CRDE

CAHIER 17-2001

THE MEASUREMENT OF DIVERSITY

Walter BOSSERT,
Prasanta K. PATTANAIK and
Yongsheng XU

Centre de recherche
et développement en économique

C.P. 6128, succursale Centre-ville
Montréal QC H3C 3J7

Téléphone : (514) 343-6557
Télécopieur : (514) 343-5831
crde@crde.umontreal.ca
http://www.crde.umontreal.ca/

Université de Montréal

CAHIER 17-2001

THE MEASUREMENT OF DIVERSITY

Walter BOSSERT[1], Prasanta K. PATTANAIK[2] and Yongsheng XU[3]

[1] Centre de recherche et développement en économique (C.R.D.E.) and Département de sciences économiques, Université de Montréal

[2] Department of Economics, University of California at Riverside

[3] Department of Economics, School of Policy Studies, Georgia State University

July 2001

The authors thank William McCausland for comments and discussions. Financial support from the Social Sciences and Humanities Research Council of Canada and the Centre de recherche et développement en économique of the Université de Montréal is gratefully acknowledged.

RÉSUMÉ

Le concept de diversité est pertinent dans différents contextes. La biodiversité d'un système écologique et la diversité des options offertes à un décideur ont été étudiées récemment. Nous caractérisons deux classes emboîtées de mesures ordinales de la diversité et un élément important de ces classes. Nous montrons que ce dernier cas particulier est équivalent au critère proposé par Weitzman.

Mots clés : diversité, mesures de distance

ABSTRACT

The notion of diversity is an issue that is of relevance in several contexts. For example, the biodiversity of a given ecological environment and the diversity of the options available to a decision maker have attracted some attention in recent research. This paper provides an axiomatic approach to the measurement of diversity. We characterize two nested classes of ordinal measures of diversity and an important member of these classes. We prove that the latter special case is equivalent to a diversity ordering proposed by Weitzman.

Key words : diversity, distance measures

1 Introduction

The biological diversity of species, diversity of languages, and cultural diversity are topics that are discussed frequently. Moreover, diversity is an issue of increasing importance in the literature on the ranking of opportunity sets, that is, sets of options available to a decison maker.

This paper presents an axiomatic approach to the measurement of diversity. Our approach is ordinal in nature: we seek to establish a ranking of sets of objects according to the diversity of their constituent elements. We will use the terms 'ordinal measure of diversity' and 'diversity ordering' synonymously in this paper.

We do not mean to suggest that diversity rankings are to be identified with measures of desirability; instead, diversity might be considered to be one of a set of criteria that may be considered relevant for the overall assessment of sets of objects such as opportunities. Furthermore, we do not examine the ethical arguments involved in discussions as to whether, and if yes, to what extent, diversity is to be considered desirable but, rather, suggest a way of measuring it in an ordinal manner. Therefore, the paper examines the descriptive aspect of diversity measurement. Given that diversity is a commonly-used term in many discussions regarding public policies, it is important to have a precise definition of what is meant when employing this term.

Weitzman (1992) suggests a measure of diversity that can be applied in a variety of contexts including the ones mentioned above. Though he discusses several desirable properties of his measure, no characterization is provided. The present paper complements Weitzman's (1992) analysis in two ways. First, we provide axiomatic characterizations of classes of ordinal measures of diversity that contain Weitzman's measure as a special case, and of Weitzman's measure itself. Second, we present an alternative formulation of his measure that may facilitate the computation of the measure for specific applications.

Pattanaik and Xu (2000) provide an approach to diversity measurement that explicitly takes into consideration that the elements of some set may be very different in nature, whereas in others, the alternatives can be considered very similar. Based on a notion of similarity that can be represented by means of a binary similarity relation, they provide a characterization of a similarity-based ordering. Weikard (1998) discusses a measure of diversity that extends Weitzman's proposed measure. Nehring and Puppe (1999) propose a multi-attribute approach to the measurement of diversity.

