Level r Consensus and Stable Social Choice*

Muhammad Mahajne[†] Shmuel Nitzan[‡] Oscar Volij[†]

December 29, 2014

Abstract

We propose the concept of level r consensus as a useful property of a preference profile which considerably enhances the stability of social choice. This concept involves a weakening of unanimity, the most extreme form of consensus. It is shown that if a preference profile exhibits level r consensus around a given preference relation, there exists a Condorcet winner. In addition, if the number of individuals is odd the majority relation coincides with the preference relation around which there is such consensus and consequently it is transitive. Furthermore, if the level of consensus is sufficiently strong, the Condorcet winner is chosen by all the scoring rules. Level r consensus therefore ensures the Condorcet consistency of all scoring rules, thus eliminating the tension between decision rules inspired by ranking-based utilitarianism and the majority rule.

Journal of Economic Literature Classification Numbers: D71.

Keywords: Social choice, unanimity, consensus, preference aggregation rules, transitivity, simple majority, scoring rules, invariance to rule selection.

---

*This paper is based on the first chapter of Mahajne's PhD thesis. We thank an editor and two anonymous referees for their useful comments.

Corresponding author's email address: ovolij@bgu.ac.il

[†]Ben-Gurion University of the Negev

[‡]Bar-Ilan University

1 Introduction

A major goal of social choice theory is to search for reasonable ways of aggregating individual preferences into a social preference relation. Arrow's [1] impossibility theorem brought a serious challenge to such aspiration by showing that any social welfare function defined over an unrestricted domain, that satisfies the unanimity and the independence of irrelevant alternatives axioms must be dictatorial. Unanimity is a weak property requiring that if all individuals share a particular preference relation, this common relation must be the social preference relation. Unanimity is such a weak and sensible requirement that its violation would render any preference aggregation rule unacceptable. As a result, the search for reasonable preference aggregation rules has focused on domain restrictions and on the weakening of the independence axiom.

Among the many attempts to find reasonable aggregation rules, one can identify four approaches which can be considered as unanimity geared. The best-known approach is based on a unanimity idea that is not about a particular preference relation, but about the pattern of preferences. Alternative forms of domain restrictions, notably single-peakedness of preference relations, impose this type of weakened, implicit unanimity. In the latter case for instance, given any three alternatives, there is a unanimous agreement that a particular alternative is never the worst alternative among the three (see Sen [24]). The second approach looks for a unanimously supported metric-based compromise. It postulates an agreed-upon metric on the set of preference relations and, given a preference profile, seeks a social preference relation that is closest to it, namely one that minimizes the sum of its distances to the individual preference relations in the profile. Baigent [3], Kemeny [10], Nitzan [15], and Nurmi [17, 18] adopt this approach. The third approach also applies a plausible metric on the set of all possible preference profiles, but seeks a social choice rule that yields an outcome which is as close as possible to be unanimously preferred. In other words, the distance between the given preference profile and a profile where the chosen alternative is unanimously supported is minimized. See Campbell and Nitzan [4], Farkas and Nitzan [5], Lerer and Nitzan [12], Nitzan [14], Nitzan [16] for instances of this approach. Finally, the fourth approach is a probabilistic one. It postulates the existence of a unanimously supported true social preference relation and assumes that the preference profile is a noisy sample of it. Specifically, it assumes that the probability that any individual's ranking of any two alternatives coincides with the true ranking is higher than 1/2, and looks for a maximum-likelihood estimator that delivers a preference relation that maximizes the probability of having induced the realized preference profile. See Young [25] for a representative of this approach.

In the present paper we propose a new unanimity-inspired approach which is based on strengthening the unanimity requirement. The reason why the unanimity axiom is very weak is that it "bites" only in those rare instances of extreme preference homogeneity where individual preference relations are identical. In this paper, we replace the notion of full homogeneity of preferences by a new and weaker one which we refer to as consensus. According to this notion, a preference profile may exhibit consensus around some preference relation even if not all individuals share the same preference and even if some of them have opposite preferences. In order for consensus around some preference relation $\succ_0$ to exist it is necessary that whenever a subset of preferences are more similar to $\succ_0$ than another disjoint subset of the same size, there are more individuals with preferences in the former than in the latter.[1] Clearly, looking for consensus around some preference relation is more challenging when preferences are heterogeneous than in the extreme event of unanimous preferences. While there is a natural consensus around a unanimous preference relation, there may still be some kind of consensus around some preference relation, even in cases of heterogeneous preferences. The proposed approach looks for preference aggregation rules that select a social preference relation around which such consensus exists.

Several levels of consensus are defined, one more stringent than the other. Consensus of level 1 is more difficult to achieve than consensus of level 2, and so on, and all of them are achieved when there is unanimity about the preference relation. The least demanding level of consensus is level $K!/2$, where K is the number of alternatives over which preferences

[1] Several other attempts have been made to formalize and measure consensus. See, for example, García-Lapresta and Pérez-Román [9].

are defined. The definition of consensus rests on a given metric on the set of preferences and thus different metrics induce different notions of consensus. Our results suggest that when applying the inversion metric, the existence of consensus of level r, for some $r \leq K!/2$, considerably enhances the stability of social choice. Specifically, it implies the existence of a Condorcet winner and, if the number of individuals is odd, the transitivity of the induced majority relation, which turns out to be the one around which consensus exists. Furthermore, the existence of a sufficiently strong level of consensus, namely for $r \leq (K-1)!$, ensures the selection of the same Condorcet winning alternative by the majority rule and by all scoring rules. In that sense, it eliminates the tension between the majority rule and decision rules inspired by ranking-based utilitarianism.

