Size invariant measures of association: characterization and difficulties

Margherita Negri\textsuperscript{a,*}, Yves Sprumont\textsuperscript{b}

\textsuperscript{a}School of Economics and Finance, University of St. Andrews, United Kingdom
\textsuperscript{b}Département de Sciences Économiques and CIREQ, Université de Montréal, Canada

Abstract

A measure of association on cross-classification tables is row-size invariant if it is unaffected by the multiplication of all entries in a row by the same positive number. It is class-size invariant if it is unaffected by the multiplication of all entries in a class (i.e., a row or a column). We prove that every class-size invariant measure of association assigns to each cross-classification table a number which depends only on the cross-product ratios of its $2 \times 2$ subtables. We submit that the degree of association should increase when mass is shifted from cells containing a proportion of observations lower than what is expected under statistical independence to cells containing a proportion higher than expected –provided that total mass in each class remains unchanged. We prove that no continuous row-size invariant measure of association satisfies this monotonicity axiom if there are at least four rows.

Keywords: association, contingency tables, margin-free measures, size invariance, monotonicity, transfer principle.

\textsuperscript{*}We thank two referees and C. Chambers for helpful comments. We also thank L. Ehlers, S. Horan, F. Maniquet, W. McCausland, H. Moulin and B. Strulovici for discussions about the topic of the paper and remarks on an earlier draft. Sprumont acknowledges support from the Fonds de Recherche sur la Société et la Culture of Québec.

\textsuperscript{*}Corresponding author. Tel. +44 (0) 1334 462419

\textit{Email addresses: mn48@st-andrews.ac.uk} (Margherita Negri),
\textit{yves.sprumont@umontreal.ca} (Yves Sprumont)
1. Introduction

This paper revisits the old problem of measuring the degree of association between two characteristics whose values cannot be quantified or even ordered in a meaningful way. Gender, professional occupation, ethnic origin, political affiliation are all common examples. Each characteristic may fall in a number of predetermined classes and the distribution of the two characteristics in the population under consideration is summarized by a cross-classification table: this is an \( m \times n \) matrix whose \((i, j)\)th entry records the mass of individuals whose first characteristic belongs to class \(i\) and whose second characteristic belongs to class \(j\). A measure of association is a function which assigns a number to each such matrix. A survey of the literature on the problem of constructing such a measure is offered in Goodman and Kruskal (1954). The measurement of racial segregation is an example of application that has received distinguished attention: see the survey by James and Taeuber (1985).

We are interested here in association measures that are invariant under multiplication of all entries in any given class—row or column—by the same positive number. Such measures are sometimes called “margin-free”. The motivation for this class-size invariance axiom is that the total mass in a class is often determined by factors that have nothing to do with the inherent association between the characteristics. Yule (1912), who originally suggested the axiom, gives the following example. Suppose we are concerned with the effect of a medical treatment on persons suffering from a potentially fatal disease: the first characteristic is whether or not the person received the treatment, the second is whether or not the person died from the disease. Data from two hospitals are recorded in the following tables.

<table>
<thead>
<tr>
<th></th>
<th>Lived</th>
<th>Died</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treated</td>
<td>84</td>
<td>4</td>
<td>88</td>
</tr>
<tr>
<td>Not treated</td>
<td>3</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>87</td>
<td>13</td>
<td>100</td>
</tr>
</tbody>
</table>

\(^1\)The problem of measuring correlation between ordinal characteristics (whose values cannot be measured but can be ordered) is different and has generated a separate literature.
These two tables differ only in the proportion of persons who received treatment and the proportion who did not: the second table may be obtained from the first by multiplying the first row by $\frac{1}{2}$ and multiplying the second by $\frac{14}{3}$. It follows that the conditional probabilities of life given treatment (non-treatment) are the same in both hospitals, namely .955 (.250). The degree of association between treatment and life should therefore be considered equal in both hospitals; the fact that the proportion of persons receiving treatment is higher in hospital 1 is irrelevant.

Class-size invariance is well understood in the case of dichotomous characteristics, i.e., when $m = n = 2$. In that case, Edwards (1963) and Goodman (1965) showed that every class-size invariant measure must be a function of the cross-product ratio of the cross-classification table, that is, the ratio between the product of its diagonal entries and the product of its off-diagonal entries.

For tables of arbitrary dimensions, Yule and Kendall (1948) observed that any function of the cross-product ratios of the $2 \times 2$ subtables defines a class-size invariant measure. Goodman (1969) showed how to use the collection of $2 \times 2$ cross-product ratios to study patterns of association. The class-size invariant approach was extended to continuous densities by Plackett (1965), Mosteller (1968), Holland and Wang (1987), and others.

Our first contribution in this paper is to prove the converse of Yule and Kendall’s observation: every class-size invariant measure of association assigns to each $m \times n$ cross-classification table a number which depends only on the cross-product ratios of its $2 \times 2$ subtables.

Class-size invariance is a demanding property. In many contexts, a one-sided version of the axiom is more compelling. In fact, in Yule’s example, the argument that the hospital’s choice of the proportion of treated patients has nothing to do with the association between treatment and life justifies row-size invariance: multiplying all entries in a row by the same number should leave the degree of association unchanged. The dual axiom of column-size invariance, on the other hand, makes little sense because “multiplying all entries in a column by the same number” is not an operation within the hos-

<table>
<thead>
<tr>
<th></th>
<th>Lived</th>
<th>Died</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treated</td>
<td>42</td>
<td>2</td>
<td>44</td>
</tr>
<tr>
<td>Not treated</td>
<td>14</td>
<td>42</td>
<td>56</td>
</tr>
<tr>
<td>Total</td>
<td>56</td>
<td>44</td>
<td>100</td>
</tr>
</tbody>
</table>
pital’s control. More generally, row-size invariance is appealing when a one-
directional causal relationship is suspected between the row characteristic
and the column characteristic.\footnote{When a one-directional causality relation exists, there is no loss of generality in assuming that it goes from the row characteristic to the column characteristic. If this is not the case, we need only transpose the table before measuring association.} The axiom has been discussed in a number of applications, notably in the literature on segregation measurement: see for instance James and Taeuber (1985) and Reardon and Firebaugh (2002). Class-size invariance may be a better axiom when there is no reason to suspect a one-directional causality between the characteristics.

Our second contribution consists in proving a disturbing incompatibility
between row-size invariance (hence, a fortiori, class-size invariance) and a
monotonicity condition which, we believe, captures the essence of what an
association measure should do. It is generally accepted that (the absolute
value of) a meaningful measure should reach its minimum when the charac-
teristics are independent, that is to say, when the proportion of observations
in any given cell \((i, j)\) of the cross-classification table is equal to the expected
proportion, namely, the product of the proportion of observations in row \(i\)
by the proportion of observations in column \(j\). We submit that an increase
in the mass of observations in cells where this mass already exceeds the ex-
pected mass, coupled with a decrease in the mass of observations in cells
where it falls short of the expected mass should, if total mass in each class
is kept unchanged, increase the degree of association.