In this paper, we examine the possibility of establishing an ordering of sets of objects when there is more information available than just a binary similarity relation. In particular, we allow for different degrees of similarity between the objects under consideration by employing a distance or dissimilarity function. We provide characterizations of classes of diversity orderings with a plausible interpretation.

Section 2 introduces diversity orderings as ordinal measures of diversity. In Section 3, we present the axioms involved in our characterization results. In Section 4, we state and prove our characterization results. The equivalence of an important member of a class of measures characterized here and that proposed by Weitzman (1992) is established in Section 5, and Section 6 provides concluding remarks.

2 Diversity Orderings

Let $\mathbb{R}_+$ denote the set of nonnegative real numbers, and let $\mathbb{R}_+^n$ be the n-fold Cartesian product of $\mathbb{R}_+$, where n is a positive integer. X is a nonempty universal set of alternatives, and $\mathcal{K}$ denotes the class of all nonempty and finite subsets of X. A typical element A of $\mathcal{K}$ can be written as $A = \{a_1, \ldots, a_{|A|}\}$, where $|A|$ denotes the cardinality of A. A set $A \in \mathcal{K}$ can, for example, be interpreted as the set of species that are present in a region or as the set of options available to a decision maker.

Let $d: X \times X \to \mathbb{R}_+$ be a function such that $d(x, y) = 0$ if and only of $x = y$ for all $x, y \in X$, and $d(x, y) = d(y, x)$ for all $x, y \in X$. d is a distance function or a dissimilarity function. It indicates the diversity (or dissimilarity) between any two alternatives in the set X. We do not require d to satisfy the triangle inequality but note that this property could be added without changing our results. An additional 'richness property' regarding X and d is required for one of our results. We will introduce it when required in Section 4.

For all $x \in X$ and for all $A \in \mathcal{K}$, let

$$d(x, A) = \begin{cases} 0 & \text{if } A = \{x\}, \\ \min\{d(x, y) \mid y \in A - \{x\}\} & \text{if } A \neq \{x\}. \end{cases}$$

This is a measure of the distance or dissimilarity between an object $x \in X$ and other elements in a set $A \in \mathcal{K}$. In slight abuse of notation, we use the same symbol to denote the distance between two objects and the distance between an object and a set. However, this should not create any ambiguity.

Let $\succeq$ be a reflexive, transitive, and complete binary relation on $\mathcal{K}$. The statement $A \succeq B$ is to be interpreted as 'the diversity offered by the set $A \in \mathcal{K}$ is greater than or equal to the diversity offered by the set $B \in \mathcal{K}$.' The asymmetric and symmetric factors of $\succeq$ are denoted by $\succ$ and $\sim$.

The class of diversity orderings we suggest is based on the distances between each element of a set and the other elements. In order to avoid 'multiple-counting,' we introduce an iterative procedure of considering the distance between a specific element of a set and the other elements of the set (if any), then eliminating this element, and repeating the procedure until the set is exhausted. Note that we only compare finite sets and, therefore, this procedure is well-defined. The crucial aspect is, of course, the choice of the element to be eliminated at each stage of the iteration. We use a lexicographic criterion with respect to the minimal distance between an element and the rest of the sets in order to capture the objective of avoiding the multiple counting of distances.

For all $A \in \mathcal{K}$ and for all $x \in A$, let $\sigma_x^A \colon \{1, \ldots, |A|\} \to A$ be a (not necessarily unique) bijection such that $d(x, \sigma_x^A(1)) \le \ldots \le d(x, \sigma_x^A(|A|))$, and let $\delta_x(A) = (d(x, \sigma_x^A(1)), \ldots, d(x, \sigma_x^A(|A|)))$. Define the ordering $\le_L^A$ on A by letting, for all $x, y \in A$, $x \le_L^A y$ if and only if $\delta_x(A) \le_{lex}^A \delta_y(A)$, where $\le_{lex}^A$ is the lexicographic ordering on $\mathbb{R}_+^{|A|}$. The asymmetric and symmetric factors of $\le_L^A$ are denoted by $<_L^A$ and $=_L^A$.

