

A Note on Approval Voting and Electing the Condorcet Loser

William V. Gehrlein

Department of Business Administration, University of Delaware
Newark, DE 19716, USA
wvg@udel.edu

Dominique Lepelley

CEMOI, Department of Economics, University of La Réunion
97715 Saint-Denis cedex 9, La Réunion, France
dominique.lepelley@univ-reunion.fr

Issoufa Moyouwou

ENS - Department of Mathematics - University of Yaounde
BP 47 Yaounde, Cameroon,
imoyouwou2@yahoo.fr

Abstract. Analytical representations are developed for the probability that Approval Voting (AV) elects the Condorcet Loser in three-alternative elections with large electorates. A comparison of AV is then made to Plurality Rule (PR) to show that AV is much less susceptible to the risk of electing the Condorcet loser than PR. All calculations in this analysis are based on IAC-like assumptions.

Keywords : Voting paradoxes, Voting Rules, Borda's Paradox, Plurality Rule, Probability

1.Introduction

Plurality Rule (PR) is the most commonly used election procedure, where each voter casts one vote for their most preferred candidate, and the candidate who receives the most total votes is declared to be the winner. Approval Voting (AV) is often proposed as a good alternative to PR in current research that is related to the evaluation of voting rules. Each voter casts a vote for as many of the candidates that they approve of with AV, and the winner is then selected as the candidate who receives the most votes of approval from the electorate. The current study focuses on a comparison of some properties of PR and AV.

Brams and Fishburn (1983) show that AV uniquely displays many good properties among voting rules when attention is restricted to scenarios in which all voters have dichotomous preferences on candidates. A voter with dichotomous preferences partitions the set of candidates into two equivalence classes of acceptable candidates and unacceptable candidates. This voter is then indifferent between any two candidates within the same equivalence class, and each candidate in the acceptable class is preferred to every candidate in the unacceptable class. However, both PR and AV can potentially suffer from serious negative occurrences without the requirement of dichotomous preferences.

To describe one of these particular negative outcomes, let $A \succ B$ denote that a voter prefers Candidate A to B for a three-candidate election on $\{A, B, C\}$. There are six possible complete preference rankings that each of n voters might have on the candidates, as shown in Figure 1.

- n_1 : number of voters with preference ranking ABC
- n_2 : number of voters with preference ranking ACB
- n_3 : number of voters with preference ranking BAC
- n_4 : number of voters with preference ranking BCA
- n_5 : number of voters with preference ranking CAB
- n_6 : number of voters with preference ranking CBA

Figure 1. The six possible complete preference rankings on three candidates.

For example, there are n_1 voters with preferences for $A \succ B$ and $B \succ C$. Intransitive voter preferences are not considered, so it follows that $A \succ C$. A specific *voting situation* is defined in terms of a particular set of the six possible n_i terms from Figure 1, with $n = \sum_{i=1}^6 n_i$.

Let AMB denote the outcome that Candidate A beats B by majority rule on that pair of candidates in a voting situation, such that more voters rank A over B than the converse. Given the definitions in Figure 1, this happens when $n_1 + n_2 + n_5 > n_3 + n_4 + n_6$. Candidate C is the *Condorcet Loser (CL)* in a voting situation if both AMC and BMC , which would make it a terrible choice for selections as the winner. When the election of a CL does result, it is referred to an occurrence of Borda's Paradox, since it was shown in Borda (1784) that this outcome can occur with PR. However, it has also been shown that this outcome can occur with AV [see e.g. Felsenthal and Maoz (1988)].

The primary focus of the current study is to determine whether or not AV performs better than PR with respect to avoiding Borda's Paradox in three-candidate elections, and to measure the

degree of any possible superiority of AV over PR. There has been only a limited amount of research conducted on this particular topic. Lepelley (1993) showed that AV is clearly superior to PR at avoiding Borda's Paradox when attention is restricted to the special case in which voters have single-peaked preferences. However, no clear results have been observed for more general assumptions regarding voters' preferences with either simulation based studies [Nurmi and Uusi-Heikkilä (1985), Plassmann and Tideman (2014)] or with analytical computations [Gehrlein and Fishburn (1978), Lepelley (1993), Gehrlein and Lepelley (1998)].

We consider a series of different models for obtaining representations for the probability that the CL is elected with AV in three-candidate elections. This is done in a manner that allows for a comparison of the performances of AV and PR in each scenario, as different frameworks are considered with the models to describe how voters' preferences are used to establish the set of acceptable candidates for AV voting. Some results are also obtained during this analysis that apply to the probability that Negative Plurality Rule (NPR) will elect the CL. Here, NPR is the voting rule that is also known as Anti-Plurality Rule, where each voter casts a vote against one candidate, and the candidate who receives the fewest negative votes is declared the winner.

A common assumption that is used in the development of our representations for the probability of observing election outcomes is the *Impartial Anonymous Culture Condition* (IAC), that applies to scenarios like those described in the development of Figure 1. In particular, IAC assumes that all possible voting situations with $n = \sum_{i=1}^6 n_i$ are equally likely to be observed for a given n . Each of the models that is presented in the current study is based on an IAC-like assumption, such that all feasible "voting situations" that are permissible under the specified assumptions of each of the respective models are assumed to be equally likely to be observed. We also assume throughout that the limiting case for voters applies with $n \rightarrow \infty$.

Notice that the IAC condition is to be contrasted with the *Impartial Culture Condition* (IC), on which most of the previous studies are based [Nurmi and Uusi-Heikkilä (1985), Gehrlein and Fishburn (1978), Gehrlein and Lepelley (1998)]¹. Under IC, the voters are assumed to vote in a completely independent way and the probability that any given voting situation will be observed is then given by a multinomial model.

2. Representations for the Probability that Borda's Paradox will be Observed

Our objective is to maintain a primary focus on the results that are obtained regarding the probability that Borda's Paradox will be observed with AV and PR from each of the models that are considered. A brief overview of the methods that are used to obtain the probability representations and results that follow in this note is therefore presented for the interested

¹ Lepelley (1993) makes use of the IAC model but only considers the case with single-peaked preferences. For their part, Plassmann and Tideman (2014) obtain their simulated frequencies by using a parametrized spatial model to generate a distribution of voting situations that corresponds to the distribution of observed voting situations in actual elections.

reader in Appendix A, and the technical details of how these procedures work will not be developed any further in the discussion that follows.

