(Running Title: Exploratory Analysis of Social Choice Rules)

EXPLORATORY ANALYSIS OF SIMILARITIES BETWEEN COMMON SOCIAL CHOICE RULES

John C. McCabe-Dansted and Arkadii Slinko

The address for correspondence:

Dr Slinko, Department of Mathematics, University of Auckland, Private Bag 92019, Auckland, NEW ZEALAND.

http://www.math.auckland.ac.nz/~slinko/

slinko@math.auckland.ac.nz

Abstract. Nurmi (1987) investigated the relationship between voting rules by determining the frequency that two rules pick the same winner. We use statistical techniques such as hierarchical clustering and multidimensional scaling to further understand the relationships between rules. We also investigate how the relationships change when elections with Condorcet winners are excluded from the frequency data, and when the homogeneity of the voting population is increased. Most common social choice rules are included, 26 in total.

Keywords: Voting, Social Choice, Homogeneity, Clusters, Multidimensional Scaling, Condorcet

JEL Classification: D7

1. Introduction

There are many possible voting rules (procedures) and the choice of voting rule does affect the outcome of votes. It has been shown that for various sets of desirable properties, no one voting rule can satisfy them all (Arrow 1963; Gibbard 1973; Satterthwaite 1975). This clearly makes the choice of voting rule an important one.

There are many feasible voting systems, with much debate as to which is best. In fact 25 existing (and one new) voting rules are considered in this paper. The behaviour of voting rules often cannot be predicted by simple inspection of their definitions.

Nurmi (1987, 1990) was the first to use simple computational techniques to determine similarities between the rules. He examined the frequency that two rules chose the same winner for several pairs of rules. In his investigation the number of agents ranged from five to 301 and the number of alternatives from three to seven. He concluded that dissimilarities between the rules are significant. It is therefore desirable to have robust methods for classifying voting rules.

Since all comparisons in Nurmi's study were pairwise, the whole picture could not be seen. In 1999 Slinko came up with the idea to use cluster analysis to understand relationships between the rules and through the construction of hierarchical trees to try to represent the whole universe of rules. The first such attempt was undertaken in Leung's thesis (2001) and published in Slinko and Leung (2003).

Their work was extended by Shah (2003), who used multidimensional scaling (see section 3.6) to analyse the differences between the rules. He found that multidimensional scaling required far more than three dimensions to fully capture the relationships between voting rules. Shah also demonstrated the effectiveness of dendrograms for visualising the results of hierarchical clustering of voting rules. He used the average clustering method to generate the hierarchical clusters.

Another attempt to understand similarities of more than two rules was undertaken in Gehrlein and Lepelley (2000) and Merlin et al. (2000), where they calculated the probability that all scoring rules, all voting rules, respectively, elect the same winner.

In this paper, we use hierarchical clustering to examine the effect of the homogeneity of the populations on the similarity of outcome of different voting rules. The effect of excluding elections with Condorcet Winners is also considered in order to understand how similar the rules are in resolving the difficulty presented by the absence of a Condorcet winner.

Our method of classification is based on Single Clustering (described below) that proved to be relatively stable when tested on our random populations

It is investigated how to visualise this data effectively. As visualisations of data can be misleading, a formal test is developed to determine if changes in the clusters produced are statistically significant. Based on this test, reports are generated that list various statistically significant changes.

2. Social Choice Rules

Let A and $\mathcal{N}$ be two finite sets of cardinality m and n respectively. The elements of A will be called alternatives, the elements of $\mathcal{N}$ agents. We assume that the agents have preferences over the set of alternatives. By $\mathcal{L} = \mathcal{L}(A)$ we denote the set of all linear orders on A; they represent the preferences of agents over A. The elements of the Cartesian product

$$\mathcal{L}(A)^n = \mathcal{L}(A) \times \ldots \times \mathcal{L}(A) \quad (n \text{ times})$$

are called n-profiles or simply profiles. They represent the collection of preferences of an n-element society of agents $\mathcal{N}$. If a linear order $R_i \in \mathcal{L}(A)$ represents the preferences of the i-th agent, then by aR_ib, where $a, b \in A$, we denote that this agent prefers a to b.

A family of mappings $F = \{F_n\}$, $n \in \mathbb{N}$,

$$F_n \colon \mathcal{L}(A)^n \to A,$$

is called a social choice function (SCF). Historically SCFs were often called rules.

2.1. **Majority Relation and Advantages.** A great number of voting rules are based on computing the majority relation. Let $R = (R_1, \ldots, R_n)$ be a profile. The majority relation M_R on A is the binary relation on A such that for any a_k and a_ℓ in A we have $a_k M_R a_\ell$ if and only if

$$\operatorname{card}(\{i \mid a_k R_i a_\ell\}) \geq \operatorname{card}(\{i \mid a_\ell R_i a_k\}).$$

When n is odd, the majority relation is a tournament on A, i.e., complete and asymmetric binary relation.

If a voting rule is based on the majority relation, then, given a profile R, one computes M_R first and then implements one or another algorithm to determine the winner using the information contained in M_R. These rules have category $C1$ in Fishburn's classification. They are also called tournament solutions (see the book Laslier (1997) devoted to them).

Some rules for their definition need a more refined information contained in the matrix $N_R = (n_{k\ell})$, where $n_{k\ell} = \operatorname{card}(\{i \mid a_k R_i a_\ell\})$. Many of them use the numbers

$$\operatorname{adv}(a_k, a_\ell) = \max(0, n_{k\ell} - n_{\ell k}),$$

which will be called advantages.

These rules have category $C2$ in Fishburn's classification. Any rule that needs more information that is contained in N_R is of category $C3$. Any rule that requires more information than is contained in R will be assigned category C4 (this is our addition to Fishburn's classifications).

2.2. **Scores.** Many social choice rules are based on calculating the vector of *scores* which are used either for selection of an alternative with the best score or repeated elimination of the alternative with the worst score.

2.2.1. *Tournament Scores.* Let R be a profile and $M = M_R$ be the matrix of its majority relation. Let the m-dimensional vector $\mathbf{1}_m = (1, 1, \ldots, 1)$ be the vector consisting of m ones. Then the vector

$$Sc_{\mathbf{c}}^{(k)} = M^k \mathbf{1}_m$$

is the vector of Copeland scores of order k. The ith coordinate of it $Sc_{\mathbf{c}}^{(k)}(a_i)$ is the Copeland score of order k of the alternative a_i. The first order Copeland score of a_i will be denoted $Sc_{\mathbf{c}}(a_i)$.

The normalised eigenvector belonging to the Frobenius eigenvalue of M is called the vector of long path scores and denoted $Sc_{\mathbf{lp}}$. Its ith component is the long path score of a_i and it can be found as

$$Sc_{\mathbf{lp}}(a_i) = \lim_{k\to\infty} \frac{Sc_{\mathbf{c}}^{(k)}(a_i)}{Sc_{\mathbf{c}}^{(k)}(a_i) \cdot \mathbf{1}_m}.$$

It is also easy to observe that the matrix

$$P = \frac{1}{m-1}\left(M + D\right),$$

where D is the diagonal matrix with $d_{ii} = Sc_{\mathbf{c}}(a_i)$, is stochastic. Then the stationary state of the corresponding Markov process is called the vector of Markov scores and its ith component is the Markov score $Sc_{\mathbf{ma}}(a_i)$.

2.2.2. *Positional Scores.* Let

$$w_1 \geq w_2 \geq \ldots \geq w_m \geq 0$$

be m real numbers to which we shall refer to as *weights*. Given the profile $R = (R_1, \ldots, R_n)$, the score of an alternative $a \in A$ is computed now as follows. Let $\mathbf{w} = (w_1, \ldots, w_m)$ and let $\mathbf{v} = (i_1, \ldots, i_m)$, be the vector, where i_k is the number of times that the alternative

a was ranked kth best by agents from $\mathcal{N}$. Then the score for this particular vector of weights $\mathbf{w}$ is calculated as follows:

$$Sc_{\mathbf{w}}(a) = \mathbf{w} \cdot \mathbf{v} = w_1 i_1 + \ldots + w_m i_m.$$

The most popular scores are:

- Plurality score $Sc_{\mathbf{p}}(a)$, where $\mathbf{p} = (1, 0, \ldots, 0)$.
- Borda score $Sc_{\mathbf{b}}(a)$, where $\mathbf{b} = (m-1, m-2, \ldots, 1, 0)$.
- Antiplurality score $Sc_{\mathbf{a}}(a)$, where $\mathbf{a} = (1, \ldots, 1, 0)$.
- k-Approval score $Sc^{(k)}_{\mathbf{ap}}(a)$, where $\mathbf{ap} = (1, \ldots, 1, 0, \ldots, 0)$ (with k ones).

2.2.3. *Advantages-based Scores.* The Dodgson score $Sc_{\mathbf{d}}(a)$ of an alternative a is the minimum number of neighbouring preferences that must be swapped to make a a Condorcet winner.

The Raynaud score $Sc_{\mathbf{r}}(a_i)$ of an alternative a_i is

$$Sc_{\mathbf{r}}(a_i) = \max_{j \neq i} \operatorname{adv}(a_i, a_j).$$

The Simpson score $Sc_{\mathbf{s}}(a_i)$ of an alternative a_i is

$$Sc_{\mathbf{s}}(a_i) = \max_{j \neq i} \operatorname{adv}(a_j, a_i).$$

The Dodgson Quick (DQ) score $Sc_{\mathbf{q}}(a_i)$ of an alternative a_i is

$$Sc_{\mathbf{q}}(a_i) = \sum_{j \neq i} \left\lceil \frac{\operatorname{adv}(a_j, a_i)}{2} \right\rceil.$$

2.3. **Rules.** We list the rules in the alphabetical order. The complete list of rules is given at the end of the section.