Most popular measures of association, such as the chi-square index or
Theil and Finizza (1971)’s entropy-based index (see Section 4 for definitions),
satisfy this monotonicity axiom. On the other hand, these measures are
known to violate row-size invariance. In the example above, for instance, the
chi-square index is approximately 46.3 for hospital 1 and 49.6 for hospital
2, while the Theil and Finizza index is approximately .16 for hospital 1 and
.29 for hospital 2. Such violations of row-size invariance persist beyond the
dichotomous case. It turns out that they are unavoidable when there are at
least four rows: we prove that no continuous measure can be both monotonic
and row-size invariant if \(m \geq 4\).
2. Framework

Let \( M = \{1, \ldots, m\} \), \( m \geq 2 \), be the set of relevant classes for the first characteristic and let \( N = \{1, \ldots, n\} \), \( n \geq 2 \), be the set of relevant classes for the second characteristic. A cross-classification table is an \( m \times n \) positive real matrix \( A \). The number \( a_{ij} \) in the intersection of row \( i \) and column \( j \) records the mass of agents whose first characteristic belongs to class \( i \) and whose second characteristic belongs to class \( j \). We denote the \( i \)th row of \( A \) by \( A_i \) and its \( j \)th column by \( A^j \). We write \( a_i = \sum_{j \in N} a_{ij} \), \( a^j = \sum_{i \in M} a_{ij} \), and \( a = \sum_{i \in M} \sum_{j \in N} a_{ij} \). We let \( \mathcal{A}(m, n) \) denote the set of positive real \( m \times n \) matrices and define an (association) measure to be a function \( F : \mathcal{A}(m, n) \to \mathbb{R}_+ \).

A number of remarks are in order.

1) The labels \( 1, \ldots, m \) attached to rows and the labels \( 1, \ldots, n \) attached to columns are used for convenience only. They have no intrinsic meaning: a lower label should not be interpreted as reflecting a “lower value” of the characteristic. No relevant order structure is assumed on the sets \( M, N \): characteristics are only “categorical”, not “ordinal”.

2) A fundamental role of an association measure is to allow ordinal comparisons across cross-classification tables. Any function \( F : \mathcal{A}(m, n) \to \mathbb{R}_+ \) generates an association ordering \( \succeq \) on \( \mathcal{A}(m, n) \) via the relation \( A \succeq B \iff F(A) \geq F(B) \): the two characteristics are “more closely associated” in table \( A \) than in table \( B \) if and only if the association measure reaches a higher value at \( A \) than at \( B \). We are primarily interested here in this ordering, which is invariant under monotonic transformations of \( F \). Cardinal measures of association are beyond the scope of this paper.

3) We assume that the number of relevant classes for each characteristic is fixed. We do so because the size invariance and monotonicity axioms do not require a variable-dimensions framework. It is straightforward to reformulate our results in such an extended context (taking \( \bigcup_{(m, n) \geq (2, 2)} \mathcal{A}(m, n) \) as the domain of an association measure) but this generalization brings no new insight into the problem we are interested in.

4) We assume that the entries in all cross-classification tables under consideration are real numbers rather than integers. This is important for the incompatibility result of Section 4, which involves a continuity argument. On the other hand, we could (at the cost of a somewhat cumbersome reformulation of the size invariance axioms) express our results in a framework where
the total mass of observations, $a$, is fixed and equal to one. Note also that we restrict ourselves to (strictly) positive matrices. Allowing zero entries creates a range of delicate issues that we want to avoid.

3. The class-size invariant measures of association

If $A \in \mathcal{A}(m,n)$, $i \in M$, $j \in N$ and $\lambda \in \mathbb{R}_+$, we denote by $(\lambda A\text{, } A_{-i})$ the matrix obtained by multiplying each entry of the $i$th row of $A$ by $\lambda$ and leaving all other entries unchanged. Likewise, $(\lambda A^j, A^{-j})$ is the matrix obtained by multiplying each entry of the $j$th column of $A$ by $\lambda$ and leaving all other entries unchanged.

**Class-Size Invariance.** For all $A \in \mathcal{A}(m,n)$, $i \in M$, $j \in N$ and $\lambda \in \mathbb{R}_+$, $F(\lambda A\text{, } A_{-i}) = F(A) = F(\lambda A^j, A^{-j})$.

An elementary submatrix of $A \in \mathcal{A}(m,n)$ (also called a “tetrad” by Yule and Kendall (1948)) is any $2 \times 2$ submatrix whose entries belong to adjacent rows and columns of $A$. The cross-product ratio of an elementary submatrix (or an elementary cross-product ratio for short) is the ratio between the product of the diagonal entries of this submatrix and the product of its off-diagonal entries. Theorem 1 below asserts that a measure is class-size invariant if and only if it assigns to each matrix a number that can be written as a function of its elementary cross-product ratios.

More formally: for all $A \in \mathcal{A}(m,n)$, all $i \in M \setminus \{m\}$ and all $j \in N \setminus \{n\}$, define

$$r_{ij}^j(A) = \frac{a_{ij}a_{i+1,j+1}}{a_{i+1,j}a_{ij}+1}$$

and let $r(A) \in \mathcal{A}(m-1, n-1)$ denote the matrix $(r_{ij}^j(A))_{i \in M \setminus \{m\}}$.

**Theorem 1.** A measure $F : \mathcal{A}(m,n) \to \mathbb{R}_+$ is class-size invariant if and only if there exists a function $f : \mathcal{A}(m-1, n-1) \to \mathbb{R}_+$ such that $F(A) = f(r(A))$ for all $A \in \mathcal{A}(m,n)$.

**Proof.** “If”. This follows from the fact that for any $\lambda > 0$, $i \in M$ and $j \in N$, we have $r(\lambda A\text{, } A_{-i}) = r(A) = r(\lambda A^j, A^{-j})$.

“Only if”. Let $F$ be a class-size invariant measure and let $A, B \in \mathcal{A}(m,n)$ be such that $r(A) = r(B)$. We must show that $F(A) = F(B)$. By multiplying each row of $A$ by the appropriate constant and then each column by the
appropriate constant, we obtain a matrix \( \hat{A} \) whose first row and first column coincide with those of \( B \), that is,

\[
\hat{a}^i_j = b^i_j \quad \text{for all } (i, j) \text{ such that } \min\{i, j\} = 1.
\] (1)

(To be explicit, the matrix \( \hat{A} \) is given by \( \hat{a}^i_j = \frac{a^1_i b^i_j}{b^1_i a^1_j a^i_j} a^i_j \) for all \( i, j \).) By construction, \( r(\hat{A}) = r(A) \), hence

\[
r(\hat{A}) = r(B).
\] (2)

By a straightforward induction argument, (1) and (2) imply \( \hat{A} = B \) : if \( 1 < k \leq \min\{m, n\} \) and \( \hat{a}^i_j = b^i_j \) for all \( (i, j) \) such that \( \min\{i, j\} \leq k - 1 \), then (2) implies \( \hat{a}^i_j = b^i_j \) for all \( (i, j) \) such that \( \min\{i, j\} = k \). It then follows from Class-Size Invariance that \( F(A) = F(\hat{A}) = F(B) \).

4. The incompatibility between Row-Size Invariance and Monotonicity

In order to construct a good measure of association, it is useful to identify circumstances under which the degree of association between two characteristics should undoubtedly be deemed to increase. The fundamental postulate of the literature is that association is nil in the case of independence, i.e., in a matrix \( A \) such that \( a^i_j a = a_i a^j \) for all \( i \in M \) and \( j \in N \). Taking independence as a benchmark, we say that \( i, j \) are positively associated (in \( A \)) if

\[
a^i_j a > a_i a^j,
\]
or equivalently \( a^i_j > \frac{a_i a^j}{a} \), i.e., the \( (i, j) \)th entry exceeds the value we would expect in case of independence. We say that \( i, j \) are negatively associated if the opposite strict inequality holds.