2.1 The first approach

The most basic model that we consider, denoted as Model I, starts with the scenario that led to the six possible voter preference rankings in Figure 1, and then adds the additional factor with AV that each voter might choose to vote only for their most preferred candidate, or they might choose to vote for the two more preferred candidates in their ranking. A voter would have no impact on the election outcome if they voted for all three candidates, so this option is ignored. If a voter chooses to vote for two candidates in this scenario, that does not imply that the voter is indifferent between the two candidates, it simply means that both are considered to be acceptable to the voter. The proportion of voters who choose to vote for two candidates when AV is used certainly could vary according to the particular ranking that is being considered. But, assume instead for this most basic scenario that this proportion, which we denote by λ , is identical for each of the six possible rankings.

It is then possible to obtain a representation for the probability that Borda's Paradox is observed by using the standard IAC framework from Figure 1. By using the admittedly strong assumption of Model I along with the assumption of a large electorate, it is easy to verify that AV will elect the same candidate as the *Weighted Scoring Rule*, denoted as *Rule*(λ). By standard definition, *Rule*(λ) takes the preference ranking for each voter and then gives one point to the most preferred candidate, λ points to the middle ranked candidate and zero points to the least preferred candidate. The winner is determined as the candidate that receives the greatest accumulated number of points from the electorate.

If we assume in addition that λ is uniformly distributed on the unit interval $[0,1]$, we can then compute the "expected value" of the probability that AV elects the CL by integrating the representation that gives the probability of electing the CL with *Rule*(λ) on λ over $[0,1]$. Similar representations have been developed in Cervone et al. (2005) and in Gehrlein et al (2013) for the limiting probability that *Rule*(λ) will elect the Condorcet Winner (CW), which is defined in the obvious fashion. The primary merit of using this "expected value" approach is that it permits a direct comparison of AV and PR in the standard IAC framework. Lepelley (1993) and Lepelley (1995) used of this "expected value" approach to study the behavior of AV, but only considered the restricted case in which voters have single-peaked preferences.

Let $CL_I(\lambda, \infty)$ denote the probability that *Rule*(λ) elects the CL for the limiting case $n \rightarrow \infty$ in this first framework. Then, with the IAC format

$$CL_I(\lambda, \infty) = \frac{3VS_{A,CL}(\lambda)}{VS_{CL}} \quad (1)$$

where VS_{CL} is the volume of the domain of all voting situations for which a CL exists and $VS_{A,CL}(\lambda)$ is the volume of the domain of all voting situations at which A is both the winner by *Rule*(λ) and the CL.

To obtain these volumes, let $x_j = n_j/n$ for the n_j terms in Figure 1. When n tends to infinity, the domain of all voting situations is the unit simplex that is defined by

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 1 \text{ with } x_j \geq 0, \text{ for } j = 1, 2, \dots, 6. \quad (2)$$

It is well known that $VS_{CL} = \frac{1}{128}$ in the context of the development in Appendix A. To determine $VS_{A,CL}(\lambda)$, we consider all voting situations for which

(i) Candidate A is the CL:
 $x_1 + x_2 - x_3 - x_4 + x_5 - x_6 < 0$ and $x_1 + x_2 + x_3 - x_4 - x_5 - x_6 < 0.$ (3)

(ii) Candidate A is selected by $Rule(\lambda)$:
 $(1 - \lambda)x_1 + x_2 - (1 - \lambda)x_3 - x_4 + \lambda x_5 - \lambda x_6 > 0,$ and (4)
 $x_1 + (1 - \lambda)x_2 + \lambda x_3 - \lambda x_4 - (1 - \lambda)x_5 - x_6 > 0.$ (5)

From these four inequalities in (2) through (5) and the non-negativity restriction on each x_j we obtain a representation for $VS_{A,CL}(\lambda)$, which is used in (1) to obtain:

$$CL_I(\lambda, \infty) = \begin{cases} \frac{(1-2\lambda)^3(12-9\lambda-2\lambda^2)}{405(1-\lambda)^3} & \text{if } 0 \leq \lambda \leq \frac{1}{2} \\ \frac{(1-2\lambda)^3(2-53\lambda+331\lambda^2-88\lambda^3+12\lambda^4)}{1620\lambda^3(1+\lambda)(1-3\lambda)} & \text{if } \frac{1}{2} \leq \lambda \leq 1. \end{cases} \quad (6)$$

This representation replicates results from Diss and Gehrlein (2012) that used very different procedures to obtain the volumes.

The ‘‘expected value’’ $ECL_I(AV, \infty)$ of the probability that AV elects the CL is therefore obtained from (6) as

$$\begin{aligned} ECL_I(AV, \infty) &= \int_0^{1/2} \frac{(1-2\lambda)^3(12-9\lambda-2\lambda^2)}{405(1-\lambda)^3} d\lambda + \int_{1/2}^1 \frac{(1-2\lambda)^3(2-53\lambda+331\lambda^2-88\lambda^3+12\lambda^4)}{1620\lambda^3(1+\lambda)(1-3\lambda)} d\lambda \quad (7) \\ &= \frac{589}{2430} - \frac{5741}{1620} \ln(2) + \frac{81}{40} \ln(3) \approx 0.0107. \end{aligned}$$

The probability that PR exhibits Borda’s Paradox is obtained from $CL_I(0, \infty)$ in (6) to obtain $4/135 = 0.0296$. Similarly, the probability that NPR exhibits Borda’s Paradox is obtained from $CL_I(1, \infty)$ and is equal to $17/540 = 0.0315$. If we assume that $1/3$ of the voters choose to vote for their two more-preferred candidates, we obtain $CL_I(1/3, \infty) = 79/29160 \approx 0.0027$; so the probability of observing Borda’s Paradox decreases rapidly as λ increases from 0 to $1/2$ (or decreases from 1 to $1/2$) and AV significantly outperforms PR (and NPR) with Model I since the risk of electing the CL is divided by three when AV replaces PR (or NPR).

2.2 The second approach

Model II takes into consideration the fact that the proportion of voters who vote for their two more preferred candidates with AV is not necessarily the same for each of the possible preference rankings. There are therefore 12 types of voter preferences for Model II, as shown

in Figure 2, where we see for example that the set of voters with the preference ranking ABC from Figure 1 is split in two subsets: those who vote only for Candidate A under AV, and those who vote for both A and B .

