Mathematical Social Sciences 45 (2003) 167–183
mathematical social sciences
www.elsevier.com/locate/econbase

Diversity and dissimilarity in lines and hierarchies

Klaus Nehring[a], Clemens Puppe[b,*]

[a]Department of Economics, University of California at Davis, Davis, CA 95616, USA
[b]Department of Economics, University of Bonn, Adenauerallee 24-42, Bonn 53113, Germany

Abstract

Within the multi-attribute framework of Nehring and Puppe [Econometrica, 70 (2002) 1155], hierarchies and lines represent the simplest and most fundamental models of diversity. In both cases, the diversity of any set can be recursively determined from the pairwise dissimilarities between its elements. The present paper characterizes the restrictions on the dissimilarity metric entailed by the two models. In the hierarchical case, this generalizes a classical result on the representation of ultrametric distance functions.

Keywords: Diversity; Dissimilarity; Conjugate Moebius inversion; Hierarchies; Ultrametric

JEL classification: C00; D00; Q20

1. Introduction

In 'A Theory of Diversity' (Nehring and Puppe, 2002a, henceforth TD), we proposed a multi-attribute approach according to which the diversity of a set of objects is determined by the number and weight of the different features ('attributes') possessed by them. In some cases, the diversity of a set can be computed recursively from the pairwise dissimilarities between its elements (plus their value as singletons). Two basic models for which this is possible are the hierarchical model studied by Weitzman (1992, 1998) in the context of biodiversity and the more general line model introduced in TD. The line model assumes that objects can be linearly ordered in such a way that any attribute possessed by two objects is also possessed by all intermediate objects. The hierarchical model in addition requires the relevant attributes to be 'nested' in the sense

*Corresponding author. Tel.: +49-228-737-996; fax: +49-228-737-900.
E-mail address: clemens.puppe@wiwi.uni-bonn.de (C. Puppe).

0165-4896/03/$ – see front matter
doi:10.1016/S0165-4896(03)00024-6

that, for any two attributes possessed by a given object, one is (unambiguously) more specific than the other. As argued in Nehring and Puppe (1999b), lines and hierarchies can serve as useful benchmark models not only in the context of biodiversity but also in the analysis of the cost structure of multi-product firms that operate under economies of scope; in this context, attributes correspond to shared inputs and costs reflect the technological diversity of a range of products.[1] The purpose of the present paper is to characterize the restrictions on the induced dissimilarity metric imposed by the line and hierarchy models. Specifically, we provide the necessary and sufficient conditions for a given dissimilarity metric to be extendable to a diversity function that is compatible with the line and with the hierarchy model, respectively. This is important for practical purposes, since the entailed restrictions on the dissmilarity metric represent a primary criterion for the applicability of a particular model.

As already observed by Weitzman (1992), the hierarchical model implies that the two greatest dissimilarities between three points are always equal if singletons are equally valued ('ultrametricity'). We show that a generalization of this condition characterizes the metric implications of the class of hierarchical models. While in some contexts ultrametricity or appropriate weakenings of it may be plausible, distance functions that arise in specific applications will only exceptionally exhibit the required property. For instance, in a biological context, there is no reason why genetic distances between species should satisfy ultrametricity. On the other hand, from the viewpoint of phylogenetic (rather than genetic) diversity a closely related but weaker condition is still applicable, as argued in Nehring and Puppe (2002b).

The main restriction entailed by the line model is a submodularity condition according to which an increase in the gap between two elements results in a smaller increase in dissimilarity the larger the gap already is. A particularly transparent special case is the translation invariant case: A given dissimilarity metric is induced by a translation invariant diversity function on a line if and only if dissimilarity is a concave transform of Euclidean distance. In terms of the underlying preferences, this translates into a 'preference for even spacing.'

The plan of the paper is as follows. Section 2 provides the necessary background from TD and from 'Diversity and the Geometry of Similarity' (Nehring, 1999, henceforth DGS). Sections 3 and 4 reconsider in more detail the hierarchy and line models, respectively. Sections 5 and 6 are devoted to the 'extension problem,' i.e. the conditions under which a given dissimilarity metric can be extended to a diversity function. In Section 5, we derive the necessary and sufficient conditions for a *given* hierarchy and a *given* line model, respectively. Section 6 addresses the extension problem then in generality. Specifically, we ask when a dissimilarity metric can be extended to a diversity function that is compatible with *some* hierarchical model and *some* line model, respectively. Generalizing a classical result on the representation of ultrametric distance functions, we give a complete answer in the hierarchical case. In the context of a line, we provide a characterization for the case of a 'sufficiently rich' family of relevant attributes. All proofs are collected in Appendix A.

[1] The line and hierarchy models are less natural in other applications of diversity theory such as, e.g., to the measurement of opportunity, see Nehring and Puppe (2002c).

2. Background

This section reviews the concepts and tools needed for our later analysis. First, we summarize the basic features of the multi-attribute model developed in TD; we then present the notion of qualitative (comparative) similarity introduced in DGS.

2.1. The multi-attribute model of diversity

Let X be a finite universe of objects. The basic idea behind the multi-attribute model is to view the diversity of a set $S \subseteq X$ of objects as being determined by the number and the value of the different features possessed by the objects in S. Throughout, we refer to features in terms of their *extension*, i.e. we identify any feature with the subset $A \subseteq X$ of those objects in the universe that possess the feature in question. For instance, the feature 'mammal' is identified with the set of all mammals in the universe. Extensionally identified features are henceforth referred to as *attributes*. Note that, given a prespecified universe X of objects, *any* conceivable feature corresponds to a particular subset $A \subseteq X$; conversely, any subset $A \subseteq X$ defines a logically possible attribute ('belonging to A'). The set of conceivable attributes is thus given by the power set 2^X. If the subset $A \in 2^X$ is interpreted as an attribute, the statement '$x \in A$' simply means 'object x possesses the attribute A.' Similarly, a set S *realizes* an attribute A if and only if $S \cap A \neq \emptyset$, i.e. if and only if there exists some object in S that possesses the attribute A.

For each attribute A, let $\lambda_A \geqslant 0$ quantify the value of the realization of A. Upon normalization, λ_A can thus be thought of as the relative importance, or *weight* of the attribute A. A function $v:2^X \to \mathbf{R}$ is called a *diversity function* if there exists a function $\lambda:2^X \to \mathbf{R}$ with $\lambda_A \geqslant 0$ for all A, such that for all $S \subseteq X$,

$$v(S) := \sum_{A \subseteq X : A \cap S \neq \emptyset} \lambda_A, \tag{2.1}$$

where, by convention, $v(\emptyset) = \lambda_\emptyset = 0$.

