3-30-2019

Working Paper Number 19006

The Will of the People: Measuring social divides in multi-dimensional choice settings

Sunanda Roy
Iowa State University, sunanda@iastate.edu

Kuan Chuen Wu
Iowa State University, kcwu@iastate.edu

Abhijit Chandra
Iowa State University, achandra@iastate.edu

Original Release Date: March 28, 2019
Latest Revision: March 30, 2019

The Will of the People: Measuring social divides in multi-dimensional choice settings

Abstract
The paper studies profile components that cause plurality scores to differ from scores under other positional methods. We focus especially on the differences between plurality tallies and the Borda Count. We show that these profile components consist of an equal number of voters supporting a specific candidate in the first and last places. In particular they have an equal number of voters supporting a specific preference order and its reverse. In light of recent attention on partisanship and political polarization, it is important to have measures of such profiles and an assessment of how much they contribute to positional tallies. The methodological contribution of the paper lies in cutting through the combinatorial complexities that crop up in analysis of rank orders over n-candidates to characterize and measure these profiles.

Keywords
Preference profiles, plurality voting, Borda Count, Condorcet cycles

Disciplines
Econometrics | Politics and Social Change

The Will of the People:

Measuring social divides in multi-dimensional choice settings

Sunanda Roy[*][†], Kuan Chuen Wu[‡], Abhijit Chandra[§]

March 30, 2019

Abstract

The paper studies profile components that cause plurality scores to differ from scores under other positional methods. We focus especially on the differences between plurality tallies and the Borda Count. We show that these profile components consist of an equal number of voters supporting a specific candidate in the first and last places. In particular they have an equal number of voters supporting a specific preference order and its reverse. In light of recent attention on partisanship and political polarization, it is important to have measures of such profiles and an assessment of how much they contribute to positional tallies. The methodological contribution of the paper lies in cutting through the combinatorial complexities that crop up in analysis of rank orders over *n*-candidates to characterize and measure these profiles.

JEL Numbers: D70, D71, D72

Keywords: Preference profiles, plurality voting, Borda Count, Condorcet cycles

[*](Corresponding Author), Department of Economics, Iowa State University; email: sunanda@iastate.edu

[†]We wish to thank the Cambridge Election Commission, MA, for providing the ballot data for this study and Lesley Waxman in particular for many useful discussions and for helping us gain access to this data. Any opinions, conclusions or recommendations expressed in the paper are, however ours and do not reflect the views of the Cambridge Election Commission. We also thank the participants of the ISU Department of Economics brown-bag seminar series, the 2014 Spring and Fall Midwest Economic Theory meetings, the 63rd Annual French Economic Association meetings, the 2015 World Congress of the Econometric Society in Montreal and Marcus Pivato and David Frankel in particular for their comments and suggestions. We are also grateful to three anonymous referees on a previous version of the work. Any remaining errors in the paper are ours.

[‡]Department of Mechanical Engineering, Iowa State University; email: kcwu@iastate.edu

[§]Department of Mechanical Engineering, Iowa State University; email: achandra@iastate.edu

1 Introduction

Voting procedures which require a voter to rank all alternatives, instead of requiring only his first choice, have recently received much attention as a way of mitigating partisan influence in elections. Election reform advocacy groups in the US, such as Fairvote.org, has long advocated the Ranked Choice Voting method or instant runoff as an alternative to plurality. A steadily increasing number of city councils in the US have adopted some version of these methods and in 2018, Maine became the first state to use one for state and federal elections. Internationally, calls to replace plurality by methods that use rank orders, such as the Single Transferable Vote and Proportional Representation have grown louder.[1] The focus on rank orders as a governance tool calls for further examination of the connection between positional tallies and a profile of such rank orders.

It is well known that various positional voting methods disagree on the social ranking obtained from a given profile of rank orders because specific procedures fail to process the information conveyed by parts of the profile that are symmetric in structure. Young (1975), Saari (2000 a,b) and Balinski and Laraki (2010) provide details of some of these arguments. The present paper studies the component profiles that cause plurality scores to differ from tallies under other positional methods. We focus especially on those which explain differences between plurality and the Borda Count - a method known to have the fewest problems with symmetric profiles in general (see Young 1975, Saari 2000b). We show that these component profiles provide useful information about the collective psyche of the electorate and explore some of their implications.

We begin by showing that all component profiles that cause plurality tallies to differ from the Borda Count have an equal number of voters supporting a specific candidate in the first and last places. With three candidates, A, B and C for example, three component profiles explain all such differences. Each profile has either A or B or C first and last ranked by an equal number of voters. Alternatively, each such profile has an equal number of voters supporting a specific preference order and its reverse. For want of a better term, we describe these component profiles as Reverse profiles to highlight this last unique feature. Although not unknown in the literature, these profiles have received little attention so far.[2] We posit that these profiles reveal social divides and identify some alternatives or candidates as *polarizing*. In light of the recent focus on partisanship and political polarization, it is important to have measures of such component profiles and an

[1] For more information on electoral reforms and alternatives to plurality, see https://www.fairvote.org, https://www.electoral-reform.org.uk., https://www.fairvote.ca.

[2] Young (1975) is amongst the earliest papers to draw attention to these profiles and interprets them as electoral ties because of their symmetry. A Condorcet cycle is another well known example of a symmetric that causes pairwise voting methods to disagree on a given profile (Saari (2000a)).

understanding of how much they may potentially contribute to positional tallies of candidates in any given situation.

Consider a hypothetical situation with three candidates. Assume that the electorate consists of two significant and equal sized groups with preference orders $A > B > C$ and $C > B > A$ and a small remaining fraction with preference given by $B > A > C$. An instant runoff elects A, who however also happens to be the last choice of nearly 50% of the voters.[3] The example also reveals a potential weakness in the argument that as the rank orders $A > B > C$ and $C > B > A$ are supported by an equal number of voters, they *cancel* each other and together constitute an electoral tie. Doing so, allows a minority of the voters with preference $B > A > C$ to determine B as the winner - who incidentally, is also the Borda winner.

Without suggesting a specific method as the ”right” one to use under these circumstances, we look towards a measure of the strength of these component profiles to offer some guidance. In other words, the weight of these component profiles in the given profile is allowed to choose the appropriate method to use under specific circumstances. With n candidates and $n!$ possible ways to rank them, uncovering these profiles is a computational challenge for a large enough n. The main methodological contribution of the paper (Theorem 1 and Corollary 2) is an implementable technique of extracting and measuring these component profiles for any number of candidates.

The paper contributes to the literature on measures of divisions within a society. Notwithstanding the attention received by issues of increasing partisanship and polarization, little attempt has been made to characterize divisions in a multi-dimensional choice setting. An existing measure, namely polarization (Esteban and Ray (1994), Duclos, Esteban and Ray (2004) and Montalvo and Reynal-Querol (2008)) is useful when divisions within a population are sought to be measured on the basis of a single characteristic or attribute, say income or a single policy position.[4] Most political elections on the other hand, especially those for high offices usually involve voters comparing available candidates along multiple cardinal and ordinal attributes. It is reasonable to expect that such a comparison results in an ordinal rank order of the candidates, in the mind of the individual voter (irrespective of what the ballot requires of him or her). A measure of social divides in a multi-dimensional choice setting must therefore accept a set of such rank

[3] Some evidence suggests that the outcome of the 2016 US Presidential Republican primaries may have been the result of precisely such a polarization amongst the primary voters. See, http://www.fairvote.org, for article posted on Mar 04, 2016 by Andrew Douglas, Rob Richie, Elliott Louthen; http://www.newsweek.com for article posted on Oct 17, 2016 by Paul Raeburn; http://thefederalist.com. for article posted on April 5, 2016 by Kyle Sammin.

[4] Existing measures of political polarization are weighted averages of measures along single-dimensional attributes. See for example, (Boxell, Gentzkow and Shapiro (2017), Mason (2015), Abramowitz and Saunders (2008), Fiorina and Abrams (2008), Fiorina, Abrams and Pope (2008) amongst others. Papers studying issues such as policy extremism, partisan voting behavior, political posturing by candidates and various socio-economic implications of polarization implicitly or explicitly adopt a framework in which voters' preferences are one-dimensional, that is, based on a single policy position. See for example, (Dellis (2009), Haimanko, Le Breton and Weber (2007), Krasa and Polborn (2014), Testa (2012) and Woo (2005).

orders as the informational input. Our paper contributes towards filling this methodological gap.

Moreover, despite its many acknowledged drawbacks, the plurality method is often credited with an ability to select the exceptional rather than the (merely) good candidate. The paper provides an assessment tool for this hypothesis. We show that plurality tallies are explained by two broad classes of component profiles (Theorem 1). On the first class of these component profiles, the same social ranking of the candidates is obtained under any aggregation method. If the given electorate consists of the first class of profiles only or consists of these profiles in large measure, the choice of the voting method will not matter.[5] With a slight abuse of the term, we may say that such a social rank order reflects the "objective" merits of the candidates. The other class of component profiles is Reverse profiles and reflect a divided jury on any candidate. Thus, the relative weight of these two classes of profiles in determining the plurality winner also provides a metric as to whether a candidate's "objective" merit (by all voting methods) or a divided jury is responsible for the plurality outcome. Several such measures are proposed in the paper in Section 5.

The paper uses the geometric or linear algebraic approach to voting theory pioneered by Saari (1992, 1999; 2000a; 2000b) to whom it also owes many of its important formal results. Hodge and Klima (2005) and Balinski and Laraki (2010) provide an exposition of this approach. A preference profile or profile for short is a distribution of voters across all possible rank orders of the available alternatives or candidates. With a n-candidate field, a profile is a vector in a $n!$ dimensional space. The linear algebraic approach consists in expressing a given profile as a linear combination of a set of orthogonal basis or component profiles. The coefficient or weight of a component profile is a measure of the extent to which the given profile can be explained by the component profile.

The usual linear algebraic method of decomposing such a vector requires characterizing a complete set of $n!$ orthogonal basis vectors, a computationally intensive perhaps impossible task for an arbitrary n. Moreover, for the decomposition to be meaningful from a social choice theoretic point of view, the component vectors must provide useful information on the collective psychology and with this in mind, a complete decomposition into $n!$ basis vectors may be unnecessary. A main result of the paper (Corollary 2) presents a novel decomposition technique that is both computationally simple and generates component profiles that are meaningful from a social choice theoretic point of view. The technique builds on a result of a previously published paper (Chandra and Roy, 2013) which showed that the weights of the first class of component profiles mentioned earlier can be extracted from the pairwise scores of all candidate pairs. Once these weights are known, the weights of the Reverse profiles of current interest can be extracted from the

[5] In deference to Saari who introduced these types of profiles to the discipline and his work, we continue to use his nomenclature and call them *Basic profiles* in the paper.

plurality tallies of the candidates.

