Multidimensional polarization for ordinal data

Martyna Kobus1 · Radosław Kurek1

Received: 7 February 2017 / Accepted: 1 October 2018 / Published online: 6 November 2018 © The Author(s) 2018

Abstract

The dominant approach to evaluating distributional features of ordinal variables (e.g. self-reported health status) has been the Allison-Foster bipolarization ordering (henceforth AF). It has not yet been extended to a multidimensional setting. Here we fill this gap. A multidimensional extension of the AF relation is characterized by a sequence of median-preserving spreads on each dimension and association-changing switches. This extension does not pay attention to the dimensions’ association. We then offer one that does and characterize it in terms of classes of polarization measures and welfare functions. Based on these two orderings we construct polarization indices and develop statistical inference for them. We measure bidimensional polarization in educational attainment and life satisfaction across OECD members. Dependence does not affect whether or not countries dominate each other bidimensionally.

Keywords Polarization · Ordinal data · Multidimensional inequality · Dominance

1 Introduction

Ordinal variables (i.e. for which there is only information on the ordering of the categories) abound in the social sciences e.g. self-reported health status, educational attainment. In recent years, they have been utilized widely to evaluate prosperity, e.g. the 2011 resolution of the United Nations General Assembly (No. 65/309), OECD Better Life Index. The problem with such data is that they do not have a natural scale. Standard summary statistics such as the mean, the variance and inequality measures point to different conclusions depending on the scale used, which is arbitrary. To account for this, an inequality measurement theory for ordinal data has been developed (Allison and Foster 2004; Kobus and
So far a notable approach is the Allison and Foster (2004) ordering for which the most unequal distribution is the most bipolarized one. We use the word polarization here too. Although this approach has been further studied (Apouey 2007; Abul Naga and Yalcin 2008; Lazar and Silber 2013; Kobus 2015) and is increasingly used in empirical research (Madden 2010; Jones et al. 2011; Dutta and Foster 2013; Arrighi et al. 2015), it has yet to be developed in some natural directions such as multidimensionality. The very few contributions in this direction (Makdisi and Yazbeck 2014; Sonne Schmidt et al. 2016) have limited applicability because they only deal with binary indicators.\(^1\) We construct two multidimensional polarization orderings that are closely related to univariate $AF$. We characterize them in terms of elementary transformations of probability mass, a class of welfare functions and a class of measures consistent with these orderings. In a companion paper (Kobus and Kurek 2018) we develop statistical inference theory for these new measures. We use these results here to rank OECD countries in terms of educational and life satisfaction polarization, which complements the study by Balestra and Ruiz (2015).

The first ordering we propose - called $mAF_1$ - requires that $AF$ holds on each dimension. Such straightforward generalization does not take into account the association between dimensions, therefore it approximates and is applicable to cases where the interdependencies between variables are not very strong. As pointed out by Fattore and Maggino (2014) in their review of different approaches to multidimensionality in social sciences: “As a matter of fact, evaluation dimensions are often weakly interdependent (...). It is intrinsic to the true multidimensionality of the concepts related to quality-of-life” (p. 201).\(^2\) We provide a type of Hardly-Littlewood-Pólya (Hardy et al. 1934) (henceforth HLP) result (Theorem 3) in which we obtain the equivalence between three notions:

(i) the unanimity of the class of welfare functions that are increasing below the median, decreasing above the median and have all cross-differences equal to zero,

(ii) an implementable criterion that compares distributions consistently with (i), namely, the $mAF_1$ relation,

(iii) elementary transformations that reflect the notion of polarization increase, which are median-preserving spreads on each dimension and association-changing, both increasing and decreasing, switches.

The second ordering we propose - called $mAF_2$ - increases according to first-order stochastic dominance for points below the multidimensional median (i.e. a vector of medians) and increases according to survival dominance for points above the median. We show that this relation is equivalent to a class of welfare functions that for two dimensions are decreasing below the median and increasing above the median and supermodular. In the general case, relevant differences are positive or negative. $mAF_2$ is also equivalent to a class of polarization measures which is continuous, decomposable by population subgroups and increases according to $mAF_2$.

\(^1\)Makdisi and Yazbeck (2014) have a model for general ordinal variables, but then they transform them into binary variables.

\(^2\)In a series of papers, these authors develop a different approach to a multidimensional aggregation of ordinal variables (Fattore et al. 2011, b, 2012, Fattore and Maggino 2014). They focus directly on the partial order imposed by ordinal variables on statistical units.
Based on the two polarization orderings, we construct polarization indices and we show their properties (Lemma 1–3). These are multidimensional analogues of the $\alpha, \beta$ family of indices by Abul Naga and Yalcin (2008) (denoted $P_{\alpha, \beta, \gamma}$) and $\alpha, b$ family of indices by Kobus and Miłoś (2012) (denoted $P_{\alpha, b, c}$). We also study a multidimensional index consistent with the $mAF2$ (denoted $P_{mAF2}$). In a companion working paper (Kobus and Kurek 2018), we obtain explicit standard errors formulae for $P_{\alpha, \beta, \gamma}, P_{\alpha, b, c}$, and $P_{mAF2}$. Finally, we apply these three measures to study educational and life satisfaction polarization in OECD countries. In empirical applications we use a generalization of the $AF$ relation from Mendelson (1987), largely overlooked by the literature. Here the mass is concentrated around any quantile, not necessarily the median. For the most equal distribution, for which all mass is in one category, there is obviously no differentiation between various quantiles. For the most unequal distribution, one still gets two distinctive groups, but they are uneven, e.g. for the first quartile groups’ sizes are 25% vs. 75%. This generalization largely extends the applicability of the polarization ordering.