The cardinal scale inherent in our concept of diversity is essential; for a rigorous decision-theoretic justification, see Section 2 of TD. An alternative, ordinal approach to the measurement of diversity is provided in Bossert et al. (2001) (see also Pattanaik and Xu, 2000).

According to (2.1), the diversity value of a set of objects is given by the total weight of all attributes realized by the set. Note especially that each attribute occurs at most once in the sum. In particular, each single object contributes to diversity the value of all those features that are not possessed by any already existing species.

Technically, the function λ that assigns to each attribute A its weight λ_A is known as the *conjugate Moebius inverse*; we will also refer to it as the *attribute weighting function*. The attribute weighting function underlying a diversity function is *uniquely* determined, as shown by the following result.

Fact 2.1. (Conjugate Moebius inversion) For any function $v:2^X \to \mathbf{R}$ with $v(\emptyset) = 0$ there exists a unique function $\lambda:2^X \to \mathbf{R}$, the *conjugate Moebius inverse*, such that $\lambda_\emptyset = 0$ and, for all S,

$$v(S) = \sum_{A: A \cap S \neq \emptyset} \lambda_A.$$

Furthermore, the conjugate Moebius inverse λ is given by the following formula. For all $A \neq \emptyset$,

$$\lambda_A = \sum_{S \subseteq A} (-1)^{\#(A \setminus S)+1} \cdot v(S^c),$$

where S^c denotes the complement of S in X.

For any set function $v: 2^X \to \mathbf{R}$ denote by

$$\Lambda := \{A: \lambda_A \neq 0\}$$

the *support* of the corresponding conjugate Moebius inverse. If v is a diversity function, the elements of Λ are those attributes that have strictly positive weight; in this case, the support is also referred to as the family of *relevant* attributes.

By Fact 2.1, the only restriction imposed on a diversity function is non-negativity of the corresponding conjugate Moebius inverse. In terms of the function v itself non-negativity of λ corresponds to the following two properties. A function $v: 2^X \to \mathbf{R}$ is *monotone* if $W \subseteq S$ implies $v(W) \leq v(S)$. Furthermore, v is called *totally submodular* if, for any collection $\{S_i\}_{i \in I}$,

$$v\left(\bigcap_{i \in I} S_i\right) \leq \sum_{J: \emptyset \neq J \subseteq I} (-1)^{\# J+1} \cdot v\left(\bigcup_{i \in J} S_i\right). \tag{2.2}$$

Fact 2.2. The function $v: 2^X \to \mathbf{R}$ has a non-negative conjugate Moebius inverse if and only if v is monotone and totally submodular.

The basic instance of total submodularity is given by the case of $\# I = 2$ in which (2.2) specializes to the following condition known as *submodularity*. For all S_1, S_2,

$$v(S_1 \cap S_2) + v(S_1 \cup S_2) \leq v(S_1) + v(S_2),$$

or equivalently, for all S, W and all x,

$$S \subseteq W \Rightarrow v(S \cup \{x\}) - v(S) \geq v(W \cup \{x\}) - v(W). \tag{2.3}$$

i.e. the marginal value of additional objects decreases with the set of objects already available. Submodularity captures the fundamental intuition that it becomes harder for an object to add to the diversity of a set the larger that set already is.

That any diversity function satisfies (2.3) follows from noting that

$$v(S \cup \{x\}) - v(S) = \sum_{A \ni x, A \cap S = \emptyset} \lambda_A, \tag{2.4}$$

which is decreasing in S due to the non-negativity of λ. By (2.4), the marginal diversity of an object x at a set S is given by the total weight of all attributes possessed by x but by no element of S. Accordingly, we will refer to the marginal diversity also as the *distinctiveness* of x from S, which we denote by

$$d(x, S) := v(S \cup \{x\}) - v(S).$$

A diversity function naturally induces a notion of pairwise *dissimilarity* between objects as follows. For all x, y,

$$d(x, y) := d(x, \{y\}) = v(\{x, y\}) - v(\{y\}). \tag{2.5}$$

The dissimilarity $d(x, y)$ from x to y is thus simply the marginal diversity of x in a situation in which y is the only other existing object. Equivalently, by (2.4), $d(x, y)$ is the weight of all attributes possessed by x but not by y. Note that, in general, d need not be symmetric, and thus fails to be a proper metric; it does, however, always satisfy the triangle inequality. To verify this, we have to show that $d(x, z) \leqslant d(x, y) + d(y, z)$, or equivalently by (2.5) and (2.4),

$$\sum_{A: x \in A, z \notin A} \lambda_A \leqslant \sum_{A: x \in A, y \notin A} \lambda_A + \sum_{A: y \in A, z \notin A} \lambda_A.$$

Consider any λ_A that occurs as a summand on the left hand side, i.e. suppose that $x \in A$ and $z \notin A$. If $y \notin A$, λ_A occurs as a summand in the first sum on the right hand side; and if $y \in A$, λ_A occurs as a summand in the second sum on the right hand side, the desired inequality thus follows from the non-negativity of λ. The function d is symmetric if and only if $v(\{x\}) = v(\{y\})$ for all $x, y \in X$, i.e. if and only if all single objects have identical diversity value. A diversity function that gives equal value to all singletons is referred to as a *uniform* diversity function.

Often it will be useful to consider the following derived notion of quantitative similarity (in contrast to the qualitative ternary similarity relation introduced below). For all x, y let

$$\sigma(x, y) := v(\{x\}) + v(\{y\}) - v(\{x, y\}) = \sum_{A \supseteq \{x, y\}} \lambda_A \tag{2.6}$$

denote the (*quantitative*) *similarity* between x and y. Note that in contrast to the dissimilarity function d the similarity function σ is always symmetric. Also observe that

$$\sigma(x, y) = v(\{x\}) - d(x, y) = v(\{y\}) - d(y, x).$$

By Facts 2.1 and 2.2, any diversity function uniquely 'reveals' the underlying collection of attributes and their weights. In particular, any diversity function uniquely determines the corresponding family Λ of relevant attributes. The major theme of TD is to exploit this basic fact in order to characterize qualitative properties of diversity functions in terms of corresponding properties of the associated family of relevant attributes. Central to this is the following notion. A non-empty family of attributes $\mathscr{A} \subseteq 2^X \setminus \{\emptyset\}$ is referred to as a *model* of diversity. A diversity function v is *compatible* with the model $\mathscr{A}$ if the corresponding set Λ of relevant attributes is contained in $\mathscr{A}$, i.e. if $\Lambda \subseteq \mathscr{A}$. A model thus represents a *qualitative* a priori restriction, namely that no attributes outside $\mathscr{A}$ can have strictly positive weight. Accordingly, a model can be interpreted as a family of *potentially* relevant attributes, in contrast to the possibly larger set 2^X of all conceivable attributes and the possibly smaller set Λ of all actually relevant attributes. The two most basic examples of models, hierarchies and lines, are studied in detail below; further examples are discussed and analyzed in TD and DGS.