For all $A \in \mathcal{K}$, let $\bar{A}_1 = A$ and $\bar{a}_1 \in \{x \in \bar{A}_1 \mid x \le_L^{\bar{A}_1} y \text{ for all } y \in \bar{A}_1\}$ and, recursively, for $j \in \{2, \ldots, |A|\}$, let $\bar{A}_j = \bar{A}_{j-1} - \{\bar{a}_{j-1}\}$ and $\bar{a}_j \in \{x \in \bar{A}_j \mid x \le_L^{\bar{A}_j} y$ for all $y \in \bar{A}_j\}$. We let $D(A) = (d(\bar{a}_1, \bar{A}_1), \ldots, d(\bar{a}_{|A|}, \bar{A}_{|A|}))$. Clearly, by definition, $d(\bar{a}_1, \bar{A}_1) \le \ldots \le d(\bar{a}_{|A|-1}, \bar{A}_{|A|-1})$ and $d(\bar{a}_{|A|}, \bar{A}_{|A|}) = 0$.

To illustrate the above definitions, consider the following example.

Example 2.1. Let $A = \{x, y, z, w\}$ and $d(x, y) = 2$, $d(x, z) = 1$, $d(x, w) = 2$, $d(y, z) = 3$, $d(y, w) = 2$, and $d(z, w) = 4$. This completely defines all pairwise distances for A; the remaining distances are determined by the symmetry of d and the property that the distance between an element and itself is equal to zero. By definition, $\bar{A}_1 = A$, and we have $\delta_x(\bar{A}_1) = (0, 1, 2, 2)$, $\delta_y(\bar{A}_1) = (0, 2, 2, 3)$, $\delta_z(\bar{A}_1) = (0, 1, 3, 4)$, $\delta_w(\bar{A}_1) = (0, 2, 2, 4)$ and, thus, $x <_L^{\bar{A}_1} z <_L^{\bar{A}_1} y <_L^{\bar{A}_1} w$. It follows that $\bar{a}_1 = x$. Next, we obtain $\bar{A}_2 = \{y, z, w\}$ and we have $\delta_y(\bar{A}_2) = (0, 2, 3)$, $\delta_z(\bar{A}_2) = (0, 3, 4)$, $\delta_w(\bar{A}_2) = (0, 2, 4)$ and, thus, $y <_L^{\bar{A}_2} w <_L^{\bar{A}_2} z$. Note that the relative ranking of the remaining alternatives y, z, w is *not* the same in $\le_L^{\bar{A}_1}$ and in $\le_L^{\bar{A}_2}$. We obtain $\bar{a}_2 = y$. Thus, we now have $\bar{A}_3 = \{z, w\}$ and $\delta_z(\bar{A}_3) = (0, 4)$ and $\delta_w(\bar{A}_3) = (0, 4)$ and, thus, $z =_L^{\bar{A}_2} w$. Therefore, there are two possibilities, namely, $\bar{a}_3 = z$ and $\bar{A}_4 = \{w\}$ or

$\bar{a}_3 = w$ and $\bar{A}_4 = \{z\}$. In the first case, we obtain $\bar{a}_4 = w$, and in the second, $\bar{a}_4 = z$. The resulting vector $D(A)$ is given by $(1, 2, 4, 0)$. ■

The classes of diversity orderings considered in this paper are based on the vectors $D(A)$. The idea is that only the distances recorded in those vectors matter in establishing the diversity ranking. This method of implicitly defining a class of orderings by focussing on the information required to rank any two sets is analogous to the median-based extension rules that can be found in nonprobabilistic decision models; see, for example, Nitzan and Pattanaik (1984) and Barberà, Bossert, and Pattanaik (2001).

Definition 2.2. A diversity ordering $\succeq$ is a *general lexicographic-distance-based ordering* if it satisfies

$$D(A) = D(B) \Rightarrow A \sim B \tag{1}$$

for all $A, B \in \mathcal{K}$. ■

Thus, the ranking of sets according to a general lexicographic-distance-based ordering depends on the vectors of distances as defined above only.