Antiplurality rule (AP) - The alternative with the best Antiplurality score is the winner.

The next three rules belong to the class of Approval Voting Rules. These rules take as input not only a profile, but also a threshold separating the approved and disapproved alternatives (Brams and Fishburn 1982). We model this threshold in three different ways and obtain three different rules accordingly:

Approval uniform (AR) - Each voter approves a random number k of her best alternatives, where k is uniformly chosen from $\{1, 2, \ldots, m-1\}$. The alternative with the most approvals is the winner.

Approval Poisson $\mu = 3$ (A3) - Each voter approves her best k alternatives where k is taken from the Poisson distribution with a mean of 3. Pritchard and Slinko (2003) suggested that out of all scoring rules, 3-approval may best approximate approval voting among scoring rules.

$\lfloor \frac{m}{2} \rfloor$*-Approval* (KA) - The alternative with the best $\lfloor \frac{m}{2} \rfloor$-approval voting score is the winner. As shown by Pritchard and Slinko (2003) it possesses certain optimal properties among all scoring rules.

Black rule (BL) - If a Condorcet winner exists, then the Black winner is the Condorcet winner. If no Condorcet winner exists, then the Black winner is the Borda winner.

Borda rule (BO) - The alternative with the greatest Borda score is the winner.

Carey rule (CA) - Eliminate those alternatives with below average plurality scores. Continue this process until one alternative remains (or a set of tied alternatives).

Coombs procedure (CO) - If one alternative's plurality score is greater than $n/2$, then that alternative is the Coombs winner. Otherwise, eliminate the alternative with the lowest Antiplurality score. Continue the process until one alternative remains.

Copeland rule (CP) - The alternative with the greatest first order Copeland score is the winner. In case of a tie, the alternative with the greatest first order Copeland score is the winner if it has the highest second order Copeland score. If several of them are tied again, then the usual tie-breaking is used.

Dictatorship of the first voter (D1)

Dodgson rule (DE) - The alternative with the minimum Dodgson score wins (Black 1958; Tideman 1987).

Dodgson Quick rule (DQ) - The alternative with the minimum DQ-score wins.

Exhaustive procedure (EX) - Eliminate the alternative with the smallest Antiplurality score. Continue until only one alternative remains. This is similar to Coombs procedure except that the procedure is repeated until all but one of the alternatives are eliminated, rather than stopping when one alternative is ranked first by more than 50% of agents, as is the case with Coombs. The distinction between the two rules is noted in Tideman (1987, 191).

Hare rule (HA) (also known as Single Transferable Vote or Alternative Vote) - If one alternative's plurality score is greater than $n/2$, then that alternative is the Hare's winner. Otherwise, eliminate the alternative with the lowest plurality score. Continue until one alternative remains.

Inverse Borda rule (IB) - Eliminate the alternative with the lowest Borda score. The last alternative remaining is the Inverse Borda winner.

Kemeny rule (KE) - Kemeni (1959). This rule uses the concept of a distance between relations to find a linear order whose sum of distances to the agents' linear orders is minimised and then takes the top alternative of that linear order. Also known as Slater rule (e.g. Laslier (1997))

Long path rule (LP) - The alternative with the greatest long path score is the winner. (Laslier 1997, 55).

Majoritarian Compromise (MC) - Let k be the smallest integer for which the set

$$\left\{a \in A \mid Sc^{(k)}_{\mathbf{ap}}(a) \geq \frac{n}{2}\right\}$$

is nonempty. Then the winner is the alternative with the greatest k-Approval score. (Sertel and Yilmaz, 1999, 620). For an odd number of alternatives the procedure was clarified by Slinko (2002).

Markov rule (MA) - The alternative with the greatest Markov score is the winner. (Laslier 1997)

Nanson's procedure (NA) - Eliminate those alternatives with below average Borda scores. Continue this process until one alternative remains or until all remaining alternatives have tied Borda scores. (Other definitions of Nanson exist (Richelson 1981, 346) in which Nanson is defined exactly the same as the Inverse Borda rule above – that definition was not used in this paper. The definition we used is the one used by Tideman (1987, 194)).

Plurality rule (PL) - The alternative with the greatest plurality score is the winner. Also known as First Past the Post.

Plurality Runoff procedure (RU) This is two-stage elimination rule. In the first round two alternatives with the best plurality scores are determined (it might be necessary to break ties for this). In the second round the winner is decided by a simple majority.

Simpson rule (SI) (also known as the minimax/maximin rule) - The alternative with the smallest Simpson score is the winner. E.g. Laslier (1997)

Top cycle rule (TO) - We define a weak linear order (sometimes called a complete preorder) $\succeq_R$ as follows. We say that $a \succeq_R b$ for two alternatives $a, b \in A$ iff there exist alternatives $a_1, \ldots, a_m$ such that

$$aM_Ra_1M_Ra_2M_R \ldots M_Ra_mM_Rb.$$

We have a partition

$$A = A_1 \cup A_2 \cup \ldots \cup A_r$$

of A into indifference classes of $\succeq_R$ and $\succeq_R$ induces a (strict) linear order

$$A_1 \succ_R A_2 \succ_R \ldots \succ_R A_r$$

on these classes. Then the Top Cycle of R is $TC(R) = A_1$ and this rule chooses the whole top cycle. Laslier (1997)

Uncovered set rule (UN) - Let $a, b \in A$. We say that a covers b and denote it $a \sqsupset_R b$, if aR_Mb and

$$\forall_{c\in A} \left[(bM_Rc) \Longrightarrow (aM_Rc)\right].$$

We define the uncovered set of R as

$$UC(R) = \{a \in A \mid \forall_{b\in A}[\,(b \sqsupset_R a) \Longrightarrow (b = a)\,]\}\,.$$

This rule chooses the whole uncovered set. Laslier (1997)

Raynaud procedure (RY) - Eliminate the alternative with the smallest Raynaud score. Repeat until only one alternative, which is the Raynaud winner, is left (Lansdowne 1997, 126).

In the table below we give two letter abbreviations for Social Choice Rules (in alphabetic order) and their respective categories.

[Insert Table 1 about here]

3. Method

3.1. Creating a Profile. For an arbitrary number of alternatives m and an arbitrary number of agents n, we wish to create random profiles of variable homogeneity that we can specify. *Homogeneity is loosely a measure of the tendency of the population to contain a sizable groups of agents, who share the same order of alternative preferences.*

Our population creation method is based on the Pólya-Eggenberger Urn Model (Norman and Samuel 1969).

3.1.1. *Pólya-Eggenberger Urn Model.* Under this model we start with a big urn containing balls, each of a different colour and a non-negative integer a. To generate each random sample, we pull a ball out of the Urn at random and note its colour. After removing each ball, we return the ball that was taken to the urn together with a additional balls of the same colour to the urn.

For generation of random voter profiles, we can replace colours with linear orders. The parameter a characterises homogeneity; for $a = 0$ we obtain the well known Impartial Culture conjecture and for $a = 1$ the so-called Impartial Anonymous Culture conjecture (Berg and Lepelly 1994). In this paper we wish the value of the parameter of homogeneity to have the same meaning for different numbers of alternatives. Therefore we use a normalized parameter $b = \frac{a}{m!}$ as our main parameter of homogeneity. For example, if $b = 1$ the second voter copies the first voter half of the time regardless of the number of alternatives.

3.2. Comparing Outcomes.

3.2.1. *Distance between voting rules.* For a randomly generated set of profiles $\mathcal{P}$, e.g. a sample of profiles created using the same homogeneity, we define distance between rules as the frequency that rules fail to pick the same alternatives. Let $p \in \mathcal{P}$ and $s = |\mathcal{P}|$, then the winning alternative chosen by $rule$ applied to profile p is $W(p, rule)$. We set

$$Distance(rule1, rule2) = \frac{1}{s} \sum_{p \in \mathcal{P}} \Delta(W(p, rule1), W(p, rule2))$$

Where $\Delta(x, x) = 0$ and $\Delta(x, y) = 1$ for $x \neq y$. The distance, so-defined, is a random variable. Thus for each set $\mathcal{P}$ we have a 26×26 matrix of distances.

3.3. **Hierarchical Clustering.** Hierarchical clustering methods group elements together into clusters, and clusters together into super-clusters. An example of a set of hierarchical clusters is the grouping of individual organisms into subspecies, subspecies into species, species into genus and so on. We use the single clustering algorithm to generate hierarchical clusters from the distances defined above.

3.3.1. *Clustering of voting rules.* We use the "Single" clustering method (Duda et al. 2001). As is discussed in section 3.4, this clustering method allows us to develop a statistical test to determine whether the difference between two clusterings could be due to random fluctuations.

Algorithm 3.1. Single clustering method

(1) Start with a minimum spanning tree.
(2) Remove the longest remaining edge e to form two sub-clusters of one of the existing parent clusters.
(3) The distance within the parent cluster is defined as the length of the removed edge e.
(4) If no edges remain then stop, otherwise return to (2).

3.4. **Clustering Methods and Statistical Significance.** Consider a sub-cluster A and its super-cluster B. Let a, b be the distances within A and B, respectively. As the lengths of the edges that the single clustering method removes are monotonically decreasing, $b \geq a$. However random fluctuations may cause the edge, whose length corresponds to a, to become longer than b. In this case the single clustering method will no longer choose A as a sub-cluster of B. We develop a Z-test with the null hypothesis that $\mu(b) \leq \mu(a)$ such that if the Z-score is greater than 3.5 we may safely conclude that the choice of A as a sub-cluster was not merely due to random fluctuations. Hence we consider such clusters to be "stable" clusters.