In practical applications, one will have to construct the diversity function from primitive data. One possibility is, of course, to first determine appropriate attribute weights and to compute the diversity function according to (2.1). Determining attribute weights is a complex task, however, since there are as many potential attributes as there are non-empty subsets of objects. An appealing alternative is to try to derive the diversity of a set from the pairwise dissimilarities between its elements, as suggested by Weitzman (1992). Say that a model $\mathscr{A}$ is *monotone in dissimilarity* if, for any compatible diversity function v and any S, the diversity $v(S)$ is uniquely determined by the value of all single species in S and the pairwise dissimilarities within S, and if, moreover, the diversity $v(S)$ is a monotone function of these dissimilarities.

The characterization of the class of models that are monotone in dissimilarity is one of the main results of TD. Say that a model $\mathscr{A}$ is *acyclic* if for no $m \geqslant 3$ there exist objects $x_1, \ldots, x_m$ and attributes $A_1, \ldots, A_m \in \mathscr{A}$ such that, for all $i = 1, \ldots, m-1$ $A_i \cap \{x_1, \ldots, x_m\} = \{x_i, x_{i+1}\}$, and $A_m \cap \{x_1, \ldots, x_m\} = \{x_m, x_1\}$. Thus, for instance in the case $m = 3$, acyclicity requires that there be no triple of objects such that each pair of them possesses an attribute that is not possessed by the third object. Theorem 6.2 in TD shows that a model of diversity is monotone in dissimilarity if and only if it is acyclic.[2] The two classes of models studied here, hierarchies and lines, are both acyclic, hence monotone in dissimilarity.

2.2. Qualitative similarity

As noted in DGS, any family $\mathscr{A} \subseteq 2^X$ of potentially relevant attributes naturally induces a *comparative similarity* relation $T_{\mathscr{A}} \subseteq X^3$ as follows. For all x, y, z, let

$$(x, y, z) \in T_{\mathscr{A}} :\Leftrightarrow [\text{for all } A \in \mathscr{A} : \{x, z\} \subseteq A \Rightarrow y \in A]. \tag{2.7}$$

In this definition, the statement '$(x, y, z) \in T_{\mathscr{A}}$' is interpreted as '$y$ is more similar than z to x,' which expresses an understanding of similarity as commonality of attributes: For y to be more similar than z to x, y must possess every attribute shared by x and z. Observe that judgments on qualitative similarity will typically change with the inclusion of further attributes. In particular, the larger the set of relevant attributes, the smaller the qualitative similarity relation, since each attribute can be viewed as a 'test' that has to be passed by any triple in $T_{\mathscr{A}}$.

For any family $\mathscr{A}$, the ternary relation $T_{\mathscr{A}}$ satisfies the following three properties:

T1 (Reflexivity) $y \in \{x, z\} \Rightarrow (x, y, z) \in T_{\mathscr{A}}$.

T2 (Symmetry) $(x, y, z) \in T_{\mathscr{A}} \Leftrightarrow (z, y, x) \in T_{\mathscr{A}}$.

T3 (Transitivity) $[(x, x', z) \in T_{\mathscr{A}}$ and $(x, z', z) \in T_{\mathscr{A}}$ and $(x', y, z') \in T_{\mathscr{A}}] \Rightarrow (x, y, z) \in T_{\mathscr{A}}$.

[2] The necessity of acyclicity hinges on a weak regularity requirement.

For any ternary relation $T \subseteq X^3$, and any $x \in X$, denote by T^x the following binary relation. For all y, z,

$$yT^x z :\Leftrightarrow (x, y, z) \in T. \tag{2.8}$$

In view of (2.7), the binary relation $T^x_{\mathscr{A}}$ describes commonality of attributes *with* x; specifically, $yT^x_{\mathscr{A}}z$ means that y shares (weakly) more attributes with x than z. Note that by T1–T3, the binary relation $T^x_{\mathscr{A}}$ is a preorder (i.e. reflexive and transitive).

A key insight of DGS is the observation that the comparative similarity relation T_Λ associated with the support Λ of a particular diversity function has an 'observational equivalent' in terms of the diversity function itself. Specifically, given any diversity function v say that x is *independent from* z *conditional on* (*the inclusion of*) y, denoted by $(x, y, z) \in T_v$, if the distinctiveness of x from any set S that includes y does not change with the addition of z to S. Formally,

$$(x, y, z) \in T_v :\Leftrightarrow \text{for all } S \ni y,\ d(x, S) = d(x, S \cup \{z\}).$$

A central result in DGS establishes that for any set function v, $T_\Lambda = T_v$, that is: y is more similar than z to x if and only if x is independent from z conditional of y.

Intuitively, the relation T_Λ $(= T_v)$ induced by a diversity function v can be viewed as the 'qualitative core' of the corresponding quantitative dissimilarity metric d. Indeed, by submodularity, the dissimilarity metric d associated with v is monotone with respect to T_Λ in the sense that

$$(x, y, z) \in T_\Lambda \Rightarrow d(x, y) \leqslant d(x, z).$$

Hence, greater qualitative dissimilarity implies greater quantitative dissimilarity (but, of course, not necessarily vice versa). In general, say that d is *monotone with respect to* T if

$$(x, y, z) \in T \Rightarrow d(x, y) \leqslant d(x, z).$$

Moreover, say that d is *adapted* to a model $\mathscr{A}$ if d is monotone with respect to $T_{\mathscr{A}}$.