Note that (1) is restricted to the comparison of sets with the same cardinality. The corresponding class of orderings can be narrowed down further by adding an appropriate requirement to (1) in order to obtain the class of *lexicographic-distance-based orderings.*

Definition 2.3. A diversity ordering $\succeq$ is a *lexicographic-distance-based ordering* if it satisfies (1) and, in addition,

$$[|A| > |B| \text{ and } D_i(A) = D_i(B) \text{ for all } i \in \{1, \ldots, |B|\}] \Rightarrow A \succ B \tag{2}$$

for all $A, B \in \mathcal{K}$. ■

The lexicographic-distance-based orderings represent a plausible subclass of the class of restricted lexicographic-distance-based orderings. In addition to depending on the vectors $D(A)$ only, the resulting orderings have a plausible monotonicity property in that they declare a larger set whose initial distances according to the iterative procedure introduced earlier are equal to the corresponding distances of a smaller set to be more diverse than this smaller set.

There is an interesting special case of the lexicographic-distance-based orderings. If the aggregation of the distances involved proceeds in an additive fashion, we obtain an ordering that turns out to be equivalent to a diversity measure suggested by Weitzman (1992). This equivalence result is established in Section 5. The ordering based on an additive procedure is defined as follows.

Definition 2.4. Define the function $U\colon \mathcal{K} \to \mathbb{R}_+$ by letting, for all $A \in \mathcal{K}$, $U(A) = \sum_{j=1}^{|A|} d(\bar{a}_j, \bar{A}_j)$. Given U, the ordering $\succeq_U$ is defined by letting, for all $A, B \in \mathcal{K}$,

$$A \succeq_U B \Leftrightarrow U(A) \geq U(B). \ \blacksquare$$

To conclude this section, we use our earlier example to illustrate how the values of the function U used in the definition of $\succeq_U$ can be calculated.

Example 2.5. Let $A = \{x, y, z, w\}$ and $d(x, y) = 2$, $d(x, z) = 1$, $d(x, w) = 2$, $d(y, z) = 3$, $d(y, w) = 2$, and $d(z, w) = 4$. As in Example 2.1, we obtain $\bar{a}_1 = x$, $\bar{a}_2 = y$, $\bar{a}_3 = z$ (or w), and $\bar{a}_4 = w$ (or z). Therefore,

$$\begin{aligned} U(A) &= d(x, A) + d(y, \{y, z, w\}) + d(z, \{z, w\}) + d(w, \{w\}) \\ &= d(x, A) + d(y, \{y, z, w\}) + d(w, \{z, w\}) + d(z, \{z\}) \\ &= 1 + 2 + 4 + 0 = 7. \ \blacksquare \end{aligned}$$

3 Axioms

In this section, we introduce the axioms that will be used in our characterization results (see Section 4).

Our first axiom requires a diversity ordering to declare all singleton sets—that is, all sets with no diversity at all—to be indifferent. The intuitive appeal of this axiom is evident.

Simple Indifference: For all $x, y \in X$,

$$\{x\} \sim \{y\}.$$

The next axiom stipulates that a set with more than one element is more diverse than a set with no diversity at all. Again, the interpretation of this condition is obvious. Note that, by assumption, the pairwise distance between any two objects in the set with more than one element is positive (see our assumptions on the properties of d) and, thus, it is very plausible to rank a set with more than one element as more diverse than a singleton.

Weak Monotonicity: For all $A \in \mathcal{K}$, for all $x \in X$, if $|A| > 1$, then

$$A \succ \{x\}.$$

Another plausible property relates the ranking of singletons and two-element sets to the ranking of the pairwise distances according to d.

Simple Monotonicity: For all $x, y, z, w \in X$,

$$\{x, y\} \succeq \{z, w\} \Leftrightarrow d(x, y) \geq d(z, w).$$

The next two axioms are independence conditions that require the addition of alternatives in accordance with the lexicographic procedure established in the previous section to two sets to be irrelevant for their relative ranking. The first of the two axioms is restricted to the comparison of sets with a fixed cardinality, and the second applies to more general situations.