We submit that the degree of association between two characteristics should increase when we shift mass from entries that are already below their expected values to entries that are already above –provided that the sum of all entries in every class (row or column) remains unchanged. Our monotonicity condition focuses on the simplest such shifts.

**Monotonicity.** Let \( i, i' \in M \) and \( j, j' \in N \) be such that \( i \neq i' \) and \( j \neq j' \). Let \( A, B \in \mathcal{A}(m, n) \). Suppose that both \( i, j \) and \( i', j' \) are positively associated
in $A$ while both $i, j'$ and $i', j$ are negatively associated. If there exists $\delta > 0$ such that $b_i^j = a_i^j + \delta$, $b_i'^{j'} = a_i'^{j'} + \delta$, $b_i'^{j} = a_i'^{j} - \delta$, $b_i^j' = a_i^j' - \delta$, and $b_k^l = a_k^l$ whenever $k \notin \{i, i'\}$ or $l \notin \{j, j'\}$, then $F(B) > F(A)$.

As an illustration, consider the matrix

$$A = \begin{pmatrix} 7 & 27 & 16 \\ 13 & 13 & 24 \\ 30 & 10 & 60 \end{pmatrix}.$$

Note that $i = 2$, $j = 2$ are positively associated in $A$ since $a_{22}^2 = 2600 > a_{2}a_{2}^2 = 2500$; the mass in cell $(2, 2)$ is above its expected value. Similarly, $i = 3$, $j = 3$ are positively associated. On the contrary, $i = 2$, $j = 3$, as well as $i = 3$, $j = 2$, are negatively associated. Shifting one unit of mass from cell $(2, 3)$ to cell $(2, 2)$ and one unit of mass from cell $(3, 2)$ to cell $(3, 3)$ yields

$$B = \begin{pmatrix} 7 & 27 & 16 \\ 13 & 14 & 23 \\ 30 & 9 & 61 \end{pmatrix}.$$

Our axiom therefore requires that $F(B) > F(A)$.

Monotonicity involves changes in no less than four distinct cells. This may seem unduly complicated. One may think that more elementary mass shifts should already be deemed to unambiguously increase the degree of association. Consider for instance the matrix

$$C = \begin{pmatrix} 7 & 27 & 16 \\ 13 & 14 & 23 \\ 30 & 10 & 60 \end{pmatrix},$$

which is obtained from $A$ by merely shifting one unit of mass from cell $(2, 3)$ to cell $(2, 2)$. Since the mass in cell $(2, 3)$ is below expectation in $A$ and since the mass in cell $(2, 2)$ is above expectation, it is tempting to ask that $F(C) > F(A)$. We think that this conclusion is unwarranted. The reason is that the change from $A$ to $C$ increases the total mass in column 2 (from $a_2^2 = 50$ to $c_2^2 = 51$) and decreases the total mass in column 3 (from $a_3^3 = 100$ to $c_3^3 = 99$). This in turn changes the expected masses in all cells belonging to columns 2 and 3. In particular, the expected mass in cell $(1, 2)$ increases from $\frac{a_1^2a_2^2}{a} = 12.5$ in $A$ to $\frac{a_1^2c_2^2}{c} = 12.75$ in $C$. As a consequence, the actual mass in cell $(1, 2)$, which remains unchanged from $A$ to $C$ (at $a_1^2 = c_1^2 = 27$)
becomes closer to its expected value (of 12.75 in C vs 12.5 in A). Similarly, the expected mass in cell (1, 3) decreases from $a_1^3 = 25$ in A to $c_1^3 = 24.75$ in C, making the actual mass in that cell ($a_1^3 = c_1^3 = 16$) closer to expectation (of 24.75 in C vs 25 in A). These induced changes should tend to decrease the degree of association between the two characteristics, counteracting the direct effect of the original changes and leaving the net effect unclear.

In fact, one checks that the most popular measure of association for cross-classification tables of constant total mass (here $a = c$), the chi-square measure, does decrease from A to C:

$$\chi^2(A) = \sum_{(i,j) \in M \times N} \frac{(a_{ij} - a_i a_j)}{a_i a_j} = 34.56$$

$$> \chi^2(C) = \sum_{(i,j) \in M \times N} \frac{(c_{ij} - c_i c_j)}{c_i c_j} \approx 34.36.$$  

In contrast, because the changes considered in the monotonicity axiom leave the total mass in every class (row or column) unchanged, the expected mass in every cell also remains constant and, as a consequence, the gap between actual and expected mass is modified only in the four cells under consideration. Since all these gaps become wider, it is compelling to impose that the degree of association should increase. It is a simple exercise to check that the chi-square measure of association does indeed satisfy Monotonicity. Another famous measure that passes the monotonicity test is the entropy-based index proposed by Theil and Finizza (1971) in the context of measuring segregation,

$$I(A) = \sum_{(i,j) \in M \times N} \frac{a_{ij}}{a_i} \log \left( \frac{a_{ij}}{a_i a_j} \right).$$

To the best of our knowledge, Monotonicity is a new condition. It is a rather weak axiom. A similar but stronger axiom, inspired by the Pigou-Dalton principle, can be found in the segregation measurement literature (see for instance James and Taeuber (1985)).

**Transfer Principle.** Let $i, i' \in M$ and $j, j' \in N$ be such that $i \neq i'$ and $j \neq j'$. Let $A, B \in A(m, n)$. Suppose that $a_i^j / a_j^i > a_i^{j'} / a_j^{i'}$ and $a_i^j / a_j^i < a_i^{j'} / a_j^{i'}$. If
there exists $\delta > 0$ such that $b_j^i = a_j^i + \delta, b_{j'}^i = a_{j'}^i + \delta, b_j^{i'} = a_j^{i'} - \delta, b_{j'}^{i'} = a_{j'}^{i'} - \delta,$ and $b_k^i = a_k^i$ whenever $k \notin \{i, i'\}$ or $l \notin \{j, j'\}$, then $F(B) > F(A)$.

This means, for instance, that if the proportion of black students is higher in school $j$ than in school $j'$ while the proportion of white students is higher in school $j'$ than in school $j$, shifting a black student from $j'$ to $j$ and a white student from $j$ to $j'$ increases segregation. This conclusion does not follow from Monotonicity, unless the proportion of black students in $j$ and the proportion of white students in $j'$ exceed their expected value and the proportion of white students in $j$ and the proportion of black students in $j'$ are below their expected value.

Our second theorem is an incompatibility result.

**Row-Size Invariance.** For all $A \in A(m, n), i \in M$ and $\lambda \in \mathbb{R}_+$, $F(\lambda A_i, A_{-i}) = F(A)$.

**Theorem 2.** If $m \geq 4$, no continuous measure $F : A(m, n) \to \mathbb{R}_+$ satisfies Monotonicity and Row-Size Invariance.