2 Definitions

Let $A = \{a_1, \ldots, a_K\}$ be a set of $K > 2$ alternatives and let $N = \{1, \ldots, n\}$ be a set of individuals. Also, let $\mathcal{R}$ be the set of binary relations on A, and $\mathcal{P}$ be the subset of complete, transitive and antisymmetric binary relations on A. We will refer to the elements of $\mathcal{P}$ as preference relations or simply as preferences. A *preference profile* or simply a *profile* is a list $\pi = (\succ_1, \ldots, \succ_n)$ of preference relations on A such that for each $i \in N$, $\succ_i$ is the preference relation of individual i. We denote by $\mathcal{P}^n$ the set of preference profiles.

Let $\pi = (\succ_1, \ldots, \succ_n)$ be a preference profile. For each preference relation $\succ \in \mathcal{P}$, $\mu_\pi(\succ) = |\{i \in N : \succ_i = \succ\}|$ is the number of individuals whose preference relation is $\succ$. More generally, for any subset $C \subseteq \mathcal{P}$ of preference relations, $\mu_\pi(C) = |\{i \in N : \succ_i \in C\}|$ is the number of individuals whose preference relations are in C.

An *aggregation rule* is a function $f : \mathcal{P}^n \to \mathcal{R}$ that assigns to each preference profile a social binary relation. An aggregation rule is a *Social Welfare Function* if its range is the subset of transitive binary relations on A.

A well-known social welfare function is the *Borda function.* In order to define it, consider a preference profile $\pi = (\succ_1, \ldots, \succ_n)$. For each individual $i = 1, \ldots, n$ and for each alternative

$a \in A$, let $S_i(a) = |\{a' \in A : a \succ_i a'\}|$ be the number of alternatives that are ranked below a according to i's preferences. The Borda function associates with π the binary relation B_π given by

$$aB_\pi b \Leftrightarrow \sum_{i=1}^{n} S_i(a) \geq \sum_{i=1}^{n} S_i(b).$$

An important example of an aggregation rule is the *Majority rule*, which we define next. Let $a, a' \in A$ be two alternatives. Denote by $C(a \to a') = \{\succ \in \mathcal{P} : a \succ a'\}$ the set of preference relations according to which a is strictly preferred to a'. The majority rule assigns to each preference profile $\pi \in \mathcal{P}$ the binary relation M_π on A defined by

$$aM_\pi a' \Leftrightarrow \mu_\pi(C(a \to a')) \geq \mu_\pi(C(a' \to a)).$$

It is well known that the majority rule does not deliver a transitive binary relation for each preference profile, and thus it is not a social welfare function. Moreover, the binary relation that the majority rule assigns to a profile may not even have a maximal element. An alternative $a \in A$ is a *Condorcet winner* for a profile π, if it is a maximal element of M_π. Namely, if $aM_\pi b$ for every alternative $b \in A$.

Let $d : \mathcal{P}^2 \to \mathbb{R}$ be a metric on $\mathcal{P}$. That is, for every $\succ, \succ', \succ'' \in \mathcal{P}$, d satisfies

- $d(\succ, \succ') \geq 0$
- $d(\succ, \succ') = 0 \ \Leftrightarrow \succ = \succ'$
- $d(\succ, \succ') = d(\succ', \succ)$
- $d(\succ, \succ'') \leq d(\succ, \succ') + d(\succ', \succ'')$

For most of our results we will use the *inversion metric*, which is defined as follows:[2] $d(\succ, \succ')$ is the minimum number of pairwise adjacent transpositions needed to obtain $\succ'$ from $\succ$. Alternatively, $d(\succ, \succ')$ is the number of pairs of alternatives in A that are ranked differently by $\succ$ and $\succ'$. Formally, the inversion metric is defined by

$$d(\succ, \succ') = \frac{|(\succ \setminus \succ') \cup (\succ' \setminus \succ)|}{2}.$$

[2] See Kemeny and Snell [11] for a characterization of this metric.

A metric defined on $\mathcal{P}$ allows us to determine which one of any two preference relations is closer to a third one. We are interested in extending this kind of comparison to equal-sized sets of preferences as well. The following definition identifies circumstances where a given set of preferences $C \subseteq \mathcal{P}$ is closer to $\succ_0$ than an alternative set $C' \subseteq \mathcal{P}$.

Definition 1 Let C and C' be two disjoint nonempty subsets of $\mathcal{P}$ with the same cardinality, and let $\succ_0 \in \mathcal{P}$ be a preference relation on A. We say that C *is at least as close to* $\succ_0$ *as* C', denoted by $C \geq_{\succ_0} C'$, if there is a one-to-one function $\phi : C \to C'$ such that for all $\succ \in C$, $d(\succ, \succ_0) \leq d(\phi(\succ), \succ_0)$. We also say that C *is closer than* C' *to* $\succ_0$, denoted by $C >_{\succ_0} C'$, if there is a one to one function $\phi : C \to C'$ such that for all $\succ \in C$, $d(\succ, \succ_0) \leq d(\phi(\succ), \succ_0)$, with strict inequality for at least one $\succ \in C$.

In other words, C is at least as close as C' to some given preference relation $\succ_0 \in \mathcal{P}$ if each preference relation $\succ'$ in C' can be paired with a preference relation $\succ$ in C that is at least as close to $\succ_0$, according to d, as $\succ'$ is. C is closer than C' to $\succ_0$ if it is at least as close to it as C' and it is not the case that C' is at least as close to $\succ_0$ as C.

