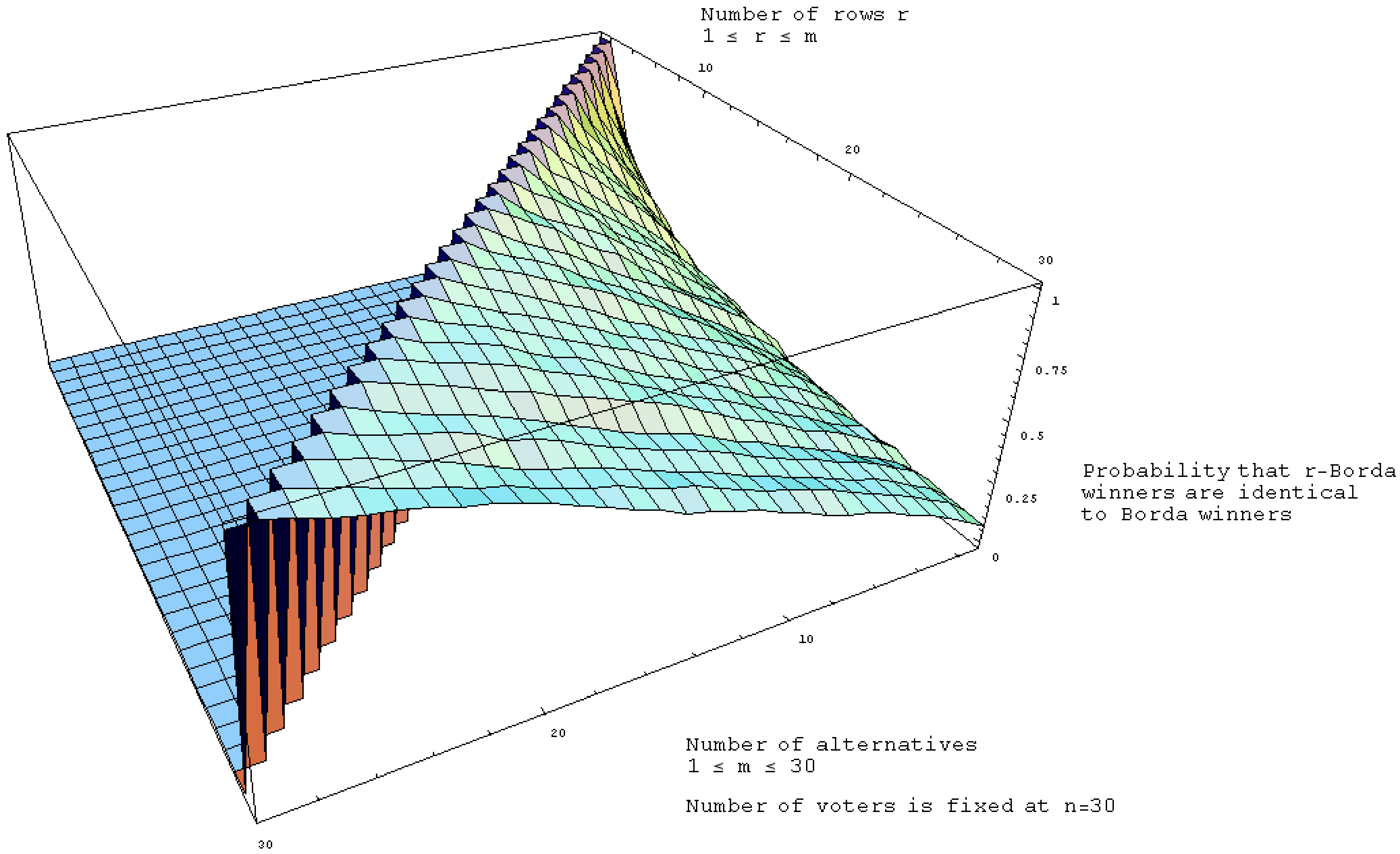

Number of rows r
1 ≤ r ≤ m
10
20
30
1
0.75
0.5
0.25
0
Probability that r-Borda
winners are identical
to Borda winners
10
20
30
Number of alternatives
1 ≤ m ≤ 30
Number of voters is fixed at n=30