3.5. **Measures of Distance between Distance Matrices and Trees.** We would like to quantify the magnitude of changes that occur in the distance matrices and dendrograms when we change homogeneity, so we need to define a distance between matrices and a distance between trees.

3.5.1. *Flattened Distance Matrix (FDM) Distance.* The flattened distance matrix is produced from the distance matrix by replacing each element with its position on a sorted list of all elements in the matrix. The distance between the two matrices is defined as the sum of the absolute differences between their elements. For example:

[Insert Table 2 about here]

As we can see from the table above, M_2 and M_3 have a vastly different magnitude. Since the smallest element of M_2 occupies the same position as the smallest element M_3, second smallest element of M_2 occupies the same position as the second smallest element M_3, and so on, their flattened matrices are the same; the FDM distance between M_2 and M_3 will be 0.

This measure of distance has the following advantages; it is not affected by changes in the overall magnitude of the distances; comparing the flattened distance matrices we lose less information than comparing the trees or clusters.

We will use FDM Distance investigating how the effect of increasing homogeneity is compared to the effect of lowering the number of agents.

3.5.2. *Tree Distance.* As the single clustering method is hierarchical, its results may be represented as a tree of clusters known as a Dendrogram.

To compare differences between trees, we define the distance between trees as follows:

Let T_1 and T_2 be two trees with the same set of leaves $l_1, \ldots, l_n$. For each tree T_i $(i = 1, 2)$ we define the edge-distance $d_e^{(i)}(l_j, l_k)$ between leaves l_j and l_k as the number of edges between l_j and l_k in T_i. Then the tree distance between T_1 and T_2 is defined as

$$treedistance(T_1, T_2) = \Sigma_{j<k} \left| d_e^{(1)}(l_j, l_k) - d_e^{(2)}(l_j, l_k) \right|$$

Tree distance has the advantage that it is not affected by changes in the overall magnitude of the distance matrix, and that the distance between tree represents differences in the topology of these trees. Its main weaknesses are that its results may depend upon the clustering method used, and as some of the information in the distance matrix is lost in the tree representation, so this measure of distance is more granular than the FDM measure of distance.

3.6. **Multidimensional Scaling.** Multidimensional scaling (see e.g. Young and Hamer 1987), also known as principle coordinates analysis (Gower 1966), takes a pairwise set of distances and forms a n dimensional map. The distances on this map are the best possible n-dimensional approximation to the original distances used as input to the algorithm.

The example used to demonstrate multidimensional scaling in the R statistical package takes a matrix of the distances between the major cities in Europe, and produces a map of the position of the cities. Because only the distances are given, the correct rotation of the map cannot be calculated. Hence the axes of maps produced are arbitrary.

Shah (2003) used multidimensional scaling to analyse the reasons why some rules are similar to others. In this paper multidimensional scaling is instead used to visualise the effect that changing the parameters of the population has on the distance between the voting rules considered. We take the distances defined in section 3.5 and map these distances onto an ordinary two dimensional page.

4. Results

4.1. **Impartial Culture.** We may visually present the hierarchical clusters that are produced by the single clustering method in the form of the Dendrogram below.

[Insert Figure 1 about here]

The results show that there is a big "Condorcet" cluster

$$\{\text{BL, CP, DE, DQ, IB, KE, LP, MA, NA, RY, SI}\},$$

which contains all $C1$ and $C2$ rules plus the Dodgson rule which has category $C3$. All of them are Condorcet efficient rules, i.e. choose the Condorcet winner if it exists. Since there is a significant percent of the profiles, which have a Condorcet winner, it is expected that these rules will have the tendency to be close to each other. This cluster has two strongly pronounced subclusters {CP, LP, MA} and {DE, DQ, IB, KE, NA, RY, SI}. We conjecture that, when $n \to \infty$ Dodgson rule and Dodgson Quick converge very fast. Simpson and Dodgson rules are also close and we expect them to coincide asymptotically, when $n \to \infty$.

Borda is the closest rule to the Condorcet cluster which reflects its special intermediate status: it is the only point-scoring rule which belongs to C2.

Three other clusters, which are much smaller, are as follows. The first one consists of the Carey, Hare and Plurality Runoff rules {CA, HA, RU}, all are multistage elimination rules which use the plurality score for elimination. The second, consists of the Coombs and Exhaustive procedures {CO, EX}, which are both multistage elimination rules which use the Antiplurality score for eliminating alternatives. The third cluster consists of the Top Cycle and the Uncovered Set {TO, UN}. Both often choose a large set of alternatives and the tie-breaking adopted in this paper might have made them look closer than they are in reality.

Al the remaining rules are "outliers."

It is also very interesting to compare these results with the results obtained with the use of other clustering methods. The R statistical language provides 7 hierarchical clustering methods as follows: 'ward', 'single', 'complete', 'average', 'mcquitty', 'median', 'centroid'. Out of these Ward's method produced the Dendrogram that looked the least like the dendrograms produced using the single clustering method.

[Insert Figure 2 about here]

We see that the clusters singled out by the single clustering method are easily detectable here too. The Dendrogram looks nicer since all rules are initially divided into two big clusters: Condorcet efficient rules and those which are not. However, we have to remember that Ward's method is strongly biased toward producing clusters with roughly the same number of observations. It is also very sensitive to outliers (Milligan 1980).

We have repeated the results of sections 4.1 to 4.3 using each of the seven clustering methods above and thus believe that these results are robust with respect to choice of clustering method.

4.2. **The Effect of Excluding Condorcet Winners.** If we were to choose an electoral system, we may decide that we always want to choose the Condorcet winner when it exists. When deciding which rule to use as a backup, we would only be interested in elections where Condorcet winners do not exist. Thus we may be interested in which clusters form when we only include simulations where no Condorcet winner occurs.

For the following graphs an Impartial Culture with 5 alternatives and 85 agents was used.

[Insert Figure 3 about here]

There does not appear to be a large difference in the topology of this tree, and the tree including Condorcet winners in Figure 1. As we would expect, excluding Condorcet winners had no effect on the clustering of Condorcet Efficient rules. The large number of profiles where Condorcet winners exist holds the Condorcet efficient rules close together. When the Condorcet winners are removed the Condorcet efficient rules become further apart, but they all become further apart by the same amount so the clustering between Condorcet efficient rules is unaffected.

What is surprising is that excluding elections with Condorcet winners has little effect on most of the clusters containing Condorcet inefficient rules such as Antiplurality (AP) and Random Approvals (AR).

Only 3 new stable (as defined in section 3.4) clusters have formed, {D1, TO, UN}, {KA, PL}, and {BL, BO}. Let us discuss the reasons for their appearance.

Dictatorship, Top Cycle and Uncovered Set (D1, TO, UN) - The top cycle and uncovered set rules can pick a very large set of tied winners if no Condorcet winner exists. Since we break ties by the preferences of the first agent, these rules will approximate a dictatorship when no Condorcet winner exists.

K-Approval and Plurality (KA, PL) - The plurality rule chooses the alternative with the most first place preferences, k-approval for 5 alternatives will pick the alternative with the most first and second place votes. It is not obvious why removing Condorcet winners would cause these two rules to cluster together.

Black rule and Borda Rule (BL, BO) - By definition Black's rule picks the Borda count winner if no Condorcet winner exists, so it is unsurprising that these rules would cluster together when we exclude Condorcet winners.

4.3. **Effect of Increasing Homogeneity.**

4.3.1. *Condorcet Efficiency.* Let $\Delta : \mathcal{L}^n \to [0,1]$ be a probability distribution on the set of all n-profiles. Given the rule, we will define the Condorcet efficiency of this rule as the conditional probability that the rule selects a Condorcet winner if there is one.

For the distribution Δ obtained by using a Pólya-Eggenberger model with parameter of homogeneity b we will denote this number as $CE_b(rule)$.

In the following table the Condorcet efficiencies were measured for 10000 profiles generated from a population with 85 agents and 5 alternatives. For each row we used a different parameter of homogeneity b to generate the population.

[Insert Table 3 about here]

Borda and Antiplurality display an interesting behaviour, for which we do not have an explanation, as $CE_b(AP)$ and $CE_b(BO)$ are not monotonic functions of b. We also note that Slinko and Leung (2003) observed a high degree of correlation between the Condorcet efficiency of the rule and its distance from the Condorcet cluster.

4.3.2. *Comparing Clustering Trees with and without Homogeneity.* Below we see two dendrograms obtained from two different sets of 10000 profiles generated using two different parameters of homogeneity $b = 0$ and $b = 1$.

Like Figure 1, the dendrogram below is produced from a set of 10000 profiles with 85 agents. Unlike Figure 1, these profiles were generated with parameter of homogeneity $b = 1$.

[Insert Figure 4 about here]

4.3.3. *Borda Cluster.* A new cluster forms in the partially homogeneous population, containing the Borda (BO) Rule, Approval with uniform random approvals (AR), Approval with Poisson ($\mu = 3$) random approvals (A3). The Borda rule is a point scoring rule: each alternative is given a number of points for each vote, depending on where the alternative is ranked on the vote. Approval voting is similar, giving a point to each alternative the voter approves. Rational agents will always approve their first k preferences, for some k (Brams and Fishburn 1982), and we assume that agents are rational in this sense. The mean approval score for AR is a linear function of the Borda score. For this reason, we conjecture that when there are a large number of agents with the same preferences and k is uniformly random, AR will approximate the Borda rule. Since there are five alternatives, the Poisson with mean 3 roughly approximates the uniform distribution on 1,2,3,4,5 and AR also approximates BO. Borda's cluster forms for $b \geq 0.5$.