An alternative way to check whether $C \geq_{\succ_0} C'$ is as follows. Let $d(C, \succ_0)$ be the list of distances $(d(\succ, \succ_0))_{\succ \in C}$ arranged in a non-decreasing order. Similarly, let $d(C', \succ_0)$ be the list $(d(\succ, \succ_0))_{\succ \in C'}$ also arranged in a non-decreasing order. Then $C \geq_{\succ_0} C' \Leftrightarrow d(C, \succ_0) \leq d(C', \succ_0)$.

Example 1 Let the set of alternatives be $A = \{a, b, c\}$. The set $\mathcal{P}$ contains six preference relations, given by

$$\begin{array}{lllllll} \succ_1 & = & a, b, c & \qquad & \succ_2 & = & a, c, b \\ \succ_3 & = & b, a, c & & \succ_4 & = & c, a, b \\ \succ_5 & = & b, c, a & & \succ_6 & = & c, b, a \end{array}$$

There are ten ways to partition $\mathcal{P}$ into two subsets with three preference relations each. One such partition is $C_1 = \{\succ_1, \succ_2, \succ_3\}$ and $\overline{C_1} = \{\succ_4, \succ_5, \succ_6\}$. Consider the preference relation

$\succ_1$. It can be checked that the distances of each preference relation in $\mathcal{P}$ to $\succ_1$, according to the inversion metric, are given by

$$
\begin{aligned}
d(\succ_1, \succ_1) &= 0 \\
d(\succ_2, \succ_1) = d(\succ_3, \succ_1) &= 1 \\
d(\succ_4, \succ_1) = d(\succ_5, \succ_1) &= 2 \\
d(\succ_6, \succ_1) &= 3
\end{aligned}
$$

It can also be checked that $C_1 >_{\succ_1} \overline{C_1}$. Indeed, $d(\succ_1, \succ_1) < d(\succ_4, \succ_1)$, $d(\succ_2, \succ_1) < d(\succ_5, \succ_1)$ and $d(\succ_3, \succ_1) < d(\succ_6, \succ_1)$. Alternatively, $d(C_1, \succ_1) = (0, 1, 1)$ and $d(\overline{C_1}, \succ_1) = (2, 2, 3)$. Therefore $d(C_1, \succ_1) < d(\overline{C_1}, \succ_1)$, which implies that $C_1 >_{\succ_1} \overline{C_1}$.

Note that any two disjoint, equal-sized subsets of preference relations contain at most $K!/2$ elements each. Taking this into account and based on the "at least as close to $\succ_0$" relation defined above, we can now define the concept of consensus.

Definition 2 Let $r \in \{1, 2, \ldots K!/2\}$, and let $\succ_0 \in \mathcal{P}$. A preference profile $\pi \in \mathcal{P}^n$ exhibits *consensus of level r around $\succ_0$* if

1. for all disjoint subsets C, C' of $\mathcal{P}$ with cardinality r, $C \geq_{\succ_0} C' \Rightarrow \mu_\pi(C) \geq \mu_\pi(C')$
2. there are disjoint subsets C, C' of $\mathcal{P}$ with cardinality r, such that $C >_{\succ_0} C'$ and $\mu_\pi(C) > \mu_\pi(C')$.

In words, given $1 \leq r \leq K!/2$, a preference profile π exhibits consensus of level r around some preference relation $\succ_0$, if whenever a subset C of r preference relations is at least as close to $\succ_0$ as another disjoint subset C' of r preference relations, the number of preference relations in π that belong to C is at least as large as the number of preference relations in π that belong to C', and in particular there exist two such subsets such that C is closer to $\succ_0$ than C' and the number of preferences in C is greater than the number of preferences in C'.

Example 2 Consider the set $\mathcal{P}$ of preference relations from Example 1. There are ten different ways to partition $\mathcal{P}$ into two subsets, $C, \overline{C}$, of cardinality 3. We have already seen that

$$C_1 = \{\succ_1, \succ_2, \succ_3\} >_{\succ_1} \{\succ_4, \succ_5, \succ_6\} = \overline{C_1}.$$

Similarly, it can be checked that

$$\begin{aligned} C_2 = \{\succ_1, \succ_2, \succ_4\} & >_{\succ_1} & \{\succ_3, \succ_5, \succ_6\} = \overline{C_2} \\ C_3 = \{\succ_1, \succ_3, \succ_4\} & >_{\succ_1} & \{\succ_2, \succ_5, \succ_6\} = \overline{C_3} \\ C_4 = \{\succ_1, \succ_2, \succ_5\} & >_{\succ_1} & \{\succ_3, \succ_4, \succ_6\} = \overline{C_4} \\ C_5 = \{\succ_1, \succ_3, \succ_5\} & >_{\succ_1} & \{\succ_2, \succ_4, \succ_6\} = \overline{C_5} \end{aligned}$$

Also, for the remaining five partitions $\{C, \overline{C}\}$, we have that neither $C \geq_{\succ_1} \overline{C}$ nor $\overline{C} \geq_{\succ_1} C$. Let π be a preference profile containing 3 copies of $\succ_1$, one copy of $\succ_3$, one copy of $\succ_4$ and 2 copies of $\succ_5$. It can be checked that $\mu_\pi(C_i) > \mu_\pi(\overline{C_i})$ for $i = 1, 2, 3, 4, 5$. Consequently, we conclude that the profile π exhibits consensus of level 3 around $\succ_1$. On the other hand, π does not exhibit consensus of level 2 around $\succ_1$. To see this, note that although $\{\succ_2, \succ_4\}$ is closer than $\{\succ_5, \succ_6\}$ to $\succ_1$ (indeed, $d(\succ_2, \succ_1) = 1 < 2 = d(\succ_5, \succ_1)$ and $d(\succ_4, \succ_1) = 2 < 3 = d(\succ_6, \succ_1)$), we have that $\mu_\pi(\{\succ_2, \succ_4\}) = 1 < 2 = \mu_\pi(\{\succ_5, \succ_6\})$.