**Proof. Step 1.** If $F : A(m, 2) \to \mathbb{R}_+$ satisfies Monotonicity and Row-Size Invariance, then $F$ satisfies the following property:

**Strong Monotonicity.** Let $i, i' \in M$ be such that $i \neq i'$ and let $A, B \in A(m, 2)$. Suppose that $a_1^1 a_2^2 > a_2^1 a_1^1$. If there exists $\delta > 0$ such that $b_1^1 = a_1^1 + \delta, b_2^2 = a_2^2 + \delta, b_2^1 = a_2^1 - \delta, b_1^2 = a_1^2 - \delta,$ and $b_k^j = a_k^j$ for all $k \notin \{i, i'\}$ and $j \in \{1, 2\}$, then $F(B) > F(A)$.

To prove this claim, let $i, i' \in M$ be such that $i \neq i'$. Without loss of generality, say $i = 1$ and $i' = 2$. Let $A, B \in A(m, 2)$. Suppose $a_1^1 a_2^2 > a_1^2 a_1^1$ and suppose there exists $\delta > 0$ such that

$$B = \begin{pmatrix}
    a_1^1 + \delta & a_1^2 - \delta \\
    a_2^1 - \delta & a_2^2 + \delta \\
    a_3^1 & a_3^2 \\
    \vdots & \vdots \\
    a_m^1 & a_m^2
\end{pmatrix}.$$ 

We must show that $F(B) > F(A)$.
For every $\varepsilon > 0$, let

$$A(\varepsilon) = \begin{pmatrix} a_1^1 & a_1^2 \\ a_2^1 & a_2^2 \\ \varepsilon a_3^1 & \varepsilon a_3^2 \\ \vdots & \vdots \\ \varepsilon a_m^1 & \varepsilon a_m^2 \end{pmatrix}.$$ 

Let $a_i^j(\varepsilon)$ denote the $(i,j)$-th entry of $A(\varepsilon)$ and write $a_i(\varepsilon) = \sum_{j \in N} a_i^j(\varepsilon)$, $a^j(\varepsilon) = \sum_{i \in M} a_i^j(\varepsilon)$, and $a(\varepsilon) = \sum_{i \in M} \sum_{j \in N} a_i^j(\varepsilon)$. We claim that for $\varepsilon$ small enough,

$$a_1(\varepsilon)a(\varepsilon) > a_1^1(\varepsilon)a_1^1(\varepsilon). \quad (3)$$

To see why, define

$$\alpha(\varepsilon) = a(\varepsilon) - (a_1^1 + a_1^2 + a_2^1 + a_2^2),$$

$$\alpha^1(\varepsilon) = a^1(\varepsilon) - (a_1^1 + a_2^1),$$

and note that $\alpha(\varepsilon) \to 0$ and $\alpha^1(\varepsilon) \to 0$ when $\varepsilon \to 0$. Moreover, we find that

$$a_1^1(\varepsilon)a(\varepsilon) - a_1^1(\varepsilon)a_1^1(\varepsilon)$$

$$= a_1^1 \left[ \alpha(\varepsilon) + (a_1^1 + a_1^2 + a_2^1 + a_2^2) \right] - a_1^1 \left[ \alpha^1(\varepsilon) + (a_1^1 + a_2^1) \right]$$

$$= (a_1^1a_2^2 - a_1^2a_2^1) + (a_1^1\alpha(\varepsilon) - a_1\alpha^1(\varepsilon)).$$

Since $a_1^1a_2^2 - a_1^2a_2^1 > 0$ and the term in the last parenthesis vanishes as $\varepsilon \to 0$, there exists $\varepsilon > 0$ such that (3) holds.

By a similar argument, the inequality $a_1^1a_2^2 > a_1^2a_2^1$ also implies that

$$a_2^2(\varepsilon)a(\varepsilon) > a_2(\varepsilon)a^2(\varepsilon) \quad (4)$$

and

$$a_1^2(\varepsilon)a(\varepsilon) < a_1(\varepsilon)a^2(\varepsilon), \quad (5)$$

$$a_2^1(\varepsilon)a(\varepsilon) < a_2(\varepsilon)a^1(\varepsilon) \quad (6)$$

for $\varepsilon > 0$ small enough.

Now pick $\varepsilon > 0$ satisfying (3), (4), (5) and (6). This means that $i = 1$, $j = 1$, as well as $i = 2$, $j = 2$, are positively associated in $A(\varepsilon)$ whereas
\[ i = 1, j = 2, \text{ as well as } i = 2, j = 1, \text{ are negatively associated. By Row-Size Invariance and Monotonicity,} \]

\[
F(A) = F(A(\varepsilon)) < F \begin{pmatrix}
    a_1^1 + \delta & a_1^2 - \delta \\
    a_2^1 - \delta & a_2^2 + \delta \\
    \varepsilon a_3^1 & \varepsilon a_3^2 \\
    \vdots & \vdots \\
    \varepsilon a_m^1 & \varepsilon a_m^2
\end{pmatrix} = F(B).
\]

**Step 2.** If \( F : A(m, 2) \rightarrow \mathbb{R}_+ \) satisfies Monotonicity and Row-Size Invariance, then \( F \) satisfies the following property:

**Local Insensitivity.** Let \( i, i', i'' \in M \) be distinct. Let \( A \in A(m, 2) \) be such that \( a_1^1 a_2^i > a_2^1 a_1^i \) and \( a_1^1 a_2^{i''} < a_2^1 a_1^{i''} \). For every \( \delta \) such that \( 0 < \delta < a_1^i \), denote by \( A^\delta \) the matrix \( B \in A(m, 2) \) such that \( b_1^1 = a_1^1 + \delta, b_2^1 = a_2^1 - \delta, \) and \( b_k^j = a_k^j \) for all \( k \neq i \) and \( j = 1, 2 \). There exists \( \delta^* \) such that \( 0 < \delta^* < a_1^{i'} \) and \( F(A^\delta) = F(A) \) for all \( \delta \) such that \( 0 < \delta < \delta^* \).

To prove this claim, fix distinct \( i, i', i'' \in M \). Without loss, say \( i = 1, i' = 2, i'' = 3 \). Let \( A \in A(m, 2) \) be such that

\[ a_1^1 a_2^2 > a_1^2 a_2^1 \]

and

\[ a_1^1 a_3^2 < a_1^2 a_3^1. \]

Let \( \delta^* = \min \left( \frac{a_1^2 a_3^1 - a_1^1 a_3^2}{a_1^1 + a_1^2 + a_3^1 + a_3^2}, a_1^i \right) \). We show that \( F(A^\delta) = F(A) \) for all \( \delta \) such that \( 0 < \delta < \delta^* \).