4.3.4. *Increasing Distances.* We would expect that the distance between of the voting rules would decrease with b. The more homogeneous the population the more likely a clear winner is to exist, which should be selected by all voting rules. These graphs mostly support that hypothesis, however there are a few exceptions to the rule. The most drastic is the Coombs (CO) and Exhaustive (EX) rules. Where the agents are impartial, these two rules are almost identical, giving the same result in all 10000 simulations. However when the parameter of homogeneity is increased to one, the rules split into two, and are not even in the same cluster.

Coombs and Exhaustive rules (CO-EX). Shah (2003) found that for 85 agents the Coombs and Exhaustive rules are identical, but they diverge as the number of agents decreases. Shah explains that both rules work by successively eliminating alternatives with the most last place votes. The only difference between the rules is that the Coombs rule will stop if an alternative has the most first place votes, while the Exhaustive rule will continue until only one alternative remains. Clearly, if no alternative achieves more than half the first place votes until the last step, both rules will choose the same winner. Shah further explains that the less agents there are, the more likely it is that an alternative will gain more than half the first place votes; causing the Coombs rule to choose a different alternative to the Exhaustive rule.

We can explain our result in a similar fashion. An increase in homogeneity will also increase the chance that an alternative gets a majority of the first place vote.

Antiplurality and other rules (AP). We can also see that the Antiplurality rule remains dissimilar to the other rules once b increases to 1. If a majority of agents votes the same way we would expect that most rules would usually select the majority's first preference. However the Antiplurality rule will only select the majority's first place vote if it also has the least last place votes. Since the majority all vote the same way, they will only lower the Antiplurality score of one of the four other alternatives. There is, therefore, no good reason to believe that the Antiplurality rule will pick the majorities first preference, even if the majority is very large. Hence we would expect a moderate increase in homogeneity to result in no substantial decrease in distance between the Antiplurality rule and other rules.

A large increase in homogeneity will decrease the distance between Antiplurality and the Dictatorship rules. We use the preferences of the first voter to break ties. In a very homogeneous population the most alternatives will be tied, having an Antiplurality score equal to 0. As we break ties according to the preferences of the first voter, the first voter's choice is likely to be chosen; further it is likely that many other agents will have copied the first voter's vote, making the first voter a part of a majority.

The complete list of distances that increased when we increased b to 1 is "BL-AP BO-AP BO-BL EX-AP CP-AP IB-AP LP-AP MC-AP NA-AP SI-AP TO-AP UN-AP RY-AP CO-AP CO-EX MA-AP DE-AP DQ-AP DQ-DE KE-AP".That is, the Black rule (BL) became less similar to Antiplurality (AP), Borda's (BO) rule became less similar to AP, and so on.

We may ask if this list would be the same or similar if we had used different assumptions about the voting population. As Table 4 shows, the list is reduced to its small subset when we reduce the number of agents to 25. However, for small b there are a number of intervals where several distances will increase, most notably the interval $0 \rightarrow 0.5$. For larger b, the distances all decrease as expected.

[Insert Table 4 about here]

In the first row of Table 4, the distances between Borda and Black (BO-BL), between Coombs and Exhaustive (CO-EX), Dodgson quick approximation and Dodson exact (DQ-DE) increase. If we had further reduced the number of agents to 5 then only BO-BL would have increased.

Black and Borda rules (BL-BO). The Black rule will pick the Condorcet winner if it exists, choosing the Borda winner otherwise. Thus we would expect the distance between the Black and Borda rules to increase as the frequency of that a Condorcet winners exist increases, and increase with Borda's Condorcet efficiency.

[Insert Table 5 about here]

As shown by the table above we find that for small b, the frequency of Condorcet winners increases rapidly with b, causing the distance to increase. For larger b, the increase in Borda's Condorcet efficiency will cause the distance to decrease with b. Shah (2003) demonstrated a similar result for varying the number of agents.

The Dodgson Quick approximation and the Dodgson rule (DQ-DE). The DQ-score approximation to the Dodgson score, is a lower bound on the Dodgson Score. This lower bound was developed in this paper as a way to quickly eliminate alternatives which could not possibly be Dodgson Winners, improving the computational efficiency of the Dodgson Exact (DE) algorithm. This approximation works well on the assumption that if $adv(b, a) > 0$ there will be at least one voter that ranks b just over a, allowing us to reduce this advantage by 2 with only one swap of neighbouring alternatives. This assumption is less likely to hold with a homogeneous population, so the approximation diverges from the exact value with an increase in homogeneity.

4.4. **Two Dimensional Map of Changes.** The map of changes in Figure 5 was generated using multidimensional scaling, described in section 3.6. From $b = 0$ to $b = 0.5$ the points, which represent clustering trees corresponding to the parameter b, move down the graph; from $b = 1$ to $b = 3$ there does not appear to be much movement; the graph changes direction and from $b = 5$ to $b = \infty$ the points move to the right.

[Insert Figure 5 about here]

The downwards movement from 0 to 0.5 may represent the changes in clustering that occur when homogeneity is added. That the points do not move significantly as b goes from 1 to 3 may suggest that no further changes occur in the clustering. Where b is 3, the population is very homogeneous, the second voter will copy the vote of the first voter three-quarters of the time. It is plausible that as the population becomes even more homogeneous, it will become impossible to distinguish between the rules, effectively the tree collapses. The change in direction of the graph may represent this different type of change.

This interpretation of the graph suggests the truth of the following statements:

- increasing b up to 0.5 will form new clusters, but increasing it further will not.
- there is little difference between the clusters produced for $b = 1, 2, 3$.
- increasing b beyond 3 will cause clusters to be lost, as they become subsumed by larger clusters.

We shall demonstrate that these statements are mostly true.

Are there new clusters after $b = 0.5$*?* Each element of Table 6 represents the number of new clusters that form as we increase the parameter b from the value at the top of the column to the value at the left of the row. Increasing b from 0 to 1/5 induces 5 new clusters; increasing b from 1/5 to 1/3 induces 3 new clusters; increasing b from 1/3 to 1/2 or from 1/2 to 1 induces one additional cluster; increasing b from 1 to any of the other values presented does not induce any more clusters. This table was generated with the `newclusters` function which returns only clusters that we can be sure did not occur just by chance (see section 3.4).

From Table 6, we see that at least one cluster was added after $b = 0.5$, so it is not strictly accurate to say that no new clusters occur after 0.5. However it does appear that most of the new clusters occur before 0.5. Also our dataset fails to demonstrates the existence of any new clusters occurring after $b = 1$. So it appears that few new clusters occur after $b = 0.5$.

[Insert Table 6 about here]

[Insert Table 7 about here]

Do any changes in the clusters occur between $b = 1$ *and* $b = 3$*?* Each element of the Table 7 represents the number of clusters we know exist for the smaller value of b but not the larger value. From this table we can see that our dataset does not demonstrate the loss of any clusters between $b = 1$ and $b = 3$. Similarly from Table 6, we can see that no new clusters are known to occur between $b = 1$ and $b = 3$. As we expected, no statistically significant changes occur within this interval.

4.5. **Magnitude of Changes.** Looking at Figure 5, the distance between $b = 0$ and the points $b = 1/2, 1, 2, 3$ appears to be around 600. The exact distances are 596, 620, 620 and 630 respectively. As discussed in section 4.4, new clusters are formed till $b \approx 1$. After $b \approx 3$ it becomes impossible to reliably distinguish some of the rules.We are primarily interested in the region where new clusters are being formed. Whether we take this region as ending at 1/2, 1, 2 or 3 we have a distance in the range 596 to 630. This is 300 times the minimum possible change and 17-20% of the maximum distance between two graphs, hence it is significant. The maximum tree distance for a tree with 25 leaves is between 3176 and 3422. The maximum is likely to be 3176 although we do not have a proof of this.

4.5.1. *Magnitude of Distance induced by Change in Number of Agents.* We may wish to compare the magnitude of the effect of varying of parameter of homogeneity b with that of varying the number of agents. As will be discussed below, varying the number of agents between reasonable values has a slightly greater effect than varying the parameter of homogeneity. From Table 8, The distance of from 7 to 1025 agents is 620, the same distance caused by varying b from 0 to 1 or 2.

[Insert Figure 6 about here]

We can see from Figure 6 that the tree with 3 agents is substantially different from the other trees. It would be intuitive to believe that this difference is at least in part due to most of the rules giving the same result when there are only three agents; however as shown in Figure 7 only two pairs of rules are identical. However the rules in the Condorcet cluster {CO, CP, DE, DQ, IB, KE, LP, MA, NA, RY, SI, TO, UN} join together at approximately the same height, making the sub-clustering very unstable.

[Insert Figure 7 about here]

As, shown in Figure 7 above, three agents is enough to generate interesting elections where most of the rules are distinct. Thus we may consider 3 a reasonable number of agents, when comparing the effect of varying b and varying the number of agents within reasonable ranges.

The elements of the matrix presented below represent the distances between the clustering trees produced from populations with different numbers of agents.

[Insert Table 8 about here]

From Table 8 above we can see that the minimum distance between the tree where $n = 3$ and the other trees is 672. This suggests that changing the number of agents from 3 to anything ≥ 5 has more effect than varying b from 0 to 3.

4.6. **Is Varying the Homogeneity Equivalent to Varying the Number of Agents?** As we increase the homogeneity of the Pólya-Eggenberger Urn Model, the number of unique votes decrease. We discovered a number of similarities between reducing the number of agents, and introducing homogeneity in the last section. For example, we found that increasing the homogeneity could increase the distance between the Coombs and Exhaustive rules (CO-EX), which mirrored Shah's (2003) finding that reducing the number of agents could increase the distance between these two rules.