The decomposition technique requires information about individual voters' rank orders of the candidates. Till date, few political elections required voters to rank all the alternatives. However, with Ranked Choice Voting spreading in the US and similar methods in use, internationally, we think the results of the paper will be practically and theoretically useful. The Cambridge (Massachusetts) City Election Commission has long required voters to rank all candidates. We use ballot data from the City Council elections to illustrate our decomposition method. There are specific features of this data that present difficulties and limits its substantive usefulness. Chief of these is that the voters are not required to rank *all* candidates but are instead required to rank at least one. Specific assumptions need to be made to tide over the difficulty posed by the incompleteness of the rank orders. The empirical results of the last section are therefore illustrative rather than conclusive. Nevertheless, the findings for the period of study provide some evidence of progression towards extreme preferences at a local level and confirm evidence in other studies.

The structure of the rest of the paper is as follows. Section 2 presents several 3-candidate examples to illustrate and explain our main ideas. Section 3 presents the necessary tools of geometric voting theory. Sections 4 and 5 present our main theoretical results and measures of social divides. Finally, in Section 6, we apply our results on ballot data from the Cambridge City Council, elections.

2 Illustrative Examples

The section presents several 3-candidate examples to illustrate and explain our main ideas.

Consider a society of V voters and three candidates, A, B and C. Assume that each voter has a strict transitive preference order across the candidates. An individual voter's rank order is thus one of six possible ways of ranking the candidates listed in the tables below. A profile is a distribution of voters across these six rank orders.

Component profiles are best understood as deviations from a profile in which the V voters are uniformly distributed over the six rank orders, as in q below. Profile q has some nice properties. Irrespective of the choice of an aggregation procedure, such a profile always yields the social ranking, $A \sim B \sim C$ - in other words, a complete tie across the candidates. Moreover, such a profile added to any other profile, x, changes the tallies of all candidates uniformly but *not* the social ranking of the candidates under any procedure. Thus, profile q represents an electorate that is *impartial* across all candidate.

Table 1:

Profile q		
	Rank order	no. of voters
(1)	$A > B > C$	$V/6$
(2)	$A > C > B$	$V/6$
(3)	$B > A > C$	$V/6$
(4)	$C > A > B$	$V/6$
(5)	$B > C > A$	$V/6$
(6)	$C > B > A$	$V/6$

EXAMPLE 1: Consider the profile p_1 obtained from profile q by shifting voters away from rank orders (3) and (4), where A is ranked in the middle, and adding them uniformly to rank orders (1), (2), (5) and (6), where A is either first or last ranked. An impartial profile of preferences, q, is thus transformed into a profile, p_1, with more extreme preferences - love or hate - towards a specific candidate, A. In particular, it has an equal number of voters supporting certain preference orders and their reverse. We name such profiles, Rverse profiles. The next section characterizes all such profiles for a n-candidate field.

Table 2:

Profile p_1		
	Rank order	no. of voters
(1)	$A > B > C$	$V/4$
(2)	$A > C > B$	$V/4$
(3)	$B > A > C$	0
(4)	$C > A > B$	0
(5)	$B > C > A$	$V/4$
(6)	$C > B > A$	$V/4$

EXAMPLE 2: Next consider another type of profile obtained by shifting voters away from rank orders (5) and (6) in profile q of Table 1, where A is last ranked, and adding them to rank orders (1) and (2), where A is first ranked. The resulting profile, p_2, is shown in Table 3. It has the property that under all sum scoring or pairwise voting procedures, the social rank order obtained is always $A > B \sim C$. That is, the social rank order is procedure independent and implies some amount of "consensus" amongst the voters that A should be elected. Note that even for such a profile, A is clearly not everybody's first choice. Saari (2000a) was the first to introduce this class of profiles to the social choice literature and named them *Basic* profiles.

Under plurality, the social ranking under both profiles p_1 and p_2 are identical - although tallies may differ - with A as the clear winner. It is likely however that the election of A will breed more resentment under p_1 than under p_2. Under p_1, half the voters rank A last whereas under p_2, A is nobody's last choice. Thus, the plurality procedure cannot distinguish the profile p_1 with more extreme preferences towards a specific candidate from the more "benign" profile p_2.

Table 3:

Profile p_1			Profile p_2		
	Rank order	no. of voters		Rank order	no. of voters
(1)	$A > B > C$	$V/4$	(1)	$A > B > C$	$V/3$
(2)	$A > C > B$	$V/4$	(2)	$A > C > B$	$V/3$
(3)	$B > A > C$	0	(3)	$B > A > C$	$V/6$
(4)	$C > A > B$	0	(4)	$C > A > B$	$V/6$
(5)	$B > C > A$	$V/4$	(5)	$B > C > A$	0
(6)	$C > B > A$	$V/4$	(6)	$C > B > A$	0

EXAMPLE 3: The third example shows that Reverse profiles, when of a significant size, are responsible for any disagreement between plurality and other positional methods, notably the Borda Count. Consider the profile x of 30 voters shown in Table 4. Under plurality, A wins the election with 13 votes followed by B in second place with 10 votes and C in third place with 7 votes. On the other hand, if Borda Count is used, B wins with 34 points and A and C are tied with 28 points each. Thus either A or B can be elected by an appropriately chosen procedure.

Table 4:

Profile x					
	Rank order	no. of voters		Rank order	no. of voters
(1)	$A > B > C$	7	(2)	$A > C > B$	6
(3)	$B > A > C$	2	(4)	$C > A > B$	0
(5)	$B > C > A$	8	(6)	$C > B > A$	7

A *decomposition* of the profile provides some insight. The profile x is actually a direct sum of two profiles, x_1 and x_2 shown in the table below.

Table 5:

Profile x_1			Profile x_2		
	Rank order	no. of voters		Rank order	no. of voters
(1)	$A > B > C$	6	(1)	$A > B > C$	1
(2)	$A > C > B$	6	(2)	$A > C > B$	0
(3)	$B > A > C$	0	(3)	$B > A > C$	2
(4)	$C > A > B$	0	(4)	$C > A > B$	0
(5)	$B > C > A$	6	(5)	$B > C > A$	2
(6)	$C > B > A$	6	(6)	$C > B > A$	1

Profile x_1 is the same as profile p_1 with $V = 24$. Profile x_2 is similar to profile s in Table (3) with $V = 6$, except that x_2 favors B as the winner under any standard procedure (instead of A, as under s), and A and C are tied for the second place. The disagreement between the plurality and the Borda rankings under profile x, is caused by the greater size (loosely speaking) of the Reverse profile x_1 relative to the Basic profile x_2 in x. If the weight of x_1 is lowered relative to the weight of x_2 in x (through having fewer voters in x_1), the social ranking under x would be more aligned with the social ranking under x_2. And as x_2 is a profile on which all standard procedures agree, we shall see more convergence of the procedures on x itself. Thus disagreement between plurality and the Borda Count happens when the weights of the Reverse profiles are significant compared to the weights of the Basic profiles.

The example serves to highlight our methodological contribution. Whilst the decomposition of x into x_1 and x_2 is easy because we have only three candidates, the standard method quickly runs into combinatorial difficulties as the number of candidates increases as such an exercise involves decomposing a profile of dimension $n!$. Corollary 2 proposes a decomposition method that is easy to implement for any number of candidates.

3 The algebra of geometric voting theory

3.1 Preference profiles

Voters have strict transitive preferences over n candidates indexed $i = 1 \ldots n$. Hence there are $n!$ different ways of ranking these candidates. Assume an electorate of a given and fixed size. A *profile* $p = (p_1 \ldots p_{n!}) \in \mathbf{R}^{n!}_{+}$ is a distribution of voters across these rank orders, with p_k equal to the number of voters with preferences given by the kth rank order of the candidates. A *profile differential* $p' \in \mathbf{R}^{n!}$ is the difference between two different profiles for an electorate of a given size, implying that p' may have negative components and that its components add up to zero.

A preference aggregation or voting procedure attempts to obtain an aggregate or social rank order of the candidates based on a given profile. *Pairwise* procedures are based on pairwise comparison of the candidates.[6] The pairwise tally of candidate i against candidate j is given by the number of voters in p, who rank i over j. The *normalized pairwise score difference* between i and j is defined as,

$$a_{ij} = \frac{(i\text{'s tally against } j - j\text{'s tally against } i)}{\text{total number of voters}} \tag{1}$$

Thus if $a_{ij} > 0$, i beats j and if $a_{ij} < 0$, then j beats i in a pairwise comparison of i and j. Further, note that, $a_{ij} = -a_{ji}$ and as a_{ij} is normalized, $-1 \leq a_{ij} \leq 1$. A pairwise procedure constructs a social ranking of the candidates based on the set of pairwise score differences $\{a_{ij}\}_{i \neq j, i<j}$.

Positional or sum-scoring methods assign fixed points to a candidate depending on his/her position in an individual's ranking. The aggregate or social ranking is obtained by the sum of these points across all individuals. Plurality and the Borda Count are commonly used examples. The plurality method involves a voter awarding one point to his/her first ranked candidate and zero to all other candidates placed in other positions in his/her preference order. The plurality tallies of the candidates are the sum of all the points awarded by all the voters. The Borda Count (BC) assigns $n-1$ points to the first ranked candidate of each voter, $n-2$ points to the second ranked candidate and so on and 0 to the last ranked candidate. The aggregate ranking is based on the sum of all points awarded by the voters.

Denote by K^n, a profile that assigns one voter for each possible ranking. Assuming that the electorate consists of V voters, the profile of a uniform distribution of voters is given by, $q = \frac{V}{n!}K^n$, for a n candidate field. It has the two nice properties noted earlier - namely, (1) it yields a tied outcome across all candidates under any preference aggregation procedure and (2) the relative rank of any two candidates in the social rank order under a given profile, p, under any procedure is not affected by adding or subtracting a scaler multiple of a K^n profile to p.

It is useful to view a given profile p as a perturbation from a $\frac{V}{n!}K^n$ profile, that is, $p = \frac{V}{n!}K^n + p'$ for some profile differential $p' \in \mathbf{R}^{n!}$. As $\frac{V}{n!}K^n$ has completely tied outcomes for all candidates, p' and p yield the same social ranking of the candidates under any procedure. Thus p' and p have the same structure but different sized elctorates. The size of the electorate for p' is normalized to zero. Mathematically, working

[6] Examples include Condorcet's "successive reversal" and the "maximal agreement" procedures, Kemeny's method and Copeland's method.

with the profile differential p' is more convenient than working with p itself. For instance, as p' is orthogonal to K^n, a decomposition of p' (rather than p) does not include neutral K^n effects, that merely changes tallies without affecting results.

The expression $p = \frac{V}{n!}K^n + p'$ has an intuitive appeal. It says that p is obtained from the uniform distribution, $\frac{V}{n!}K^n$, by moving voters away from specific rank orders and adding them to others. The negative components of p' indicate which rank orders suffer depletion whereas the positive components indicate which ones experience gains. In other words, any given profile p can be viewed as a result of "padding" and "thinning" of specific rank orders of an uniform distribution of voters. The profile $x = (7,6,2,0,8,7)$ of 30 voters of Example 3 can be expressed as $5K^3 + x'$ where $x' = (2,1,-3,-5,3,2)$. It is obtained from a $5K^3$ profile by moving voters away from the rank orders (3) and (4) - under which A is middle ranked - and adding them to the other rankings - under which A is first or last ranked.