For every \( \delta \) such that \( 0 < \delta < a_1^1 \), consider the matrix

\[
A^\delta = \begin{pmatrix}
    a_1^1 + \delta & a_1^2 - \delta \\
    a_2^1 & a_2^2 \\
    \vdots & \vdots \\
    a_m^1 & a_m^2
\end{pmatrix}.
\]
For all $\lambda > \frac{\delta}{a_1^2}$, Row-Size Invariance and Strong Monotonicity imply

$$F(A) = F\left( \begin{array}{ccc} a_1 & a_2 \\ \lambda a_2 & \lambda a_2 \\ a_3 & a_3 \\ \vdots & \vdots \\ a_m & a_m \end{array} \right)$$

$$< F\left( \begin{array}{ccc} a_1 + \delta & a_2^2 - \delta \\ \lambda a_2^2 - \delta & \lambda a_2^2 + \delta \\ a_3 & a_3 \\ \vdots & \vdots \\ a_m & a_m \end{array} \right)$$

$$= F\left( \begin{array}{ccc} a_1 + \delta & a_2^2 - \delta \\ a_1 - \delta & a_2^2 + \delta \\ a_3 & a_3 \\ \vdots & \vdots \\ a_m & a_m \end{array} \right).$$

Since $F$ is continuous, letting $\lambda \to +\infty$ implies

$$F(A) \leq F(A^\delta). \quad (7)$$

Next, fix $\delta$ such that $0 < \delta < \delta^*$. Observe that for all $\lambda \geq 1$,

$$\delta < \delta^* \leq \frac{a_2^2 a_3 - a_1^2 a_3}{a_1^2 + a_2^2 + a_3^2 + a_3^2} \leq \frac{\lambda(a_2^2 a_3 - a_1^2 a_3)}{(a_1^2 + a_2^2) + \lambda(a_3^2 + a_3^2)},$$

hence, after rearranging,

$$(a_1 + \delta)(\lambda a_3^2 + \delta) < (a_1^2 - \delta)(\lambda a_3^2 - \delta).$$

For all $\lambda > \max(1, \frac{\delta}{a_1^2})$, Row-Size Invariance and Strong Monotonicity there-
fore imply

\[
F(A) = F \begin{pmatrix}
    a_1^1 & a_2^1 \\
    a_1^2 & a_2^2 \\
    \lambda a_3^1 & \lambda a_3^2 \\
    a_4^1 & a_4^2 \\
    \vdots & \vdots \\
    a_m^1 & a_m^2
\end{pmatrix}
\]

\[
> F \begin{pmatrix}
    a_1^1 + \delta & a_1^1 - \delta \\
    a_2^1 & a_2^2 \\
    \lambda a_3^1 - \delta & \lambda a_3^2 + \delta \\
    a_4^1 & a_4^2 \\
    \vdots & \vdots \\
    a_m^1 & a_m^2
\end{pmatrix}
\]

\[
= F \begin{pmatrix}
    a_1^1 + \delta & a_1^1 - \delta \\
    a_2^1 & a_2^2 \\
    a_3^1 - \delta & a_3^2 + \delta \\
    a_4^1 & a_4^2 \\
    \vdots & \vdots \\
    a_m^1 & a_m^2
\end{pmatrix}
\]

Since \( F \) is continuous, letting \( \lambda \to +\infty \) yields \( F(A) \geq F(A^\delta) \). Combining this inequality with (7) gives \( F(A) = F(A^\delta) \), as claimed.

**Step 3.** No continuous measure \( F : \mathcal{A}(m, 2) \to \mathbb{R}_+ \) satisfies Monotonicity and Row-Size Invariance.

Let \( m \geq 4 \) and suppose that \( F : \mathcal{A}(m, 2) \to \mathbb{R}_+ \) is a continuous measure satisfying Monotonicity and Row-Size Invariance. We derive a contradiction.

Let \( A \in \mathcal{A}(m, 2) \) be a matrix such that

\[ a_1^1 a_2^2 > a_1^2 a_2^1, \quad (8) \]
\[ a_1^1 a_3^2 < a_1^2 a_3^1, \quad (9) \]
\[ a_2^2 a_3^1 > a_1^2 a_3^2, \quad (10) \]
\[ a_2^2 a_3^1 < a_1^2 a_3^2, \quad (11) \]
\[ a_1^1 a_4^2 > a_1^2 a_4^1. \quad (12) \]
The above inequalities are compatible: take for instance
\[
\begin{pmatrix}
  a_1^1 & a_2^1 \\
  a_1^2 & a_2^2 \\
  a_3^1 & a_3^2 \\
  a_4^1 & a_4^2 \\
\end{pmatrix} = \begin{pmatrix}
  3 & 2 \\
  1 & 2 \\
  2 & 1 \\
  2 & 3 \\
\end{pmatrix}.
\]

For each \( \delta > 0 \) sufficiently small, define
\[
A^\delta = \begin{pmatrix}
  a_1^1 + \delta & a_1^2 - \delta \\
  a_1^2 & a_2^2 \\
  \vdots & \vdots \\
  a_m^1 & a_m^2 \\
\end{pmatrix}.
\]

By Step 2, \( F \) satisfies Local Insensitivity. Because of inequalities (8) and (9) we may use that property with \( i = 1, i' = 2, i'' = 3 \) to conclude
\[
F(A) = F(A^\delta)
\] (13)
for all \( \delta > 0 \) sufficiently small.

For each \( \delta > 0 \) sufficiently small, define
\[
A_5^\delta = \begin{pmatrix}
  a_1^1 + \delta & a_1^2 - \delta \\
  a_1^2 & a_2^2 \\
  a_3^1 & a_3^2 \\
 a_4^1 - \delta & a_4^2 + \delta \\
 a_5^1 & a_5^2 \\
 \vdots & \vdots \\
 a_m^1 & a_m^2 \\
\end{pmatrix}
\]
and note that (10) and (11) imply
\[
(a_4^4 - \delta)a_2^2 > (a_4^4 + \delta)a_2^2.
\]
\[
(a_4^4 - \delta)a_3^2 < (a_4^4 + \delta)a_3^2.
\]
Because of these inequalities, we may use Local Insensitivity with \( i = 4, i' = 2, i'' = 3 \) to conclude that
\[
F(A_5^\delta) = F(A^\delta)
\]
for each $\delta > 0$ sufficiently small, which, combined with (13), implies that

$$F(A) = F(A_\delta)$$

for each $\delta > 0$ sufficiently small. Because of inequality (12), this contradicts Strong Monotonicity.

**Step 4.** Let $m \geq 4$ and let $n \geq 3$. Suppose, by way of contradiction, that there exists a continuous measure $F : A(m, n) \rightarrow \mathbb{R}_+$ that satisfies Monotonicity and Row-Size Invariance. We show that there must then exist a continuous measure $\tilde{F} : A(m, 2) \rightarrow \mathbb{R}_+$ that satisfies Monotonicity and Row-Size Invariance, contradicting Step 3.

For each matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ \vdots & \vdots \\ a_{m1} & a_{m2} \end{pmatrix}$$

in $A(m, 2)$, define $\tilde{A} \in A(m, n)$ by

$$\tilde{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1} \\ a_{21} & a_{22} & \cdots & a_{2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{m} \end{pmatrix}.$$ 

Define $\tilde{F} : A(m, 2) \rightarrow \mathbb{R}_+$ by

$$\tilde{F}(A) = F(\tilde{A}).$$

Because $F$ is continuous, so is $\tilde{F}$. Because $F$ satisfies Row-Size Invariance, so does $\tilde{F}$: considering for instance a rescaling of the first row, we have, for any $\lambda > 0$,

$$\tilde{F} \left( \begin{pmatrix} \lambda a_{11} & \lambda a_{12} \\ a_{11} & a_{12} \\ \vdots & \vdots \\ a_{m1} & a_{m2} \end{pmatrix} \right).$$
\[
F = \begin{pmatrix}
\lambda a_1^1 & \lambda a_2^2 & \lambda a_1 & \cdots & \lambda a_1^1 \\
a_1^1 & a_2^2 & a_2 & \cdots & a_2 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
a_m^1 & a_m^2 & a_m & \cdots & a_m \\
\end{pmatrix}
\]

\[
= \begin{pmatrix}
a_1^1 & a_2^2 & a_1 & \cdots & a_1 \\
a_1^2 & a_2^2 & a_2 & \cdots & a_2 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
a_m^1 & a_m^2 & a_m & \cdots & a_m \\
\end{pmatrix}
\]

\[
= \begin{pmatrix}
a_1^1 \\
\vdots \\
a_m^1 \\
\end{pmatrix}
\]

\[\tilde{F} = \begin{pmatrix}
a_1^1 \\
\vdots \\
a_m^1 \\
\end{pmatrix}.\]