We may see from the following graph that varying the homogeneity has a substantially different effect to varying the number of agents. It uses multi-dimensional scaling to represent the flattened distance matrix distances as closely as possible on a two dimensional surface. The labels on the vertices are of the format b:v, where b is the parameter of homogeneity and v is the number of agents. If varying b were equivalent to varying v, we would expect that the graph would form a line. However, as we can see the graph forms a clearly defined grid. Changing the homogeneity, is represented by a horizontal change, varying the number of agents results in a primarily vertical change. Thus, though similar, the effect of the parameter of homogeneity is not the same as the effect of varying the number of agents.

[Insert Figure 8 about here]

5. Conclusions

The Pólya Eggenberger Urn is an effective model of partially homogeneous voting populations. Introducing homogeneity from the Urn model has some similarities to reducing the number of agents, for example both may cause the Coombs and Exhaustive rules to diverge. There are also a number of differences between introducing homogeneity and reducing the number of agents. For example, introducing moderate homogeneity causes the Borda rule to cluster with the Random Approval rule, but reducing the number of agents does not. For a fixed parameter of homogeneity a the Urn model will not generate a consistent degree of homogeneity. This weakness can be remedied by defining a parameter of homogeneity $b = fracam!$, which is adjusted for the number of alternatives.

The most interesting range of b is 0 to 0.5. In general the distance between most rules decreases as we increase the homogeneity. Within the range 0 to 0.5, increasing the homogeneity actually increases the distance between a number of rules.

Removing elections in which a Condorcet winner occurs resulted in little difference in the clustering of the social choice rules. When Condorcet winners were removed only three new clusters formed.

We have found the single clustering method an effective tool for studying social choice rules. Combined with the test described in section 3.4, the single clustering method can be used to quickly find significant differences between the clustering of different populations, differences that we can know are not due to chance. This certainty provides a foundation for the reports developed for this paper, such as the number of new clusters that form.

We have also found that multidimensional scaling is an effective tool for visualising the effect that varying the parameters, of a randomly generated population, has on the relationships between the voting rules.

6. Acknowledgements

We would like to thank Matthew Jackson for suggesting we study the effect of excluding elections with Condorcet Winners; Vincent Merlin for attracting our attention to the paper by Fishburn (1977) which classified voting rules into the categories C1, C2 and C3, and Hannu Nurmi for helpful comments on the earlier version of this paper.

7. Appendix

Sample dendrograms with increasing homogeneity b follow:

[Insert Figure 9 about here]

[Insert Figure 10 about here]

[Insert Figure 11 about here]

[Insert Figure 12 about here]

[Insert Figure 13 about here]

References

Arrow, K. J.: 1951, 1963, *Social choice and individual values*. New York: Wiley.

Berg, S. and D. Lepelly: 1994, 'On Probability Models in Voting Theory'. *Statist. Neerlandica* **48**, 133–146.

Black, D.: 1958, *Theory of commitees and elections*. Cambridge: Cambridge University Press.

Brams, S. and P. Fishburn: 1982, *Approval Voting*. Birkhauser.

Duda, R. O., P. E. Hart, and D. G. Stork: 2001, *Pattern Classification*. New York: Wiley, 2nd edition.

Fishburn, P. C.: 1977, 'Condorcet Social Choice Functions'. *SIAM Journal on Applied Mathematics* **33**, 3:469–489.

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

Gibbard, A.: 1973, 'Manipulation of Voting Schemes: A General Result'. *Econometrica* **41**.

Gower, J. C.: 1966, 'Some distance properties of latent root and vector methods used in multivariate analysis'. *Biometrika* **53**, 325–328.

Kemeni, J.: 1959, 'Mathematics without numbers'. *Daedalus* **88**, 577–91.

Lansdowne, Z. F.: 1997, 'Outranking methods for multicriterion decision making: Arrow's and Raynaud's conjecture'. *Social Choice and Welfare* **14**, 125–128.

Laslier, J.-F.: 1997, *Tournament solutions and majority voting*. Berlin - New York: Springer.

Leung, W.: 2001, 'A Statistical Investigation of Social Choice Rules'. Master's thesis, Auckland University.

Merlin, V., M. Tataru, and F. Valognes: 2000, 'On the probability that all decision rules select the same winner'. *Journal of Mathematical Economics* **33**, 183–207.

Milligan, G. W.: 1980, 'An Examination of the Effect of Six Types of Error Perturbation on Fifteen Clustering Algorithms'. *Psychometrika* pp. 325–342.

Norman, L. J. and K. Samuel: 1969, *Discrete Distributions*. Boston: Houghton Mifflin.

Nurmi, H.: 1987, *Comparing voting systems*. Reidel.

Nurmi, H.: 1990, 'Computer Simulation of Voting Systems'. *Contemporary Issues in Decision Making* **77:1**, 391–405. Elsevier Science B. V.

Pritchard, G. and A. Slinko: 2003, 'On the Average Minimum Size of a Manipulating Coalition'.
Available http://www.math.auckland.ac.nz/deptdb/dept_reports/507.pdf.

Richelson, J. T.: 1981, 'A Comparative Analysis of Social Choice Functions IV'. *Behavioral Science* **26**, 346–353.

Satterthwaite, M. A.: 1975, *Strategy-proofness and Arrow's Conditions: Existence and Correspondence.*

Sertel, M. R. and B. Yilmaz: 1999, 'The Majoritarian Compromise is Majoritarian-Optimal and Subgame-Perfect Implementable'. *Social Choice and Welfare* **16**, 615–627.

Shah, R.: 2003, 'Statistical Mappings of Social Choice Rules'. Master's thesis, Stanford University.

Slinko, A.: 2002, 'The Majoritarian Compromise in Large Societies'. *Review of Economics Design* **7(3)**, 343–350.

Slinko, A. and W. Leung: 2003, 'Exploratory data Analysis of Common Social Choice Functions II'. *International Conference on the Problems of Control 17-19 June 2003 in Moscow* **1**, 224–228.

Tideman, T. N.: 1987, 'Independence of Clones as a Criterion for Voting Rules'. *Social Choice and Welfare* **4**, 185–206.

Young, F. W. and R. M. Hamer: 1987, *Multidimensional scaling–history, theory, and applications*. Hillsdale, N.J.: L. Erlbaum Associates.

List of Figures

List of Tables

Figure 1. Dendrogram for Impartial Culture with 5 Alternatives and 85 Agents

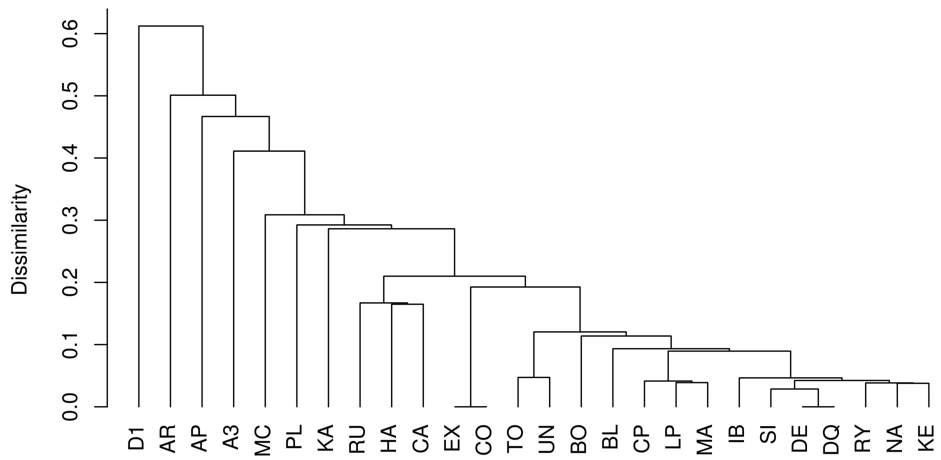


Figure 2. Ward Method Dendrogram for Impartial Culture with 5 Alternatives and 85 Agents

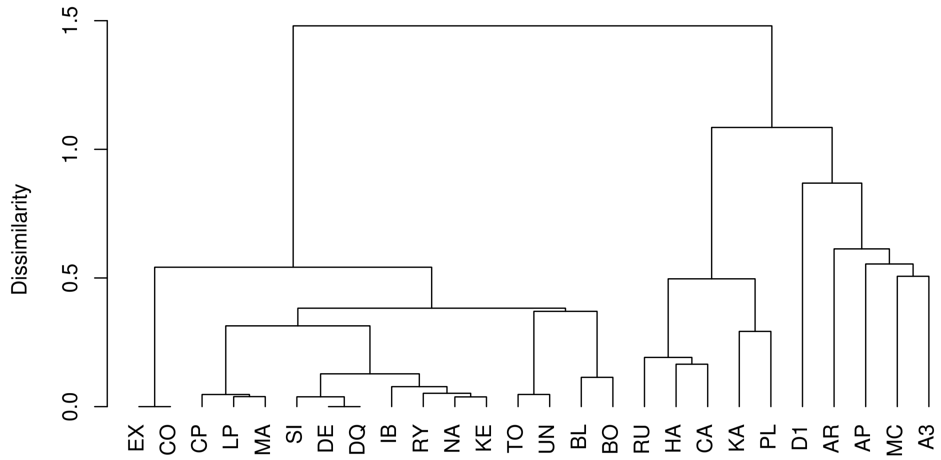


Figure 3. Dendrogram Excluding Condorcet Winners, $b = 0$, $m = 5$ Alternatives, $n = 3$ agents