3.2 Basic profiles

Fix a candidate, say i. Take a K^n profile and shift a voter from each ranking which has i last ranked and add the voter to a ranking which has i first ranked, taking care not to add more than one such voter to a ranking. The profile thus obtained is the n-candidate analogue of profile p_2 of Table 3. The associated profile differential $p'(i)$ has one voter for each ranking that has i top ranked, (-1) voter for each ranking that has i bottom ranked and 0 voter for each ranking that has i ranked somewhere in the middle. Following Saari's nomenclature we call this a Basic profile and denote it by B_i^n.

An electorate consisting of only B_i^n has the i-th candidate top ranked and everyone else tied for the second place under *any* voting procedure. There are $(n-1)$ linearly independent Basic profiles in a n candidate field, each favoring a specific candidate in the social ranking while having the others tied for second place. By construction, $\sum_{i=1}^{n} B_i^n = 0$. The nice aggregation properties of B_i^n extend to linear combinations of these profiles as well. A social ranking of the candidates obtained on a linear sum of Basic profiles is the *same* under any aggregation procedure or in other words, *procedure invariant*.

The 4-candidate Basic profiles favoring candidates A and B are illustrated in Table (6).

Table 6:

Rank order	B_1^4	B_2^4	Rank order	B_1^4	B_2^4
1. $A > B > C > D$	(1)	(0)	13. $D > C > B > A$	(-1)	(0)
2. $A > B > D > C$	(1)	(0)	14. $C > D > B > A$	(-1)	(0)
3. $A > C > B > D$	(1)	(0)	15. $D > B > C > A$	(-1)	(0)
4. $A > C > D > B$	(1)	(-1)	16. $B > D > C > A$	(-1)	(1)
5. $A > D > C > B$	(1)	(-1)	17. $B > C > D > A$	(-1)	(1)
6. $A > D > B > C$	(1)	(0)	18. $C > B > D > A$	(-1)	(0)
7. $B > A > C > D$	(0)	(1)	19. $D > C > A > B$	(0)	(-1)
8. $B > A > D > C$	(0)	(1)	20. $C > D > A > B$	(0)	(-1)
9. $C > A > B > D$	(0)	(0)	21. $D > B > A > C$	(0)	(0)
10. $C > A > D > B$	(0)	(-1)	22. $B > D > A > C$	(0)	(1)
11. $D > A > B > C$	(0)	(0)	23. $C > B > A > D$	(0)	(0)
12. $D > A > C > B$	(0)	(-1)	24. $B > C > A > D$	(0)	(1)

Saari (2000a) shows that the social ranking obtained on a linear combination of Basic profiles satisfies a strong and useful property. Consider such a combination, specifically, $p = \frac{V}{n!}K^n + \sum_{i=1}^{n-1} a_i B_i^n$, where the a_i's are given constants. The pairwise score difference, a_{ij}, for the candidate pair (i, j), is shown to be equal to $a_i - a_j$, the difference in the weights or coefficients of the B_i^n and B_j^n profiles. Thus in a social ranking obtained by any pairwise procedure, the relative rank of i and j is determined by the sign of $a_i - a_j$ alone and not affected by the coefficient of any other Basic profile. If $a_i - a_j > 0$, i is ranked above j, if $a_i - a_j < 0$, i is ranked below j and if $a_i - a_j = 0$, i and j are tied in the social rank order under any pairwise procedure. Thus, the social rank order obtained has the same linear (transitive) order as the set of coefficients $\{a_i\}$. Moreover, the social ranking is consistent over *subsets* of candidates. Furthermore, as the same social ranking is obtained under any procedure, pairwise or positional, on linear sums of Basic profiles, the preceding statements apply for positional methods as well. Saari (2000a) describes this strong property as the *additive transitive* property of Basic profiles.

3.3 Reverse profiles

Definition 1 *Fix an integer k, such that $2 \leq k \leq \frac{n+1}{2}$. (1) If $k < \frac{n+1}{2}$, a k-shift, i-inclined Reverse profile, $R_i^n(k)$, has 1 voter for each ranking in which the i-th candidate is first ranked and the reversal of this ranking, (-1) voter for each ranking in which he/she is k-th ranked and to the reversal of this ranking and 0 voters for all other rankings. (2) If $k = \frac{n+1}{2}$, $R_i^n(k)$ has 1 voter for each ranking in which the candidate is first ranked and the reversal of this ranking, (-2) voters for each ranking in which the candidate is k-th ranked and 0 voters for all other rankings.*

To understand the structure of the profile $R_i^n(k)$, assume $n > 3$ and $k = 2$. $R_i^n(2)$ has the same structure and tallies as a K^n profile with a voter moved from each ranking in which i is either 2nd or $(n-1)$th ranked and added to a ranking in which i is first or last ranked. In words, such a profile is obtained by padding the rankings in which i is placed at the two extremes and thinning the rankings in which i is placed in an intermediate position, namely in the 2nd or $(n-1)$th place.

The value of k in $R_i^n(k)$ determines which rankings with i in an intermediate position are thinned and the voters moved to the extremes. If $k = 2$, voters are shifted away from the rankings in which i is in the 2nd or $(n-1)$th place. If $k = 3$ voters are shifted away from rankings in which i is in the 3rd or $(n-2)$th place and so on. Adding a K^n profile to a $R_i^n(k)$ profile shows that each rank order that has i first and last ranked are supported by the same number of voters. Alternatively, each rank order with i at the top and the reverse of this rank order are supported by the same number of voters.

For a given and fixed k, there are $(n-1)$ linearly independent $R_i^n(k)$ profiles in a n-candidate field. Specifically, for a given and fixed k, $\sum_{i=1}^{n} R_i^n(k) = 0$. Table (7) illustrates the $R_1^4(2)$ and $R_2^4(2)$ profiles for a 4-candidate field. With only 4 candidates, as $2 \leq k \leq \frac{n+1}{2}$, the only possible value for k is $k = 2$. Thus, there are only three distinct Reverse profiles in a 4-candidate field, two of which, $R_1^4(2)$ and $R_2^4(2)$, are shown in the table.

Table 7:

Rank order	$R_1^4(2)$	$R_2^4(2)$	Rank order	$R_1^4(2)$	$R_2^4(2)$
1. $A > B > C > D$	(1)	(-1)	13. $D > C > B > A$	(1)	(-1)
2. $A > B > D > C$	(1)	(-1)	14. $C > D > B > A$	(1)	(-1)
3. $A > C > B > D$	(1)	(-1)	15. $D > B > C > A$	(1)	(-1)
4. $A > C > D > B$	(1)	(1)	16. $B > D > C > A$	(1)	(1)
5. $A > D > C > B$	(1)	(1)	17. $B > C > D > A$	(1)	(1)
6. $A > D > B > C$	(1)	(-1)	18. $C > B > D > A$	(1)	(-1)
7. $B > A > C > D$	(-1)	(1)	19. $D > C > A > B$	(-1)	(1)
8. $B > A > D > C$	(-1)	(1)	20. $C > D > A > B$	(-1)	(1)
9. $C > A > B > D$	(-1)	(-1)	21. $D > B > A > C$	(-1)	(-1)
10. $C > A > D > B$	(-1)	(1)	22. $B > D > A > C$	(-1)	(1)
11. $D > A > B > C$	(-1)	(-1)	23. $C > B > A > D$	(-1)	(-1)
12. $D > A > C > B$	(-1)	(1)	24. $B > C > A > D$	(-1)	(1)

REMARK 1: For a fixed n and a fixed i, a given k defines a class of $R_i^n(k)$ profiles by Definition 1. For example, for $n = 5$, there are two classes of Reverse profiles. There are four distinct $R_i^5(2)$ profiles and four

distinct $R_i^5(3)$ profiles. So far as the results of this paper are concerned however (see Section 4, Proposition 1), the relevant properties of the $R_i^n(k)$ profiles are independent of the specific choice of k. Most importantly, plurality tallies cannot identify different classes of these Reverse profiles from one another for a given n and a given i. Plurality tallies of the candidates are the same for all $R_i^n(k)$ profiles for all possible values of k, for a given n and i. Thus, for the rest of the paper we may without loss of generality assume $k = 2$ and describe $R_i^n(2)$ simply as Reverse profiles.

REMARK 2: $R_i^n(k)$ profiles are similar but not identical in structure to *Symmetric profiles* defined in Saari (2000b). Hence, we use a different name, although both types of profiles share common properties. Saari's construction appears to be driven by a specific need to express positional tallies as deviations from the Borda Count (2000b, Proposition 1). Such a step in turn is necessary if the weights of the Basic profiles are not known for a given profile and the Borda Count is used as a *surrogate* for the Basic profiles.[7] We have a technique, however, to extract the weights of the Basic profiles directly (see Section 5) and can consequently adopt an alternative and simpler characterization of these problem profiles.

3.4 Condorcet profiles

Condorcet profiles do not play a central role in the present paper but they influence the pairwise score differences between two candidates which in turn plays a part in our proposed decomposition methodology. To characterize Condorcet profiles, specify a *reference ranking* of the candidates, say $1 > 2 > 3 \ldots > n$, and index this ranking as (1). Consider the two sets of cyclic rankings generated by the reference ranking (1), listed in Table (8). The set of rankings in the first column of the table, $c^n_{(1)}$ is described as the *Condorcet n-tuple* generated by the reference ranking (1). The second set of rankings, listed in column 2 of the table, $\rho(c_{(1)})^n$, is another set of cyclic rankings that is the reverse of the first set. The sets $c^n_{(1)}$ and $\rho(c_{(1)})^n$ are thus two specific sets of cyclic rankings of the candidates with the feature that each ranking in the set $\rho(c_{(1)})^n$ is a reversal of a ranking in $c^n_{(1)}$.

[7]Borda scores may be used as surrogates for Basic profiles because Borda scores are influenced *only* by Basic profiles.

Table 8:

$c^n_{(1)}$	$\rho(c_{(1)})^n$
$1 > 2 > 3 \ldots > n$	$n > n-1 > n-2 \ldots > 1$
$2 > 3 > 4 \ldots > 1$	$n-1 > n-2 > n-3 \ldots > n$
$3 > 4 > 5 \ldots > 2$	$n-2 > n-3 > n-4 \ldots > n-2$
$\ldots$	$\ldots$
$n > 1 > 2 \ldots > n-1$	$1 > n-1 > n-2 \ldots > 2$

A *Condorcet profile* $C^n_{(1)}$ associated with the reference ranking (1), is a profile that has one voter for each ranking in $c^n_{(1)}$ and (-1) voter for each ranking in $\rho(c^n_{(1)})$ and zero voter for each remaining ranking in the profile. The first or reference ranking of the set $c^n_{(1)}$ uniquely characterizes a Condorcet profile. A n candidate field has thus $\frac{(n-1)!}{2}$ distinct Condorcet profiles because of $\frac{(n-1)!}{2}$ distinct ways of constructing the reference rank order. Condorcet profiles thus vastly outnumber the Basic and Reverse profiles. One of the reasons why constructing a set of $n!$ orthogonal basis profiles for a standard decomposition of a given profile is computationally infeasible is the difficulty of characterizing these distinct Condorcet profiles, for any arbitrary n.