Restricted Independence: For all $A, B \in \mathcal{K}$, for all $x \in X - A$, for all $y \in X - B$, if $|A| = |B|$, $x \leq_L^{A\cup\{x\}} a$ for all $a \in A$, $y \leq_L^{B\cup\{y\}} b$ for all $b \in B$, and $d(x, A) = d(y, B)$, then

$$A \succeq B \Leftrightarrow A \cup \{x\} \succeq B \cup \{y\}.$$

Independence: For all $A, B \in \mathcal{K}$, for all $x \in X - A$, for all $y \in X - B$, if $x \leq_L^{A\cup\{x\}} a$ for all $a \in A$, $y \leq_L^{B\cup\{y\}} b$ for all $b \in B$, and $d(x, A) = d(y, B)$, then

$$A \succeq B \Leftrightarrow A \cup \{x\} \succeq B \cup \{y\}.$$

To introduce our final axiom, we first define the notion of a *link element.* Although not identical, our formulation of a link element is related to Weitzman's *link property* (see Weitzman, 1992, p. 378). Weitzman postulates, for every set $A \in \mathcal{K}$, the existence of a 'link species' defined as an element x of A such that the value of a representation of the diversity ordering at A is equal to the sum of the value of this representation at $A - \{x\}$ and the distance between x and $A - \{x\}$. Note that this property of a representation uses more than just ordinal features of the diversity ordering, which is one reason why our formulation is different.

For all $x, y, a, b \in X$ with $x \neq y$ and $a \neq b$, we say that $c \in X - \{a, b\}$ is a *link element of* $\{a, b\}$ *relative to* $\{x, y\}$ if and only if $\max\{d(a, c), d(b, c)\} \leq d(a, b)$ and $d(x, y) = d(a, b) + d(c, \{a, b\})$. For given sets of elements, $\{x, y\}$ and $\{a, b\}$, an element c is a link element of $\{a, b\}$ relative to $\{x, y\}$ if the distance between c and $\{a, b\}$ is no greater than the distance between a and b, and if the distance between

x and y is the sum of the distance between a and b and the distance between c and $\{a, b\}$.

We can now state our last axiom.

Link Indifference: For all $x, y, a, b, c \in X$ with $x \neq y$ and a, b, c being pairwise distinct, if $\{x, y\} \succ \{a, b\}$ and c is a link element of $\{a, b\}$ relative to $\{x, y\}$, then

$$\{x, y\} \sim \{a, b, c\}.$$

4 Characterizations

Our first characterization result provides an axiomatization of the class of general lexicographic-distance-based orderings.

Theorem 4.1. A diversity ordering $\succeq$ satisfies simple indifference and restricted independence if and only if $\succeq$ is a general lexicographic-distance-based ordering.

Proof. That the general lexicographic-distance-based orderings satisfy the required axioms is straightforward to verify. Now suppose $\succeq$ satisfies simple indifference and restricted independence. Let $A, B \in \mathcal{K}$ be such that $D(A) = D(B)$. Clearly, this implies $|A| = |B|$, and we can apply restricted independence as many times as required to obtain

$$A \succeq B \Leftrightarrow \{\bar{a}_{|A|}\} \succeq \{\bar{b}_{|B|}\}. \tag{3}$$

By simple indifference, $\{\bar{a}_{|A|}\} \sim \{\bar{b}_{|B|}\}$, and (3) implies $A \sim B$, which establishes (1) and completes the proof. ■

Next, we narrow down the class characterized in the previous theorem by axiomatizing the lexicographic-distance-based orderings.

Theorem 4.2. A diversity ordering $\succeq$ satisfies simple indifference, weak monotonicity, and independence if and only if $\succeq$ is a lexicographic-distance-based ordering.