- n_1 : number of voters with ranking ABC and approving A only
- n_2 : number of voters with ranking ABC and approving A and B
- n_3 : number of voters with ranking ACB and approving A only
- n_4 : number of voters with ranking ACB and approving A and C
- n_5 : number of voters with ranking BAC and approving B only
- n_6 : number of voters with ranking BAC and approving B and A
- n_7 : number of voters with ranking BCA and approving B only
- n_8 : number of voters with ranking BCA and approving B and C
- n_9 : number of voters with ranking CAB and approving C only
- n_{10} : number of voters with ranking CAB and approving C and A
- n_{11} : number of voters with ranking CBA and approving C only
- n_{12} : number of voters with ranking CBA and approving C and B .

Figure 2. Twelve possible voter preference rankings on three candidates for Model II.

Voting situations are now defined in this scenario by a combination of n_i terms from Figure 2 with $\sum_{i=1}^{12} n_i = n$, and each of these voting situations is assumed to be equally likely to be observed.

In the framework of Model II and the definitions in Figure 2, Candidate A is the CL iff

$$n_1 + n_2 + n_3 + n_4 + n_9 + n_{10} < n/2, \text{ and} \quad (8)$$

$$n_1 + n_2 + n_3 + n_4 + n_5 + n_6 < n/2.$$

The number of voting situations with a CL can be represented by a quasi-polynomial of degree 11 in n and the coefficient of the n^{11} term is a constant that can be shown to be equal to:

$$\frac{79}{10218700800}. \quad (9)$$

This coefficient can be interpreted in Appendix A as the volume of the domain of all voting situations at which a CL exists in the Model II framework when $n \rightarrow \infty$.

Candidate A is the AV winner with the definitions from Figure 2 iff:

$$n_1 + n_3 + n_4 + n_{10} > n_5 + n_7 + n_8 + n_{12}, \text{ and} \quad (10)$$

$$n_1 + n_2 + n_3 + n_6 > n_8 + n_9 + n_{11} + n_{12}.$$

The number of voting situations with A as the CL and the AV winner is also a quasi-polynomial of degree 11 in n and the coefficient of n^{11} is a constant that can be shown to be equal to:

$$\frac{179}{840853094400}. \quad (11)$$

The limiting probability that AV elects the CL follows from (9) and (11):

$$CL_{II}(AV, \infty) = \frac{\frac{179}{840853094400}}{\frac{79}{10218700800}} = \frac{1253}{45504} \approx 0.0275. \quad (12)$$

In this same framework, A is the PR winner iff:

$$n_1 + n_2 + n_3 + n_4 > n_5 + n_6 + n_7 + n_8, \text{ and} \quad (13)$$

$$n_1 + n_2 + n_3 + n_4 > n_9 + n_{10} + n_{11} + n_{12}$$

The volume associated to the domain of voting situations where Candidate A is both the CL and the PR winner is given as:

$$\frac{5111}{20113468784640}. \quad (14)$$

Based on (9) and (14),

$$CL_{II}(PR, \infty) = \frac{\frac{5111}{20113468784640}}{\frac{79}{10218700800}} = \frac{51110}{1554957} \approx 0.0329. \quad (15)$$

The representation that is associated with NPR is obtained in the same fashion with:

$$CL_{II}(NPR, \infty) = \frac{\frac{277775}{1029809601773568}}{\frac{79}{10218700800}} = \frac{6944375}{199034496} \approx 0.0349. \quad (16)$$

It turns out that in this second framework, AV performs slightly better than both PR and NPR on its ability to avoid Borda's Paradox.

It is possible to refine Model II to obtain a better assessment of how $CL_{II}(AV, \infty)$ responds to the number of voters who approve of their two more preferred candidates. With the above notation from Figure 2, the number of voters who approve of their two more preferred candidates is $k = n_2 + n_4 + n_6 + n_8 + n_{10} + n_{12}$. It is then possible to obtain a representation for the probability that AV elects the CL as a function of n and k . Then k is replaced with βn , so that β is the proportion of voters who approve of their two more preferred candidates. By considering the coefficients of the terms of highest degree in n in the resulting representation, we then obtain a representation for $CL_{II}(AV(\beta), \infty)$ in the limiting case as $n \rightarrow \infty$ that is a function of β . This resulting limiting representation is the following:

$$\begin{aligned} CL_{II}(AV(\beta), \infty) &= -\frac{1037\beta^5 - 1290\beta^4 - 855\beta^3 + 2160\beta\alpha^2 - 1260\beta + 252}{135(42\beta^5 - 274\beta^4 + 603\beta^3 - 624\beta^2 + 315\beta - 63)}, \text{ for } 0 \leq \beta \leq 1/2 \quad (17) \\ &= \frac{\left[\frac{4125184\beta^{10} - 27617280\beta^9 + 82103040\beta^8 - 142295040\beta^7 + 158659200\beta^6 - 118468224\beta^5 + 59754240\beta^4 - 20021760\beta^3 + 4245120\beta^2 - 510660\beta + 26199}{69120(1-\beta)^5(42\beta^5 + 64\beta^4 - 73\beta\alpha^3 + 39\beta^2 - 10\beta + 1)} \right]}{60(42\alpha^5 + 64\alpha^4 - 73\alpha^3 + 39\alpha^2 - 10\alpha + 1)}, \text{ for } 1/2 \leq \beta \leq 3/4 \\ &= \frac{60\beta^5 + 380\beta^4 - 850\beta^3 + 870\beta^2 - 415\beta + 74}{60(42\alpha^5 + 64\alpha^4 - 73\alpha^3 + 39\alpha^2 - 10\alpha + 1)} \text{ for } 3/4 \leq \beta \leq 1. \end{aligned}$$

Note that $CL_{II}(AV(0), \infty) = CL_{II}(PR, \infty) = 4/135$ and $CL_{II}(AV(1), \infty) = CL_{II}(NPR, \infty) = 17/540$, in accordance with Lepelley (1993).

The representation in (17) is used to obtain some calculated values of $CL_{II}(AV(\beta), \infty)$ for various β that are given in Table 1. The results from Table 1 indicate that the value of $CL_{II}(AV(.5), \infty)$ is very similar to the value of $CL_{II}(AV, \infty)$ that was computed above in (12).