To clarify the concept of consensus, let's consider the case of $r = 1$. In this case, a preference profile π exhibits consensus of level 1 around some preference relation $\succ_0$, if for any two preference relations $\succ$ and $\succ'$, whenever $\succ$ is at least as close as $\succ'$ to $\succ_0$, there are at least as many individuals in the profile whose preferences are given by $\succ$ than those whose preferences are given by $\succ'$. That is, no matter which two preferences are chosen, there are at least as many individuals with the one that is closer to $\succ_0$ than those with the other one. Note that if there is unanimity of preferences, there is also consensus of level 1 around the unanimous preference relation. In general, π exhibits consensus of level r around some preference relation $\succ_0$, if no matter which two disjoint subsets of r preferences we choose, if one of them is at least as close to $\succ_0$ as the other, then π also contains at least as many individuals whose preferences are in this subset than in the other one.

In the next section we will show that consensus of level r implies consensus of level $r+1$. Therefore, consensus of level 1 is the most difficult to achieve. Nevertheless, there are known models that induce consensus of level 1. For instance Mallows's [13] ϕ model is one of them. According to this model, there is a "true" social preference relation $\succ_0$ and each individual's preference is an unbiased estimate of it. Under certain assumptions about the individual preference generation, the probability that a given individual's preference relation is $\succ \in \mathcal{P}$ is given by $P(\succ) = Ce^{-\lambda d(\succ,\succ_0)}$ for some $\lambda > 0$ and where d is the inversion metric and C is a normalizing constant. If the preference profile is distributed according to P, namely $\mu_\pi/n = P$, then it exhibits consensus of level 1 around $\succ_0$. Indeed, for any $\succ, \succ' \in \mathcal{P}$, $d(\succ,\succ_0) \leq d(\succ',\succ_0) \Rightarrow P(\succ) \geq P(\succ') \Rightarrow \mu_\pi(\succ) \geq \mu_\pi(\succ')$. To put it another way, if individuals' preferences are distributed according to Mallows's model, the profile will exhibit consensus of level 1 around some preference relation.

3 Consensus and Majority Rule

In this section we show some striking implications of the existence of consensus around some preference relation. But before turning to this task, we first show that there is a hierarchy in the levels of consensus: they are ordered by strength, the strongest being consensus of level 1 and the weakest consensus of level $K!/2$. Example 2 shows that this hierarchy may be strict.

Proposition 1 Let r be an integer between 1 and $K!/2-1$. If $\pi \in \mathcal{P}^n$ exhibits consensus of level r around $\succ_0$, then it exhibits consensus of level $r+1$ as well around $\succ_0$.

Proof : Assume $\pi \in \mathcal{P}^n$ exhibits consensus of level r around $\succ_0$. We need to show that conditions 1, and 2 in Definition 2 are satisfied.

1. Let $C = \{\succ_1, \ldots, \succ_{r+1}\}$ and $C' = \{\succ'_1, \ldots, \succ'_{r+1}\}$ be two disjoint subsets of $\mathcal{P}$ with cardinality $r+1$ such that $C \geq_{\succ_0} C'$. Then, there is a one-to-one function $\varphi : C \to C'$ such that $d(\succ_i, \succ_0) \leq d(\varphi(\succ_i), \succ_0)$ for all $i = 1, \ldots, r+1$. Assume, without loss of generality, that $\varphi(\succ_i) = \succ'_i$ for all $i = 1, \ldots, r+1$. We need to show that $\mu_\pi(C) \geq \mu_\pi(C')$.
Assume by contradiction that

$$\mu_\pi(C) < \mu_\pi(C') \tag{1}$$

Then, there must be two preference relations $\succ_i \in C$ and $\succ'_i \in C'$ such that $\mu_\pi(\succ_i) < \mu_\pi(\succ'_i)$. Assume without loss of generality that these two preferences are $\succ_{r+1}$ and $\succ'_{r+1}$. Namely that

$$\mu_\pi(\succ_{r+1}) < \mu_\pi(\succ'_{r+1}). \tag{2}$$

Assume also without loss of generality that

$$\mu_\pi(\succ_1) - \mu_\pi(\succ'_1) \geq \mu_\pi(\succ_i) - \mu_\pi(\succ'_i) \qquad \forall i = 1, \ldots, r. \tag{3}$$

Consider the subsets $C_{-1} = \{\succ_2, \ldots, \succ_{r+1}\}$ and $C'_{-1} = \{\succ'_2, \ldots, \succ'_{r+1}\}$. Since $C \cap C' = \emptyset$, and since C is at least as close as C' to $\succ_0$, we have that $C_{-1} \cap C'_{-1} = \emptyset$, and C_{-1} is at least as close as C'_{-1} to $\succ_0$ as well. Since π exhibits consensus of order r around $\succ_0$, we must have

$$\mu_\pi(C_{-1}) \geq \mu_\pi(C'_{-1}). \tag{4}$$

Since $\mu_\pi(C_{-1}) = \sum_{k=2}^{r+1} \mu_\pi(\succ_k)$ and $\mu_\pi(C'_{-1}) = \sum_{k=2}^{r+1} \mu_\pi(\succ'_k)$, there must be some $k = 2, \ldots, r+1$ such that $\mu_\pi(\succ_k) \geq \mu_\pi(\succ'_k)$. Further, given (2) this k cannot be $r+1$. As a result, using (3)

$$\mu_\pi(\succ_1) - \mu_\pi(\succ'_1) \geq 0. \tag{5}$$

But then, using (4) and (5)

$$\mu_\pi(C) = \mu_\pi(C_{-1}) + \mu_\pi(\succ_1) \geq \mu_\pi(C'_{-1}) + \mu_\pi(\succ'_1) = \mu_\pi(C')$$

which contradicts (1).