It remains to be checked that \(\tilde{F}\) is monotonic. Let \(A \in \mathcal{A}(m, 2)\) and assume that, say,

\[a_1^1 a > a_2^1, \quad (14)\]
\[a_2^2 a > a_2 a^2, \quad (15)\]
\[a_1^1 a < a_1^2, \quad (16)\]
\[a_2^2 a < a_2 a^1. \quad (17)\]

Let \(\delta\) be such that \(0 < \delta < \min(a_1^2, a_2^1)\) and define

\[A(\delta) = \begin{pmatrix}
a_1^1 + \delta & a_1^2 - \delta \\
a_2^1 - \delta & a_2^2 + \delta \\
\vdots & \vdots \\
a_1^1 & a_2^2 \\
\end{pmatrix}.\]

We must prove that \(\tilde{F}(A) < \tilde{F}(A(\delta))\). From the definition of \(\tilde{A}\) and inequalities (14) to (17), we get

\[\tilde{a}_1^2 \tilde{a}_1 = (n-1)a_1^1 a > (n-1)a_1 a^1 = \tilde{a}_1 \tilde{a}_1^1,\]
\[\tilde{a}_2^2 \tilde{a}_2 = (n-1)a_2^2 a > (n-1)a_2 a^2 = \tilde{a}_2 \tilde{a}_2^2,\]
\[\tilde{a}_1^1 \tilde{a}_1 = (n-1)a_1^1 a < (n-1)a_1 a^2 = \tilde{a}_1 \tilde{a}_1^2,\]
\[\tilde{a}_2^1 \tilde{a}_2 = (n-1)a_2^2 a < (n-1)a_2 a^1 = \tilde{a}_2 \tilde{a}_2^1.\]
By definition of $\tilde{F}$ and monotonicity of $F$,

$$
\tilde{F}(A) = F(\tilde{A})
$$

$$
< F \begin{pmatrix}
\tilde{a}^1_{1} + \delta & \tilde{a}^2_{1} - \delta & \tilde{a}^3_{1} & \cdots & \tilde{a}^n_{1} \\
\tilde{a}^1_{2} - \delta & \tilde{a}^2_{2} + \delta & \tilde{a}^3_{2} & \cdots & \tilde{a}^n_{2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\tilde{a}^1_{m} & \tilde{a}^2_{m} & \tilde{a}^3_{m} & \cdots & \tilde{a}^n_{m}
\end{pmatrix}
$$

$$
= F \begin{pmatrix}
a^1_{1} + \delta & a^2_{1} - \delta & a_1 & \cdots & a_1 \\
a^1_{2} - \delta & a^2_{2} + \delta & a_2 & \cdots & a_2 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
a^1_{m} & a^2_{m} & a_m & \cdots & a_m
\end{pmatrix}
$$

$$
= \tilde{F} \begin{pmatrix}
a^1_{1} + \delta & a^2_{1} - \delta \\
a^1_{2} - \delta & a^2_{2} + \delta \\
\vdots & \vdots \\
a^1_{m} & a^2_{m}
\end{pmatrix}
$$

$$
= \tilde{F}(A(\delta)),
$$
as desired. \qed

The whole analysis above is ordinal. We could have worked directly with association orderings (as defined in Section 2) rather than association measures on $\mathcal{A}(m, n)$. In fact, this is worth doing because some association orderings cannot be represented by an association measure.

The suitable reformulation of Class-Size Invariance, Row-Size Invariance and Monotonicity for association orderings is obvious. The reformulation of Theorem 1 is equally straightforward: an association ordering $\succeq$ on $\mathcal{A}(m, n)$ is class-size invariant if and only if there exists an ordering $\succeq_0$ on $\mathcal{A}(m - 1, n - 1)$ such that, for all $A, B \in \mathcal{A}(m, n)$, $A \succeq B$ if and only if $r(A) \succeq_0 r(B)$. As for Theorem 2, call an ordering $\succeq$ on $\mathcal{A}(m, n)$ continuous if, for every $A \in \mathcal{A}(m, n)$, the sets $\{B \in \mathcal{A}(m, n) \mid B \succeq A\}$ and $\{B \in \mathcal{A}(m, n) \mid A \succeq B\}$ are closed in the standard (Euclidean) topology of $\mathcal{A}(m, n)$. Using this definition, Theorem 2 may be reformulated as follows: if $m \geq 4$, no continuous association ordering $\succeq$ on $\mathcal{A}(m, n)$ satisfies Monotonicity and Row-Size Invariance.
We do not know whether monotonic and row-size invariant (necessarily discontinuous) association orderings exist on $A(m, n)$ when $m$ and $n$ are arbitrary. Continuity plays an important role in the proof of Theorem 2.

5. Some possibility results

This section discusses a few special cases in which possibility results can be established.

1) Monotonic, class-size invariant and continuous orderings exist when $m = 2$ and $n = 2$. Indeed, the ordering $A ≿ B \iff \max \left( \frac{a_1^1 a_2^2}{a_1^2 a_2^1}, \frac{a_1^2 a_2^1}{a_1^1 a_2^2} \right) \geq \max \left( \frac{b_1^1 b_2^2}{b_1^2 b_2^1}, \frac{b_1^2 b_2^1}{b_1^1 b_2^2} \right)$ satisfies the three properties. In fact, it follows from (the reformulation of) Theorem 1 that this is the only ordering on $A(2, 2)$ that is monotonic, class-size invariant and guarantees that $A \sim B$ whenever $B$ is obtained by permuting the rows (or the columns) of $A$.

2) Monotonic and class-size invariant (but discontinuous) orderings exist when $m = 2$ or $n = 2$. This is of some interest because dichotomous characteristics are common in practice: a good deal of the segregation measurement literature focuses on that case. Here is an example when, say, $m = 2$ and $n$ is arbitrary. For every $A \in A(2, n)$ and every each pair $\{j, l\} \in 2^N$ define

$$\rho(\{j, l\})(A) = \max \left( \frac{a_1^j a_2^l}{a_1^l a_2^j}, \frac{a_1^l a_2^j}{a_1^j a_2^l} \right).$$

Let $\rho(A) \in \mathbb{R}^{(n-1)n}$ be the vector $\rho(A) = (\rho(\{j, l\})(A))_{(j, l) \in 2^N}$. Let $\succsim^L$ denote the lexicmax ordering\(^3\) on $\mathbb{R}^{(n-1)n}$ and define the ordering $\succsim$ on $A(2, n)$ by

$$B \succsim A \iff \rho(B) \succsim^L \rho(A).$$

---

\(^3\)For any $x \in \mathbb{R}^{(n-1)n}$, let $x^* \in \mathbb{R}^{(n-1)n}$ denote the vector obtained by reordering the coordinates of $x$ in nondecreasing order. The lexicmax ordering $\succsim^L$ on $\mathbb{R}^{(n-1)n}$ is defined by letting $x \succsim^L y$ if and only if there exists $j \in \left\{1, \ldots, \frac{(n-1)n}{2} \right\}$ such that $x_i^* = y_i^*$ for all $i > j$ and $x_i^* > y_i^*$ (in which case we write $x \succsim^L y$) or $x_i^* = y_i^*$ for all $i \in \left\{1, \ldots, \frac{(n-1)n}{2} \right\}$ (in which case we write $x \sim^L y$).
We prove in the Appendix that $\succeq$ is class-size invariant and monotonic.