β	$CL_{II}(AV(\beta), \infty)$	β	$CL_{II}(AV(\beta), \infty)$
0	0.0296	0.6	0.0289
0.1	0.0292	0.7	0.0308
0.2	0.0264	0.75	0.0314
0.3	0.0252	0.8	0.0315
0.4	0.0252	0.9	0.0315
0.5	0.0264	1.0	0.0315

Table 1. Computed Values of $CL_{II}(AV(\beta), \infty)$.

We also consider a further refinement of Model II that assumes that voters' preferences are single-peaked, which we refer to as Model IISP. The concept of single-peaked preferences is based on the premise that all voters have agreement on a common ranking of candidates from left to right along some common dimension, and that the voters then form their preferences based on this common ranking. Suppose that A is the leftist candidate, B the centrist candidate and C the rightist candidate in this common ranking. Given the notion of single-peakedness and this common ranking from left to right, Candidate B can never be ranked as least preferred by any voter. The possible preference rankings in this scenario follow from Figure 2, and they are listed in Figure 3 as follows:

- n_1 : number of voters with ranking ABC and approving A only
- n_2 : number of voters with ranking ABC and approving A and B
- n_3 : number of voters with ranking BAC and approving B only
- n_4 : number of voters with ranking BAC and approving B and A
- n_5 : number of voters with ranking BCA and approving B only
- n_6 : number of voters with ranking BCA and approving B and C
- n_7 : number of voters with ranking CBA and approving C only
- n_8 : number of voters with ranking CBA and approving C and B

Figure 3. Eight possible voter preference rankings on three candidates for Model IISP.

Voting situations with Model IISP are defined by a combination of n_i terms from Figure 2 with $\sum_{i=1}^8 n_i = n$, and each of these voting situations is assumed to be equally likely to be observed. Given the candidate preference rankings in Figure 3, Candidate B cannot be the CL. Proceeding as above, we obtain a representation for $CL_{IISP}(AV, \infty)$:

$$CL_{IISP}(AV, \infty) = \frac{\frac{1}{5806080}}{\frac{1}{10080}} = \frac{1}{576} \approx 0.0017. \quad (18)$$

Applying the same approach to PR, we obtain:

$$CL_{II SP}(PR, \infty) = \frac{\frac{37}{14696640}}{\frac{1}{10080}} = \frac{37}{1458} \approx 0.0254. \quad (19)$$

The results in (18) and (19) confirm that the restriction of single-peakedness significantly reduces the probability of electing the CL with AV and that the same conclusion does not hold for PR. The improved performance of AV under SP is directly connected to the fact that NPR can never elect the CL when preferences are single-peaked [Lepelley (1989)].

2.3 A third approach with voter indifference allowed.

Model III considers a framework that directly allows the possibility of actual indifference between two candidates in a voter's preferences. By introducing such indifference in a three-candidate election, we are effectively requiring that voter to have dichotomous preferences; which we already know from the results of Brams and Fishburn (1983) that were mentioned above will generally have a very positive impact on the overall performance of AV. The set of possible voter preferences on candidates are shown in Figure 4 for this scenario:

- n_1 : number of voters with preference ranking ABC
- n_2 : number of voters with preference ranking ACB
- n_3 : number of voters with preference ranking BAC
- n_4 : number of voters with preference ranking BCA
- n_5 : number of voters with preference ranking CAB
- n_6 : number of voters with preference ranking CBA
- n_7 : number of voters with preference ranking $(AB)C$
- n_8 : number of voters with preference ranking $(AC)B$
- n_9 : number of voters with preference ranking $(BC)A$
- n_{10} : number of voters with preference ranking $A(BC)$
- n_{11} : number of voters with preference ranking $B(AC)$
- n_{12} : number of voters with preference ranking $C(AB)$

Figure 4. Twelve possible voter preference rankings on three candidates for Model III.

The notation (AB) in Figure 4 indicates that a voter is indifferent between the pair of candidates A and B . This model has been recently used by Gehrlein and Lepelley (2014) to study the probability that various voting rules, including AV, will elect the *Condorcet Winner*. We ignore the possibility of complete voter indifference between A , B and C , since a voter who casts a ballot for all three candidates with AV would have absolutely no impact on the election outcome.

In this framework, Candidate A is the CL iff:

$$-n_1 - n_2 + n_3 + n_4 - n_5 + n_6 - n_8 + n_9 - n_{10} + n_{11} > 0, \text{ and} \quad (20)$$

$$-n_1 - n_2 - n_3 + n_4 + n_5 + n_6 - n_7 + n_9 - n_{10} + n_{12} > 0.$$

Candidate A is the AV winner iff:

$$n_1 + n_2 - n_3 - n_4 + n_8 - n_9 + n_{10} - n_{11} > 0, \text{ and} \quad (21)$$

$$n_1 + n_2 - n_5 - n_6 + n_7 - n_9 + n_{10} - n_{12} > 0.$$

Model III applies directly to the implementation of AV, but some accommodation must be made for applying both PR and NPR when indifference is allowed on a pair of candidates. We follow here the notion of *Extended Scoring Rules* from Diss *et al.* (2010). When the two more preferred candidates are viewed with indifference in an individual preference, PR gives each of these candidates the average $\frac{1}{2}$ vote. Similarly, under NPR, with indifference between the two less preferred candidates, the top ranked candidate receives one vote and the two other candidates each receive $\frac{1}{2}$ vote.

Candidate A is the PR winner iff:

$$2n_1 + 2n_2 - 2n_3 - 2n_4 + n_8 - n_9 + 2n_{10} - 2n_{11} > 0, \text{ and} \quad (22)$$

$$2n_1 + 2n_2 - 2n_5 - 2n_6 + n_7 - n_9 + 2n_{10} - 2n_{12} > 0.$$

Candidate A is the NPR winner iff:

$$2n_2 - 2n_4 + 2n_5 - 2n_6 + 2n_8 - 2n_9 + n_{10} - n_{11} > 0, \text{ and} \quad (23)$$

$$2n_1 + 2n_3 - 2n_4 - 2n_6 + 2n_7 - 2n_9 + n_{10} - n_{12} > 0.$$

Our basic assumption is that all the possible voting situations with $\sum_{i=1}^{12} n_i = n$ from Figure 4 are equally likely to occur. The resulting limiting conditional probabilities of electing the CL with AV, PR and NPR, given that a CL exists are given respectively in (24), (25) and (26):

$$CL_{III}(AV, \infty) = \frac{1192}{79389} \approx 0.0150 \quad (24)$$

$$CL_{III}(PR, \infty) = \frac{962332829}{38583054000} \approx 0.0249 \quad (25)$$

$$CL_{III}(NPR, \infty) = \frac{20150629}{771661080} \approx 0.0261. \quad (26)$$

We conclude once again that the election of the CL is less probable with AV than with PR (and NPR). But, this conclusion is reached by assuming that, on average, one-half of the voters in the electorate have dichotomous preferences, which is disputable.