3. Hierarchies

A model $\mathscr{H} \subseteq 2^X$ is called a (*taxonomic*) *hierarchy* if the elements of $\mathscr{H}$ are nested in the sense that, for all $A, B \in \mathscr{H}$,

$$A \cap B \neq \emptyset \Rightarrow [A \subseteq B \text{ or } B \subseteq A].$$

Accordingly, we will refer to a diversity function v, as well as to the associated attribute weighting function λ, as *hierarchical* if the support Λ of relevant attributes forms a hierarchy. Theorem 3.1 in TD shows that a diversity function is hierarchical if and only if, for all x and S,

$$v(S \cup \{x\}) - v(S) = \min_{y \in S} [v(\{x, y\}) - v(\{y\})], \tag{3.1}$$

or, equivalently,

$$d(x, S) = \min_{y \in S} d(x, y).$$

By (3.1), the entire diversity function can be recursively determined from its values on the set $\mathscr{B} := \{S : \#S \leq 2\}$ of all *binary sets* of X containing at most two elements. Note that the restriction of v to $\mathscr{B}$ contains the same information as the associated dissimilarity metric plus the value of singletons.

For a hierarchical family $\mathscr{H}$, the associated binary relation $T^x_{\mathscr{H}}$ according to (2.7) and (2.8) is complete (hence, a weak order) for all $x \in X$. This follows at once from the observation that the family $\mathscr{H}_x := \{A \in \mathscr{H} : A \ni x\}$ of all attributes in $\mathscr{H}$ that contain x is totally ordered by set-inclusion, i.e. a chain. Completeness of all $T^x_{\mathscr{H}}$ in fact characterizes the hierarchical model, as shown in DGS.

In a hierarchy, attribute weights are determined by the dissimilarity metric in a simple way, as shown by the following result.

Proposition 3.1. (Conjugate Moebius inverse on a hierarchy) *Let v be a diversity function with attribute weighting function λ. If v is hierarchical, then for all $A \in \Lambda$ and all $x \in A$,*

$$\lambda_A = \min_{z \in A^c} d(x, z) - \max_{z \in A} d(x, z). \tag{3.2}$$

Conversely, suppose that for all $A \in \Lambda$ and all $x \in A$, $\lambda_A = d(x, A^c) - \max_{z \in A} d(x, z)$. Then λ is hierarchical.

Observe that positivity of the term (3.2) says that any x in an attribute A is less similar to any element outside A than to any element in A. In the hierarchical case, attributes are thus 'similarity clusters' of objects.

Despite its specific structure, the hierarchical model is quite flexible, as illustrated by the following two degenerate cases: the class of all additive diversity functions of the form $v(S) = \Sigma_{x \in S} v(\{x\})$, and the class of all functions of the form $v(S) = \max_{x \in S} v(\{x\})$. As is easily verified, the first class of 'additive counting' is characterized by the property that all relevant attributes are singletons, i.e. $\Lambda \subseteq \{\{x\} : x \in X\}$; the second class, in which only the object richest in attributes counts, is characterized by attributes being totally ordered, i.e. by the property that Λ forms a chain.

4. Lines

As another simple and fundamental model of diversity, we consider the 'line model' introduced in TD. Assume that the universe X is ordered by some given linear (i.e. complete, transitive and antisymmetric) ordering $\geq$. For instance, objects may be ordered according to size, mass, age, etc. For any $x, z \in X$ with $x \leq z$, denote by $[x, z] := \{y : x \leq y \leq z\}$ the *interval* spanned by x and z; furthermore, denote by $\mathscr{L}$ the family of all intervals with respect to the ordering $\geq$. We refer to $\mathscr{L}$ as the *line model* associated with $\geq$. A diversity function v is *line compatible* if $\Lambda \subseteq \mathscr{L}$. Note that any hierarchical diversity function is line compatible, since to any hierarchical attribute family one can associate a (non-unique) linear ordering such that all attributes are intervals. Line compatibility is more general in that it allows attributes to (non-trivially) overlap.

The induced qualitative similarity relation according to (2.7) is the canonical betweenness relation on a line. Indeed, for all x, y, z,

$$(x, y, z) \in T_{\mathscr{L}} \Leftrightarrow [x \leqslant y \leqslant z \text{ or } z \leqslant y \leqslant x],$$

i.e. y is more similar than z to x if and only if y is between x and z.

By Theorem 3.2 in TD, diversity in a line model is characterized by the following simple formula, referred to as the 'line equation.' A diversity function v is compatible with a line model if and only if, for all $y_1 < y_2 < \cdots < y_m$,

$$v(\{y_1, \ldots, y_m\}) = v(\{y_1\}) + \sum_{i=2}^{m} d(y_i, y_{i-1}). \tag{4.1}$$

By the line Eq. (4.1), a diversity function on a line is again determined by its values on the family $\mathscr{B}$ of all binary sets.

As in the hierarchical case, one obtains a simple formula for the conjugate Moebius inverse on a line. For notational convenience, we henceforth identify $(X, \geqslant)$ with the set $\{1, \ldots, n\}$ of natural numbers endowed with the standard ordering, where $n = \#X$.

Proposition 4.1. (Conjugate Moebius inverse on a line) *Let $v{:}2^X \to \mathbf{R}$ be a set function with conjugate Moebius inverse λ. If $\Lambda \subseteq \mathscr{L}$, then for all $x, z \in X$ with $1 < x \leqslant z < n$,*

$$\lambda_{[x,z]} = v(\{z+1, x\}) - v(\{z, x\}) - [v(\{z+1, x-1\}) - v(\{z, x-1\})]. \tag{4.2}$$

Furthermore, for $1 \leqslant z < n$ and $1 < x \leqslant n$,

$$\lambda_{[1,z]} = v(\{z+1, 1\}) - v(\{z, 1\}) - [v(\{z+1\}) - v(\{z\})],$$

$$\lambda_{[x,n]} = v(\{x\}) - v(\{n, x\}) - [v(\{x-1\}) - v(\{n, x-1\})],$$

$$\lambda_{[1,n]} = v(\{1\}) - v(\{n, 1\}) + v(\{n\}).$$

5. The extension problem for a given model

As observed above, hierarchical and line compatible diversity functions are determined by their values on the family $\mathscr{B}$ of all binary sets with at most two elements, or, equivalently, by their values on singletons together with the induced dissimilarity metric d. A natural question is: What restrictions do the hierarchical model and the line model impose on the induced dissimilarity metric, respectively? In the present section, we will study these restrictions for a *given* hierarchy and a *given* line model. Formally, we will study the conditions under which a function $v^{\mathscr{B}}{:}\mathscr{B} \to \mathbf{R}$ is the restriction of some diversity function v with $\Lambda \subseteq \mathscr{H}$ for some given hierarchy $\mathscr{H}$, respectively $\Lambda \subseteq \mathscr{L}$ for some given line model $\mathscr{L}$. In Section 6 below, the corresponding results will be used to achieve a more ambitious goal, namely to characterize compatibility of a dissimilarity metric with *some* line structure and *some* hierarchy, respectively.