2. We need to show that there are two disjoint subsets $D, D' \in \mathcal{P}$ with cardinality $r+1$, such that $D >_{\succ_0} D'$ and $\mu_\pi(D) > \mu_\pi(D')$. Since π exhibits consensus of level r

around $\succ_0$, there are two disjoint subsets $C = \{\succ_1, \ldots, \succ_r\}$ and $C' = \{\succ'_1, \ldots, \succ'_r\}$ such that $C >_{\succ_0} C'$ and $\mu_\pi(C) > \mu_\pi(C')$. In particular, there is a one-to-one function $\varphi : C \to C'$ such that $d(\succ_i, \succ_0) \leq d(\varphi(\succ_i), \succ_0)$ for $i = 1, \ldots, r$, with strict inequality for at least one $i \in \{1, \ldots, r\}$. Without loss of generality assume that $\varphi(\succ_i) =\succ'_i$. Let $\succ_{r+1}$ and $\succ'_{r+1}$ two preference relations not in C nor in C', and assume without loss of generality that $d(\succ_{r+1}, \succ_0) \leq d(\succ'_{r+1}, \succ_0)$. We will show that $D = C \cup \{\succ_{r+1}\}$ and $D' = C' \cup \{\succ'_{r+1}\}$ are the two subsets we are looking for. By construction, $D >_{\succ_0} D'$. Therefore it remains to be shown that $\mu_\pi(D) > \mu_\pi(D')$. Assume by contradiction that

$$\sum_{i=1}^{r+1} \mu_\pi(\succ_i) \leq \sum_{i=1}^{r+1} \mu_\pi(\succ'_i). \tag{6}$$

Note that for $j = 1, \ldots, r$, $(D \setminus \{\succ_j\}) \geq_{\succ_0} (D' \setminus \{\succ'_j\})$. Therefore, since π exhibits consensus of level r around $\succ_0$,

$$\sum_{\substack{i=1 \\ i \neq j}}^{r+1} \mu_\pi(\succ_i) \geq \sum_{\substack{i=1 \\ i \neq j}}^{r+1} \mu_\pi(\succ'_i) \qquad j = 1, \ldots, r. \tag{7}$$

Consequently, it follows from (6) and (7) that

$$\mu_\pi(\succ_j) \leq \mu_\pi(\succ'_j) \qquad j = 1, \ldots r, \tag{8}$$

which implies that

$$\sum_{i=1}^{r} \mu_\pi(\succ_i) \leq \sum_{i=1}^{r} \mu_\pi(\succ'_i).$$

This contradicts the fact that $\mu_\pi(C) > \mu_\pi(C')$. Therefore, we conclude that $\mu_\pi(D) > \mu_\pi(D')$ and the proof is complete. □

Proposition 1 implies that if a profile exhibits consensus of any level around a given preference relation then it also exhibits consensus of level $K!/2$ around that relation. For that reason, whenever a profile exhibits consensus of level $K!/2$ around a preference relation we will simply say that it exhibits consensus around it.

We now turn to the implications of consensus on the outcomes of the majority rule. The next theorem shows that despite not being a social welfare function, if a profile with an odd number of agents exhibits any possible level of consensus with respect to the inversion metric, the majority rule associates with it a transitive binary relation. In fact, in that case the preference relation associated with majority rule is the only one around which there is consensus.

Theorem 1 Let $\succ_0 \in \mathcal{P}$ be a preference relation and let a_1 be the alternative that is ranked first according to $\succ_0$. Let $\pi \in \mathcal{P}^n$ be a preference profile that exhibits consensus of level $r \in \{1, \ldots, K!/2\}$ around $\succ_0$ with respect to the inversion metric. Then, a_1 is a Condorcet winner. Furthermore, if n is odd $\succ_0$ is the unique preference relation in $\mathcal{P}$ around which there is consensus, and M_π, the binary relation assigned by the majority rule to π, coincides with $\succ_0$.