The $C^n_{(1)}$ profile can be obtained from a K^n profile by moving a voter away from each of the $\rho(c^n_{(1)})$ rankings and adding it to each of the $c^n_{(1)}$ rankings. Alternatively, a Condorcet profile is obtained from a K^n profile by adding a voter to each ranking in the set of cyclic rankings that form a Condorcet n-tuple and taking a voter away from the reversal of this ranking.

A Condorcet profile has the feature that each candidate is placed in each position by exactly the same number of voters. Thus under any positional method, such a profile produces a complete tie. However it generates an intransitive ranking of the candidates under pairwise methods - a phenomenon often described as a *Condorcet paradox*.

3.5 Profile decomposition approach

The set of Basic, Condorcet and Reverse profiles do not describe a *complete* system of orthogonal basis vectors for the space of preference profiles, when $n > 3$. The span of the set of all such distinct profiles is a subspace of lower dimension than $n!$, when $n > 3$ (For reference, the 2016 US Republican Presidential primaries began with 17 candidates). On a more positive note, however, most of the residual component profiles are not informative from a social choice theoretic point of view. Moreover, the present paper has a limited objective for which a complete decomposition of a profile of dimension $n!$ is not necessary.[8]

Our goal is to characterize and measure only those component profiles that explain plurality tallies of the candidates with an eye towards explaining their differences with outcomes of other positional methods. Specifically, we desire to establish the algebraic relationship between the plurality tallies and the coefficients of these component profiles. Section 4 provides some of the necessary results and shows amongst other things that plurality tallies can be entirely explained by only Basic and Reverse profiles. The conclusions of Section 4 in conjunction with two existing results provide a pathway for a meaningful profile decomposition.

The first of these two existing results, Saari (2000a), shows that for any given profile p, the pairwise scores for any candidate pair, (i, j), are fully determined by the weights or coefficients of the Basic and the Condorcet profiles embedded in p. Thus p may be a sum of many different types of component profiles but only Basic and Condorcet profiles contribute to the pairwise scores across all candidate pairs. Other types of profiles contribute nothing. The second result, Chandra and Roy (2013), shows that it is possible to extract the component explained by Condorcet profiles from pairwise scores. In other words, the coefficients of the Basic profiles can be identified. This in turn implies that the measure of the Reverse profiles can be identified from the plurality tallies.

4 Reverse profiles, Basic profiles and plurality tallies

The following proposition describes the properties of the $R_i^n(k)$ profiles that are most important for us, for a given and fixed k, specifically, the pairwise and plurality tallies generated by these profiles.

The set of positional tallies of all candidates under any sum-scoring or positional method is a n dimensional vector in R_+^n, with each component of the vector representing the total points received by a candidate under the specific positional method, for the given profile. Mathematically, a vector of positional tallies maps a profile, $p \in R^{n!} \longrightarrow R_+^n$. The PV procedure involves a voter awarding one point to his/her first ranked candidate and zero to all other candidates placed in other positions in his/her preference order. Mathematically, the PV procedure is a linear map, represented by a $n \times n!$ matrix of 0 's and 1 's. Appendix 7.2 illustrates the map for the 3-candidate field.

Proposition 1 *For a given and fixed k,*

1. *The set of $\{R_i^n(k)\}_{i=1\ldots n}$ profiles are not pairwise orthogonal to each other and span a $(n-1)$ dimensional subspace of the profile space.*

[8] For the interested reader, Saari (2000b) describes a complete set of orthogonal basis vectors that explain a given profile for general n. A complete decomposition is necessary to explain all possible voting paradoxes.

2. *The set of* $\{R_i^n(k)\}_{i=1\ldots n}$ *profiles are pairwise orthogonal to the set of* $\{B_i^n\}_{i=1\ldots n}$ *profiles.*

3. *The plurality tallies of* B_i^n *and* $R_i^n(k)$ *profiles are identical, with candidate i receiving* $(n-1)!$ *points and every other candidate receiving* $-(n-2)!$ *points each. In particular, these tallies do not depend on the specific choice of k. The pairwise scores for each candidate pair under a* $R_i^n(k)$ *profile is a complete tie.*

Proof: See Appendix 7.1 for a formal proof.

REMARK 3: Part 2 of Proposition 1 shows that for any given k, the $\{R_i^n(k)\}_{i=1\ldots n}$ profiles are orthogonal to the Basic profiles in $R^{n!}$. This provides the foundation for our main result, Theorem 1, which shows furthermore that the *only* profiles that are both orthogonal to Basic profiles *and* contributes to plurality tallies are the $\{R_i^n(k)\}_{i=1\ldots n}$ profiles. Moreover, part 3 of Proposition 1 shows that the plurality tallies for both the $R_i^n(k)$ (for a fixed k) and the B_i^n profiles lie in an identical direction (are collinear) in R_+^n. The statement formally proves for n candidates our earlier observation in Example 2. Plurality tallies cannot distinguish between B_i^n and $R_i^n(k)$ profiles.

REMARK 4: Most importantly, part 3 of Proposition 1 shows that the plurality tallies of candidates under $R_i^n(k)$ profiles for a given k, do not depend on the value of k. This is not surprising. We saw in Section 3 that Reverse profiles are obtained by thinning rank orders where a candidate is placed in the middle and padding the rank orders where that candidate is placed first and last. So far as plurality tallies of the candidates are concerned, it does not matter which rank orders are thinned. We may therefore without loss of generality assume that plurality tallies contributed by $\{R_i^n(k)\}_{i=1\ldots n}$ profiles for any given k, comes from the $\{R_i^n(2)\}_{i=1\ldots n}$ profiles. In other words and without loss of generality, we may designate the set, $\{R_i^n(2)\}_{i=1\ldots n}$, as the set of generic Reverse profiles and denote it simply by $\{R_i^n\}_{i=1\ldots n}$.

REMARK 5: Part 3 of Proposition 1 further shows that pairwise scores across all candidate pairs produce a complete tie across the candidates on a R_i^n profile. In other words, R_i^n profiles do not influence pairwise scores differences a_{ij}, for any candidate pair (i,j). Combined with other existing results on pairwise scores, this provides the foundation of our decomposition technique.

The following theorem is one of the main results of the paper and provides a foundation for all subsequent ones.

Theorem 1 *Let p be any given profile. The difference in the plurality tallies of any two candidates, for p, can be fully explained by Basic and (generic) Reverse profile components.*

Proof: Basic profile components of p influence the difference in the plurality tallies of any two candidates. The residual difference in the plurality tallies can therefore be explained by profile components orthogonal to Basic profiles that influence plurality tallies. To prove our result we need to characterize these orthogonal profiles. Besides orthogonality (with Basic profiles) we require these profiles to satisfy the following condition. Each such component profile (like a Basic profile) must affect the plurality tally of a specific candidate but not others - that is, for each such component profile, all candidates other than a specified one must be tied under plurality. Profiles orthogonal to Basic profiles that affect the plurality tally of a specific candidate must have an equal number of voters for each ranking with the candidate in the first and last places. The structure of the remaining rankings (which have the specific candidate in other positions) does not matter, so long as for each such component profile two conditions are satisfied. (1) The other candidates are tied under plurality, implying that the distribution of voters across rank orders with other candidates in the first place is symmetric. (2) The component profile is a profile differential, implying that the total number of voters is zero. Such component profiles are therefore fully characterized by generic Reverse profiles. Δ.

Theorem 1 shows that differences in plurality tallies of all candidate pairs, for any profile p, can be explained by its Basic and Reverse profile components. No other type of structured component profile need be considered to explain these differences. Thus, a vector of plurality tallies for any given profile p may be fully explained by assuming that they have been generated by a linear combination of a K^n profile, $n-1$ independent Basic profiles and $n-1$ independent R_i^n profiles.

Amongst all positional methods, the Borda Counts is known to be least affected by symmetric profiles. In particular Reverse profiles contribute nothing to the Borda Count which has been shown to follow Basic profile rankings (see Saari, 2000a, 2000b). Theorem 1 implies the following relationship between plurality and the Borda Count.

Corollary 1 *Disagreements between the Plurality Voting (PV) and the Borda Count (BC) procedures can be entirely explained by the presence of Reverse profile components within a given profile.*

Proof: The BC follows the ranking obtained by the combination of the Basic profiles only within the given profile. The plurality ranking follows the combination of the Basic and Reverse profiles. Any disagreement between the two rank orders is therefore solely due to the Reverse profiles. Δ.

The next result formally decomposes the plurality tallies for a given profile, p, into components contributed by Basic and Reverse profiles for a normalized electorate of size $V = 1$. We denote by $\tau = (\tau_1 \dots \tau_n)$, the vector of the plurality tallies of the candidates.

Corollary 2 *The plurality tally differences between any two candidates, $\tau_i - \tau_j$, can be uniquely decomposed into two components, one determined by Basic profiles and the other by Reverse profiles.*

Proof: By Theorem 1, we can assume without loss of generality that $p = \sum_{i=1}^{n} a_i B_i^n + \sum_{i=1}^{n} r_i R_i^n + \frac{1}{n!} K^n$, as no other type of component profile contributes to plurality tallies. We also note that only $(n-1)$ of the Basic profiles and $(n-1)$ of the Reverse profiles are independent, that is, $\sum_{i=1}^{n} B_i^n = 0$ and $\sum_{i=1}^{n} R_i^n = 0$.

From Proposition 1, the tallies of B_i^n and R_i^n are in identical direction, for all i. Define $\mathbf{t_i}$ as a vector with $(n-1)!$ as its i-th component and $-(n-2)!$ as all the other components. Denote $\mathbf{1} = (1, 1, \dots 1)$, a n-component vector. Then the plurality tallies for p are given by (see Appendix 7.3)

$$\tau = \sum_{i=1}^{n} (a_i + r_i)\mathbf{t_i} + \frac{1}{n}\mathbf{1} \tag{2}$$

Denote by $\alpha = \sum_{i=1}^{n} (a_i + r_i)(n-2)! - \frac{1}{n}$. A simple manipulation yields,

$$\tau = n(n-2)!\omega - \alpha\mathbf{1} \tag{3}$$

where the n-dimensional vector $\omega = \{a_i + r_i\}_{i=1}^{n}$. The difference in the plurality tallies for the (i, j) pair is therefore given by,

$$\tau_i - \tau_j = n(n-2)!((a_i - a_j) + (r_i - r_j)) \tag{4}$$

where the component $(a_i - a_j)$ is contributed by the Basic profiles and the component $(r_i - r_j)$ is contributed by the Reverse profiles. Δ.

5 Isolating Basic and Reverse profiles

5.1 Basic profile weights and pairwise scores

Equation (4) shows that if the differences $(a_i - a_j)$ are known, then the coefficients $\{r_i\}_{i=1}^{n}$ can be obtained from plurality tallies, subject to a normalization. We begin with the result by Saari (2000a) mentioned earlier.