As mentioned, empirical application complements the study by Balestra and Ruiz (2015) (henceforth BR), who use the univariate $AF$ relation. They acknowledge the need for the multidimensional extension of $AF$ but their study lacks statistical analysis. We provide both extensions. We find that while univariate dominances are fairly common, bidimensional dominance is rare. There are only 20 cases of bidimensional polarization dominance and 14 cases of bidimensional welfare dominance - i.e. when welfare is understood as two univariate first-order stochastic dominances occurring together. As a region Northern Europe performs best in terms of low bidimensional polarization. The results are highly significant. Interestingly, the results are the same no matter which polarization measure we use. Since some measures neglect dependence between dimensions and others do not, this suggests that dependence does not play an important role when comparing joint distributions of education and happiness in OECD countries.

Although the $AF$ bipolarization ordering is most often used to measure inequality in ordinal variables, it is important to note that in a cardinal setting, polarization and inequality are different concepts (Esteban and Ray 1994). In particular, polarization may increase after an inequality-decreasing transfer that improves group homogeneity. In an ordinal setting, it is difficult to conceptualize inequality as a deviation from a perfectly equal distribution, because there are as many of them as there are categories. With recent contributions, however, it seems that the differentiation between inequality and polarization for ordinal data is possible. In particular, Gravel et al. (2015) propose the Hammond transfer (Hammond 1976) as a defining concept for inequality for ordinal variables. A Hammond transfer moves two individuals closer in the distribution of an ordinal indicator irrespective of whether the gain equals the loss. Cowell et al. (2017) use similar transfers. Cowell and Flachaire (2017) consider various reference points, not only the median (e.g. the highest category).

This paper is organized as follows. In Section 2, we cover the basic definitions and notation. In Section 3, we introduce the $mAF1$ relation and in Section 4 the $mAF2$ relation. We define axioms and formulate main results concerning these relations (Theorems 1–6). In Section 5 we study measures based on two polarization orderings and their properties (Lemma 1–3). In Section 6 we measure bidimensional polarization in educational attainment and life satisfaction in OECD countries. Finally, in the concluding remarks we give examples of other extensions of $AF$. This will be the subject of further research. Empirical results are collected in Appendix A and proofs in Appendix B.
2 Basic definitions and notation

A relation \( \preccurlyeq \) on a set of probability distributions is a partial ordering if and only if it satisfies

(REFLEX) \( p \preccurlyeq p \)

(ANTISYM) \( p_1 \preccurlyeq p_2 \) and \( p_2 \preccurlyeq p_1 \) implies \( p_1 = p_2 \)

(TRANSI) \( p_1 \preccurlyeq p_2 \) and \( p_2 \preccurlyeq p_3 \) implies \( p_1 \preccurlyeq p_3 \).

In what follows we will also use relations that fulfill only (REFLEX) and (TRANSI), which are quasi-orderings. An ordering is a partial ordering in which all elements are comparable (a chain). Each such ordering has an associated indifference relation (being the equivalence relation) defined as \( p_1 \sim p_2 \) if and only if \( p_1 \preccurlyeq p_2 \) and \( p_2 \preccurlyeq p_1 \). We call an element \( p \) maximal (resp. minimal) in \( \preccurlyeq \) if there exists no element \( \tilde{p} \) such that \( p \preccurlyeq \tilde{p} \) (resp. \( \tilde{p} \preccurlyeq p \)) and \( p \not\sim \tilde{p} \).

We define \( I := \{1, \ldots, n_1 \} \times \{1, \ldots, n_2 \} \times \ldots \times \{1, \ldots, n_k \} \) which is endowed with the usual partial order: \( (i_1, \ldots, i_k) \preceq (i'_1, \ldots, i'_k) \) if and only if \( i_j \leq i'_j \) for all \( j \in \{1, \ldots, k\} \). \( I \) gives labeling of ordinal categories; the results are the same if such labeling is transformed monotonically. For example, we have \( I := \{1, 2, 3, 4, 5\} \times \{1, 2, 3, 4\} \times \{1, 2, 3\} \). Let \( \bar{i} = (i_1, \ldots, i_k) \) denote the element of \( I \) e.g. \( \bar{i} = (2, 4, 1) \). Throughout the article \( I, k, n_i \) are fixed unless we explicitly state otherwise. Now let \( p \) be a probability distribution on the set \( I \).

Obviously we require
\[
\sum_{i \in I} p(i) = 1 \quad \text{and} \quad p(i) \geq 0 \quad \forall i \in I.
\]

Let \( p \) be a probability distribution on \( I \) as above. For \( j \in \{1, 2, \ldots, k\} \) we define
\[
p^j(i) := \sum_{\bar{i} \in I \text{ such that } i_j = l} p(\bar{i}), \quad l \in \{1, 2, \ldots, n_j\}.
\]

Let us take third dimension and the following probability distribution on it \( p^3(1) = 0.15, p^3(2) = 0.55, p^3(3) = 0.30 \).