We start with the line model, and then derive the result in the hierarchical case as a corollary. As before, let $X=\{1,\ldots,n\}$ and let $\mathscr{L}$ be the line model corresponding to the standard ordering of $\{1,\ldots,n\}$. Using (4.1), it is easily verified that *any* function $v^{\mathscr{B}}:\mathscr{B}\to\mathbf{R}$ can be uniquely extended to a set function $v:2^X\to\mathbf{R}$ with $\Lambda\subseteq\mathscr{L}$. It thus remains to find the conditions under which the extension v is a diversity function. Since diversity functions are characterized by non-negativity of the corresponding conjugate Moebius inverse, these conditions follow at once from Proposition 4.1. To make these intuitively more transparent, we reformulate them in terms of dissimilarity.

Let $v^{\mathscr{B}}:\mathscr{B}\to\mathbf{R}$ be a function on the family of all binary sets with associated dissimilarity metric d. Say that d is *bounded* if, for all x, y,

$$d(x,y)\leq v^{\mathscr{B}}(\{x\}).$$

Furthermore, say that d is *submodular with respect to T* if the following condition holds. For all x_1, x_2, x_3, x_4 such that $(x_i, x_j, x_l)\in T$ whenever $1\leq i<j<l\leq 4$,

$$d(x_4,x_1)-d(x_3,x_1)\leq d(x_4,x_2)-d(x_3,x_2).$$

Finally, say that d is *line-submodular* if it is submodular with respect to $T_{\mathscr{L}}$. Line-submodularity says that increasing a gap between two elements results in a smaller increase in dissimilarity the larger the gap already is. Observe that line-submodularity entails the triangle inequality by taking $x_2=x_3$.

Theorem 1. (Line extension theorem) *Any $v^{\mathscr{B}}:\mathscr{B}\to\mathbf{R}$ can be uniquely extended to a set function $v:2^X\to\mathbf{R}$ with $\Lambda\subseteq\mathscr{L}$. The extension v is a diversity function if and only if the dissimilarity metric d associated with $v^{\mathscr{B}}$ is bounded, and monotone and submodular with respect to $T_{\mathscr{L}}$.*

Theorem 1 follows as an immediate corollary of Proposition 4.1 above. Indeed, rewriting (4.2) in terms of dissimilarity one obtains, for $1<x\leq z<n$,

$$\lambda_{[x,z]}=d(z+1,x)-d(z,x)-[d(z+1,x-1)-d(z,x-1)].$$

Non-negativity of λ at interior intervals thus follows from line-submodularity. Similarly, for $z<n$ and $1<x$, one has $\lambda_{[1,z]}=d(1,z+1)-d(1,z)$, $\lambda_{[x,n]}=d(n,x-1)-d(n,x)$, and $\lambda_{[1,n]}=v(\{1\})-d(1,n)$. Hence, non-negativity at all other intervals follows from monotonicity with respect to $T_{\mathscr{L}}$ and boundedness, respectively.

By Theorem 1, the metric implications of the line model are somewhat stronger than the minimal ‘canonical’ implications given by adaptedness (i.e. monotonicity with respect to $T_{\mathscr{L}}$) and the triangle inequality. To further illustrate the crucial condition of line-submodularity, consider the benchmark case in which the diversity of a set only depends on its internal shape and not on its location within the line. Specifically, say that a diversity function on a line is *translation invariant* if, for all $S\subseteq X$ and all integers t,

$$v(S)=v(S+t),$$

whenever $S + t: = \{x + t: x \in S\} \subseteq X = \{1, \ldots, n\}$. As is easily verified, v is translation invariant if and only if the associated dissimilarity metric d is translation invariant in the sense that $d(x, z)$ only depends on the Euclidean distance $|z - x|$. In this case, line-submodularity is equivalent to *concavity* of d, i.e. to the condition that, for all x, z with $z \geqslant x$, the difference $d(x, z + 1) - d(x, z)$ is decreasing in z. The following result is thus an immediate corollary of Theorem 1.

Corollary 5.1. *A function $v^{\mathscr{B}}: \mathscr{B} \to \mathbf{R}$ can be extended to a translation invariant diversity function $v: 2^X \to \mathbf{R}$ with $\Lambda \subseteq \mathscr{L}$ if and only if d is translation invariant, bounded, monotone with respect to $T_{\mathscr{L}}$ and concave.*

Concavity of the dissimilarity metric implies the following ordinal 'preference for even spacing.' For all x, y_1, y_2, z with $x < y_i < z$, and $S_1: = \{x, y_1, z\}$, $S_2 := \{x, y_2, z\}$,

$$v(S_1) \geqslant v(S_2) \Leftrightarrow y_1 \text{ is closer than } y_2 \text{ to the midpoint } (z - x)/2, \tag{5.1}$$

where $v(S_1)$ and $v(S_2)$ are defined from d via the line Eq. (4.1). For $X = \mathbf{N}$, one can show that, conversely, condition (5.1) implies concavity of d (cf. Nehring and Puppe, 1999a, Section 4.5).

In hierarchies line-submodularity is automatically satisfied. By consequence, the metric implications of the hierarchical model are canonical, as shown by the following result.

Theorem 2. (Hierarchy extension theorem) *Let $\mathscr{H}$ be a hierarchy. A function $v^{\mathscr{B}}: \mathscr{B} \to \mathbf{R}$ can be uniquely extended to a diversity function $v: 2^X \to \mathbf{R}$ with $\Lambda \subseteq \mathscr{H} \cup \{X\}$ if and only if the associated dissimilarity metric d is bounded and adapted to $\mathscr{H}$.*

The proof of Theorem 2 in Appendix A is based on the fact that adaptedness to a hierarchy implies line-submodularity with respect to any ordering such that all elements of the hierarchy are intervals. Theorem 2 thus also follows as a corollary from the above line extension theorem.

6. The general extension problem

In the previous section, we have characterized the restrictions on a dissimilarity metric for it to be consistent with a *given* line model associated with a particular linear ordering of the object space. A more fundamental question addresses the restrictions of the line model as such: Under what conditions on a dissimilarity metric d is there *some* linear ordering $\geqslant$ of the object space such that d is the dissimilarity associated with a diversity function v that is compatible with the betweenness induced by $\geqslant$? This is a non-trivial problem, and we provide an answer for the two polar and most interesting cases of hierarchies and 'exact' lines, i.e. the case in which $T_\Lambda = T_{\mathscr{L}}$ for an appropriate line model $\mathscr{L}$.