Saari (Proposition 5, 2000a): $a_{ij} = a_{ij}^T + a_{ij}^C$ where the component a_{ij}^T is determined by the weights of the Basic profiles and the component a_{ij}^C is determined by the weights of all the Condorcet profiles. Other types of profiles contribute nothing towards these values.

The curse of dimensionality makes a direct application of this powerful result difficult, for any arbitrary n. There are $\frac{(n-1)!}{2}$ distinct Condorcet profiles and pairwise score differentials associated with them.

Chandra and Roy, (2013) discusses a technique of extracting the weights of the Basic profiles from pairwise scores that is easily implementable. Let $\{a_{ij}^{(0)}\}_{i \neq j}$ denote a set of initial and given pairwise scores across the candidate pairs, From the set of intital scores we form a set of *revised* scores, $\{a_{ij}^{(1)}\}$, using the following algorithm.

$$a_{ij}^{(1)} = a_{ij}^{(0)} + \frac{1}{2} \sum_{k \neq i,j} (a_{ik}^{(0)} + a_{kj}^{(0)}), \forall i, j \tag{5}$$

We take a given pairwise score difference $a_{ij}^{(0)}$, between candidates i and j and revise it by adding a weighted average of all pairwise score differences between i and other candidates (except j) and between j and other candidates (except i). The main result of Chandra and Roy (2013) shows that the scores, $\{a_{ij}^{(1)}\}$, are free of the Condorcet components and are linearly related to the weights of the Basic profiles. Formally,

Chandra and Roy (2013): $a_{ij}^{(1)} = (1 + \frac{1}{2}(n-2))(a_i - a_j)$, or alternatively, $a_i - a_j = \frac{a_{ij}^1}{(1+\frac{1}{2}(n-2))}$.

The interested reader is referred to the paper for a complete proof and explanation of the main result. The cancellation of the Condorcet components of the pairwise scores follows from a specific mathematical property of Condorcet profiles, namely, $\sum_{j \neq i} a_{ij}^C = 0$ for all i. For each candidate i, the Condorcet components of the scores against all other candidates must sum to zero. The property is also responsible for the intransitive social ranking of candidates that Condorcet profiles are known for. With three candidates, indexed i, j, and k for example, if $a_{ij}^C + a_{ik}^C = 0$ with $a_{ij}^C > 0$, the property implies $a_{ik}^C < 0$ and $a_{jk} > 0$. If i beats j, then j must beat k (to satisfy the equation for j) and k must beat i.

Pairwise scores may lead to intransitive social ranking because by their nature, such scores do not use the full information contained in the entire profile. Instead, they use partial information about the profile in the form of selective binary components of the rank orders. The iteration (5) essentially attempts to restore some of these lost information contained in the original multilateral rank orders. The score difference for the candidate pair (i, j) is revised by placing positive weights on the scores of i against other candidates (the a_{ik} 's) and the scores of j against other candidates (the a_{kj}). Thus, the difference between i and j is *re-assessed* by using all possible indirect evidence concerning i and j against other candidates. If i won against j but lost against k, whereas j won against k, the iteration (5) reduces the margin by which i won against j. In the process, the cyclic components $\sum_{j \neq i} a_{ij}^C = 0$ cancel out. The computational advantage of the algorithm (5) is that forming the revised scores, $\{a_{ij}^{(1)}\}$, require only simple addition operations.

5.2 Normalized Basic profile weights

The Chandra-Roy algorithm enables us to extract the differences $a_i - a_j$ for all (i, j) pairs from the given pairwise scores but not the coefficients a_i themselves, as there are only $(n-1)$ linearly independent Basic profiles. In principle, any one of the coefficients a_i, $i = 1 \dots n$ may be normalized to zero to obtain the remaining $(n-1)$ coefficients. From our point of view, however, it is convenient to choose the normalizing coefficient (the zero coefficient) in such a way that the weights of the remaining $(n-1)$ independent Basic profiles are non-negative. We use the following steps to select the normalizing coefficient.

As the differences $(a_m - a_n)$ are ordered, choose the (m,n) pair for which this difference is maximized. Suppose that $\max_{(m,n)}(a_m - a_n) = (a_i - a_j)$. Note that $(a_i - a_j) \geq 0$ and therefore $(a_j - a_i) = \min_{(m,n)}(a_m - a_n) \leq 0$. Select a_j to be the normalizing coefficient, that is, set $a_j = 0$. Obtain the values of all the remaining a_is from the differences.

Note that $a_j = 0$ implies $a_i \geq 0$. Note that $\forall m \neq j$, $(a_m - a_j) = (a_m - a_i) + (a_i - a_j)$. Since $(a_j - a_i) \leq (a_m - a_i) \leq (a_i - a_j)$, it follows that $(a_m - a_j) \geq 0$, implying $a_m \geq a_j = 0$. Thus all other coefficients are positive.

5.3 Normalized Reverse profile weights

Once the differences $a_i - a_j$ are obtained from pairwise scores using our technique, the differences $r_i - r_j$ can be obtained from equation (4) and plurality tally differences $\tau_i - \tau_j$. As with the Basic profile weights, any one of the r_i coefficients can be normalized to zero to obtain the other coefficients. To obtain a set of non-negative r_i's therefore, we use the same steps as above. Note however, that, in general, the value of j for which a_j is normalized to zero may not be the same as the value of k for which r_k is normalized to zero. In other words, the two normalizations are independent.

We illustrate our decomposition technique with the following 3-candidate example.

EXAMPLE 4: The following preference-profile reported in Balinski and Laraki (2010) reflects voters' preferences for the position of the President of the *Social Choice and Welfare Society*.

Table 9:

Rankings	No. of voters	Rankings	No.of voters
1. $A > B > C$	13	4. $C > B > A$	8
2. $A > C > B$	11	5. $C > A > B$	11
3. $B > C > A$	9	6. $B > A > C$	0

It is straightforward to check that C is the Condorcet winner and the majority rule ranking is $C > A > B$. The Borda ranking is $A > C > B$. The plurality tallies are $A = 24, B = 9$ and $C = 19$ inducing the plurality ranking $A > C > B$. Thus there is no disagreement between the Borda and the plurality rankings. However both conflict with the majority rule ranking. The decomposition methodology provides some additional insight.

The normalized profile is given by $(13/52, 11/52, 9/52, 8/52, 11/52, 0)$.The pairwise scores are $a_{12} = 9/26$, $a_{13} = -1/13$ and $a_{23} = -2/13$. Applying the Chandra-Roy algorithm, the revised pairwise scores are $a_{12}^{(1)} = \frac{10}{26}$, $a_{13}^{(1)} = \frac{1}{52}$, and $a_{23}^{(1)} = -\frac{19}{52}$. The weights of the Basic profiles are obtained from the differences, $\hat{a}_{12} = a_1 - a_2 = \frac{20}{78}$, $\hat{a}_{13} = a_1 - a_3 = \frac{1}{78}$, $\hat{a}_{23} = a_2 - a_3 = -\frac{19}{78}$. As $\max(a_i - a_j) = a_1 - a_2$, we normalize $a_2 = 0$, implying $a_1 = \frac{20}{78}$ and $a_3 = \frac{19}{78}$. The normalized plurality tally differences are $\tau_1 - \tau_2 = 15/52$, $\tau_2 - \tau_3 = -10/52$ and $\tau_1 - \tau_3 = 5/52$. Applying equation (4), $r_1 - r_2 = -25/156$, $r_1 - r_3 = 3/156$ and $r_2 - r_3 = 28/156$. As the maximum difference is $r_2 - r_3$, we set $r_3 = 0$ and obtain $r_1 = 3/156$ and $r_2 = 28/156$.

In this example, $a_1 - a_3$ and $r_1 - r_3$ have the same sign - both are positive. We explain in the next subsection why this observation and the relatively low values of the ratios, $r_1/a_1 = 3/40$ and $r_3/a_3 = 0$ imply that extreme preferences are *not* a problem with this profile. The low values of the Reverse profile coefficients relative to the Basic profile coefficients also explain why the plurality and the Borda Count agree on the social rank order of the candidates.

A more striking feature of this profile is the fact that the Condorcet winner is different from the Borda and plurality winner. Although a detailed discussion of this feature is beyond the scope of this paper, the decomposition results above show that a significant presence of Condorcet profiles is responsible for this. The Condorcet components obtained from the pairwise scores and the revised pairwise scores are, $a_{12}^c = a_{12} - a_{12}^{(1)} = 9/26 - 10/26 = -1/26$; $a_{13}^c = a_{13} - a_{13}^{(1)} = -1/13 - 1/52 = -5/52$ and $a_{23}^c = a_{23} - a_{23}^{(1)} = -2/13 + 19/52 = 11/52$. These account for the difference between the Borda ranking and the majority rule ranking, specifically the switch from A to C as the winning candidate.

REMARK 6: These coefficients a_i, r_i have interpretations that are similar to those assigned to the coefficients of explanatory variables in an ordinary least square decomposition. The coefficient r_i or a_i represents the "pull" of the given profile vector in the direction of R_i^n or B_i^n. Thus it measures the extent to which the given profile vector is influenced or determined by the R_i^n or B_i^n profile. A coefficient r_i bigger than the coefficient a_i implies that the component profile R_i^n has more explanatory power than the component profile B_i^n with respect to the given profile. Alternatively, as any given profile is a padded and thinned $\frac{V}{n!}K^n$ profile, these coefficients measure the "thickness" of the padding and thinning performed relative to the $\frac{V}{n!}$ profile, to obtain the given profile. If $p = \frac{V}{n!}K^n + a_iB_i^n$, a value of $a_i = \frac{V}{n!}$, and $a_j = 0$ for $j \neq i$, the entire profile has the same structure as a B_i^n profile. Thus, when a given profile is a linear combination of many Basic and Reverse profiles, the relative values of the a_i and r_i coefficients provide direct measures of the importance or strength of these component profiles relative to the others. The above normalizations are equivalent to setting the weakest of the Basic and Reverse profiles to have zero weights.

5.4 Measuring social divisions in a multi-dimensional choice setting

The weights $\{r_i\}$ and $\{a_i\}$ can be used in a variety of ways to measure how socially divided an electorate is. Some of these constructs are presented below.

1. CANDIDATE SPECIFIC r_i, $\frac{r_i}{a_i}, a_i \neq 0$

The weight r_i and in particular the ratio $\frac{r_i}{a_i}$ for $a_i \neq 0$ provide useful information about individual candidate i. The higher the value of r_i or $\frac{r_i}{a_i}$, the more polarizing the candidate i is, as it implies a greater influence of the Reverse profile relative to the Basic profile in determining the plurality tally of candidate i. A value of $\frac{r_i}{a_i} > 1$ implies that the Reverse profile favoring i has a greater influence than the Basic profile favoring i. Thus, if candidate i is the winner, he/she wins because of divided electorate that simultaneously loves and hates him/her. If $a_i = 0$ and $r_i > 0$, the plurality tallies are entirely contributed by a profile of voters who either love or hate candidate i.