We notice that \( p^j \) is a unidimensional distribution for which we define the cumulative distribution function
\[
P^j(i) = \sum_{h \leq i} p^j(h), \quad j \in \{1, 2, \ldots, k\}.
\]

For example, we get the value \( P^3(2) = 0.15 + 0.55 = 0.70 \). Let
\[
\tilde{P}^j(i) = \sum_{h \geq i} p^j(h), \quad j \in \{1, 2, \ldots, k\}.
\]

denote the survival function for dimension \( j \)-th.

In a similar manner we define a multidimensional cumulative distribution function by
\[
\mathbb{P}(\bar{i}) = \sum_{\bar{h} \leq \bar{i}} p(\bar{h}).
\]

By focusing on probability distributions instead of actual individuals (e.g. see Apouey (2007)), the usual Anonymity and Population Principle axioms are assumed.

Probability values here are arbitrary.
and a multidimensional survival function by
\[ \mathbb{P}(i) = \sum_{h \geq i} \mathbb{P}(h). \] (5)

Continuing our example we have, for example, \( \mathbb{P}(2, 4, 1) = 0.05 \). Let \( \lambda, \Lambda \) denote, respectively, the set of all probability distributions and cumulative distribution functions.

For each dimension \( j \) we define a median \( m_j \) which is the number of the category for which \( P^j(m_j - 1) < 1/2 \) and \( P^j(m_j) \geq 1/2 \). Let \( m = (m_1, \ldots, m_k) \) denote the vector of unidimensional medians. We often call such a defined multidimensional median simply the median. It is unique. This assumption can be relaxed but it is mostly technical.\(^5\) Further, for each dimension \( j \), \( \tau_j \) is the number of the category for which \( P^j(\tau_j - 1) < \pi \) and \( P^j(\tau_j) \geq \pi \), where \( \pi \in [0, 1] \). So \( \tau_j \) is a quantile of distribution \( P^j \) and \( \tau_j = m_j \) for \( \pi = 0.5 \). Let \( \tau = (\tau_1, \ldots, \tau_k) \) denote a multidimensional quantile. When there is no confusion that we are considering unidimensional objects, we omit superscript \( j \).

Finally, let the polarization index be denoted by \( \mathcal{P} : \Lambda \rightarrow \mathbb{R} \). Its properties are defined in Section 5. We also consider a class of social welfare functions that are additively separable and symmetric with respect to individuals as it is typically assumed in multidimensional dominance literature (Atkinson and Bourguignon 1982, p. 190). Let \( \mathcal{U} \) denote the set of utility functions \( u : I \rightarrow \mathbb{R} \) which can be both interpreted as a general evaluation function (which includes a utility function too) and also as a cardinal scale itself (see Yalonetzky (2013) or (Cowell and Flachaire 2017) for the same interpretation). It measures the contribution of an individual (or other unit) to total welfare and is evaluated at ordinal categories. Formally, let \( W : \lambda \times \mathcal{U} \rightarrow \mathbb{R} \) be a social welfare function such that
\[ W(p, u) = \sum_{i \in I} u(i)p(i). \]

Allison and Foster (2004) postulate that inequality in ordinal data increases when probability mass is moved away from the median, that is, when the so-called median-preserving spread occurs. They introduce a particular relation on the space of distributions (Definition 1) that embodies the notion of one distribution being more equal than the other.

**Definition 1 Unidimensional AF (AF)** Let \( p_1, p_2 \) be two distributions and let \( m \) denote the median. We write \( p_1 \preceq_{AF} p_2 \) if and only if the following conditions hold:

\[(AF1) \quad p_1, p_2 \text{ have a unique and common median } m, \]
\[(AF2) \quad P_1(i) \leq P_2(i) \text{ for any } i < m, \]
\[(AF3) \quad P_1(i) \geq P_2(i) \text{ for any } i \geq m, \]

The interpretation of the AF ordering is intuitive. In particular, we have that \( p_1 \preceq_{AF} p_2 \) when \( p_1 \) is more concentrated (i.e. when there is more probability mass) around the median than \( p_2 \). The most bipolarized distribution, that is, the one that has half of the mass in the lowest category and half of the mass in the highest category, is the most unequal according to this relation.

\(^5\)Kobus (2015) shows how the definition of AF can be extended to cover the case of several medians and only one common.
3 Dependence neutral multidimensional polarization ordering: a characterization theorem

In a cardinal setting, Atkinson’s theorem (Atkinson 1970) states that Lorenz ordering is the largest ordering compatible with each symmetric index fulfilling the Pigou-Dalton transfer axiom. We ask essentially the same question in the multidimensional ordinal framework: for which transfers is the multidimensional AF the largest equivalent ordering? And also, what is the class of indices in which all indices are consistent with the multidimensional AF relation? Finally, what are the comparisons of the distributions in terms of polarization over which all social welfare functions in a given class agree? By answering these questions we get a type of Hardly-Littlewood-Pólya result that combines transfers with a dominance ordering and a class of measures and welfare functions. Such results are typically difficult to obtain in a multidimensional framework (Gravel et al. 2015).

We take the natural extension of the unidimensional AF relation, which we call \( m\text{AF}_1 \). Then we define two types of transfers: multidimensional spread and association-changing switch. They are related to \( m\text{AF}_1 \) (Theorem 1). Then we define a class of polarization indices and a class of social welfare functions and show that they too are related to \( m\text{AF}_1 \) (Theorem 2). Together these results imply our main result (Theorem 3). We keep the notation of Allison and Foster (2004), so the dominating distribution is worse in the sense of the polarization relation.