Let $v^{\mathscr{B}}: \mathscr{B} \to \mathbf{R}$ be given, and denote by d the corresponding dissimilarity metric. The

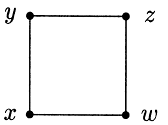

Fig. 1. The 2-hypercube.

following metric betweenness relation introduced in DGS will play a crucial role in the present analysis. For all x, y, z,

$$(x, y, z) \in T_d :\Leftrightarrow [d(x, y) \leqslant d(x, z) \text{ and } d(z, y) \leqslant d(z, x)]. \tag{6.1}$$

Furthermore, say that a ternary relation T is *line-asymmetric* if, for all x, y, z, w such that $y \notin \{x, z, w\}$ and $(x, y, z) \in T$,

$$(w, y, z) \in T \Leftrightarrow (x, y, w) \notin T.$$

Theorem 3. *A function $v^{\mathscr{B}}: \mathscr{B} \to \mathbf{R}$ can be uniquely extended to a diversity function v on 2^X such that $T_\Lambda = T_{\mathscr{L}}$ for $\mathscr{L}$ associated with some linear order $\geqslant$ on X if and only if T_d is line-asymmetric, and d is bounded and submodular with respect to T_d.*

The condition that drives the result is line-asymmetry. Line-asymmetry as a condition on T_Λ is already quite restrictive. For instance, it is neither satisfied in hierarchies nor in multi-dimensional models. Typical examples are the following two graphs. Fig. 1 depicts the graph corresponding to the 2-dimensional hypercube model (see Section 3 of TD), while Fig. 2 depicts the '3-star' tree (see TD, Section 5).

In both cases, a natural specification of Λ leads to the following qualitative similarity relation: $(x, y, z) \in T_\Lambda$ if and only if y lies on a shortest path that connects x and z. Thus, in Fig. 1 one has $(x, y, z) \in T_\Lambda$, but neither $(w, y, z) \in T_\Lambda$ nor $(x, y, w) \in T_\Lambda$ in violation of line-asymmetry.

In Fig. 2, one has again $(x, y, z) \in T_\Lambda$; now the violation of line-asymmetry occurs since $(w, y, z) \in T_\Lambda$ as well as $(x, y, w) \in T_\Lambda$.

Line-asymmetry is even more powerful when applied to the less regular T_d. Consider, for instance, the case $X = \{x, y, z, w\}$ with $\Lambda = 2^X \setminus \{\emptyset\}$. As is easily verified, the corresponding betweenness T_Λ is 'vacuous' in the sense that no other element is between two given elements. By consequence, T_Λ vacuously satisfies line-asymmetry. On the other hand, giving each attribute in Λ unity weight, say, results in a diversity function such that the induced T_d violates line-asymmetry.

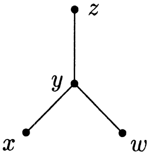

Fig. 2. The 3-star tree.

In order to prove Theorem 3 one wishes to make use of the line extension theorem for a *given* line structure. To make use of this result one needs to find conditions on a ternary relation that are necessary and sufficient for its being the betweenness of a linear ordering. These can be found in the literature. Specifically, by Fishburn (1985), Theorem 6, a betweenness relation T is a line-betweenness ($T = T_{\mathscr{L}}$ for some line $\mathscr{L}$) if and only if T is symmetric (in the sense of T2 above), line-asymmetric, and *triple-connected* in the sense that, for all x, y, z, at least one of the three triples (x, y, z), (y, z, x), (z, x, y) is an element of T. An additional contribution of Theorem 3 consists in obtaining triple-connectedness from the definition of T_d. The insight of the result is thus the appropriateness of T_d as the 'right' notion of betweenness derived from d in the context of the line structure.

For the class of all hierarchical models, we have the following result.

Theorem 4. *Let* $v^{\mathscr{B}}:\mathscr{B} \to \mathbf{R}$ *be given with associated dissimilarity metric* d *and quantitative similarity* σ. *The following statements are equivalent.*

(i) $v^{\mathscr{B}}$ *can be uniquely extended to a hierarchical diversity function on* 2^X.
(ii) d *is non-negative, bounded and, for all* x, T_d^x *is complete.*
(iii) σ *is non-negative, bounded* (*i.e.* $\sigma(x, y) \leqslant v^{\mathscr{B}}(\{x\})$) *and, for all* x, y, z,

$$\text{mid}\{\sigma(x, y), \sigma(y, z), \sigma(x, z)\} = \min\{\sigma(x, y), \sigma(y, z), \sigma(x, z)\}. \tag{6.2}$$

Theorem 4 generalizes a classical result (Johnson, 1967; Benzecri et al., 1973) on the representation of ultrametric distance functions by not assuming symmetry of d. A symmetric distance function d is called *ultrametric* if, for all x, y, z,

$$\text{mid}\{d(x, y), d(y, z), d(x, z)\} = \max\{d(x, y), d(y, z), d(x, z)\}, \tag{6.3}$$

i.e. if the two greatest distances between any three points are equal. In the symmetric case, one has $v(\{x\}) = v(\{y\})$ for all x, y, hence (6.2) and (6.3) are equivalent.

Acknowledgements

It is our great pleasure to dedicate this paper to our cherished colleague Prasanta Pattanaik; his unique combination of valuable attributes within the scientific community greatly enhances its total diversity value.

Appendix A. Proofs

Proof of Proposition 3.1. Suppose that Λ is a hierarchy. Let $x \in A \in \Lambda$, and define $z^* := \arg\max_{z \in A} d(x, z)$. Since $\Lambda_x = \{B \in \Lambda : x \in B\}$ is a chain, one has $B \subset A \Leftrightarrow z^* \notin B$ for all $B \in \Lambda_x$, where '$\subset$' denotes the *proper* subsethood relation. Hence, using (3.1),

$$\begin{aligned}&\min_{z\in A^c} d(x,z)-\max_{z\in A} d(x,z)\\&\quad = v(\{x\}\cup A^c)-v(A^c)-d(x,z^*)\\&\quad = \lambda(\{B{:}x\in B\subseteq A\})-\lambda(\{B{:}x\in B, z^*\not\in B\})\\&\quad = \lambda_A+\lambda(\{B{:}x\in B\subset A\})-\lambda(\{B{:}x\in B, z^*\not\in B\})\\&\quad = \lambda_A.\end{aligned}$$