Proof: Let $\pi \in \mathcal{P}^n$ be a preference profile that exhibits consensus of level $r \in \{1, \ldots, K!/2\}$ around $\succ_0$ with respect to the inversion metric. By Proposition 1, π exhibits consensus of level $K!/2$. Let $a, b \in A$ be two alternatives. We will show first that $a \succ_0 b \Rightarrow aM_\pi b$. This will immediately imply that a_1 is a Condorcet winner. So assume that $a \succ_0 b$. Partition $\mathcal{P}$ into the two sets $C(a \to b)$ and $C(b \to a)$. These sets contain $K!/2$ elements each. Consider the one-to-one function $\varphi : C(a \to b) \to C(b \to a)$ defined as follows: for each $\succ \in C(a \to b)$, let $\varphi(\succ) \in \mathcal{P}$ be the preference relation that is obtained from $\succ$ by switching a and b in the ranking. Consequently, since $a \succ_0 b$, $d(\succ, \succ_0) < d(\phi(\succ), \succ_0)$ for all $\succ \in C(a \to b)$, where d is the inversion metric. In other words, according to the inversion metric, $C(a \to b)$ is closer to $\succ_0$ than $C(b \to a)$ is. Since there is consensus of level $K!/2$ around $\succ_0$, this implies that $\mu_\pi(C(a \to b)) \geq \mu_\pi(C(b \to a))$, which means that $aM_\pi b$. This shows that the alternative that is ranked first according to $\succ_0$ is a Condorcet winner. When n is odd and $a \neq b$, $aM_\pi b$ means that $\mu_\pi(C(a \to b)) > \mu_\pi(C(b \to a))$. It follows that $a \succ_0 b$ since otherwise, by the previous argument we would have that $\mu_\pi(C(b \to a)) \geq \mu_\pi(C(a \to b))$. This shows that when n is odd, $\succ_0 = M_\pi$ and, consequently, that $\succ_0$ is the only preference relation around which there is consensus. For if π exhibited consensus around $\succ_1$ as well, we would have $\succ_1 = M_\pi = \succ_0$. □

Theorem 1 shows that, under the inversion metric, if a profile with an odd number of individuals exhibits consensus of any level around $\succ_0$, then $\succ_0$ coincides with the binary relation assigned to π by the majority rule. In this sense we can say that majority rule is rationalizable by consensus.

It is clear that not every preference profile exhibits consensus around some preference relation. For instance, a profile containing one copy of each of the $K!$ preference relations on A does not exhibit consensus around any of them. On the other hand, the proof of Theorem 1 shows that if there is consensus around $\succ_0$, then $a \succ_0 b$ implies that $aM_\pi b$. This suggests a way to look for preference relations around which there is consensus. One should check preference relations $\succ$ that are close to the binary relation associated with majority rule in the sense that $a \succ b$ implies that $aM_\pi b$.

The next example shows that if n is even, there can be consensus around more than one preference relation.

Example 3 Let $\mathcal{P}$ be again the set of preference relations from Example 1 and consider the preference profile $\pi = (\succ_1, \succ_2)$. It can be checked that π exhibits consensus of level 3 around both $\succ_1$ and $\succ_2$. Indeed, every partition $\{C, C'\}$ of $\mathcal{P}$ into two sets of cardinality 3 such that $C \geq_{\succ_1} C'$ satisfies that $\succ_1 \in C$, and consequently $\mu_\pi(C) \geq 1 \geq \mu_\pi(C')$. Furthermore there are partitions $\{C, C'\}$ such that $\mu_\pi(C) = 2 > 0 = \mu_\pi(C')$. One such partition is $\{\{\succ_1, \succ_2, \succ_3\}, \{\succ_4, \succ_5, \succ_6\}\}$. This shows that π exhibits consensus of level 3 around $\succ_1$. A similar argument shows that it also exhibits consensus of level 3 around $\succ_2$.

The next example shows that the fact that a preference profile π exhibits consensus of level $K!/2$ around some preference relation does not imply that the Borda function will assign this preference relation to π.

Example 4 We have seen in Example 2 that profile π exhibits consensus of level 3 around $\succ_1$. Consistent with Theorem 1, the majority rule applied to π yields $\succ_1$. In contrast, it can be verified that the preference relation assigned by the Borda function is $\succ_3$, whose top-ranked alternative is not even a Condorcet winner.

4 Consensus and Scoring Rules

Sometimes one is not interested in the social preference relation but only in its maximal elements. In that case, instead of focusing on social welfare functions one should concentrate on social choice rules. A *social choice rule* is a function $g : \mathcal{P}^n \to 2^A$ that assigns to each preference profile a set of chosen alternatives.

A special class of social choice rules consists of *scoring rules*, also known as positional voting rules. Each scoring rule is characterized by a list $S = \{S_1, S_2, \cdots, S_K\}$ of K non-negative scores with $S_1 \geq S_2 \geq \cdots \geq S_K$ and $S_1 > S_K$. Given a preference profile $\pi = (\succ_1, \ldots, \succ_n)$, each individual $i = 1, \ldots, n$ assigns S_k points, for $k = 1, \ldots, K$, to the alternative that is ranked k-th in his preference relation, $\succ_i$. That is, each agent assigns S_1 points to his most preferred alternative, S_2 points to the second best alternative and so on. The scoring rule associated with the sores in S, denoted by V_S, chooses the alternatives with the maximum total score.

Many social choice rules are instances of scoring rules. For example, the *plurality* rule is the scoring rule associated with the scores $(1, 0, \ldots, 0)$. The *inverse plurality* rule is the scoring rule associated with $(1, \ldots, 1, 0)$. More generally, for $1 \leq t \leq K - 1$, the *t-approval voting method*, denoted V_t, is the scoring rule associated with $S_t = (\underbrace{1, \ldots, 1}_{t}, \underbrace{0, \ldots, 0}_{K-t})$. Lastly, the Borda social choice rule is the scoring rule associated with $S_B = (K - 1, K - 2, \ldots, 0)$.

The t approval voting rules, for $t = 1, \ldots, K - 1$, play a central role in the theory of scoring rules since any list of scores $S = \{S_1, S_2, \cdots, S_K\}$ can be written as a non-negative linear combination $S = \sum_{t=1}^{K-1} \alpha_t S_t$ of the $K - 1$ approval voting scores. Based on this fact, Saari [21] showed that if all approval voting methods choose alternative a, then this same alternative is chosen by all the scoring methods. Formally, if $a \in V_t(\pi)$ for $t = 1, \ldots, K - 1$, then $a \in V_S(\pi)$ for all scores S.