Definition 2 Multidimensional AF (mAF1)
Let \( \mathbb{P}_1, \mathbb{P}_2 \) be two probability distributions with a unique and common median \( \mathbf{m} \). We say that \( \mathbb{P}_1 \preceq_{m\text{AF}_1} \mathbb{P}_2 \) if and only if \( p_{1j} \preceq_{\text{AF}} p_{2j} \) for all \( j \in \{1, 2, \ldots, k\} \).

In the definition above, \( p_{1j}, p_{2j} \) are the marginals of \( \mathbb{P}_1, \mathbb{P}_2 \) given by Eq. 1. Note that \( \mathbf{m} \) denotes the vector of medians \((m_1, m_2)\). If the probability mass on each marginal is concentrated in one category, then joint distribution is concentrated in one category too. The opposite is true as well, that is, if the joint probability mass is concentrated in one (multidimensional) category (which is then also the median \( \mathbf{m} \)), then so is the probability mass on each marginal. This is the least polarized distribution according to \( \preceq_{m\text{AF}_1} \).

Definition of the multidimensional spread is the following: \textit{ceteris paribus} probability mass is moved away from the median on one dimension.

Definition 3 Multidimensional Spread
Let \( \mathbb{P}_1, \mathbb{P}_2 \) be probability distributions with a unique and common median \( \mathbf{m} \). We say that \( \mathbb{P}_2 \) was obtained from \( \mathbb{P}_1 \) via a multidimensional spread if and only if for some \( j \leq k, \epsilon > 0 \) one of the following two conditions holds:

1. There exist \( i_1 < i_2 < m_j \) and \( \hat{i} \in \mathbb{I} \) such that \( p_{2\hat{i}}(\hat{i}) = p_{1\hat{i}}(\hat{i}) + \epsilon, \mathbb{P}_2(\hat{i}) = \mathbb{P}_1(\hat{i}) - \epsilon \), where \( \hat{i} \) is \( \hat{i} \) except for \( j \)-th coordinate where we put \( i_1 \) and analogously for \( \hat{i} \) where we put \( i_2 \). \( \mathbb{P}_1, \mathbb{P}_2 \) coincide on other coordinates.

2. There exist \( i_1 > i_2 \geq m_j \) and \( \hat{i} \in \mathbb{I} \) such that \( p_{2\hat{i}}(\hat{i}) = p_{1\hat{i}}(\hat{i}) + \epsilon, \mathbb{P}_2(\hat{i}) = \mathbb{P}_1(\hat{i}) - \epsilon \), where \( \hat{i}, \hat{i} \) are the same as in condition 1 and \( \mathbb{P}_1, \mathbb{P}_2 \) coincide on other coordinates.

Figure 1 gives examples of the transfer described in Definition 3. The first picture refers to point 1 in the definition and the second picture refers to point 2. The solid lines show medians on both dimensions. Another type of transfer that is a necessary constraint on distributions consistent with \( m\text{AF}_1 \) is the following.
Definition 4 Association-changing Switch

Let $p_1, p_2$ be probability distributions on $I$. We say that $p_2$ was obtained from $p_1$ via an association-changing switch if and only if for some $i \in I$ and $j_1, j_2 \leq k$ there exist $i_1, i_2$ on dimension $j_1$ and $h_1, h_2$ on dimension $j_2$ and $\epsilon \in \mathbb{R}$ such that

$$p_2(i_1) = p_1(i_1) + \epsilon,$$
$$p_2(i_4) = p_1(i_4) + \epsilon,$$
$$p_2(i_2) = p_1(i_2) - \epsilon,$$
$$p_2(i_3) = p_1(i_3) - \epsilon,$$

where $i_1$ is $i$ except that on $j_1$–th coordinate we put $i_1$ and on $j_2$–th coordinate we put $h_1$. Analogously in $i_2$ we put $i_2, h_1$, in $i_3$ we put $i_1, h_2$ and in $i_4$ we put $i_2, h_2$.

The type of transfers described in Definition 4 either increase or decrease association by putting more mass on the diagonal or counter-diagonal, respectively. Please note that there is no relation between $i_1, i_2$ and between $h_1, h_2$. Association-increasing switches are described in Tchen (1980), Epstein and Tanny (1980) and Tsui (1999). Figure 1 (the third picture) illustrates this concept. Here, there are only two dimensions so $j_1 = 1$ and $j_2 = 2$. We choose $i_1 = 2, i_2 = 4, h_1 = 2, h_2 = 5$ and thus the shift takes place between points $p(2, 2), p(2, 5), p(4, 2), p(4, 5)$. We transfer the probability mass $\epsilon$ from $p(2, 5)$ to $p(2, 2)$, which is then compensated by a transfer from $p(4, 2)$ to $p(4, 5)$. Alternatively, one could take $\epsilon < 0$ in Definition 4 in which case a transfer would be from $p(4, 2)$ to $p(2, 2)$ and from $p(2, 5)$ to $p(4, 5)$. In our example both transfers cross the median, but in general this does not have to be the case. In a three (and more) dimensional space (Fig. 2), an association-changing switch takes place between two dimensions with the third dimension (and more dimensions) fixed.