Conversely, suppose that Λ is not a hierarchy, i.e. suppose there exist $A, C\in\Lambda$ such that $A\cap C$, $A\backslash C$, and $C\backslash A$ are all non-empty. Let $x\in A\cap C$. Without loss of generality we may assume that A is a minimal element of Λ satisfying $x\in A$ and $A\backslash C\neq\emptyset$, i.e. for no proper subset A' of A, $x\in A'\in\Lambda$ and $A'\backslash C\neq\emptyset$. Let $y\in A\backslash C$. By construction one has $\{B\in\Lambda{:}x\in B, B\subset A\}\subset\{B\in\Lambda{:}x\in B, y\not\in B\}$ since C belongs to the latter but not to the former set. Since by assumption, $\lambda_C>0$, this implies $\lambda(\{B{:}x\in B\subset A\})-\lambda(\{B{:}x\in B, y\not\in B\})<0$. Therefore,

$$\begin{aligned}&d(x,A^c)-\max_{z\in A} d(x,z)\\&\quad = v(\{x\}\cup A^c)-v(A^c)-\max_{z\in A} d(x,z)\\&\quad \leqslant v(\{x\}\cup A^c)-v(A^c)-d(x,y)\\&\quad = \lambda(\{B{:}x\in B\subseteq A\})-\lambda(\{B{:}x\in B, y\not\in B\})\\&\quad = \lambda_A+\lambda(\{B{:}x\in B\subset A\})-\lambda(\{B{:}x\in B, y\not\in B\})\\&\quad < \lambda_A.\end{aligned}$$

Proof of Proposition 4.1. Let v be compatible with the line model. We prove the formula for the case $1<x\leqslant z<n$. The other cases follow along the same lines. Since $\Lambda\subseteq\mathscr{L}$,

$$\begin{aligned}&v(\{z+1,x\})-v(\{z,x\})\\&\quad = \lambda(\{A{:}z+1\in A, z\not\in A\})-\lambda(\{A{:}z\in A, A\subseteq[x,z], x\not\in A\}),\end{aligned}\tag{A.1}$$

and

$$\begin{aligned}&v(\{z+1,x-1\})-v(\{z,x-1\})\\&\quad = \lambda(\{A{:}z+1\in A, z\not\in A\})-\lambda(\{A{:}z\in A, A\subseteq[x,z]\}).\end{aligned}\tag{A.2}$$

Subtracting (A.2) from (A.1) one obtains (4.2).

Proof of Theorem 1. Sufficiency of the stated conditions has already been established in the main text. The proof of necessity is straightforward.

Proof of Theorem 2. Necessity of the stated conditions is straightforward. In order to prove their sufficiency, we want to make use of Theorem 1. Let $\geqslant$ be a linear ordering of X such that all elements of the hierarchy $\mathscr{H}$ are intervals with respect to $\geqslant$. First, note that since all elements of $\mathscr{H}$ are intervals, adaptedness to $\mathscr{H}$ in particular implies adaptedness to the line $(X, \geqslant)$.

Next, we show that adaptedness to $\mathscr{H}$ also implies line-submodularity. To do so, it will be convenient to reformulate both conditions in terms of the similarity function σ

associated to $v^{\mathscr{B}}$. Recall that $\sigma(x, y) = v^{\mathscr{B}}(\{x\}) - d(x, y)$, and observe that, in contrast to d, σ is always a symmetric function. As is easily verified, line-submodularity is equivalent to

$$\sigma(x_4, x_1) + \sigma(x_3, x_2) \geqslant \sigma(x_4, x_2) + \sigma(x_3, x_1),$$

whenever $x_1 \leqslant x_2 \leqslant x_3 \leqslant x_4$. Let d be adapted to $\mathscr{H}$. In terms of similarity, this is easily seen to be equivalent to

$$(x, y, z) \in T_{\mathscr{H}} \Rightarrow \sigma(x, z) \leqslant \min\{\sigma(x, y), \sigma(y, z)\}.$$

Consider $x_1 \leqslant x_2 \leqslant x_3 \leqslant x_4$. Suppose first that $(x_2, x_1, x_4) \in T_{\mathscr{H}}$; by adaptedness, this implies $\sigma(x_4, x_2) \leqslant \sigma(x_4, x_1)$. Since $(x_1, x_2, x_3) \in T_{\mathscr{H}}$, one also has $\sigma(x_3, x_1) \leqslant \sigma(x_3, x_2)$, again by adaptedness. Together, this implies line-submodularity. If, on the other hand, $(x_2, x_1, x_4) \not\in T_{\mathscr{H}}$, one must have $(x_1, x_4, x_3) \in T_{\mathscr{H}}$ since $\mathscr{H}$ is a hierarchy. By adaptedness, this implies $\sigma(x_3, x_1) \leqslant \sigma(x_4, x_1)$. Since $(x_2, x_3, x_4) \in T_{\mathscr{H}}$, one has $\sigma(x_4, x_2) \leqslant \sigma(x_3, x_2)$ by adaptedness, hence the line-submodularity condition follows also in this case.

By Theorem 1 there exists a unique extension $v: 2^X \to \mathbf{R}$ satisfying line compatibility with respect to $\geqslant$. It remains to show that $\lambda_A = 0$ for all intervals $A \not\in \mathscr{H} \cup \{X\}$. Assume, by way of contradiction, that $\lambda_{[x,z]} > 0$ for some interval $[x, z]$ not contained in $\mathscr{H} \cup \{X\}$. Consider the set $A_{\{x,z\}} := \cap \{A \in \mathscr{H} \cup \{X\}: A \supseteq [x, z]\}$. Since $\mathscr{H}$ is a hierarchy, one has $A_{\{x,z\}} \in \mathscr{H} \cup \{X\}$. Since by assumption $A_{\{x,z\}} \neq [x, z]$, there must exist $y \not\in [x, z]$ with $y \in A_{\{x,z\}}$, i.e. with $(x, y, z) \in T_{\mathscr{H}}$. Without loss of generality, assume $y > z$. This would imply $d(x, y) > d(x, z)$, which is not possible by adaptedness to $\mathscr{H}$.