2. RATIO $\frac{\bar{r}}{\bar{a}}$

Let $\sum_{i=1}^{n} r_i = \bar{r}$ and $\sum_{i=1}^{n} a_i = \bar{a}$. The ratio $\frac{\bar{r}}{\bar{a}}$ provides a measure of how strong the Reverse profiles are relative to the Basic profiles *on average*. The higher this ratio, the more divided the electorate is into groups that simultaneously love or hate specific candidates. The ratio also indicates to what extent Reverse profiles, on average, play a part relative to Basic profiles in determining the social ranking of the candidates under plurality with $\frac{\bar{r}}{\bar{a}} > 1$ indicating that Reverse profiles play a greater role.

3. RATIO $\frac{r_i - r_j}{a_i - a_j}$ FOR CANDIDATE PAIRS (i, j)

From Corollary 2, the plurality tally difference $(\tau_i - \tau_j)$ for the candidate pair (i, j) and $(a_i - a_j)$ have the same sign if $\frac{r_i - r_j}{a_i - a_j} > -1$ and opposite signs if $\frac{r_i - r_j}{a_i - a_j} < -1$, assuming $(a_i - a_j) \neq 0$. Thus $\frac{r_i - r_j}{a_i - a_j} < -1$ implies that the relative plurality ranking of the (i, j) pair is a *strong* reversal of the relative ranking according to the pair's Basic profiles. Since the Borda ranking of the candidates is equivalent to ranking them according to their a_i's, values of $\frac{r_i - r_j}{a_i - a_j} < -1$ for all or some (i, j) pairs imply disagreement between plurality and the Borda Count. Moreover, the disagreement can be imputed to strong Reverse profiles favoring one of the candidates.

If $(a_i - a_j) = 0$ but $(r_i - r_j) \neq 0$, the candidates i and j are equal according to their Basic profiles and the relative plurality rank is determined by their relative Reverse profiles. Thus when $(a_i - a_j) = 0$ but $(r_i - r_j) \neq 0$, the relative plurality ranking of the (i, j) pair is a *weak* reversal of the relative ranking according to their Basic profiles (which is a equality), in particular of the relative ranking according to the Borda method. Once again the disagreement is to be imputed to a strong Reverse profile favoring one of the candidates.

In light of our previous discussion, a count of such strong or weak reversals (along with the ratio $\frac{\bar{r}}{\bar{a}}$) may be used as an assessment of the plurality method itself, under the specific circumstances.

6 Results from the Cambridge City Council Elections

In this section, we test our decomposition method and measures on ballot data from the Cambridge (Massachussetts) City Council elections over the period 1997-2013. These elections are held every two years providing us with nine years of data.

The nine members of the City Council are elected under a proportional representation (PR) method over several counts of the ballots. A candidate is elected if he/she wins a certain proportion of the votes, called a quota. The quota is determined by dividing the total number of valid ballots by ten (the number of candidates to be elected plus one) and adding one to the result. The first count involves determining the plurality tallies of all the candidates. All candidates who reach the quota after the first count are declared elected. Surplus votes (beyond the quota) received by them are transferred to the second choice candidates on the surplus ballots (A formula determines which ballots are selected as surplus ballots). After surplus votes are transferred, candidates who have fewer than fifty votes are eliminated and their votes are transferred to the second choice on these ballots. A new ranking is established of the continuing candidates, after this. The candidate with the lowest number of tallies after the two transfers is declared defeated and his/her ballots are transferred to the next continuing candidate marked on each ballot. Once a candidate reaches the quota, no more ballots are transferred to him/her. The process continues till all nine members are elected.

The present paper does *not* attempt a comprehensive critique of this specific voting procedure. Our objective is to uncover the weights of the Basic and Reverse profiles from the ballot data (which consists of individual voters' rankings of the candidates) and obtain measures of extreme preferences within a local context. To the extent that Reverse profiles influence the plurality tallies of the first count and thereby influences the final outcome, oue analysis throws some incidental light on the final results as well.

The data set has some limitations chief of which is the following. The Cambridge City electoral laws do not require voters to rank all candidates. Voters must rank at least one of them for the first place and are free to rank as many of the others as they like. On an average there are 18 or19 *official* candidates, on the ballot, out of which 9 city council members are elected. Most voters rank about only 4 or 5 candidates. Thus the major limitation of the data set is that the individual voters' rank orders are not complete, as required by our model.

A second minor limitation of the data set is that under the electoral laws, voters also have the right to vote for unofficial candidates (hereafter described as *write-in* candidates) of their choice by writing their names in a designated space on the ballot. [9] We exclude the few ballots where a write-in candidate is ranked first from our analysis. We do this essentially to reduce the number of candidates, so that the ratio $\frac{\tau_i - \tau_j)}{n(n-2)!}$ does not vanish at the specified place after decimal, in our numerical calculations. The exclusion does not affect our qualitative results.

A third minor limitation of the data set is that, for many of the elections prior to 2005, we found several ballots with multiple candidates ranked in the same position (”overvotes”). The problem of overvotes is significantly less beginning with 2005, because of a new practice put in place by the Election Commission that automatically ejects all such ballots and gives the voters another chance to redo their ballots. In keeping with the traditional model of social choice (that assumes strict transitive ranking), we excluded all ballots

[9] These candidates, appear in most cases to be people well known within the very small group of voters who have ranked them but not widely known outside this circle. There is, however, one exception to this observation that happened in the year 2009. In the year 2009, a popular candidate who was successfully elected multiple times previously, failed to file her nomination papers on time and hence was not included in the official list of candidates. The candidate participated as a write-in candidate and ended up being elected. For our analysis for the year 2009, we treated this candidate as an official rather than as a write-in candidate. With the exception of 2009, every election year, there are typically 7-9 write-in candidates who are ranked (anywhere on the ballot) only by a very small set of individual voters.

with multiple candidates placed in the same position. Thus, for the years 1997-2003, on an average about 8-9% of the total ballots were discarded. For the years 2005-2011, this percentage is about 1-2%. The discards (for all of the reasons listed so far) account for some slight differences between our plurality tallies for the candidates and the official plurality tallies of the candidates, after the first count. The working paper version (Roy, Wu and Chandra, 2015) of this work provides more details.

As our decomposition method requires pairwise score differences for *all* candidate pairs, assumptions need to be made about how the candidates who are not ranked by a specific voter, stand relative to each other in that voter's preferences. We make the following two assumptions about the official candidates to overcome this difficulty with the data set. First, we assume that if a voter has not ranked a specific candidate, A (say), then the voter strictly prefers all the candidates that he or she has actually ranked to the candidate A. In other words, unranked candidates are ranked *below* the ranked candidates for any voter. Secondly, if a voter has not ranked two candidates A and B, we assume that the voter prefers A to B with probability half and B to A with probability half. So far as calculating pairwise score differences are concerned, this is equivalent to distributing all voters who have not ranked a specific pair (A,B), equally between A and B. These assumptions do reduce the accuracy of the pairwise score differences that we use to extract the Basic profile weights and to that extent affect the estimates of the Reverse profile weights. However, we consider these assumptions to be most reasonable under the circumstances and present our results as a first attempt to apply the methods and measures discussed in the paper.

In the interest of brevity, the estimated Reverse and Basic profile coefficients for all candidates for all election years are not included here. All such related tables are published in the working paper version of this work and the interested reader is referred to it (see Roy, Wu and Chandra, 2015).

Table 10 presents the values obtained for the average measures, $\bar{r}$ and $\frac{\bar{r}}{\bar{a}}$, discussed in Section 5. We also include two other variants of these measures, $\bar{r_w}$ and $\frac{\bar{r_w}}{\bar{a_w}}$, which indicate these averages amongst the winning candidates (that is, the 9 candidates who were finally elected to the Council). Finally, Ψ denotes the proportion of candidate pairs (in the total number of distinct pairs) that experienced strong or weak rank reversals under plurality, on the basis of the estimated Basic and Reverse profile coefficients. Table 11 selectively reports some of the final outcome oddities that may be attributed to strong Reverse profile components. Specifically, these are cases in which a candidate who was Borda lower ranked by the ballot data got elected but a candidate who was Borda higher ranked, didn't. The complicated system of surplus vote transfers described earlier is undoubtedly partly responsible for these oddities. However, a strong Reverse profile component also helped the candidates who were Borda lower ranked boost their tallies in the initial plurality count and edge past the candidates who were Borda higher ranked. Figure 1 plots the

Table 10 values. The main findings are summarized below.

Summary results

1. The most significant finding is that with the exception of 1997, the values of $\bar{r}$, $\frac{\bar{r}}{\bar{a}}$, $\bar{r_w}$ and $\frac{\bar{r_w}}{\bar{a_w}}$ are generally higher for the period 2005-2013 than for the period 1999-2003. The ratio, $\frac{\bar{r}}{\bar{a}}$, $\bar{r_w}$, in particular shows steady increase from 2007. Moreover, after 2005, extreme preferences seem to have played a generally bigger role in determining the set of winners (compared to before 2005) as evidenced by the values of $\bar{r}_w$ and the ratio $\frac{\bar{r_w}}{\bar{a_w}}$.

2. The measure Ψ does not show a trend over the years but is generally significant at an average of 36% across all years. The interested reader is referred to the working paper version of this paper for details of the specific rank reversals for every year.

Amongst the Table 11 entries, the year 2009 is specially interesting. A candidate who was Borda last ranked (21st) edged out Borda 4th and 9th ranked to get elected. This candidate was 1st ranked according to the r_i coefficients and 5th ranked by the initial plurality counts. Thus the candidate got elected because of a strong r_i profile working in his/her favor. There were no such rank reversals during the 2001 and 2003 elections. All the first nine Borda ranked candidates got elected to the Council. Interestingly enough, the measures $\frac{\bar{r}}{\bar{a}}$, $\bar{r_w}$ and $\frac{\bar{r_w}}{\bar{a_w}}$ were also noticeably lower during these two years than during the others. Furthermore, in 2013, four major rank reversals are estimated - higher than the number for any other year. Also noticeably, the values of $\bar{r}$, $\frac{\bar{r}}{\bar{a}}$ and $\bar{r_w}$ are significantly higher for this year than for the other years.