This shows that $mAF_1$ relation and Definitions 3 and 4 are related.

Theorem 1 Let $p_1, p_n$ be probability distributions with a unique and common median $m$. The following statements are equivalent:

(i) $p_1 \succ_{mAF_1} p_n$

(ii) There exists a sequence $p_i, i = 1, \ldots, n$ such that $p_i$ differs from $p_{i+1}$ by either a multidimensional spread or an association-changing switch.

A class of indices consistent with Definition 3 increases following a multidimensional spread.
Definition 5 An index $\mathcal{P}$ is increasing with respect to a multidimensional spread if
$$\mathcal{P}(\mathbb{P}_1) \geq \mathcal{P}(\mathbb{P}_2),$$
where $\mathbb{P}_1$ was obtained from $\mathbb{P}_2$ via a multidimensional spread.

A class of indices consistent with Definition 4 does not change following an association-decreasing or increasing switch.

Definition 6 An index $\mathcal{P}$ is neutral with respect to an association-changing switch if
$$\mathcal{P}(\mathbb{P}_1) = \mathcal{P}(\mathbb{P}_2)$$
where $\mathbb{P}_1$ was obtained from $\mathbb{P}_2$ via an association-changing switch.

Similarly to polarization indices, a class of welfare functions that is consistent with Definitions 3 and 4 increases with respect to the unidimensional $AF$ relation on each marginal and treats dependence structure as irrelevant. Let us define the discrete analogue of a differential operator
$$D_{lj}(i_1, i_2, \ldots, i_k) = u(i_1, i_2, \ldots, i_l + 1, \ldots, i_k) - u(i_1, i_2, \ldots, i_l, \ldots, i_k)$$
We consider the following class of utility functions
$$\mathcal{U}_1 := \{u : D_{lj}(\hat{i}) \geq 0 \text{ for all } i_j < m_j \text{ and all } j \in \{1, \ldots, k\},$$
$$D_{lj}(\hat{i}) \leq 0 \text{ for all } i_j \geq m_j \text{ and all } j \in \{1, \ldots, k\},$$
$$D_j D_{lj}(\hat{i}) = 0 \text{ for all } i \in I \text{ and all } j < l, j, l \in \{1, \ldots, k\}\}. \}
Utility functions that belong to $\mathcal{U}_1$ are increasing for points below the median, decreasing for points above the median and have all cross-differences equal to zero. This class of welfare functions reminds one of ALEP neutrality (Kannai 1980). These are functions that
increase with the ordering on each marginal distribution and do not pay attention to the dependence.

**Theorem 2** Let \( P_1, P_n \) be probability distributions with a unique and common median \( m \). The following statements are equivalent:

(i) \( P_1 \preceq_{mAF1} P_n \)
(ii) \( P(P_1) \leq P(P_n) \) for all \( P \) satisfying Definitions 5 and 6.

(iii) \( W(P_1) \geq W(P_n) \) for all \( u \in U^1 \).

Combining Theorems 1 and 2 gives us the main result that characterizes the multidimensional polarization relation in the HLP spirit.

**Theorem 3** Let \( P_1, P_n \) be probability distributions with a unique and common median \( m \). The following statements are equivalent:

(i) \( P_1 \preceq_{mAF1} P_n \)
(ii) There exists a sequence \( P_i \) such that \( P(P_i) \leq P(P_{i+1}) \) and \( P_i \) differs from \( P_{i+1} \) by a multidimensional spread or association-changing switch.

(iii) \( P(P_1) \leq P(P_n) \) for all \( P \) satisfying Definitions 5 and 6.

(iv) \( W(P_1) \geq W(P_n) \) for all \( u \in U^1 \).

### 4 Dependence increasing multidimensional polarization ordering: a characterization theorem

Ignoring dependence is a significant limitation of \( mAF1 \), although as mentioned in the Introduction, ordinal indicators used in social sciences are often weakly interdependent. Here we propose a relation which is a different multidimensional generalization of \( AF \) and increases with dependence. The more interdependent attributes are, the more polarized the joint distribution is.

**Definition 7** Multidimensional AF (\( mAF2 \))

Let \( P_1, P_2 \) be two probability distributions with a unique and common median \( m \). We say that \( P \preceq_{mAF2} \) if and only if the following two conditions hold

1. \( P(i) \leq Q(i) \) for \( i < m \)
2. \( \bar{P}(i) \leq \bar{Q}(i) \) for \( i \geq m \) where \( \bar{P}, \bar{Q} \) denote survival functions of \( P \) and \( \eta \), respectively.

In other words, for \( i < m \) relation \( mAF2 \) increases according to first-order stochastic dominance, and for \( i \geq m \) relation \( mAF2 \) increases according to survival dominance. Figure 3 helps to explain the intuition behind \( mAF2 \). Consider the types of bidimensional transfers of probability mass that are consistent with \( mAF1 \). These transfers increase bipolarization on each dimension. That is, for \( i < m \) probability mass is moved towards lower categories and for \( i \geq m \) probability mass is moved towards higher categories. However, when \( i_1 \leq m_1 \) and \( i_2 \geq m_2 \) or when \( i_1 \geq m_1 \) and \( i_2 \leq m_2 \), then mass is moved in reversed direction on both dimensions. Transfers that move mass in the same direction are
consistent with either first-order stochastic dominance or with survival dominance, which are both well-known partial orderings on distributions. However, transfers that move mass in opposite directions are not characterized by any known relation on distributions. Therefore, we do not impose anything on how $mAF_2$ behaves in such cases.