Figure 1

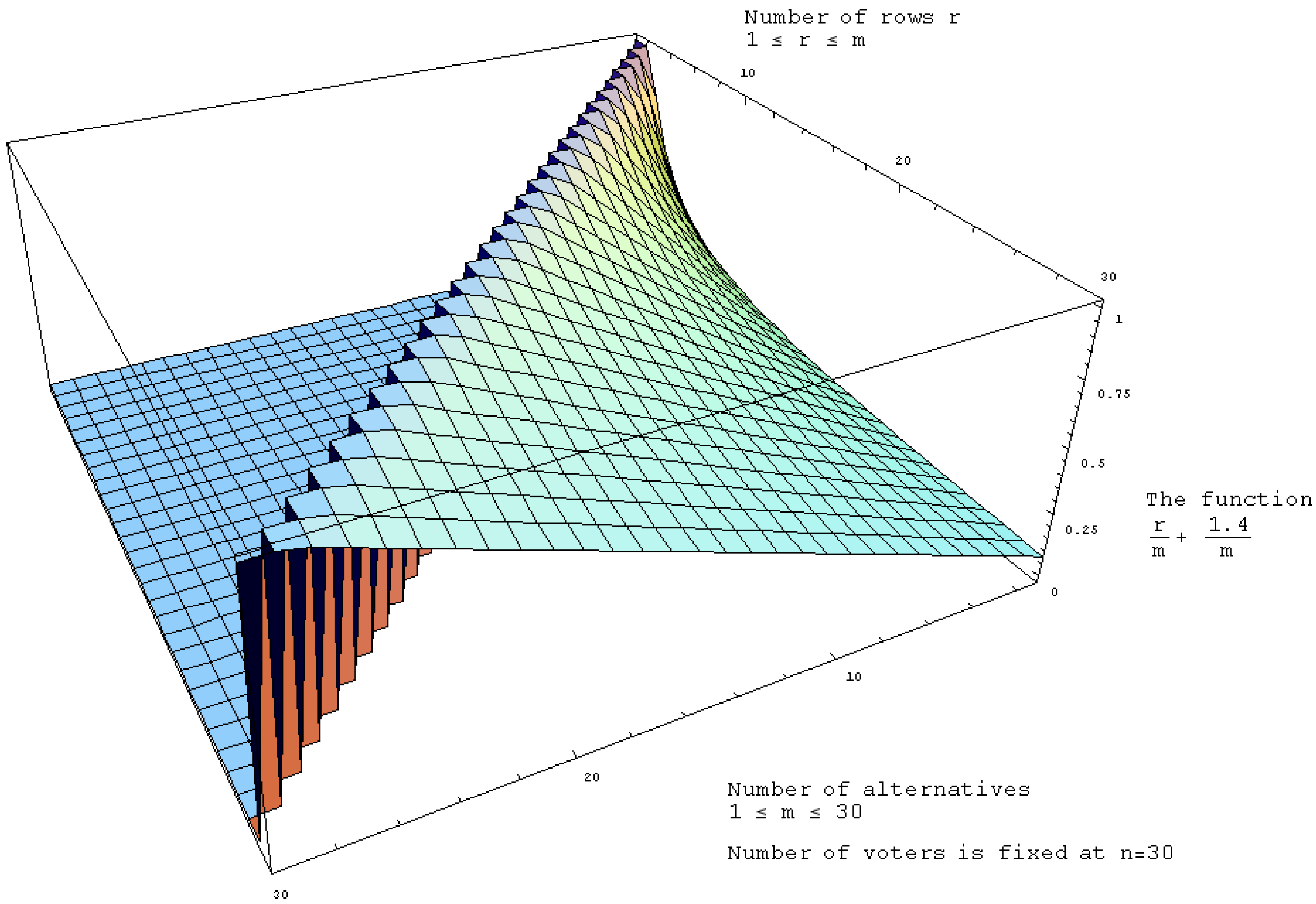