A refinement of Model III is now considered, mirroring the development that led to $CL_{II}(AV(\beta), \infty)$ above. Let $h = n_7 + n_8 + n_9 + n_{10} + n_{11} + n_{12}$ be the number of voters who are indifferent between two candidates. We wish to obtain limiting representations for $CL_{III}(AV(\alpha), \infty)$; the limiting probability of electing the CL, given that such a candidate exists, with AV as a function of the proportion $\alpha = h/n$ of voters that have indifference in their preference rankings, and hence dichotomous preferences. The resulting representation is:

$$CL_{III}(AV(\alpha), \infty) = -\frac{3(4389\alpha^5 - 5080\alpha^4 - 7540\alpha^3 + 15360\alpha^2 - 8960\alpha + 1792)}{5(33073\alpha^5 - 176904\alpha^4 + 361260\alpha^3 - 362880\alpha^2 + 181440\alpha - 36288)}, \quad (27)$$

for $0 \leq \alpha \leq 1/3$

$$= -\frac{[14527269\alpha^{10} - 63397080\alpha^9 + 122211180\alpha^8 - 137324160\alpha^7 + 99701280\alpha^6 - 48843648\alpha^5 + 16329600\alpha^4 - 3680640\alpha^3 + 537840\alpha^2 - 46080\alpha + 1760]}{135\alpha^5(33073\alpha^5 - 176904\alpha^4 + 361260\alpha^3 - 362880\alpha^2 + 181440\alpha - 36288)}, \text{ for } 1/3 \leq \alpha \leq 1/2$$

$$\begin{aligned}
&= - \frac{\left[\begin{array}{l} 5569317\alpha^{10} - 30219480\alpha^9 + 70785900\alpha^8 - 92534400\alpha^7 + 72122400\alpha^6 \\ -32006016\alpha^5 + 5443200\alpha^4 + 1814400\alpha^3 - 1218240\alpha^2 + 264960\alpha - 21568 \end{array} \right]}{\left[\begin{array}{l} 27(890357\alpha^{10} - 5277480\alpha^9 + 13729500\alpha^8 - 20707200\alpha^7 + 20180160\alpha^6) \\ -13390272\alpha^5 + 6158880\alpha^4 - 1929600\alpha^3 + 388800\alpha^2 - 45480\alpha + 2343 \end{array} \right]}, \text{ for } 1/2 \leq \alpha \leq 2/3 \\
&= \frac{2880(\alpha-1)^7(\alpha^3-7\alpha^2+5\alpha-1)}{\left[\begin{array}{l} 822064\alpha^{10} - 6532320\alpha^9 + 22880880\alpha^8 - 46477440\alpha^7 + 60651360\alpha^6 \\ -53234496\alpha^5 + 31943520\alpha^4 - 13000320\alpha^3 + 3447360\alpha^2 - 538200\alpha + 37593 \end{array} \right]}, \text{ for } 2/3 \leq \alpha \leq 3/4 \\
&= - \frac{20(\alpha-1)^2(\alpha^3-7\alpha^2+5\alpha-1)}{1573\alpha^5 - 1385\alpha^4 + 2770\alpha^3 - 2450\alpha^2 + 985\alpha - 149}, \text{ for } 3/4 \leq \alpha \leq 1.
\end{aligned}$$

The corresponding representations for $CL_{III}(\text{PR}(\alpha), \infty)$ and $CL_{III}(\text{NPR}(\alpha), \infty)$ are rather complex and are given in Appendix B. Computed values of $CL_{III}(\text{AV}(\alpha), \infty)$, $CL_{III}(\text{PR}(\alpha), \infty)$ and $CL_{III}(\text{NPR}(\alpha), \infty)$ are displayed in Table 2 for each $\alpha = 0(.1)1$.

α	$CL_{III}(\text{AV}(\alpha), \infty)$	$CL_{III}(\text{PR}(\alpha), \infty)$	$CL_{III}(\text{NPR}(\alpha), \infty)$
0	.02963	.02963	0.0315
.1	.02916	.02931	0.0314
.2	.02760	.02838	0.0313
.3	.02476	.02710	0.0304
.4	.02067	.02605	0.0295
.5	.01547	.02541	0.0274
.6	.00962	.02404	0.0240
.7	.00476	.02178	0.0188
.8	.00177	.02044	0.0137
.9	.00036	.02029	0.0103
1.0	0	.02037	0.0093

Table 2. Computed values of $CL_{III}(\text{VR}(\alpha), \infty)$ for $\text{VR} \in \{\text{AV}, \text{PR}, \text{NPR}\}$.

For each of the three voting rules, the probability of electing the CL decreases as α increases. But, $CL_{III}(\text{AV}(\alpha), \infty)$ tends to zero when α tends to one, whereas $CL_{III}(\text{PR}(\alpha), \infty)$ always remains higher than 2% and $CL_{III}(\text{NPR}(\alpha), \infty)$ always remains higher than 0.9%. The replacement of PR or NPR with AV significantly reduces the risk of electing the CL, particularly when the proportion of voters with dichotomous preferences is high.

2.4 Election of a Strong Condorcet Loser

The notion of Strong, or Absolute, Condorcet Loser, denoted by SCL, is introduced in Lepelley (1989) [see also Felsenthal (2012) or Plassmann and Tideman (2014)] where a SCL is defined as a candidate who is ranked as least preferred by more than 50% of the voters. A SCL is obviously a CL but the converse is not true. We compute the conditional probability of having the SCL elected under AV and PR, given that such a candidate exists, by considering successively each of the three above approaches.