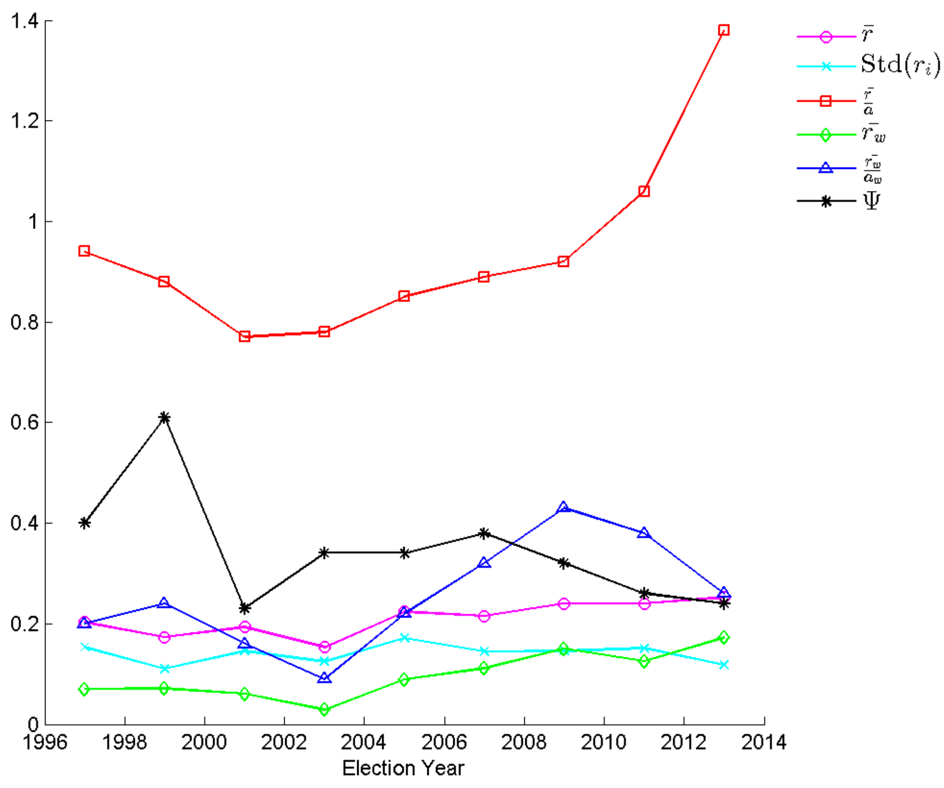
Table 10: SUMMARY RESULTS FOR THE PERIOD 1997-2013

Years	$\bar{r}$	Std(r_i)	$\frac{\bar{r}}{\bar{a}}$	$\bar{r_w}$	$\frac{\bar{r_w}}{\bar{a_w}}$	Ψ
1997	0.203	0.153	0.94	0.070	0.20	0.40
1999	0.173	0.110	0.88	0.071	0.24	0.61
2001	0.193	0.146	0.77	0.061	0.16	0.23
2003	0.153	0.125	0.78	0.029	0.09	0.34
2005	0.224	0.171	0.85	0.089	0.22	0.34
2007	0.215	0.145	0.89	0.111	0.32	0.38
2009	0.240	0.146	0.92	0.150	0.43	0.32
2011	0.240	0.151	1.06	0.126	0.38	0.26
2013	0.253	0.118	1.38	0.172	0.26	0.24

Table 11: SPECIFIC ELECTION ODDITIES

Years	Description
1997	Borda and a_i ranked 9th was edged out by Borda and a_i ranked 11th
1999	Borda and a_i ranked 2nd and 9th were edged out by Borda and a_i ranked 10th and 11th
2001	None
2003	None
2005	Borda and a_i ranked 7th was edged out by Borda and a_i ranked 11th
2007	Borda and a_i ranked 9th was edged out by Borda and a_i ranked 10th
2009	Borda and a_i ranked 4th and 9th were edged out by Borda and a_i ranked 10th and 21st
2011	Borda and a_i ranked 4th and 8th were edged out by Borda and a_i ranked 10th and 11th
2013	Borda and a_i ranked 2nd, 4th 8th and 9th were edged out by Borda and a_i ranked 11th

Figure 1: Summary results for the period 1997-2013



7 Appendix

7.1 Proof of Proposition 1

For this proof, we index the candidates by the lower case letters, $i = 1 \dots n$, as we do in the text and name the candidates with upper case letters, $A, B \dots N$, whenever necessary for exposition. Thus the ith candidate is named I and the jth candidate J.

Part 1: Assume $k = 2$ to start with. Also, without loss of generality, consider the pair (R^n_1, R^n_2). R^n_1 has non-zero voters for A in the 1-st, 2-nd, $(n-1)$-th and n-th places. R^n_2 has non-zero voters for B in the 1-st, 2-nd, $(n-1)$-th and n-th places. The inner product of $(R^n_1)^T$ and R^n_2 have non-zero components for all rankings in which (1) A is in the 1-st place and B is in the 2-nd, $(n-1)$-th or n-th place (2) A is in the 2-nd place and B is in the 1-st, $(n-1)$-th or n-th place (3) A is in the $(n-1)$-th place and B is in the 1-st, 2-nd or n-th place and (4) A is in the n-th place and B is in the 1-st, 2-nd or $(n-1)$-th place. In each of these cases (a total of twelve cases), A and B can be placed in their positions in $(n-2)!$ ways. The relevant components of R^n_1 and R^n_2 belong to the set $\{1, -1\}$. The non-zero components of the inner product equal

$$
\begin{aligned}
-(n-2)! - (n-2)! + (n-2)! - (n-2)! + (n-2)! - (n-2)! - (n-2)! + (n-2)! \\
-(n-2)! + (n-2)! - (n-2)! - (n-2)! = -4(n-2)!
\end{aligned}
$$

Hence R^n_1 and R^n_2 are not orthogonal. By way of illustration, for $n = 3$ and $n = 4$, $(R^3_1)^T R^3_2 = -6$ and $(R^4_1)^T R^4_2 = -8$. The argument extends to all pairs of R^n_i profiles for $k = 2$.

Next note that all the previous steps of the proof apply directly without any changes to any $k < \frac{n+1}{2}$. Now suppose we choose $k = \frac{n+1}{2}$ which can only happen if n is odd. Note that the candidate, I, can be in the k-th place in $(n-1)!$ rankings and that half of these rankings are reversals of the other half. Each such ranking has (-2) voters by construction. The inner product of $(R^n_1)^T$ and R^n_2 have non-zero components for all rankings in which (1) A is in the 1-st place and B is in the $\frac{n+1}{2}$-th or n-th place (2) A is in the $\frac{n+1}{2}$-th place and B is in the 1-st, or n-th place (3) A is in the n-th place and B is in the 1-st, $\frac{n+1}{2}$-th place. The non-zero components of the inner product equal

$$
-2(n-2)! + (n-2)! - 2(n-2)! - 2(n-2)! + (n-2)! - 2(n-2)! \quad = \quad -6(n-2)!
$$

which is not 0. Hence the non-orthogonality claim is true for any k and for all pairs of generic Reverse profiles.

Consider the sum $\sum_{i=1}^{n} R_i^n$ for $k=2$. Only four out of these n profiles at a time contribute non-zero voters for each ranking. Two of the profiles contribute (1) voter each for the first and last places. The other two profiles contribute (-1) each for the 2-nd and $(n-1)$-th places. Hence the sum is 0. Using similar argument, it is clear that the sum of any $(n-1)$ profiles out of the n profiles is not 0. Hence the set spans a $(n-1)$ dimensional subspace, for $k=2$. The steps apply directly without any changes to any $k<\frac{n+1}{2}$. When $k=\frac{n+1}{2}$, three out of these profiles contribute non-zero voters for each ranking at a time. Two of the profiles contribute (1) voter each for the first and last places. The profile contributes (-2) each for the $\frac{n+1}{2}$-th place. Hence the sum is 0

Part 2: Consider the inner product of $(R_i^n)^T$ and B_i^n, for any given k. This has non-zero components for all rankings in which (1) candidate I is in the 1-st place and (2) candidate I is in the n-th or last place. As there are $(n-1)!$ rankings in which candidate I is 1-st ranked and another $(n-1)!$ rankings in which he/she is last ranked, the non-zero components equal $(n-1)!.(1).(1)-(n-1)!.(1).(-1)=0$. Hence this pair is orthogonal to each other.

Next assume that $k=2$ and consider the inner product of $(R_i^n)^T$ and B_j^n, where $i \neq j$. This has non-zero components for all rankings in which (1) candidate J is in the 1-st place and I is in the 2-nd place (2) candidate J is in the 1-st place and I is in the $(n-1)$-th place (3) candidate J is in the n-th place and I is in the 2-nd place and (4) candidate J is in the n-th place and I is in the $(n-1)$-th place. The non-zero components equal $-(n-2)!-(n-2)!+(n-2)!+(n-2)!=0$. Hence these two vectors are orthogonal and the claim is true.

Again, the arguments extend directly without any changes for any $k<\frac{n+1}{2}$. When $k=\frac{n+1}{2}$, the inner product has non-zero components for all rankings in which (1) candidate J is in the 1-st place and I is in the $\frac{n+1}{2}$-th place (2) candidate J is in the n-th place and I is in the $\frac{n+1}{2}$-th place. The non-zero components equal $-2(n-2)!+2(n-2)!=0$. Hence claim is true for any given k.

Part 3: Under a B_i^n profile, candidate I is ranked first $(n-1)!$ times and hence receives as many points. Candidate J receives non-zero votes only for rankings in which he/she is ranked first and candidate I is ranked last. There are $(n-2)!$ such rankings each with (-1) voter. Thus every other candidate receives $-(n-2)!$ points. Under a R_i^n profile, with $k=2$, candidate I is ranked first $(n-1)!$ times and receives as many points. Candidate J receives non-zero votes for every ranking in which (1) J is first ranked and I is second ranked (2) J is first ranked and I is $(n-1)$-th ranked (3) J is first ranked and I is n-th ranked. There are $(n-2)!$ rankings in each category. J receives (-1) for each ranking in the first two categories and (1) for each ranking in the last category. Hence J receives $-(n-2)!$ points.

These tallies remain unchanged for any $k < \frac{n+1}{2}$. For $k = \frac{n+1}{2}$, candidate J receives non-zero votes for every ranking in which (1) J is first ranked and I is $\frac{n+1}{2}$-th ranked (2) J is first ranked and I is n-th ranked. There are $(n-2)!$ rankings in each category. J receives (-2) for each ranking in the first category and (1) for each ranking in the last category. Hence J receives $-(n-2)!$ points.

The total number of voters in a $B_i^n + K^n$ profile is $2(n-1)! + (n-2)(n-1)! = n!$. The total number of voters in a $R_i^n + K^n$ profile is $2(n-1)! + 2(n-1)! + (n-4)(n-1)! = n!$ for $n > 3$. The normalized plurality scores can be derived using the previous steps. Under a R_i^n profile, each ranking and its reversal has the same number of voters. Hence pairwise scores are a complete tie for each candidate pair.

7.2 Profile decomposition for a 3-candidate field

In a 3-candidate election, there are $3! = 6$ possible rankings of the candidates A, B and C who are indexed 1, 2 and 3, respectively. The rank orders are indexed according to the convention of Table 1.

Denote A as candidate 1, B as candidate 2 and C as candidate 3. The Basic profile differentials are given by, $B_1^3 = (1,1,0,0,-1,-1)$, $B_2^3 = (0,-1,1,-1,1,0)$, and $B_3^3 = (-1,0,-1,1,0,1)$. Note that $B_1^3 + B_2^3 + B_3^3 = 0$ and hence only two of these profile differentials are independent. We choose them to be B_1^3 and B_2^3.

Each B_i^3 has the same election outcomes as $B_i^3 + K^3$ where $K^3 = (1,1,1,1,1,1)$. Under the profile $B_1^3 + K^3 = (2,2,1,0,0,1)$, A wins over B or C under any pairwise or positional voting procedure. B and C are tied under any procedure. Thus everyone in this group of voters likes A best and is indifferent between the others.