3) Continuous, monotonic and class-size invariant measures do exist when $m = 2$ and $n = 3$ (or $m = 3$ and $n = 2$). Two examples, as the reader may check, are $F(A) = \max \left( \rho^{(1,2)}(A), \rho^{(1,3)}(A), \rho^{(2,3)}(A) \right)$ and $F(A) = \rho^{(1,2)}(A) \rho^{(1,3)}(A) \rho^{(2,3)}(A)$. Both generate the same ordering.

4) A matrix $A \in \mathcal{A}(m,n)$ is isotropic (Yule and Kendall (1948)) if $r_{i,k}^{j,l}(A) \geq 1$ for all $i,k \in M$ such that $i < k$ and all $j,l \in N$ such that $j < l$. Moreover, $r_{i,k}^{j,l}(A) \leq r_{i',k'}^{j',l'}(A)$ if $i' \leq i < k' \leq k$ and $j' \leq j < l \leq l'$.

A rearrangement of a matrix $A \in \mathcal{A}(m,n)$ is any matrix $\tilde{A} \in \mathcal{A}(m,n)$ obtained by permuting the rows and columns of $A$. We let $\mathcal{A}^{iso}(m,n)$ denote the subset of matrices in $\mathcal{A}(m,n)$ that possess an isotropic rearrangement. This corresponds to the case where, although the characteristics are only categorical, their values can be “endogenously” ordered in a way that reveals a positive correlation between them.

Let $\bar{\tau}(A) = (r_{i,k}^{j,l}(A))_{1 \leq i,j,k,l \leq m} \in \mathbb{R}^+ \mu$, where $\mu$ is the number of $2 \times 2$ submatrices of an $m \times n$ matrix and let $\succeq^L$ be the lexicmax ordering on $\mathbb{R}^+ \mu$. If $\bar{A}, \tilde{A}$ are two isotropic rearrangements of a matrix $A \in \mathcal{A}^{iso}(m,n)$, it is easy to see that $\bar{\tau}(\bar{A}) \sim^L \bar{\tau}(\tilde{A})$. We may therefore define the cpr-leximax ordering $\succeq$ on $\mathcal{A}^{iso}(m,n)$ by

$$B \succeq A \iff \bar{\tau}(\tilde{B}) \succeq^L \bar{\tau}(\tilde{A})$$

for any isotropic rearrangements $\tilde{A}, \tilde{B}$ of $A, B$.

**Proposition.** The cpr-leximax ordering $\succeq$ on $\mathcal{A}^{iso}(m,n)$ is class-size invariant and monotonic.
show that these inequalities imply \( i < k \) of generality, assume

\[ a_i \cdot a_j > a_i \cdot d^j, a_k \cdot a_i > a_k \cdot d^i, a_j \cdot a_i > a_j \cdot d^i. \]

Proof. Class-size invariance is obvious. To establish Monotonicity, let \( A, B \in A^\text{iso}(m, n) \), let \( i, k \in M \) and \( j, l \in N \) be such that \( i \neq k \) and \( j \neq l \), and suppose that \( i, j \) and \( k, l \) are positively associated in \( A \) whereas \( i, l \) and \( k, j \) are negatively associated, i.e., \( a_i \cdot a_j > a_i \cdot d^j, a_k \cdot a_l > a_k \cdot d^l, a_j \cdot a_l > a_j \cdot d^l \).

These inequalities imply \( \frac{a_i}{a_k} \cdot \frac{a_j}{a_l} > 1 \). Let \( \tilde{A}, \tilde{B} \) be isotropic rearrangements of \( A, B \). To simplify notation, suppose \( \tilde{A} = A \) and \( \tilde{B} = B \). Since \( A \) is isotropic, it follows that either (i) \( i < k \) and \( j < l \) or (ii) \( i > k \) and \( j > l \). Without loss of generality, assume \( i < k \) and \( j < l \).

Suppose there is \( \delta > 0 \) such that such that \( b_i^j = a_i^j + \delta \), \( b_i^k = a_i^k - \delta \), \( b_i^j = a_i^j + \delta \), \( b_i^k = a_i^k + \delta \), and all other entries of \( A \) and \( B \) coincide. We must show that \( \tau(\tilde{B}) \succ \tau(\tilde{A}) \).

Without loss of generality, assume that \( \delta \) is small enough to guarantee that for all \( i', j', k', l' \in M \) and \( j', j'', l', l'' \in N \) such that \( i' < k', j'' < l', \),

\[ r_{i',j',k',l'}(A) < r_{i'',j'',l',l''}(B) \Rightarrow r_{i',j',k',l'}(B) \leq r_{i'',j'',l',l''}(B) \]

and

\[ r_{i',j',k',l'}(B) < r_{i'',j'',l',l''}(B) \Rightarrow r_{i',j',k',l'}(A) \leq r_{i'',j'',l',l''}(A). \]

If \( \delta \) violates this assumption, we can write it as a sum of increments that do satisfy the assumption and apply the argument below to each of these increments.

In order to prove that \( \tau(\tilde{B}) \succ \tau(\tilde{A}) \), it suffices to show that if there exist \( i', k' \in M \) and \( j', l' \in N \) such that \( i' < k', j' < l' \), and

\[ r_{i',j',k',l'}(B) < r_{i',j',k',l'}(A), \]

then there exist \( i'', k'' \in M \) and \( j'', l'' \in N \) such that \( i'' < k'', j'' < l'' \), and

\[ r_{i'',j'',k'',l''}(B) > r_{i'',j'',k'',l''}(B), \]

(22)

Because \( A \) is isotropic, inequality (21) only holds in the following cases: (i) \( (i, j) = (i', l') \), (ii) \( (i, j) = (k', j') \), (iii) \( (i, l) = (i', j') \), (iv) \( (i, l) = (k', l') \), (v) \( (k, j) = (k', l') \), (vi) \( (k, j) = (k', j') \), (vii) \( (k, l) = (i', l') \), (viii) \( (k, l) = (k', j') \).