Let us consider the following class of utility functions.

$$U^2 := \{ u : \forall 1 \leq s \leq k \forall 1 \leq t_1 < t_2 < ... < t_s \leq k (-1)^s D_{t_1} D_{t_2} ... D_{t_s} u(i_1, i_2, \ldots, i_k) \leq 0 \text{ for } i < m, \ D_{t_1} D_{t_2} ... D_{t_s} u(i_1, i_2, \ldots, i_k) \leq 0 \text{ for } i \geq m, \ u_i = 0 \text{ for } i \notin \{i < m \} \cup \{i > m\} \}. $$

Utility functions that belong to $U^2$ are the following. When $j$ is an odd number, differences of the $j$-th order are lower than zero for categories below the median and greater than zero for categories above the median. When $j$-th is an even number they are all greater than zero.

For two dimensions, such functions are decreasing (increasing) below (above) the median, i.e. $u(i_2, j_2) - u(i_1, j_2) - u(i_2, j_1) + u(i_1, j_1) \geq 0$ for $i_2 > i_1$ and $j_2 > j_1$.

The $mAF_2$ relation is the largest (in the sense of inclusion) relation on distributions that is equivalent to the class of welfare functions for which utility functions belong to $U^2$.

**Theorem 4** Let $\Pi_1, \Pi_2$ be probability distributions with a unique and common median $m$. The following statements are equivalent:

(i) $\Pi_1 \preceq_{mAF_2} \Pi_2$
(ii) $W(\Pi_1) \geq W(\Pi_2)$ for all $u \in U^2$.

The $mAF_2$ relation is also consistent with a class of polarization measures. We state this result here, because it gives another characterization of the $mAF_2$ relation, however, the axioms used in Theorem 5 will be defined in the next section. This result will also be useful in studying properties of polarization measures related to the $mAF_2$ relation.
Theorem 5 Let $\mathbb{P}_1, \mathbb{P}_2$ be two probability distributions with a unique and common median $m$. The following statements are equivalent:

(i) $\mathbb{P}_1 \preceq_{mAF_2} \mathbb{P}_2$
(ii) $P(\mathbb{P}_1) \leq P(\mathbb{P}_2)$ for all $P$ satisfying CON, NORM, DECOMP and EQUAL2.

Combining Theorems 4 and 5 provides a characterization result for the $mAF_2$ relation.

Theorem 6 Let $\mathbb{P}_1, \mathbb{P}_2$ be two probability distributions. The following statements are equivalent:

(i) $\mathbb{P}_1 \preceq_{mAF_1} \mathbb{P}_2$
(ii) $W(\mathbb{P}_1) \geq W(\mathbb{P}_2)$ for all $u \in U^2$.
(iii) $P(\mathbb{P}_1) \leq P(\mathbb{P}_2)$ for all $P$ satisfying CON, NORM, DECOMP and EQUAL2.

5 Multidimensional polarization measures

In Sections 3 and 4 we characterized polarization relations $mAF_1$ and $mAF_2$. These are partial orderings, so we will now consider polarization measures related to them. Measures ensure linear ordering on the set of distributions, hence they are always conclusive. The following axioms will be imposed on polarization indices.

CON $P : \lambda \rightarrow \mathbb{R}$ is a continuous function.
NORM The range of $P(Ran(\mathbb{P}))$ is the closed interval $[0, 1]$.
DECOMP There exist $f : Ran(\mathbb{P}) \times Ran(\mathbb{P}) \times (0, 1) \rightarrow \mathbb{R}$ continuous and strictly increasing with respect to the first two coordinates such that for any $\mathbb{P}_1, \mathbb{P}_2 \in \lambda$, $\alpha \in (0, 1)$

$$P(\alpha \mathbb{P}_1 + (1 - \alpha) \mathbb{P}_2) = f(P(\mathbb{P}_1), P(\mathbb{P}_2), \alpha),$$

where $\alpha \mathbb{P}_1 + (1 - \alpha) \mathbb{P}_2$ is a weighted sum of probability distributions, i.e. if $\mathbb{P}_1$ assigns mass $\mathbb{P}_1(i)$ to category $i$ and $\mathbb{P}_2$ assigns mass $\mathbb{P}_2(i)$, then the probability mass attributed to $i$ in $\alpha \mathbb{P}_1 + (1 - \alpha) \mathbb{P}_2$ is $\alpha \mathbb{P}_1(i) + (1 - \alpha) \mathbb{P}_2(i)$.

ATTRDECOMP There exist $f : Ran(\mathbb{P})^k \rightarrow \mathbb{R}$ continuous and strictly increasing with respect to each coordinate such that for any $\mathbb{P} \in \lambda$ we have

$$P(\mathbb{P}) = f(P(p^1), P(p^2), \ldots, P(p^k))$$

where

$$P(p^j) = P(p^j \otimes p^j \otimes \cdots \otimes p^j)$$

where $p \otimes q$ denotes product measure of $p$ and $q$.

ADDSEP Let $P$ fulfill ATTRDECOMP. Thus $P(\mathbb{P}) = f(P(p^1), P(p^2), \ldots, P(p^k))$.