Figure 2

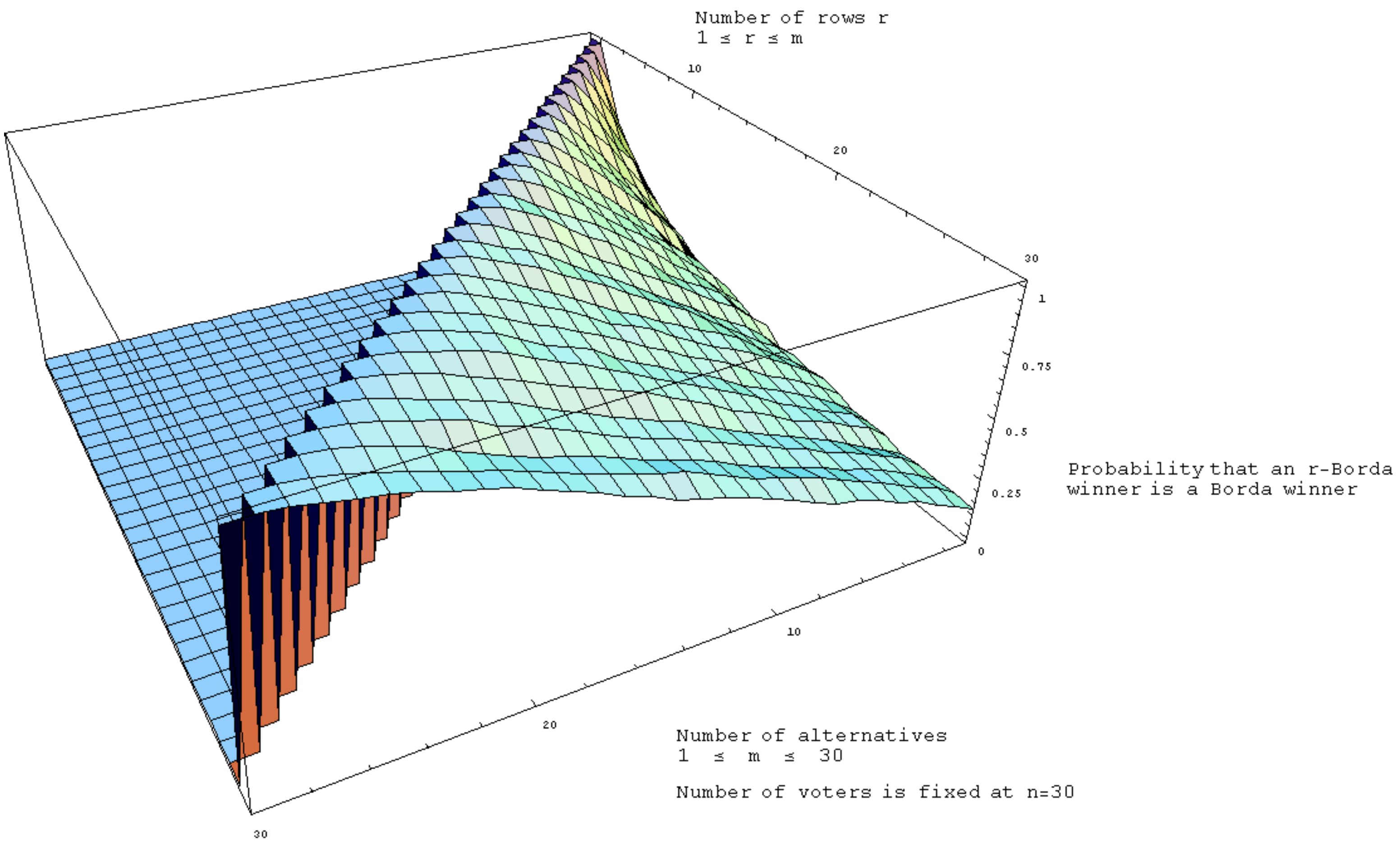