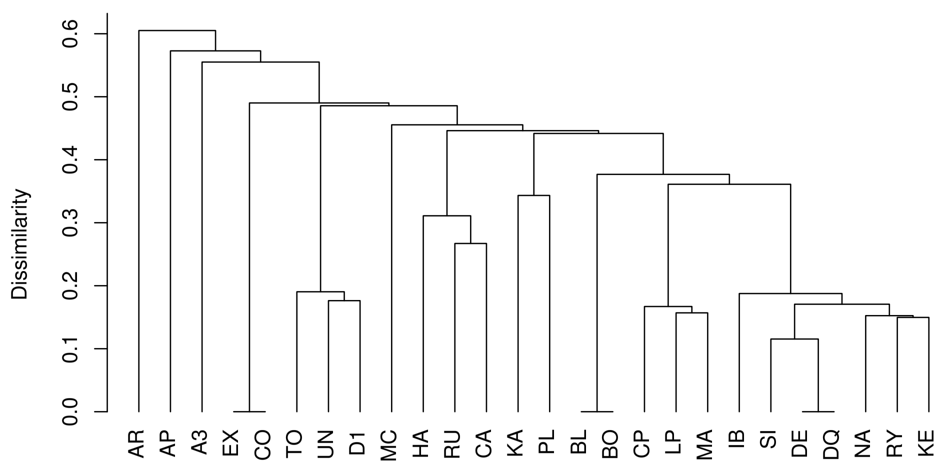


Figure 4. Partially Homogeneous Culture ($b = 1$), with 5 Alternatives and 85 Agents

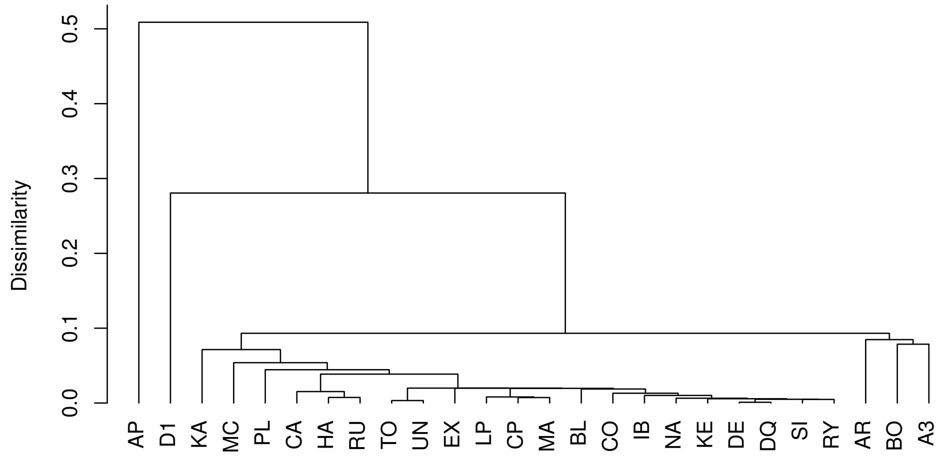
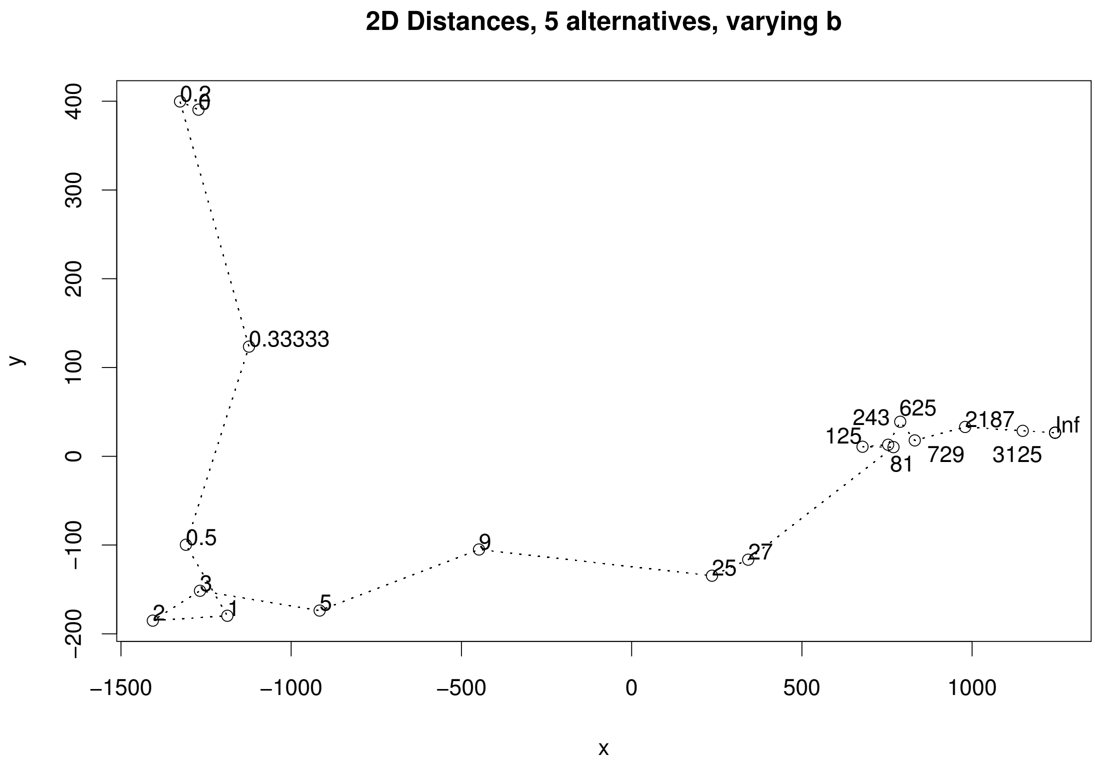
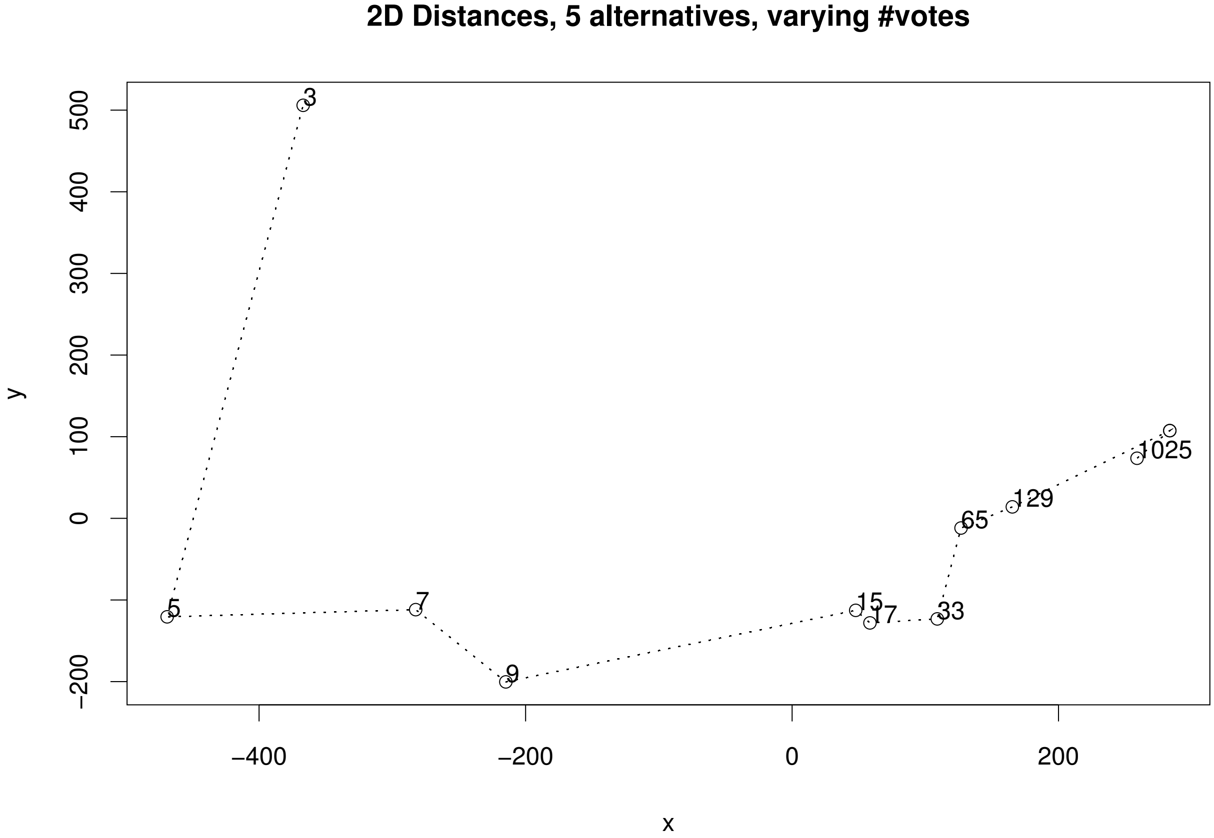


Figure 5. Map of Distances between Clustering Trees, Varying Homogeneity



The x, y axes are unlabelled as they are arbitrary, see section 3.6
Only the distance between points is meaningful

FIGURE 6. Map of Tree Distances, Varying number of Agents, $b = 0$



The x, y Axes are unlabelled as they are arbitrary
Only the distance between points is meaningful

FIGURE 7. Dendrogram for Impartial Culture ($b = 0$), $m = 5$ Alternatives, $n = 3$ agents