Let us assume (i); all other cases are treated similarly. If \( k' \neq k \), we have

\[ r_{i',j',k',l'}(B) = \frac{a_i^j \cdot a_k^l}{a_i^j \cdot (a_k^l - \delta)} > \frac{a_i^j \cdot a_k^l}{a_i^j \cdot a_i^k} = r_{i',j',k',l'}(A) \geq \frac{a_i^j \cdot a_k^l}{a_i^j \cdot a_i^k} > \frac{a_i^j \cdot a_k^l}{a_i^j \cdot (a_i^k + \delta)} = r_{i',j',k',l'}(B), \]

21
where the second inequality is an application of property (20). If \( k' = k \), we have

\[
r_{j',k'}(B) = \frac{a_{j'}(a_k^l + \delta)}{a_k^l(a_i^l - \delta)} > \frac{a_{j'} a_k^l}{a_k^l a_i^l} = r_{j',k}(A) \geq \frac{a_{j'} a_k^l}{a_k^l a_i^l} > \frac{a_{j'}(a_k^l - \delta)}{a_k^l(a_i^l + \delta)} = r_{j',k'}(B).
\]

In both cases, (22) holds for \( i'' = i', k'' = k', j'' = j', l'' = l \).

6. Appendix

Let \( \succ \) be the ordering on \( A(2, n) \) defined in (19). We prove that \( \succ \) is class-size invariant and monotonic.

Class-Size Invariance follows immediately from the observation that if \( B \) is obtained by multiplying a row or a column of \( A \), then \( \rho(A) = \rho(B) \), hence, \( A \sim B \).

As for Monotonicity, let \( A, B \in A(2, n) \). Note that by definition \( \rho^{(j,l)}(A) \geq 1 \) for every \( \{j, l\} \in 2^N \). Assume, without loss of generality, that \( i = 1, j = 1 \) as well as \( i = 2, j = 2 \) are positively associated whereas \( i = 1, j = 2 \) as well as \( i = 2, j = 1 \) are negatively associated in \( A \). These assumptions imply that \( a_1^1 a_2^2 > a_1^2 a_2^1 \), hence by (18),

\[
\rho^{(1,2)}(A) = \frac{a_1^1 a_2^2}{a_1^2 a_2^1} > 1.
\]

Suppose there is \( \delta > 0 \) such that \( b_1^1 = a_1^1 + \delta, b_2^1 = a_2^1 - \delta, b_1^2 = a_1^2 - \delta, b_2^2 = a_2^2 + \delta \) and \( b_i^j = a_i^j \) for \( i = 1, 2 \) and all \( j \in N \setminus \{1, 2\} \). We must check that \( B \succ A \).

Assume that \( \delta \) is small enough to guarantee that for all \( \{j, l\}, \{j', l'\} \),

\[
\rho^{(j,l)}(A) < \rho^{(j',l')} (A) \Rightarrow \rho^{(j,l)} (B) \leq \rho^{(j',l')} (B)
\]

and

\[
\rho^{(j,l)} (B) < \rho^{(j',l')} (B) \Rightarrow \rho^{(j,l)} (A) \leq \rho^{(j',l')} (A).
\]

This is without loss of generality because if \( \delta \) violates this assumption, we can write it as a sum of increments that do satisfy the assumption and apply the argument below to each of these increments.

By definition of \( \succ \), it is necessary and sufficient to prove that \( \rho(B) \succ_L \rho(A) \). Note first that \( \rho^{(j,l)} (B) = \rho^{(j,l)} (A) \) whenever \( j, l \notin \{1, 2\} \). Moreover,
\[ \rho^{(1,2)}(B) = \frac{(a_1^2 + \delta)(a_2^2 + \delta)}{(a_1^2 - \delta)(a_2^2 - \delta)} > \frac{a_1^2 a_2^2}{a_1^2 a_2^2} = \rho^{(1,2)}(A). \] This means that \( \rho^{(j,l)}(B) < \rho^{(j,l)}(A) \) only if (i) \( \{j,l\} = \{1,l\} \) and \( l \neq 2 \) or (ii) \( \{j,l\} = \{2,l\} \) and \( l \neq 1 \).

Suppose first that there exists \( l \neq 2 \) such that
\[ \rho^{(1,l)}(B) < \rho^{(1,l)}(A). \]
This can only occur if \( \rho^{(1,l)}(A) = \frac{a_1^2 a_2^2}{a_1^2 a_2^2} \) and \( \rho^{(1,l)}(B) = \frac{(a_1^2 - \delta)a_2^2}{(a_1^2 + \delta)a_2^2} \). In that case, however, we have
\[ \frac{a_1^2 a_2^2}{a_2^2 a_1^2} = \frac{a_1^2 a_2^2}{a_1^2 a_2^2} = \rho^{(1,l)}(A)\rho^{(1,l)}(A). \]

Since \( \rho^{(1,l)}(A) \geq 1 \) and \( \rho^{(1,2)}(A) > 1 \), it follows that \( \frac{a_1^2 a_2^2}{a_1^2 a_2^2} > 1 \), hence,
\[ \rho^{(2,l)}(A) = \frac{a_1^2 a_2^2}{a_2^2 a_1^2} > \rho^{(1,l)}(A). \] Furthermore, \( \rho^{(2,l)}(B) = \frac{(a_1^2 - \delta)a_2^2}{a_2^2 a_1^2} > \frac{a_1^2 a_2^2}{a_2^2 a_1^2} = \rho^{(2,l)}(A). \) Hence,
\[ \rho^{(1,l)}(B) < \rho^{(1,l)}(A) < \rho^{(2,l)}(A) < \rho^{(2,l)}(B). \] (23)

Suppose next that there exists \( l \neq 1 \) such that
\[ \rho^{(2,l)}(B) < \rho^{(2,l)}(A). \]
This can only occur if \( \rho^{(2,l)}(A) = \frac{a_1^2 a_2^2}{a_2^2 a_1^2} \) and \( \rho^{(2,l)}(B) = \frac{(a_1^2 - \delta)a_2^2}{(a_1^2 + \delta)a_2^2} \). In that case, however, we have
\[ \frac{a_1^2 a_2^2}{a_1^2 a_2^2} = \frac{a_1^2 a_2^2}{a_1^2 a_2^2} = \rho^{(2,l)}(A)\rho^{(1,l)}(A). \]

Since \( \rho^{(2,l)}(A) \geq 1 \) and \( \rho^{(2,l)}(A) > 1 \), it follows that \( \frac{a_1^2 a_2^2}{a_1^2 a_2^2} > 1 \), hence,
\[ \rho^{(1,l)}(A) = \frac{a_1^2 a_2^2}{a_2^2 a_1^2} > \rho^{(2,l)}(A). \] Furthermore, \( \rho^{(1,l)}(B) = \frac{a_1^2 (a_1^2 + \delta)}{a_1^2 (a_1^2 - \delta)} > \frac{a_1^2 a_2^2}{a_1^2 a_2^2} = \rho^{(1,l)}(A). \) Hence,
\[ \rho^{(2,l)}(B) < \rho^{(2,l)}(A) < \rho^{(1,l)}(A) < \rho^{(1,l)}(B). \] (24)

Equations (23) and (24) say that if a coordinate of the \( \rho(X) \) vector decreases from \( X = A \) to \( X = B \), some other coordinate whose value is greater than the value of the decreasing coordinate must increase. By definition of the lexicrom ordering, this implies that \( \rho(B) \succ_L \rho(A) \).
References


URL http://www.jstor.org/stable/2774451

URL http://www.jstor.org/stable/2775611

URL http://www.jstor.org/stable/2281536

URL http://dx.doi.org/10.1080/03610928708829408


URL http://www.jstor.org/stable/2283825

URL http://www.jstor.org/stable/2282685

URL http://www.jstor.org/stable/3186153