Proof. Again, it is straightforward to verify that the lexicographic-distance-based orderings satisfy the required axioms. Now suppose $\succeq$ satisfies simple indifference, weak monotonicity, and independence. That (1) is satisfied follows from the previous theorem. To prove (2), suppose $|A| > |B|$ and $D_i(A) = D_i(B)$ for all $i \in \{1, \ldots, |B|\}$. By applying independence as many times as required, we obtain

$$A \succeq B \Leftrightarrow \{\bar{a}_{|B|}, \ldots, \bar{a}_{|A|}\} \succeq \{\bar{b}_{|B|}\}. \tag{4}$$

Because $|A| > |B|$, we have $|\{\bar{a}_{|B|}, \ldots, \bar{a}_{|A|}\}| > 1$, and weak monotonicity implies $\{\bar{a}_{|B|}, \ldots, \bar{a}_{|A|}\} \succ \{\bar{b}_{|B|}\}$. By (4), we obtain $A \succ B$, which completes the proof. ■

Although the characterizations of Theorems 4.1 and 4.2 are not very deep, they have some use because they clarify the conditions under which the vectors $D(A)$ for all $A \in \mathcal{K}$ are the sole determinants of the diversity ranking to be established.

We now move on to a more substantial characterization result involving the specific diversity ordering $\succeq_U$. To do so, we need an additional richness assumption regarding X and d. This is necessary because, in order to be able to invoke the axiom link indifference, it has to be ensured that link elements exist in certain circumstances. This richness assumption is satisfied, for example, if X is a Euclidean space and d is the Euclidean metric. We obtain

Theorem 4.3. Suppose that, for all $s, t \in \mathrm{I\!R}_+$ with $s \geq t$, there exist $x, y, z \in X$ such that $t = d(x, y) \leq d(x, z) \leq d(y, z) = s$. A diversity ordering $\succeq$ satisfies simple monotonicity, independence, and link indifference if and only if $\succeq = \succeq_U$.

Proof. It is straightforward to verify that $\succeq_U$ satisfies the required axioms. Now suppose $\succeq$ is an ordering on $\mathcal{K}$ satisfying simple monotonicity, independence, and link indifference. First, we show that

$$\sum_{j=1}^{|A|} d(\bar{a}_j, \bar{A}_j) = d(x, y) \Rightarrow A \sim \{x, y\} \tag{5}$$

for all $A \in \mathcal{K}$ and for all $x, y \in X$.

Step 1: If $|A| \leq 2$, (5) follows immediately from simple monotonicity.

Step 2: Suppose now that A contains three elements so that $A = \{a, b, c\}$. Let $a \leq_L^A b$ and $a \leq_L^A c$ and $d(a, A) + d(b, c) = d(x, y)$. Clearly, $d(a, b) \leq d(b, c)$ and $d(a, c) \leq d(b, c)$. Note that $d(a, A) + d(x, y) = d(a, \{b, c\}) + d(x, y)$. Therefore, a is a link element of $\{b, c\}$ relative to $\{x, y\}$. By link indifference, we obtain $A \sim \{x, y\}$.

Step 3: Suppose that, for all positive integers k, for all $A \in \mathcal{K}$ with $|A| \leq k$, and for all $x, y \in X$, (5) is satisfied. We show that this implies (5) for all $A \in \mathcal{K}$ with $|A| = k + 1$ and for all $x, y \in X$.

Let $A = \{a_1, a_2, \ldots, a_{k+1}\} \in \mathcal{K}$ and $x, y \in X$ be such that $a_1 \leq_L^A a$ for all $a \in A$ and $d(x, y) = d(a_1, A) + d(\bar{a}_2, \bar{A}_2) + \ldots + d(\bar{a}_k, \bar{A}_k) + d(\bar{a}_{k+1}, \bar{A}_{k+1}) = d(a_1, A) + \alpha$, where $\alpha = d(\bar{a}_2, \bar{A}_2) + \ldots + d(\bar{a}_k, \bar{A}_k) + d(\bar{a}_{k+1}, \bar{A}_{k+1})$. It is then clear that $\alpha \geq d(a_1, A)$. By our richness assumption stated in the theorem, there exist $z, w, u \in X$ such that $d(z, w) = d(a_1, A) \leq d(w, u) \leq d(z, u) = \alpha$. By construction, $w \leq_L^{\{z,w,u\}} z$ and