In Model I, Candidate A is the SCL iff

$$n_4 + n_6 > n/2. \quad (28)$$

With Model II, Candidate A is the SCL iff

$$n_7 + n_8 + n_{11} + n_{12} > n/2. \quad (29)$$

With Model III that allows indifference, there is no unique way to define Candidate A as the SCL. We choose the interpretation in which a weight of one is given whenever A is the only candidate ranked as least preferred, and with a weight of 1/2 whenever A is tied with another candidate as one of the two less preferred candidates. Candidate A is then the SCL iff:

$$n_4 + n_6 + n_9 + \frac{1}{2}(n_{11} + n_{12}) > n/2. \quad (30)$$

We then obtain representations for the conditional probability $SCL_x(\text{VR}, \infty)$ that voting rule VR elects the SCL, given that such a candidate exists, when Model x applies:

$$SCL_I(\text{AV}, \infty) = \frac{40}{81} \ln(2) - \frac{247}{729} \approx 0.00348 \text{ and } SCL_I(\text{PR}, \infty) = \frac{2}{81} \approx 0.02469 \quad (31)$$

$$SCL_{II}(\text{AV}, \infty) = \frac{1}{1044} \approx 0.00096 \text{ and } SCL_{II}(\text{PR}, \infty) = \frac{10954}{570807} \approx 0.01919 \quad (32)$$

$$SCL_{III}(\text{AV}, \infty) = \frac{34}{16767} \approx 0.00203 \text{ and } SCL_{III}(\text{PR}, \infty) = \frac{24955677137}{3520265184000} \approx 0.00709. \quad (33)$$

The results of (31) through (33) indicate that replacing PR by AV divides the risk of electing the SCL by a factor that ranges from 3.5 to 20, according to the approach that is used. However, NPR never elects the SCL when such a candidate exists [see Lepelley (1989)].

3. Conclusion

Over the range of scenarios that have been considered in this study, AV consistently has a reduced risk of electing the CL when compared to PR. In a number of the cases that have been considered, this risk reduction is significant. AV also has a reduced risk of electing the CL compared to NPR in most of the cases that were considered. All these results are based on IAC-like assumptions.

It is worth noticing that these results contrast with those obtained by Gehrlein and Lepelley (1998) who show that, under a model derived from the Impartial Culture (IC) condition, AV, PR and NPR have exactly the same probability of electing the Condorcet Loser in three-candidate elections. As mentioned earlier, the IC model assumes -in addition to impartiality with respect to candidates- that votes are independent whereas IAC introduces some level of correlation between individual preferences [see Berg (1985)]. This observation suggests that introducing some degree of dependence or homogeneity in voters' preferences is sufficient to significantly reduce the propensity of AV to elect the CL, while the impact of homogeneous preferences on PR performance is much more limited (note that the same conclusion applies when homogeneity of preferences is introduced in a very different way *via* the notion of single-peakedness).

Finally, when the number of candidates is higher than three, Gehrlein and Lepelley (1998) prove that the likelihood of electing the CL is strictly higher with AV than with PR under IC. As IC and IAC tend to coincide when the number of candidates increases, the superiority of AV over PR for avoiding Borda's Paradox can be considered as a rather robust theoretical conclusion.

References

- Berg S. (1985). Paradox of voting under an urn model: the effect of homogeneity. *Public Choice* 43, 377-387
- Borda J.C. (1784), Mémoires sur les élections au scrutin, in *Mémoires de l'Académie Royale des Sciences*. Paris
- Brams S.J. and P.C. Fishburn (1983). *Approval Voting*, Birkhäuser Publishers, Boston
- Cervone D., Gehrlein W.V. and Zwicker W. (2005), Which scoring rules maximizes Condorcet efficiency under IAC? *Theory and Decision* 58, 145-185
- Cohen J. and Hickey T. (1979), Two algorithms for determining volumes of convex polyhedra, *Journal of the ACM* 26, 401-414
- Diss M. and Gehrlein W.V. (2012), Borda's Paradox with Weighted Scoring Rules, *Social Choice and Welfare* 38, 121-136
- Diss M., Merlin V. and Valognes F. (2010), On the Condorcet Efficiency of Approval Voting and Extended Scoring Rules for Three Alternatives, in J.F. Laslier and M.R. Sanver (eds), *Handbook of Approval Voting*, Springer
- Felsenthal D.S. (2012), Review of Paradoxes Afflicting Procedures for Electing a Single Candidate, in D. Felsenthal and M. Machover (eds), *Electoral Systems: Paradoxes, Assumptions and Procedures*, Springer
- Felsenthal D.S. and Maoz Z. (1988), A Comparative Analysis of Sincere and Sophisticated Voting Under the Plurality and Approval Voting Procedures, *Behavioral Science*, 33, 116-130
- Gehrlein W.V. and Fishburn P.C. (1978), Probabilities of elections outcomes for large electorates, *Journal of Economic Theory* 19, 38-49
- Gehrlein W.V. and Lepelley D. (1998), The Condorcet efficiency of approval voting and the probability of electing the Condorcet loser, *Journal of Mathematical Economics* 29, 271-283
- Gehrlein W.V. and Lepelley D. (2014), The Condorcet Efficiency advantage that voter indifference gives to Approval Voting over some other Voting Rules, *Group Decision and Negotiation* 24 (2), 243-269
- Gehrlein W.V., Lepelley, D. and Moyouwou, I. (2014), Voters' preference diversity, concepts of agreement and Condorcet's paradox. *Quality & Quantity* (forthcoming)

Gehrlein W.V., Moyouwou I. and Lepelley D. (2013), The impact of voters' preference diversity on the probability of some electoral outcomes, *Mathematical Social Sciences* 66, 352-365

Lepelley D. (1989), Contribution à l'analyse des procédures de décision collective, Université de Caen, unpublished doctoral dissertation

Lepelley D. (1993), On the probability of electing the Condorcet loser, *Mathematical Social Sciences* 25, 105-116

Lepelley D. (1995), Condorcet efficiency of positional voting rules with single-peaked preferences, *Economic Design* 1, 289-299

Nurmi H. and Uusi-Heikkilä Y. (1985), Computer simulations of Approval Voting and Plurality Voting : the frequency of weak Pareto violations and Condorcet loser choices in Impartial Cultures, *European Journal of Political Economy* 2, 47-59

Plassmann F. and Tideman N. (2014), How Frequently Do Different Voting Rules Encounter Voting Paradoxes in Three-candidate elections? *Social Choice and Welfare* 42, 31-75

Verdoolaege, S., Woods, K., Bruynooghe, M. and Cools, R. (2005), Computation and manipulation of enumerators of integer projections of parametric polytopes. Technical Report CW 392, Katholieke Universiteit Leuven, Department of Computer Sciences, Celestijnenlaan 200A-B-3001 Heverlee

Wilson M.C. and Pritchard, G. (2007), Probability Calculations under the IAC Hypothesis, *Mathematical Social Sciences* 54, 244-256

Appendix A

Under our IAC-based assumptions with large electorates, obtaining the desired probabilities is accomplished by the computation of volumes that are associated with convex polyhedra that are defined by a linear system of inequalities.