The next theorem establishes that the existence of r consensus, for $r \leq (K - 1)!$, guarantees the invariance of the chosen alternative across all scoring rules.

Theorem 2 Suppose that $r \leq (K - 1)!$. Also, let $\pi \in \mathcal{P}^n$ be a preference profile, $\succ_0 \in \mathcal{P}$

a preference relation, and $a \in A$ the alternative that is ranked first according to $\succ_0$. If π exhibits consensus of level r around $\succ_0$ according to the inversion metric, then $a \in V_S$ for all scoring rules V_S.

Proof : Let S be a list of scores. Given that S can be written as a non-negative linear combination of the $K-1$ t-approval voting scores S_t, it is enough to show that $a \in V_t(\pi)$ for $t = 1, \ldots, K-1$.

Fix $t \in \{1, \ldots, K-1\}$, and let $b \in A \setminus \{a\}$. Denote by $C(a \overset{t}{\to} b)$ the set of preference relations in $\mathcal{P}$ such that a ranks t-th or above, and b ranks strictly below the t-th place. Similarly, denote by $C(b \overset{t}{\to} a)$ the set of preference relations in $\mathcal{P}$ such that b ranks t-th or above, and a ranks strictly below the t-th place. Since b is a fixed but otherwise arbitrary alternative different from a, in order to show that $a \in V_t(\pi)$ we must show that $\mu_\pi(C(a \overset{t}{\to} b)) \geq \mu_\pi(C(b \overset{t}{\to} a))$. By definition, $C(a \overset{t}{\to} b) \cap C(b \overset{t}{\to} a) = \emptyset$. Furthermore, these two sets have equal cardinality, which we denote by c. Therefore, in order to show that $\mu_\pi(C(a \overset{t}{\to} b)) \geq \mu_\pi(C(b \overset{t}{\to} a))$ it is enough to show that π exhibits consensus of level c around $\succ_0$ and that $C(a \overset{t}{\to} b)$ is closer than $C(b \overset{t}{\to} a)$ to $\succ_0$.

Note that there are $\binom{K-2}{t-1}$ ways to partition the K alternatives into two subsets, one containing t alternatives, one being a, and the other containing $K-t$ alternatives, one of them being b. Therefore the cardinality of $C(a \overset{t}{\to} b)$ (and similarly of $C(b \overset{t}{\to} a)$) is

$$\begin{aligned} c &= \binom{K-2}{t-1} t!(K-t)! \\ &= \frac{(K-2)!}{(K-t-1)!(t-1)!} t!(K-t)! \\ &= (K-1)! \frac{t(K-t)}{K-1} \end{aligned}$$

But since $\frac{t(K-t)}{K-1} \geq 1$ if and only if $(t-1)(K-1-t) \geq 0$ and since $1 \leq t \leq K-1$ we have that $\frac{t(K-t)}{K-1} \geq 1$. Therefore $c \geq (K-1)! \geq r$. Consequently, since π exhibits consensus of level r around $\succ_0$, Proposition 1 implies that π exhibits consensus of level c around $\succ_0$ as well.

In order to show that $C(a \overset{t}{\to} b) >_{\succ_0} C(b \overset{t}{\to} a)$, let $M_i(a \overset{t}{\to} b)$, for each $i = 1, \ldots, t$, be the set of preference relations such that alternative a is ranked i-th and alternative b is ranked strictly below the t-th place. Similarly, let $M_i(b \overset{t}{\to} a)$ be the set of preference relations such that alternative b is ranked ith and alternative a is ranked strictly below the t-th place. Let $\phi : M_i(a \overset{t}{\to} b) \to M_i(b \overset{t}{\to} a)$ be the one-to-one function that maps each preference relation $\succ \in M_i(a \overset{t}{\to} b)$ into the preference relation that is obtained from $\succ$ by switching alternatives a and b in the preference ranking. Clearly, since a is ranked first in $\succ_0$, we have that $d(\succ, \succ_0) < d(\phi(\succ), \succ_0)$ for all $\succ \in M_i(a \overset{t}{\to} b)$. Noting that $C(a \overset{t}{\to} b) = \cup_{i=1}^{t} M_i(a \overset{t}{\to} b)$ and $C(b \overset{t}{\to} a) = \cup_{i=1}^{t} M_i(b \overset{t}{\to} a)$, we conclude that $C(a \overset{t}{\to} b) >_{\succ_0} C(b \overset{t}{\to} a)$. □

For a given preference profile, different scoring rules may result in the selection of any of the K alternatives (see, for instance, Fishburn [8], Saari [19, 20, 23]). It is also possible that an alternative, and even a Condorcet winning alternative, will not be selected by any scoring rule (see, Fishburn [6, 7] and Saari [22]). These findings are balanced by the results of Baharad and Nitzan [2] and Saari [21] that specify necessary and sufficient conditions for the selection of the same alternative by all scoring rules. Theorem 2 shows that level r consensus is another sufficient condition for the selection of the same Condorcet winning alternative by all scoring rules. We have seen in Example 4 that even if the majority rule yields a transitive preference relation and even if there is consensus of level $K!/2$ around it, a scoring rule may not select the Condorcet winner. The following result shows, however, that level r consensus for $r \leq (K-1)!$ is a sufficient condition for all scoring rules to be Condorcet consistent.