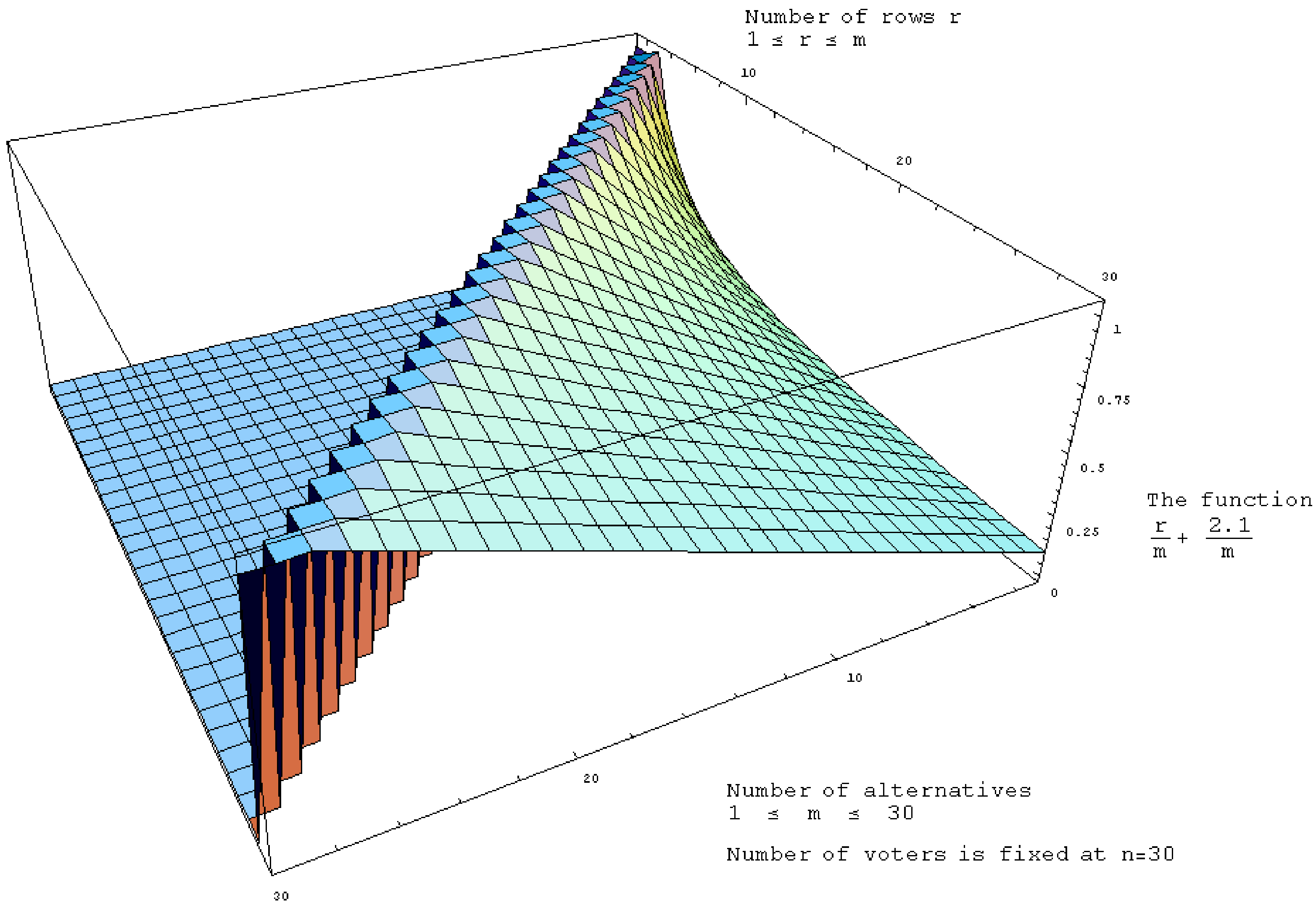