There is a unique reference ranking and only one distinct Condorcet profile for a 3-candidate field. The Condorcet 3-tuple $c_{(1)}^3$ and its reversal set $\rho(c_{(1)})^3$ are the set of rankings in the table below.

Table 12:

$c_{(1)}^3$	$\rho(c_{(1)})^3$
$A > B > C$	$C > B > A$
$B > C > A$	$A > C > B$
$C > A > B$	$B > A > C$

The Condorcet profile for a 3-candidate field is described by the vector, $C^3 = (1,-1,-1,1,1,-1)$.

There are three possible pairs of candidates in a 3-candidate field and hence three possible pairwise scores differences. Thus the set of normalized pairwise score differences (as defined in Section 3) is a 3-dimensional cube with each side given by the interval $[-1,1]$. A vector of pairwise score differences in this cube is represented as $a = (a_{12}, a_{13}, a_{23})$, where a_{12} is the pairwise score difference between A and B, a_{13} is the pairwise score difference between A and C and a_{23} is the pairwise score difference between B and C.

The vector of pairwise score differences generated by the three Basic profiles are $T_1^3 = (1,1,0)$, $T_2^3 = (-1,0,1)$, and $T_3^3 = (0,-1,-1)$. Under B_1^3, A unanimously beats B and C who are tied. Hence in T_1^3, $a_{12} = a_{13} = 1$ and $a_{23} = 0$. Under B_2^3, B unanimously beats A and C who are tied in turn. Hence in T_2^3, $a_{12} = -1$, $a_{13} = 0$ and $a_{23} = 1$. Under B_3^3, C unanimously beats A and B who are tied in turn. Hence in T_3^3, $a_{12} = 0$, $a_{13} = -1$ and $a_{23} = -1$.

The unique Condorcet profile generates the normalized pairwise score differences vector $q = (a_{12}, a_{13}, a_{23}) = (1,-1,1)$. q illustrates the well known Condorcet paradox - under the social pairwise rank order, A unanimously beats B, B unanimously beats C and C unanimously beats A. As the number of distinct Condorcet profiles increase very rapidly with the candidates, characterizing the distinct profiles as well as the directional vectors for the normalized pairwise score differences for each profile becomes difficult.

With three candidates, Definition 1 implies $k = 2$. Of the three possible Reverse profiles, two can be chosen to be independent and we choose them to be, $R_1^3 = (1,1,-2,-2,1,1)$ and $R_2^3 = (-2,1,1,1,1,-2)$.

Note that the set of vectors $(K^3, B_1^3, B_2^3, R_1^3, R_2^3, C^3)$ form an orthogonal basis for any given profile in $R^{n!}$. Assume V voters as before. Then a given profile can be expressed as $p = \frac{V}{6}K^3 + aB_1^3 + bB_2^3 + cR_1^3 + dR_2^3 + eC^3$.

The number of voters favoring each possible ranking within the profile turns out as follows:

Table 13:

ranking	no.of voters	ranking	no. of voters
$A > B > C$	$(V/6 + a + c - 2d + e)$	$C > A > B$	$(V/6 - b - 2c + d + e)$
$A > C > B$	$(V/6 + a - b + c + d - e)$	$B > C > A$	$(V/6 - a + b + c + d + e)$
$B > A > C$	$(V/6 + b - 2c + d - e)$	$C > B > A$	$(V/6 - a + c - 2d - e)$

The pairwise election tallies for the pair (A,B) turn out to be $(A:B) = ((V/2 + 2a - 2b + e) : (V/2 + 2b - 2a - e))$. Similarly, $(A:C) = ((V/2 + 2a - e) : (V/2 - 2a + e))$, and $(B:C) = ((V/2 + 2b + e) : (V/2 - 2b - e))$. Note that the pairwise tallies depend on the relative weights of the two relevant Basic profiles and the weight of the Condorcet profile. The weights of the Reverse profiles do not contribute to these tallies because such profiles produce a complete tie under pairwise comparisons. The pairwise score differences turn out to be, $a_{12} = \frac{4a-4b+2e}{V}$, $a_{13} = \frac{4a-2e}{V}$ and $a_{23} = \frac{4b+2e}{V}$. Each pairwise score can be broken up into two components - one attributable to Basic profiles and another to Condorcet profiles. For example, for a_{12}, the component attributable to the Basic profiles is $4a - 4b$ and the component attributable to the

Condorcet profile is $2e$.

Mathematically, the plurality tallies for a 3-candidate field is a linear map, $p \in R^6 \longrightarrow R^3_+$ given by the matrix,

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

The plurality tallies of A, B and C under p are given by the vector,

$$Ap = \begin{bmatrix} (V/3+2a+2c-b-d) \\ (V/3-a+2b-c+2d) \\ (V/3-a-b-c-d) \end{bmatrix}$$

.

Note that these tallies are determined by the weights of the Basic and Reverse profiles only. In particular the Condorcet profile contributes nothing. The difference in the plurality tallies of any two candidates can be decomposed into two components - one contributed by the Basic profiles and another by the Reverse profiles. Thus the plurality tally difference between A and B is given by, $3(a-b)+3(c-d)$ with the component, $3(a-b)$, contributed by the Basic profiles and the component, $3(c-d)$, contributed by the Reverse profiles.

Very importantly, note that the plurality tallies obtained under the Basic profile, aB^3_1, point in the same direction as those obtained under the Reverse profile, cR^3_1, as,

$$AaB^3_1 = \begin{bmatrix} 2a \\ -a \\ -a \end{bmatrix}, \; AcR^3_1 = \begin{bmatrix} 2c \\ -c \\ -c \end{bmatrix}$$

.

Under both profiles, A is ranked higher than B and C who are tied. Similarly, the plurality tallies obtained under the Basic profile, bB^3_2, point in the same direction as those obtained under the Reverse profile, dR^3_2.

$$AbB^3_2 = \begin{bmatrix} -b \\ 2b \\ -b \end{bmatrix}, \; AdR^3_2 = \begin{bmatrix} -d \\ 2d \\ -d \end{bmatrix}$$

.

Under both profiles, B is ranked higher than A and C who are tied. The last two observations make clear why plurality tallies by themselves can be confounding and point towards a need for a decomposition technique that will identify the weights a, b, c and d.

7.3 Step details of Corollary 1

The profile is given by $p = \sum_{i=1}^{n} a_i B_i^n + \sum_{i=1}^{n} r_i R_i^n + \frac{V}{n!} K^n$. The plurality tally of candidate i, is $(a_i + r_i)(n-1)! - \sum_{j \neq i, j=1}^{n} (a_j + r_j)(n-2)! + \frac{V}{n}$. Arranging terms and using the definitions of $\mathbf{t_i}$ and $\mathbf{1}$, we then get equation(2).

Next note that,

$$(a_i + r_i)(n-1)! - \sum_{j \neq i, j=1}^{n} (a_j + r_j)(n-2)! + \frac{V}{n}$$

$$= (n-1)(n-2)!(a_i + r_i) - \sum_{j \neq i, j=1}^{n} (a_j + r_j)(n-2)! + \frac{V}{n}$$

$$= n(n-2)! - (\sum_{j=1}^{n} (a_j + r_j)(n-2)! - \frac{V}{n})$$

The definitions of ω and α then yields equation (3).

References

ABRAMOWITZ, A. I., AND K. L. SAUNDERS (2008): "Is Polarization a Myth?," *The Journal of Politics*, 70.

BALINSKI, M., A. JENNINGS, AND R. LARAKI (2009): "Monotonic Incompatibility Between Electing and Ranking," *Economics Letters*, 105.

BALINSKI, M., AND R. LARAKI (2007): "A Theory of Measuring, Electing and Ranking," *Proceedings of the National Academy of Sciences*.

——— (2010): *Majority Judgement*. The MIT Press.

BOXELL, L., M. GENTZKOW, AND J. M. SHAPIRO (2017): "Is the Internet Causing Political Polarization? Evidence from Demographics," *NBER Working Paper Series*, (23258).

CHANDRA, A., AND S. ROY (2013): "On Removing Condorcet Effects from Pairwise Elections Data," *Social Choice and Welfare*, 40(4).

DELLIS, A. (2009): "Would letting people vote for multiple candidates yield policy moderation?," *Journal of Economic Theory*, 144.

DUCLOS, J., J. ESTEBAN, AND D. RAY (2004): "Polarization: Concepts, Measurement, Estimation," *Econometrica*, 72, 1737–1772.

ESTEBAN, J., AND D. RAY (1994): "On the Measurement of Polarization," *Econometrica*, 62, 819–851.

FIORINA, M., AND S. J. ABRAMS (2008): "Political Polarization in the American Public," *The Annual Review of Political Science*, 11.

FIORINA, M., S. J. ABRAMS, AND J. C. POPE (2008): "Polarization in the American Public: Misconceptions and Misreadings," *The Journal of Politics*, 70.

HAIMANKO, O., M. LEBRETON, AND S. WEBER (2007): "The Stability Threshold and Two Facets of Polarization," *Economic Theory*, 30, 415–430.

HODGE, J. K., AND R. E. KLIMA (2005): *The Mathematics of Voting and Elections: A Hands-On Approach*. American Mathematical Society.

KRASA, S., AND M. POLBORN (2014): "Policy Divergence and Voter Polarization in a Structural Model of Elections," *Journal of Law and Economics*, 57, 31–76.

MASON, L. (2015): ""I Disrespectfully Agree" : The Differential Effects of Partisan Sorting on Social and Issue Polarization," *American Journal of Political Science*, 59.

MONTALVO, J. G., AND M. REYNAL-QUEROL (2008): "Discrete Polarization with an Application to the Determinants of Genocides," *Economic Journal*, 118, 1835–65.

ROY, S., K. C. WU, AND A. CHANDRA (2015): "Uncovering the "Will of the People": Measuring Preference Polarization among Voters," Dept. of Economics Working paper 15001.

SAARI, D. G. (1992): "Millions of election outcomes from a single profile," *Social Choice and Welfare*, 9, 277–306.

——— (1999): "Explaining All Three Voting Outcomes," *Journal of Economic Theory*, 87, 313–355.

——— (2000a): "Mathematical Structure of Voting Paradoxes I. Pairwise Votes," *Economic Theory*, 15, 1–53.

——— (2000b): "Mathematical Structure of Voting Paradoxes II. Positional Methods," *Economic Theory*, 15, 55–102.

TESTA, C. (2012): "Is Polarization Bad?," *European Economic Review*, 56, 1104–1118.

WOO, J. (2005): "Social Polarization, Fiscal Instability and Growth," *European Economic Review*, 49, 1451–1477.

YOUNG, H. P. (1974): "A Note on Preference Aggregation," *Econometrica*, 42, 1129–1131.

——— (1975): "Social Choice Scoring Functions," *SIAM Journal of Applied Mathematics*, 28, 824–838.

——— (1988): "Condorcet's Theory of Voting," *American Political Science Review*, 82, 1231–1244.

YOUNG, H. P., AND A. LEVENGLICK (1978): "A Consistent Extension of Condorcet's Election Principle," *SIAM Journal of Applied Mathematics*, 35, 285–300.