ADDSEP further requires that function $f$ has the form $f(x) = \sum_{j=1}^{k} f_j(x_j)$.

EQUAL Let $P$ be consistent with $AF$ relation on a single dimension i.e. $p_1 \prec_{AF} p_2 \implies P(p_1) < P(p_2)$, where $P(p)$ is as in ATTRDECOMP.

EQUAL1 Let $P$ be consistent with $mAF_1$ relation i.e. $\mathbb{P}_1 \prec_{mAF_1} \mathbb{P}_2 \implies P(\mathbb{P}_1) < P(\mathbb{P}_2)$.

EQUAL2 Let $P$ be consistent with $mAF_2$ relation i.e. $\mathbb{P}_1 \prec_{mAF_2} \mathbb{P}_2 \implies P(\mathbb{P}_1) < P(\mathbb{P}_2)$.

CON is a natural technical assumption. The index fulfills NORM when it attaches value 0 to the least polarized distribution (i.e. all mass is in one multidimensional category) and when it
attains value 1 for the most polarized distribution (i.e. the probability mass is equally divided by points \((1, \ldots, 1)\) and \((n_1, \ldots, n_k)\)). \text{DECOMP} means that a given measure is decomposable by population subgroups. The original notion comes from Shorrocks (1984) and in the ordinal setting it was first defined by Kobus and Miłoś (2012). It works in the following way. If \(P_1, P_2\) is the distribution of, respectively, men and women, and \(\alpha\) is the population size of the men’s group, then \(\alpha P_1 + (1 - \alpha) P_2\) is the whole distribution, and the polarization measure which is decomposable represents the polarization in the whole distribution as some function of the polarization values in men’s and women’s distribution. \text{ATTRDECOMP} requires that a multidimensional polarization index can be represented as a function of unidimensional polarization indices. It is a desired property of inequality, polarization or poverty measures as it allows the evaluation of the contribution of each dimension to overall inequality/polarization. This definition was introduced by Abul Naga and Yalcin (2008). This can also be treated as a form of obtaining a multidimensional measure which is alternative to, for example, first computing individual welfare levels and then computing the overall index. Zhong (2009) applies the Abul Naga-Geoffard decomposition in a health-income context and Crocci Angelini and Michelangeli (2012) to study the evolution of well-being inequality in some EU countries. Kobus and Miłoś (2012) characterizes attribute decomposability in the strongest form, namely when association is not taken into account, and provides a simple proof to check the decomposability of specific indices. A strong form of \text{ATTRDECOMP} is the one used here. ADDSEP makes the index additive. \text{EQUAL}, \text{EQUAL1} and \text{EQUAL2} ensure consistency with, respectively, the \(\text{AF} \), \(\text{mAF} \) and \(\text{mAF2} \) relation.

Before we study measures related to our polarization orderings, we note the following relationship between axioms.

\text{Remark 1} Let \(P(p^j)\) be defined as in \text{ATTRDECOMP}. Then, \(\text{EQUAL1} \implies \text{EQUAL}\).

\text{Remark 2} If \(P\) fulfills \text{ATTRDECOMP} and \text{EQUAL}, then \(P\) fulfills \text{EQUAL1}.

We will now offer three families of multidimensional polarization measures. In Lemmas 1–3 we show what properties they satisfy. The following index is consistent with the \(\text{mAF2} \) relation, i.e. it fulfills \text{EQUAL2}. This is a simple index that adds probability mass for points below the multidimensional median and subtracts probability mass for points above the multidimensional median. The denominator is for normalization.

\text{Lemma 1}

\[ P_{\text{mAF2}}(\mathcal{P}) = \frac{\sum_{i < m} \mathcal{P}(i) + \sum_{i \geq m} \mathcal{P}(i) - 1}{\# \{i : i < m\} + \# \{i : i \geq m\} + 1} - 1 \]  

fulfills \text{CON}, \text{NORM}, \text{DECOMP} and \text{EQUAL2}.

The following index is consistent with the \(\text{mAF1} \) relation.

\text{Lemma 2}

\[ P_{\alpha, \beta, \gamma}(\mathcal{P}) = \left[ \left( P_{\alpha_1, \beta_1} (\mathcal{P}^1) \right)^\gamma + \left( P_{\alpha_2, \beta_2} (\mathcal{P}^2) \right)^\gamma + \cdots + \left( P_{\alpha_k, \beta_k} (\mathcal{P}^k) \right)^\gamma \right]^{\frac{1}{\gamma}}, \alpha, \beta \geq 1. \]  

\(\square\) Springer
where $P(p)$ is the Abul Naga and Yalcin (2008) $\alpha, \beta$ index, namely

$$P_{\alpha,\beta}(p) = \frac{\sum_{i < m} P(i)^{\alpha} - \sum_{i \geq m} P(i)^{\beta} + n + 1 - m}{(m - 1)(\frac{1}{2})^{\alpha} - 1 + (n - m)(\frac{1}{2})^{\beta}} + (n + 1 - m), \quad \alpha, \beta \geq 1$$

fulfills CON, NORM, EQUAL1 and ATTRDECOMP.