These volumes are obtained in our first approach by using a method derived from the well-known Cohen and Hickey algorithm for triangulating a polytope [Cohen and Hickey(1979)]. Let P be a given d -dimensional polytope described by some non-redundant linear inequalities $E_j : c_j y \leq b_j; j = 1, 2, \dots, m$. Each facet F_j of P corresponds to at most one equation $c_j y = b_j$ with $j = 1, 2, \dots, m$. Each vertex can then be attached to the subset of facets it belongs to. Choosing a vertex, said v_l , a dissection of P is obtained by considering all pyramids with apex v_l and bases F_j such that v_l is out of F_j . This operation is then applied recursively to find a triangulation of P into simplices, each containing $d+1$ points that are affine independent. Finally the volume of P is the sum of the volumes of simplices obtained in its triangulation using the following formula of the d -dimensional volume of a simplex $\Delta(a_0, a_1, \dots, a_d)$:

$$vol(\Delta(a_0, a_1, \dots, a_d)) = \frac{|\det(a_1 - a_0, a_2 - a_0, \dots, a_d - a_0)|}{d!} vol_0$$

where each a_j is a vertex of $\Delta(a_0, a_1, \dots, a_d)$ and the notation det stands for the determinant and vol_0 is a constant that depends on the Cartesian coordinate system used for vertices. Since each probability in this paper is a ratio, one can simply set $vol_0 = 1$. For a detailed illustration of this method, the reader is referred to Gehrlein *et al.* (2014).

The volumes in our second and third approaches are easier to calculate and are obtained *via* Barvinok's algorithm that computes the number of integer points in a polytope [see Wilson and Pritchard (2007)]. Using the program of Verdoolaege *et al.* (2005) for implementing this algorithm, we derive the quasi polynomial representing the total number of integer points that satisfy the set of constraints defining the event under consideration as a function of n , the total number of voters. The corresponding volume is the coefficient of the leading term in n in that representation. When two parameters (e.g. n and h) are in order, the volume is the coefficient of the leading term in n when h is replaced by αn in the quasi polynomial.

Appendix B

The representations for PR are as follows:

$$\begin{aligned}
CL_{III}(\text{PR}(\alpha), \infty) &= -\frac{2(24401\alpha^5-23845\alpha^4-39690\alpha^3+72000\alpha^2-40320\alpha+8064)}{15(33073\alpha^5-176904\alpha^4+361260\alpha^3-362880\alpha^2+181440\alpha-36288)}, \text{ for } 0 \leq \alpha \leq 1/3 \\
&= -\frac{\begin{bmatrix} 855944415\alpha^{10}-3500745480\alpha^9+6330348045\alpha^8-6649442280\alpha^7+4483240650\alpha^6 \\ -2020849236\alpha^5+613533690\alpha^4-123304680\alpha^3+15761385\alpha^2-1152270\alpha+36293 \end{bmatrix}}{10935\alpha^5(33073\alpha^5-176904\alpha^4+361260\alpha^3-362880\alpha^2+181440\alpha-36288)}, \text{ for } 1/3 \leq \alpha \leq 1/2 \\
&= \frac{\begin{bmatrix} 5262736293\alpha^{10}-32182113240\alpha^9+85904583615\alpha^8-132396518520\alpha^7+130899145230\alpha^6 \\ -86977348668\alpha^5+39412969470\alpha^4-12040432920\alpha^3+2374088535\alpha^2-272960865\alpha+13894067 \end{bmatrix}}{\begin{bmatrix} 10935(890357\alpha^{10}-5277480\alpha^9+13729500\alpha^8-20707200\alpha^7+20180160\alpha^6) \\ -13390272\alpha^5+6158880\alpha^4-1929600\alpha^3+388800\alpha^2-45480\alpha+2343 \end{bmatrix}}, \text{ for } 1/2 \leq \alpha \leq 3/5 \\
&= -\frac{\begin{bmatrix} 25586873082\alpha^{10}-175696793010\alpha^9+536877838260\alpha^8-962470356480\alpha^7+1122046823520\alpha^6 \\ -889545488832\alpha^5+486049319280\alpha^4-180872650080\alpha^3+43901136540\alpha^2-6279706665\alpha+402224236 \end{bmatrix}}{\begin{bmatrix} 10935(890357\alpha^{10}-5277480\alpha^9+13729500\alpha^8-20707200\alpha^7+20180160\alpha^6) \\ -13390272\alpha^5+6158880\alpha^4-1929600\alpha^3+388800\alpha^2-45480\alpha+2343 \end{bmatrix}}, \text{ for } 3/5 \leq \alpha \leq 2/3 \\
&= -\frac{\begin{bmatrix} 59266739664\alpha^{10}-448064278560\alpha^9+1503540280080\alpha^8-2946730907520\alpha^7+3730328108640\alpha^6 \\ -3180766361472\alpha^5+1844802948240\alpha^4-715836489120\alpha^3+176908842180\alpha^2-24955649535\alpha+1506645364 \end{bmatrix}}{\begin{bmatrix} 273375(822064\alpha^{10}-6532320\alpha^9+22880880\alpha^8-46477440\alpha^7+60651360\alpha^6) \\ -53234496\alpha^5+31943520\alpha^4-13000320\alpha^3+3447360\alpha^2-538200\alpha+37593 \end{bmatrix}}, \\
&\hspace{15em} \text{for } 2/3 \leq \alpha \leq 3/4 \\
&= \frac{\begin{bmatrix} 449530466688\alpha^{10}-3820726091520\alpha^9+14576795199360\alpha^8-32861250339840\alpha^7+48461073050880\alpha^6 \\ -48836566324224\alpha^5+34049117854080\alpha^4-16211192567040\alpha^3+5041737412560\alpha^2-924352518720\alpha+75833892463 \end{bmatrix}}{314928000(1-\alpha)^5(1573\alpha^5-1385\alpha^4+2770\alpha^3-2450\alpha^2+985\alpha-149)}, \\
&\hspace{15em} \text{for } 3/4 \leq \alpha \leq 5/6 \\
&= \frac{\begin{bmatrix} 965866624\alpha^{10}-8359496960\alpha^9+32498801280\alpha^8-74737720320\alpha^7+112598698240\alpha^6-116130391552\alpha^5 \\ +83041091840\alpha^4-40653265920\alpha^3+13040274880\alpha^2-2474912560\alpha+211054449 \end{bmatrix}}{144000(1-\alpha)^5(1573\alpha^5-1385\alpha^4+2770\alpha^3-2450\alpha^2+985\alpha-149)}, \text{ for } 5/6 \leq \alpha \leq 7/8 \\
&= \frac{33711\alpha^5-155115\alpha^4+291990\alpha^3-270870\alpha^2+123835\alpha-22319}{45(1573\alpha^5-1385\alpha^4+2770\alpha^3-2450\alpha^2+985\alpha-149)}, \text{ for } 7/8 \leq \alpha \leq 1.
\end{aligned}$$