Number of rows r
1 ≤ r ≤ m
10
20
30
1
0.75
0.5
0.25
0
Probability that an r-Borda
winner is a Borda winner
10
20
30
Number of alternatives
1 ≤ m ≤ 30
Number of voters is fixed at n=30

Figure 3



Number of rows r
1 ≤ r ≤ m
10
20
30
1
0.75
0.5
0.25
0
The function
r/m + 2.1/m
10
20
30
Number of alternatives
1 ≤ m ≤ 30
Number of voters is fixed at n=30

Figure 4

APPENDIX

Given m and n, we compute the probability that $GCD(m!,n) = 1$, i.e. $m!$ and n are relatively prime. We assume that m is fixed and n is allowed to vary. As in the most cases with probabilities on integers, this probability is actually the *density*. Let S be a subset of the integers. The density of S (if it exists) is

$$\lim_{n \to \infty} \frac{S_n}{n}$$

where S_n is the number of elements in the set $S \cap \{1,2,..,n\}$. Thus, when we say that the probability of an integer to be even is ½, we mean that the density of even numbers is ½.

Proposition Suppose m is given. If $p_k \leq m < p_{k+1}$ where p_k denotes the k-th prime number, then

$$Pr\,[GCD(m!,n) = 1] = \prod_{i=1}^{k}(1 - \frac{1}{p_i})$$

Proof Let S be the set of positive integers relatively prime to $m!$. We calculate S_n, which is the number of these less than or equal to n. Since $p_k \leq m < p_{k+1}$, the prime factors of $m!$ are $p_1, p_2, ..., p_k$. Therefore, S_n is the number of positive integers less than or equal to n which are not divisible by any of these primes. This is a textbook example of an inclusion-exclusion problem in which the alternating sum of the inclusion-exclusion formula simplifies to

$$n\prod_{i=1}^{k}(1 - \frac{1}{p_i}).$$

Hence, the probability is as stated in the proposition.

The following table gives these probabilities for small values of m:

m	Probability that $GCD(m!,n) = 1$
2	$(1-\frac{1}{2}) = \frac{1}{2}$
3	$(1-\frac{1}{2})\,(1-\frac{1}{3}) = \frac{1}{6}$
4	$(1-\frac{1}{2})\,(1-\frac{1}{3}) = \frac{1}{6}$
5	$(1-\frac{1}{2})\,(1-\frac{1}{3})\,(1-\frac{1}{5}) = \frac{2}{15}$
6	$(1-\frac{1}{2})\,(1-\frac{1}{3})\,(1-\frac{1}{5}) = \frac{2}{15}$

7	$(1-\frac{1}{2})(1-\frac{1}{3})(1-\frac{1}{5})(1-\frac{1}{7}) = \frac{4}{35}$
8	$(1-\frac{1}{2})(1-\frac{1}{3})(1-\frac{1}{5})(1-\frac{1}{7}) = \frac{4}{35}$
9	$(1-\frac{1}{2})(1-\frac{1}{3})(1-\frac{1}{5})(1-\frac{1}{7}) = \frac{4}{35}$
10	$(1-\frac{1}{2})(1-\frac{1}{3})(1-\frac{1}{5})(1-\frac{1}{7}) = \frac{4}{35}$
11	$(1-\frac{1}{2})(1-\frac{1}{3})(1-\frac{1}{5})(1-\frac{1}{7})\ (1-\frac{1}{11})$ $= \frac{8}{77}$

These probabilities get smaller for increasing *m*, and, in fact, the limit is zero. To see this, note that

$$\prod_{i=1}^{\infty}(1-\frac{1}{p_i})^{-1} = \infty$$

since the left hand side is another way of writing the harmonic series. However, it is also the expression for the reciprocal of the limiting probability we are considering. Therefore, the probability that *m!* and *n* are relatively prime goes to zero as *m* increases.