Corollary 1 Let $\succ_0 \in \mathcal{P}$ be a preference relation and $a \in A$ the alternative that is ranked highest according to $\succ_0$. If π exhibits consensus of level $r \leq (K-1)!$ around $\succ_0$ according to the inversion metric, then a is a Condorcet winner and it is chosen by all scoring rules.

Proof : By Theorem 2, a is chosen by all scoring rules. By Theorem 1, a is a Condorcet winner. □

5 Concluding Remarks

In this paper we have proposed the concept of level r consensus and showed that its existence in its mildest form has significant implications. It ensures stability of one of the most extensively studied aggregation rules, namely the simple majority rule. Specifically, we show that under the inversion metric, when a preference profile with an odd number of agents exhibits level r consensus around a given preference relation, this preference relation is the one assigned by the majority rule to that profile which furthermore turns out to be transitive. The corresponding social choice function therefore selects the Condorcet winning alternative. Additionally, if the level of consensus is strong enough ($r \leq (K-1)!$), this chosen alternative is also the choice of all scoring rules. In other words, not only does the existence of r consensus ensure stability under simple majority, it also ensures the Condorcet consistency of all scoring rules. That is, it eliminates the tension between the simple majority rule and the scoring rules (in particular, the Borda rule). The existence of r consensus thus simultaneously resolves two of the major problems in social choice theory.

The two unanimity geared metric approaches mentioned in the introduction, the ones used in Farkas and Nitzan [5] and in Kemeny [10] respectively, are different from our level r consensus approach. Whereas the latter is based on a new preference domain restriction, the former two approaches do not impose any domain restriction; in fact one of their notable advantages is that they can be applied to any given preference profile. Interestingly, the simple majority rule is rationalized by the level r consensus approach, provided that one applies the inversion metric. This is in contrast to the outcome obtained under the two alternative metric approaches. Indeed, under the first one, for any given preference profile, the application of the inversion metric results in the rationalization of the Borda rule, and under the second one, the application of the inversion metric need not result in either the simple majority rule nor the Borda rule. However, as mentioned above, if a preference profile exhibits consensus of any level r, then there exists a Condorcet winner which is selected by Kemeny's rule (see Nurmi [17]), and if the consensus is sufficiently strong, the Borda rule is also Condorcet consistent.

References

[1] Arrow, K. J. *Social choice and individual values*, vol. 12. Yale university press, 2012.

[2] Baharad, E., and Nitzan, S. On the selection of the same winner by all scoring rules. *Social Choice and Welfare 26*, 3 (2006), 597–601.

[3] Baigent, N. Preference proximity and anonymous social choice. *The Quarterly Journal of Economics* (1987), 161–169.

[4] Campbell, D. E., and Nitzan, S. I. Social compromise and social metrics. *Social Choice and Welfare 3*, 1 (1986), 1–16.

[5] Farkas, D., and Nitzan, S. The borda rule and pareto stability: a comment. *Econometrica 47*, 5 (1979), 1305–06.

[6] Fishburn, P. C. Paradoxes of voting. *The American Political Science Review* (1974), 537–546.

[7] Fishburn, P. C. Condorcet social choice functions. *SIAM Journal on Applied Mathematics 33*, 3 (1977), 469–489.

[8] Fishburn, P. C. Inverted orders for monotone scoring rules. *Discrete Applied Mathematics 3*, 1 (1981), 27–36.

[9] García-Lapresta, J. L., and Pérez-Román, D. Measuring consensus in weak orders. In *Consensual Processes*. Springer, 2011, pp. 213–234.

[10] Kemeny, J. G. Mathematics without numbers. *Daedalus 88*, 4 (1959), 577–591.

[11] Kemeny, J. G., and Snell, J. L. *Mathematical models in the social sciences*, vol. 9. Ginn Boston, 1962.

[12] Lerer, E., and Nitzan, S. Some general results on the metric rationalization for social decision rules. *Journal of Economic Theory 37*, 1 (1985), 191–201.

[13] Mallows, C. L. Non-null ranking models. i. *Biometrika* (1957), 114–130.

[14] Nitzan, S. Some measures of closeness to unanimity and their implications. *Theory and Decision 13*, 2 (1981), 129–138.

[15] Nitzan, S. More on the preservation of preference proximity and anonymous social choice. *The Quarterly Journal of Economics* (1989), 187–190.

[16] Nitzan, S. *Collective preference and choice.* Cambridge University Press, 2009.

[17] Nurmi, H. *Voting procedures under uncertainty.* Springer, 2002.

[18] Nurmi, H. A comparison of some distance-based choice rules in ranking environments. *Theory and Decision 57*, 1 (2004), 5–24.

[19] Saari, D. G. The ultimate of chaos resulting from weighted voting systems. *Advances in Applied Mathematics 5*, 3 (1984), 286–308.

[20] Saari, D. G. A dictionary for voting paradoxes. *Journal of Economic Theory 48*, 2 (1989), 443–475.

[21] Saari, D. G. Explaining all three-alternative voting outcomes. *Journal of Economic Theory 87*, 2 (1999), 313–355.

[22] Saari, D. G. Mathematical structure of voting paradoxes. *Economic Theory 15*, 1 (2000), 1–53.

[23] Saari, D. G. *Chaotic elections!: A mathematician looks at voting.* American Mathematical Soc., 2001.

[24] Sen, A. *Collective Choice and Social Welfare.* Holden Day, 1970.

[25] Young, H. P. Condorcet's theory of voting. *American Political Science Review 82*, 04 (1988), 1231–1244.