The representations for NPR are as follows:

$$\begin{aligned}
CL_{III}(\text{NPR}(\alpha), \infty) &= \frac{2(373\alpha^5-12420\alpha^4+27810\alpha^3-28320\alpha^2+14280\alpha-2856)}{5(33073\alpha^5-176904\alpha^4+361260\alpha^3-362880\alpha^2+181440\alpha-36288)}, \text{ for } 0 \leq \alpha \leq 1/4, \\
&= -\frac{\begin{bmatrix} 215563232\alpha^{10}-728887680\alpha^9+1030069440\alpha^8-824540160\alpha^7+430590720\alpha^6 \\ -158113536\alpha^5+40158720\alpha^4-6428160\alpha^3+667440\alpha^2-40680\alpha+1107 \end{bmatrix}}{10000\alpha^5(33073\alpha^5-176904\alpha^4+361260\alpha^3-362880\alpha^2+181440\alpha-36288)}, \text{ for } 1/4 \leq \alpha \leq 1/3 \\
&= -\frac{\begin{bmatrix} 205007700384\alpha^{10}-1394098076160\alpha^9+3743158625280\alpha^8-5494225489920\alpha^7+4985727344640\alpha^6 \\ -2947589471232\alpha^5+1153741760640\alpha^4-297001105920\alpha^3+48789611280\alpha^2-4638647160\alpha+194509009 \end{bmatrix}}{21870000\alpha^5(33073\alpha^5-176904\alpha^4+361260\alpha^3-362880\alpha^2+181440\alpha-36288)}, \text{ for } 1/3 \leq \alpha \leq 1/2 \\
&= \frac{\begin{bmatrix} 2266329515616\alpha^{10}-13198343523840\alpha^9+34057299294720\alpha^8-51277728430080\alpha^7+49897241775360\alpha^6 \\ -32798068224768\alpha^5+14748226839360\alpha^4-4476665854080\alpha^3+876650373720\alpha^2-99923037840\alpha+5027378741 \end{bmatrix}}{4374000\begin{bmatrix} 890357\alpha^{10}-5277480\alpha^9+13729500\alpha^8-20707200\alpha^7+20180160\alpha^6 \\ -13390272\alpha^5+6158880\alpha^4-1929600\alpha^3+388800\alpha^2-45480\alpha+2343 \end{bmatrix}}, \\
&\hspace{15em} \text{for } 1/2 \leq \alpha \leq 2/3
\end{aligned}$$

$$\begin{aligned}
&= \frac{\left[\begin{array}{l} 23430687912\alpha^{10} - 144266495040\alpha^9 + 385172340480\alpha^8 - 577209369600\alpha^7 + 519529489920\alpha^6 \\ -268870129920\alpha^5 + 54488064960\alpha^4 + 20459848320\alpha^3 - 16668461880\alpha^2 + 4363181520\alpha - 428995097 \end{array} \right]}{34992 \left[\begin{array}{l} 822064\alpha^{10} - 6532320\alpha^9 + 22880880\alpha^8 - 46477440\alpha^7 + 60651360\alpha^6 \\ -53234496\alpha^5 + 31943520\alpha^4 - 13000320\alpha^3 + 3447360\alpha^2 - 538200\alpha + 37593 \end{array} \right]}, \text{ for } 2/3 \leq \alpha \leq 3/4 \\
&= \frac{\left[\begin{array}{l} 5483923884\alpha^{10} - 39661993440\alpha^9 + 126675239040\alpha^8 - 233823075840\alpha^7 + 273546927360\alpha^6 \\ -208408152960\alpha^5 + 101322174240\alpha^4 - 28596525120\alpha^3 + 3168624420\alpha^2 + 401212500\alpha - 108352859 \end{array} \right]}{2519424(\alpha-1)^5(1573\alpha^5 - 1385\alpha^4 + 2770\alpha^3 - 2450\alpha^2 + 985\alpha - 149)}, \text{ for } 3/4 \leq \alpha \leq 4/5 \\
&= - \frac{\left[\begin{array}{l} 4008263616\alpha^{10} - 36275506560\alpha^9 + 146699760960\alpha^8 - 349376924160\alpha^7 + 542933072640\alpha^6 - 575412647040\alpha^5 \\ + 421225025760\alpha^4 - 210282194880\alpha^3 + 68494991580\alpha^2 - 13141410900\alpha + 1127568731 \end{array} \right]}{2519424(1-\alpha)^5(1573\alpha^5 - 1385\alpha^4 + 2770\alpha^3 - 2450\alpha^2 + 985\alpha - 149)}, \\
& \hspace{20em} \text{for } 4/5 \leq \alpha \leq 5/6 \\
&= \frac{17587\alpha^5 - 64015\alpha^4 + 94860\alpha^3 - 71560\alpha^2 + 27785\alpha - 4433}{18(1573\alpha^5 - 1385\alpha^4 + 2770\alpha^3 - 2450\alpha^2 + 985\alpha - 149)}, \text{ for } 5/6 \leq \alpha \leq 1.
\end{aligned}$$