$w \leq_L^{\{z,w,u\}} u$. From Step 2, noting that $d(w, \{z, u\}) + d(z, u) = d(a_1, A) + \alpha = d(x, y)$, we obtain $\{z, w, u\} \sim \{x, y\}$. Because $w \leq_L^{\{z,w,u\}} z$, $w \leq_L^{\{z,w,u\}} u$, $a_1 \leq_L^A a$ for all $a \in A$ and $d(w, \{z, u\}) = d(a_1, A)$, independence implies

$$\{z, w, u\} \succeq A \Leftrightarrow \{z, u\} \succeq A - \{a_1\}.$$

But $d(z, u) = \alpha = \sum_{j=2}^{k+1} d(\bar{a}_j, \bar{A}_j)$. Because $|A - \{a_1\}| = k$,

$$\{z, u\} \sim A - \{a_1\}$$

follows immediately from our hypothesis. Hence, $\{z, w, u\} \sim A$. By transitivity, from $\{z, w, u\} \sim \{x, y\}$, we obtain $A \sim \{x, y\}$. This completes the proof of (5) for all $A \in \mathcal{K}$ and all $x, y \in X$.

Let $A, B \in \mathcal{K}$ be arbitrary. First, we note that, from the richness assumption, for all $t \in \mathbb{R}_+$, there exist $x, y \in X$ such that $d(x, y) = t$. Therefore, there exist $x_A, y_A, x_B, y_B \in X$ such that

$$\sum_{j=1}^{|A|} d(\bar{a}_j, \bar{A}_j) = d(x_A, y_A) \text{ and } \sum_{j=1}^{|B|} d(\bar{b}_j, \bar{B}_j) = d(x_B, y_B). \tag{6}$$

By (5), $A \sim \{x_A, y_A\}$ and $B \sim \{x_B, y_B\}$. Therefore,

$$A \succeq B \Leftrightarrow \{x_A, y_A\} \succeq \{x_B, y_B\}.$$

By simple monotonicity,

$$\{x_A, y_A\} \succeq \{x_B, y_B\} \Leftrightarrow d(x_A, y_A) \geq d(x_B, y_B).$$

Combining the last two equivalences and using (6), we obtain

$$A \succeq B \Leftrightarrow U(A) \geq U(B). \blacksquare$$

5 An Equivalence Result

In this section, we show that our ordering $\succeq_U$ (see Definition 2.4) is equivalent to an ordering which was proposed by Weitzman (1992) and which we denote by $\succeq_V$. The ordering $\succeq_V$ is based on a representation V of this ordering, introduced by Weitzman (1992).

Definition 5.1. Define the function $V: \mathcal{K} \to \mathbb{R}_+$ recursively by letting $V(A) = 0$ for all $A \in \mathcal{K}$ such that $|A| = 1$, and $V(A) = \max_{a_i \in A}\{V(A - \{a_i\}) + d(a_i, A)\}$ for all $A \in \mathcal{K}$ such that $|A| \geq 2$. Given V, the ordering $\succeq_V$ is defined by letting, for all $A, B \in \mathcal{K}$,

$$A \succeq_V B \Leftrightarrow V(A) \geq V(B). \ \blacksquare$$

Strictly speaking, Weitzman's formulation of V is slightly more general because he allows for V assigning any value d_0 to singleton sets. However, it is clear that our normalization does not involve any substantive restriction.

The following theorem shows that the functions U and V are identical and, thus, the diversity ordering $\succeq_U$ is equal to the ordering $\succeq_V$.