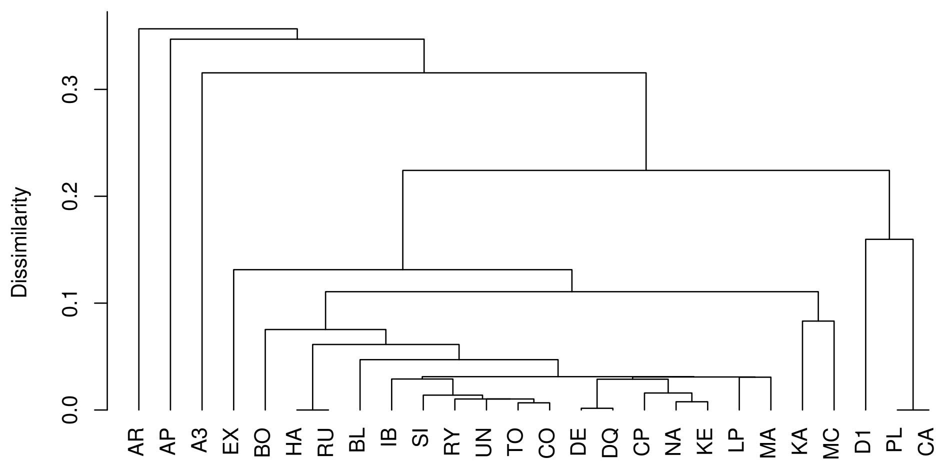


FIGURE 8. Multidimensional Scaling Plot of Parameter of Homogeneity b v.s. Number of Agents