Proof of Theorem 3. We only prove sufficiency of the stated conditions. By Fishburn (1985, Theorem 6), a ternary relation T can be represented as $T = T_{\mathscr{L}}$ for $\mathscr{L}$ associated to some unique (up to reversal) linear order $\geqslant$ on X if and only if T is symmetric, triple-connected and line-asymmetric. Clearly, T_d is symmetric, and line-asymmetry holds by assumption; triple-connectedness follows at once from the observation that, for all x, y, z,

$$(x, y, z) \in T_d \Leftrightarrow \sigma(x, z) \leqslant \min\{\sigma(x, y), \sigma(y, z)\}. \tag{A.3}$$

Let $\geqslant$ be the linear order on X such that $T_d = T_{\mathscr{L}}$. The dissimilarity metric d satisfies all requirements in order to apply Theorem 1. Hence, there exists a unique extension of $v^{\mathscr{B}}$ to a diversity function v on X that is compatible with $\geqslant$. By Fact A.3 below, $T_{\Lambda} = T_d = T_{\mathscr{L}}$.

For the proof of Theorem 4, we need the following results from DGS.

Fact A.1. If T satisfies T1–T3, then $T = T_{\mathscr{A}}$, where

$$\mathscr{A} := \{A \subseteq X: \text{for all } (x, y, z) \in T: \{x, z\} \subseteq A \Rightarrow y \in A\}.$$

Fact A.2. $\mathscr{H}$ is a hierarchy if and only if, for all x, $T^x_{\mathscr{H}}$ is complete.

Fact A.3. If v is line compatible with respect to some ordering $\geqslant$, then $T_\Lambda = T_d$.

Fact A.4. If $\mathscr{H}$ is a hierarchy and $T_{\mathscr{H}} = T_{\mathscr{H}'}$ for some $\mathscr{H}'$, then $\mathscr{H} \cup \{X\} = \mathscr{H}' \cup \{X\}$.

Proof of Theorem 4. The equivalence of (ii) and (iii) is easily established. Necessity of (ii) and (iii) for the existence of a hierarchical extension is also obvious. We now show the sufficiency of (ii) using Facts A.1–4 above. Clearly, T_d satisfies conditions T1 (reflexivity) and T2 (symmetry). Right below we show that under condition (ii), T_d also satisfies T3 (transitivity). Consider

$$\mathscr{H} := \{A \subseteq X: \text{for all } (x, y, z) \in T_d: \{x, z\} \subseteq A \Rightarrow y \in A\},$$

and note that $X \in \mathscr{H}$ by definition. By Fact A.1, $T_d = T_{\mathscr{H}}$. Hence, by the assumed completeness of T_d^x and Fact A.2, $\mathscr{H}$ is a hierarchy. By definition, d is monotone with respect to $T_d = T_{\mathscr{H}}$, hence adapted to $\mathscr{H}$. By Theorem 2, there exists a unique extension $v: 2^X \to \mathbf{R}$ of $v^{\mathscr{B}}$ such that $\Lambda \subseteq \mathscr{H}$. This demonstrates existence.

To verify uniqueness (of $\mathscr{H}$), let v' be an extension of $v^{\mathscr{B}}$, and let $\mathscr{H}'$ denote the (hierarchical) support of its attribute weighting function. By Fact A.3, $T_{\mathscr{H}'} = T_d = T_{\mathscr{H}}$, hence by Fact A.4, $\mathscr{H} = \mathscr{H}'$ up to the inclusion of the universal attribute X. However, the weight of X is uniquely determined by $\lambda_X = \min_{x,y \in X} \sigma(x, y)$, which is non-negative by the boundedness of d.

It remains to be shown that T_d satisfies the transitivity condition T3. This is done in two steps. First, we show that completeness of T_d^x implies (standard) transitivity of T_d^x, i.e. all relations T_d^x are weak orders. We then show that, for any symmetric ternary relation T such that all T^x are weak orders, T satisfies the transitivity condition T3. Thus, suppose that $(x, y, z) \in T_d$ and $(x, z, w) \in T_d$. We have to show that $(x, y, w) \in T_d$. Using (A.3), one has $\sigma(x, w) \leqslant \sigma(x, z)$ (from $(x, z, w) \in T_d$) and $\sigma(x, z) \leqslant \sigma(x, y)$ (from $(x, y, z) \in T_d$), hence $\sigma(x, w) \leqslant \sigma(x, y)$. Now assume that $(x, y, w) \not\in T_d$; by completeness of T^x, this implies $(x, w, y) \in T_d$, in particular $\sigma(x, y) \leqslant \sigma(w, y) = \sigma(y, w)$. Thus, $\sigma(x, w) \leqslant \sigma(x, y)$ and $\sigma(x, w) \leqslant \sigma(y, w)$, which by (A.3) implies $(x, y, w) \in T_d$, a contradiction.

We now show that symmetry, completeness and transitivity of all T^x together imply transitivity of T. Take x, x', y, z, z' such that $(x, x', z) \in T$, $(x, z', z) \in T$ and $(x', y, z') \in T$. By completeness, $(x, x', z') \in T$ or $(x, z', x') \in T$; without loss of generality, assume $(x, x', z') \in T$. By symmetry, $(z', x', x) \in T$ as well as $(z', y, x') \in T$. By transitivity of $T^{z'}$, $(z', y, x) \in T$, hence by symmetry, $(x, y, z') \in T$. Finally, by transitivity of T^x, $(x, y, z) \in T$.

References

Benzecri, J.P. et al., 1973. La Taxinomie. L'Analyse des Données, Vol. 1. Dunod, Paris.
Bossert, W., Pattanaik, P., Xu, Y., 2001. The measurement of diversity. Mimeo.
Fishburn, P., 1985. Interval Orders and Interval Graphs. Wiley, New York.
Johnson, S.C., 1967. Hierarchical clustering schemes. Psychometrika 32, 241–254.
Nehring, K., 1999. Diversity and the geometry of similarity. Mimeo.

Nehring, K., Puppe, C., 1999a. A theory of diversity. SFB 303 Discussion Paper A-605, University of Bonn.
Nehring, K., Puppe, C., 1999b. Modelling economies of scope. SFB 303 Discussion Paper A-607, University of Bonn.
Nehring, K., Puppe, C., 2002a. A theory of diversity. Econometrica 70, 1155–1198.
Nehring, K., Puppe, C., 2002b. Modelling phylogenetic diversity. Resource & Energy Economics, forthcoming.
Nehring, K., Puppe, C., 2002c. Diversity and the metric of opportunity. In preparation.
Pattanaik, P., Xu, Y., 2000. On diversity and freedom of choice. Mathematical Social Sciences 40, 123–130.
Weitzman, M., 1992. On diversity. Quarterly Journal of Economics 107, 363–405.
Weitzman, M., 1998. The Noah's ark problem. Econometrica 66, 1279–1298.