Theorem 5.2. For all $A \in \mathcal{K}$,

$$U(A) = V(A).$$

Proof. We proceed by induction. Clearly, $V(A) = U(A)$ for all $A \in \mathcal{K}$ such that $|A| \leq 2$. Now suppose $V(A) = U(A)$ is true for all $A \in \mathcal{K}$ such that $|A| \leq n$ with $n \geq 2$. Let $A \in \mathcal{K}$ be such that $|A| = n + 1$. From the induction hypothesis,

$$V(A - \{\bar{a}_1\}) = U(A - \{\bar{a}_1\}). \tag{7}$$

Because, by definition of U,

$$U(A) = d(\bar{a}_1, A) + U(A - \{\bar{a}_1\}),$$

it follows that

$$U(A) = d(\bar{a}_1, A) + V(A - \{\bar{a}_1\}). \tag{8}$$

From the definition of $\bar{a}_1$,

$$d(\bar{a}_1, A) \leq d(a, a') \text{ for all } a, a' \in A \text{ such that } a \neq a'. \tag{9}$$

Weitzman's fundamental representation theorem (see Weitzman (1992, p. 384)) states that $V = W$, where $W\colon \mathcal{K} \to \mathbb{R}_+$ is defined as follows. Let $A \in \mathcal{K}$. If $|A| = 1, W(A) = 0$, and if $|A| \geq 2$, $W(A)$ is the solution of the dynamic programming recursion

$$W(A) = d(g, h) + \max\{W(A - \{g\}), W(A - \{h\})\},$$

where g and h are such that $d(g, h) \leq d(a, A)$ for all $a \in A$. Because $W(A - \{g\}) = V(A - \{g\})$ and $W(A - \{h\}) = V(A - \{h\})$ by Weitzman's theorem, it follows that $U(A - \{g\}) = W(A - \{g\})$ and $U(A - \{h\}) = W(A - \{h\})$. From (9), $\bar{a}_1 \in \{g, h\}$. Suppose, without loss of generality, $\bar{a}_1 = g$. Then $\max\{U(A - \{g\}), U(A - \{h\})\} =$

$\max\{U(A-\{\bar{a}_1\}), U(A-\{h\})\}$. By definition of $\bar{a}_1$, $d(\bar{a}_1, A-\{h\}) \leq d(h, A-\{\bar{a}_1\})$ and hence $U(A-\{\bar{a}_1\}) \geq U(A-\{h\})$. Therefore, $\max\{U(A-\{\bar{a}_1\}), U(A-\{h\})\} = U(A-\{\bar{a}_1\})$. Hence, noting (7), we have

$$V(A) = d(\bar{a}_1, A) + V(A-\{\bar{a}_1\}).$$

Thus, together with (8), it follows that $U(A) = V(A)$. ■

6 Concluding Remarks

The measurement of diversity is an issue that is of importance in many areas. The ordinal diversity measures suggested in this paper have some useful properties and, moreover, a special case turns out to be an ordering that has been advocated in the earlier literature. We have shown that this special case is equivalent to the diversity ordering proposed by Weitzman (1992). An interesting question is whether there are other members of the classes of orderings presented here that can be axiomatized by means of plausible properties.

References

Barberà, S., W. Bossert, and P.K. Pattanaik (2001), "Ranking sets of objects," in: S. Barberà, P. Hammond, and C. Seidl (eds.), *Handbook of Utility Theory, Vol. 2: Extensions,* Kluwer, Dordrecht, forthcoming.

Nehring, K. and C. Puppe (1999), "A theory of diversity," *Discussion Paper* **A-605**, University of Bonn.

Nitzan, S. and P.K. Pattanaik (1984), "Median-based extensions of an ordering over a set to the power set: an axiomatic characterization," *Journal of Economic Theory* **34**, 252–261.

Pattanaik, P.K. and Y. Xu (2000), "On diversity and freedom of choice," *Mathematical Social Sciences* **40**, 123–130.

Weikard, H.-P. (1998), "On the measurement of diversity," *Discussion Paper* **9801**, University of Graz, Institute of Public Economics.

Weitzman, M. (1992), "On diversity," *Quarterly Journal of Economics* **107**, 363–405.