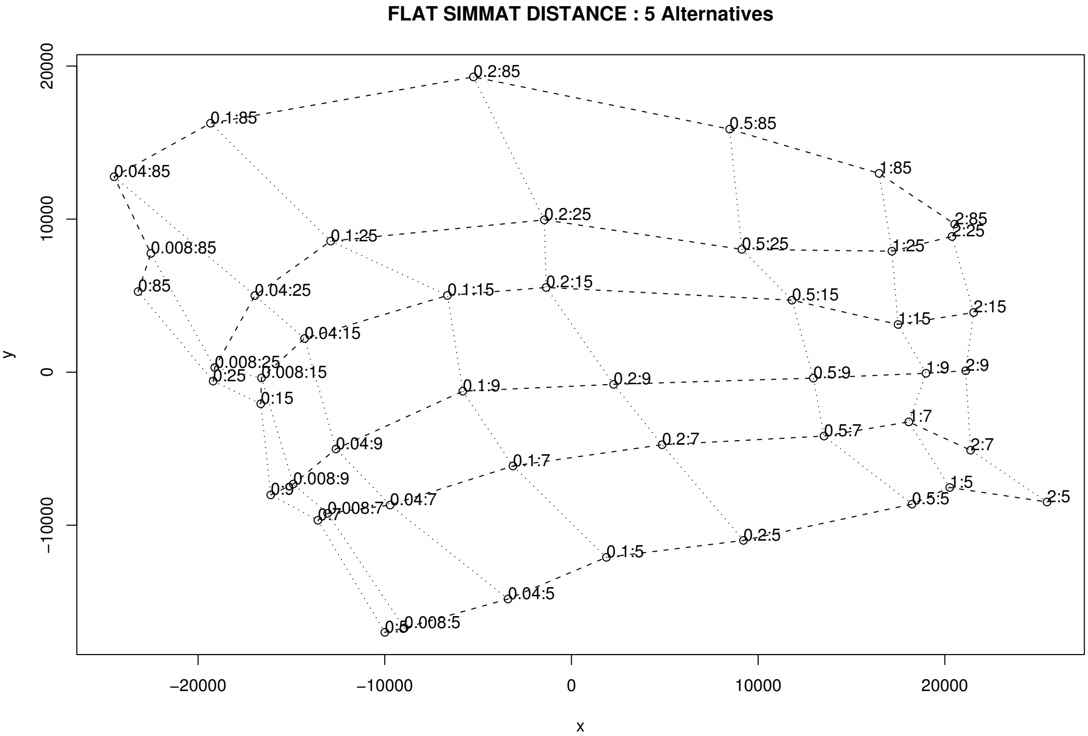


FIGURE 9. $b = 0$ (Impartial Culture), 85 Agents, 7 Alternatives, 10000 samples

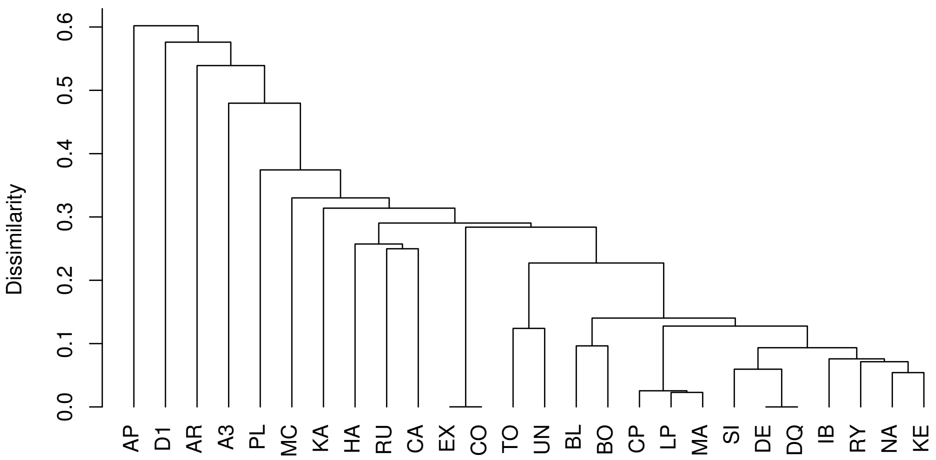


FIGURE 10. $b = 0.1$, 85 Agents, 7 Alternatives, 10000 samples

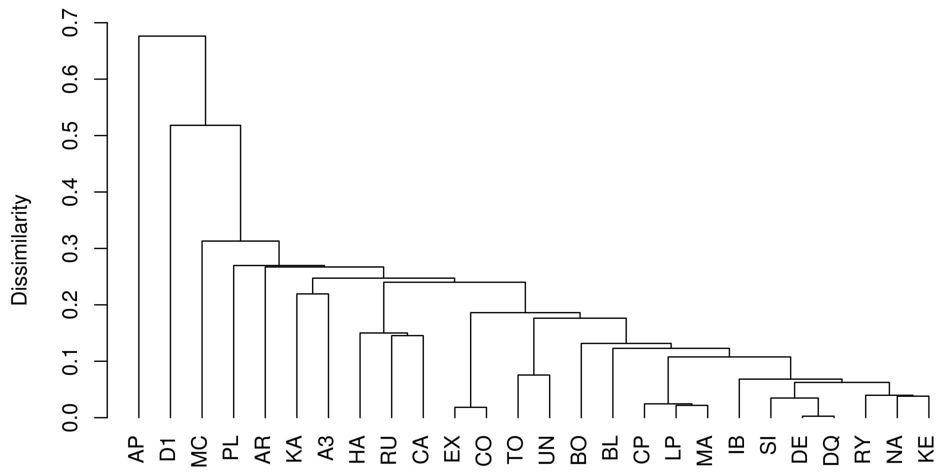


FIGURE 11. $b = 0.2$, 85 Agents, 7 Alternatives, 10000 samples

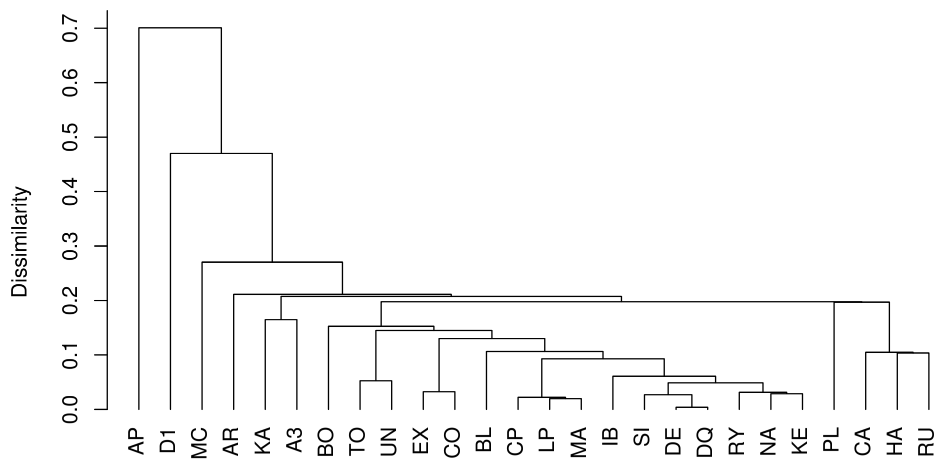


FIGURE 12. $b = 0.5$, 85 Agents, 7 Alternatives, 10000 samples

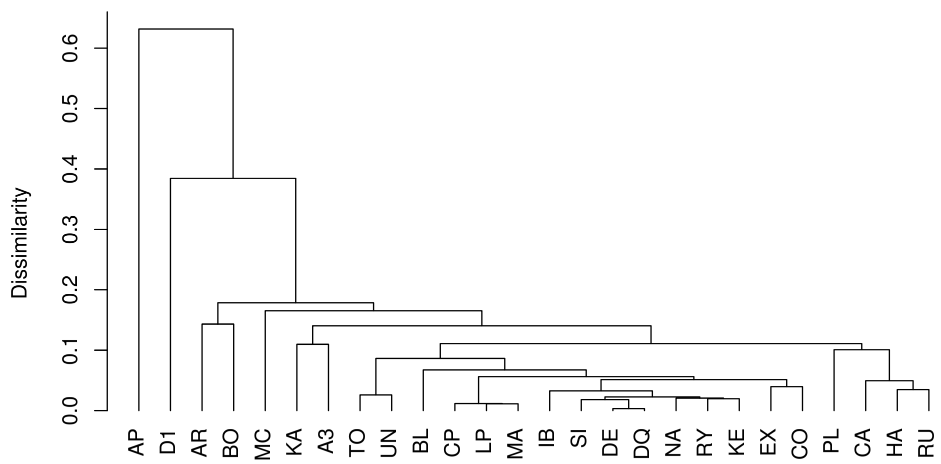


Figure 13. $b = 1$, 85 Agents, 7 Alternatives, 10000 samples

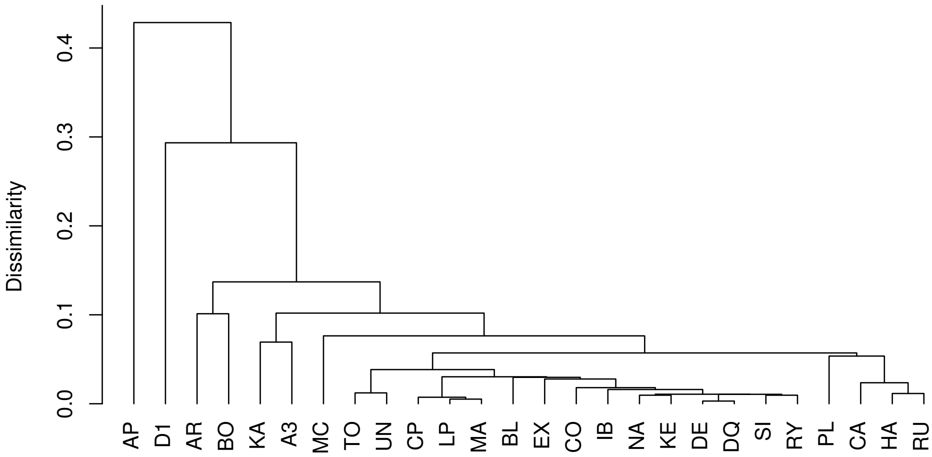


Table 1. The two letter abbreviations for the Social Choice Rules (Summary)

SCR	Social Choice Rule	Cat	SCR	Social Choice Rule	Cat
A3	Approval Poisson $\mu = 3$	C4	IB	Inverse Borda	C2
AR	Approval Uniform	C4	KA	$\lfloor \frac{m}{2} \rfloor$-Approval	C3
AP	Antiplurality	C3	KE	Kemeny	C1
BL	Black	C2	LP	Long Path	C1
BO	Borda	C2	MA	Markov	C1
CA	Carey	C3	MC	Majoritarian Compromise	C3
CO	Coombs	C3	NA	Nanson	C2
CP	Copeland	C1	PL	Plurality	C3
D1	Dictatorship	C3	RU	Plurality Runoff	C3
DE	Dodgson Rule	C3	RY	Raynaud	C2
DQ	Dodgson Quick	C2	SI	Simpson	C2
EX	Exhaustive	C3	TO	Top Cycle	C1
HA	Hare	C3	UN	Uncovered Set	C1

Table 2. Example of forming a Flattened Distance Matrix (FDM)

	M_1	M_2	M_3
Distance M	$\begin{array}{ccc} 1 & & \\ 0.03 & 1 & \\ 0.02 & 0.01 & 1 \end{array}$	$\begin{array}{ccc} 1 & & \\ 0.01 & 1 & \\ 0.02 & 0.03 & 1 \end{array}$	$\begin{array}{ccc} 1 & & \\ 0.1 & 1 & \\ 0.4 & 0.9 & 1 \end{array}$
Flattened M	$\begin{array}{ccc} X & & \\ 3 & X & \\ 2 & 1 & X \end{array}$	$\begin{array}{ccc} X & & \\ 1 & X & \\ 2 & 3 & X \end{array}$	$\begin{array}{ccc} X & & \\ 1 & X & \\ 2 & 3 & X \end{array}$
$\lvert M_i - M_2 \rvert$	$\begin{array}{ccc} X & & \\ 2 & X & \\ 0 & 2 & X \end{array}$		$\begin{array}{ccc} X & & \\ 0 & X & \\ 0 & 0 & X \end{array}$
Distance	4		0

Table 3. Condorcet Efficiencies of Rules for different homogeneities

	$b = 0$	0.04	0.1	0.2	0.3	0.4	0.5	1
A3	0.6086	0.7701	0.8172	0.8449	0.8610	0.8668	0.8769	0.9061
AP	0.5314	0.4556	0.4056	0.3780	0.3616	0.3507	0.3606	0.4521
AR	0.5183	0.6952	0.7566	0.7842	0.8014	0.8036	0.8111	0.8552
BL	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
BO	0.8487	0.8523	0.8490	0.8373	0.8425	0.8440	0.8469	0.8736
CA	0.8226	0.8496	0.8723	0.8911	0.9178	0.9292	0.9352	0.9688
CO	0.9055	0.9220	0.9484	0.9690	0.9854	0.9890	0.9935	0.9988
CP	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
D1	0.2441	0.3181	0.3790	0.4637	0.5159	0.5617	0.5932	0.7104
DE	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
DQ	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
EX	0.9055	0.9167	0.9285	0.9400	0.9513	0.9531	0.9584	0.9791
HA	0.8872	0.9010	0.9096	0.9185	0.9371	0.9464	0.9487	0.9746
IB	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
KA	0.7339	0.8029	0.8433	0.8696	0.8859	0.8963	0.9107	0.9423
KE	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
LP	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
MA	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
MC	0.6959	0.7275	0.7194	0.7855	0.8414	0.8749	0.9052	0.9659
NA	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
PL	0.5868	0.6858	0.7494	0.8153	0.8568	0.8830	0.8983	0.9494
RU	0.8140	0.8634	0.8878	0.9112	0.9309	0.9439	0.9480	0.9745
RY	0.9983	0.9973	0.9994	0.9998	0.9988	0.9988	0.9992	0.9994
SI	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
TO	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
UN	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 4. Distances that Increase with Homogeneity

#alter	#voter	b	distances that increase
5	25	$0 \rightarrow 1$	BO-BL CO-EX DQ-DE
5	25	$0 \rightarrow 0.5$	TO-AP UN-AP RY-AP CO-AP CO-EX MA-AP DE-AP DQ-AP DQ-DE KE-AP BL-AP BO-AP BO-BL EX-AP CP-AP HA-AP IB-AP LP-AP MC-AP NA-AP SI-AP
5	25	$0.5 \rightarrow 1$	*none increase*
5	25	$0.2 \rightarrow 0.5$	BO-BL

Table 5. Analysis of distance between Borda and Black

#alter	#voter	b	%Condorcet Winners	BO Condorcet Efficiency	BL-BO Distance
5	25	0	76%	86%	0.11
5	25	0.2	84%	84%	0.13
5	25	0.5	91%	85%	0.14
5	25	1	95%	87%	0.12
5	25	2	98%	90%	0.09

Table 6. Number of New Clusters that Form as b Increases

b	0	1/5	1/3	1/2	1	2	3	5	9	25	27	81	125	243	625	729
1/5	5															
1/3	7	3														
1/2	6	2	1													
1	6	2	2	1												
2	5	2	2	1	0											
3	4	1	1	1	0	0										
5	4	2	2	1	0	0	0									
9	3	1	1	1	0	0	0	0								
25	3	1	1	1	0	0	0	0	0							
27	3	1	1	1	0	0	0	0	0	0						
81	2	0	0	0	0	0	0	0	0	0	0					
125	1	0	0	0	0	0	0	0	0	0	0	0				
243	0	0	0	0	0	0	0	0	0	0	0	0	0			
625	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
729	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Inf	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b	0	1/5	1/3	1/2	1	2	3	5	9	25	27	81	125	243	625	729

Table 7. Number of Clusters Lost as b Increases

b	0	1/5	1/3	1/2	1	2	3	5	9	25	27	81	125	243	625	729
1/5	5															
1/3	7	1														
1/2	7	2	1													
1	9	4	2	0												
2	10	6	6	3	0											
3	9	6	4	3	0	0										
5	10	7	7	4	1	0	0									
9	9	6	6	4	0	0	0	0								
25	13	10	11	8	5	4	3	1	0							
27	13	10	11	9	5	4	3	1	0	0						
81	13	10	10	9	5	4	3	1	0	0	0					
125	12	10	11	9	6	5	3	2	0	0	0	0				
243	13	10	11	9	5	4	3	1	0	0	0	0	0			
625	13	10	11	9	6	5	3	2	0	0	0	0	0	0		
729	13	10	11	9	7	6	4	3	1	1	1	0	0	0	0	
Inf	13	12	13	11	9	8	6	5	3	3	3	2	1	0	0	0
b	0	1/5	1/3	1/2	1	2	3	5	9	25	27	81	125	243	625	729

Table 8. Distance Matrix, Varying # Agents, 5 Alternatives, impartial culture ($b = 0$)

	3	5	7	9	15	17	33	129	1025
3	0	716	672	744	754	770	794	758	786
5	716	0	406	466	624	632	672	690	774
7	672	406	0	298	482	436	482	514	620
9	744	466	298	0	368	378	430	490	592
15	754	624	482	368	0	136	148	346	376
17	770	632	436	378	136	0	106	348	362
33	794	672	482	430	148	106	0	296	354
129	758	690	514	490	346	348	296	0	172
1025	786	774	620	592	376	362	354	172	0