Here $\alpha, \beta$ are vectors, therefore we put them in bold. When $\alpha_j \to 1$, then the index becomes more sensitive to inequality below the median and it abstracts from it when $\alpha_j \to \infty$. Similarly for $\beta$. $P$ increases in $\gamma$. When $\gamma \to -\infty$ the index places more weight on the dimension with the smallest polarization; on the other hand, when $\gamma \to \infty$ the index takes into account the dimensions with the highest polarization.

Here is another index consistent with $mAF_1$.

Lemma 3

$$P_{a,b,c}(p) = \frac{c_1P_{a_1,b_1}(p^1) + c_2P_{a_2,b_2}(p^2) + \cdots + c_kP_{a_k,b_k}(p^k)}{\sum_{i=1}^k c_i},$$

(8)

where $P(p)$ is the Kobus and Miłoś (2012) $a, b$ index, namely

$$P_{a,b}(p) = \frac{a \sum_{i < m} P(i) - b \sum_{i \geq m} P(i) + b(n + 1 - m)}{a(m-1) + b(n-m)}; \quad a, b \geq 0.$$

fulfills CON, NORM, DECOMP, ATTRDECOMP, EQUAL1 and ADDSEP.

Here $a, b, c$ are vectors. If $a_j = 1$ and $b_j = 1$ for all $j$, then we get $P_{1,1}$ which is the multidimensional version of the absolute value index introduced by Abul Naga and Yalcin (2008). When $a_j > b_j$ the index is more sensitive to polarization below the median on the $j$-th dimension, whereas the opposite is true if $a_j < b_j$ and more weight is attached to polarization above the median.

### 6 Educational and life satisfaction polarization among OECD countries

In the empirical analysis below we use the Mendelson (1987) definition which allows for crossing points $\tau$ other than the median $m$. In our dataset we mostly find dominances for which $\tau \neq m$. We recall the original definition of Mendelson.

**Definition 8 Unidimensional $AF_\tau$ ($AF_\tau$)** Let $p_1, p_2$ be two probability distributions and let $\tau$ denote the quantile. We write $p_1 \prec_{AF_\tau} p_2$ if and only if the following conditions hold:

- **(AF\_1)** $p_1, p_2$ have a unique and common quantile $\tau$,
- **(AF\_2)** $P_1(i) \leq P_2(i)$ for any $i < \tau$,
- **(AF\_3)** $P_1(i) \geq P_2(i)$ for any $i \geq \tau$.

Distributions can cross at several quantiles at the same time, but then the values of cdfs are the same for all these quantiles (Lemma 4). Therefore we take $AF_\tau$ with the lowest $\alpha$.

**Lemma 4** If $p \preceq_{AF} q$ for quantiles $\tau_1 < \tau_2$ then $P(i) = Q(i)$ for $\tau_1 \leq i < \tau_2$. 
Proof From $p \preceq_{AF} q$ ($\tau_1$) we get that $P(i) \geq Q(i)$ for $i \geq \tau_1$ and from $p \preceq_{AF} q$ ($\tau_2$) we get that $P(i) \leq Q(i)$ for $i < \tau_2$, so by combining both we obtain $P(i) = Q(i)$ for $\tau_1 \leq i < \tau_2$.

We generalize $AF_\tau$ in the same way as we constructed $mAF_1$.

**Definition 9 Multidimensional $AF_\tau$ ($mAF_\tau$)** Let $p_1, p_2$ to be two probability distributions with a unique and common multidimensional quantile $\tau = (\tau_j)_{j=1}^k$. We say that $p_1 \preceq_{mAF_\tau} p_2$ if and only if $p_1^{(j)} \preceq_{AF_\tau} p_2^{(j)}$ for all $j \in \{1, \ldots, k\}$.

In words, $p_1 \preceq_{mAF_\tau} p_2$ when $p_2$ is less concentrated around the $\tau$-quantile than $p_1$ and is therefore considered more polarized. It should be noted that $\tau$ may include different quantiles for marginal distributions, that is, it might be that $\tau_j \neq \tau_l$ for $j \neq l$. With this definition, after replacing the median by $\tau$, Theorem 3 is still valid. Coming back to measures, when $AF_\tau$ replaces $AF$, then the Abul Naga and Yalcin index becomes

$$P_{\alpha,\beta}(p) = \frac{\sum_{i<m} P(i)^\alpha - \sum_{i\geq m} P(i)^\beta + n + 1 - m}{(m-1)\tau^\alpha - \left[1 + (n - m)\tau^\beta\right] + (n + 1 - m)}, \alpha, \beta \geq 1$$

where now $m$ denotes the $\tau$-th quantile. That is, although Lemma 4 holds, the values of the index may change when the parameter changes from the median to any quantile. The Kobus and Mikoš index changes to

$$P_{a,b}(p) = \frac{a \sum_{i<m} P(i) - b \sum_{i\geq m} P(i) + b(n + 1 - m)}{a(m-1)\tau + b(n - m)(1 - \tau)}; \ a, b \geq 0.$$ 

where again $m$ denotes the $\tau$-th quantile.

We are now ready to state the results of our empirical analysis. It complements the study made recently by Balestra and Ruiz (2015) who use the $AF$ approach to compare OECD countries in terms of polarization in education and life satisfaction. However, they themselves mention that the shortcoming of their study is that they treat the two dimensions separately. We fill this gap. We use the World Values Survey Wave 5 (collected between years 2005 and 2009) and Wave 6 (2010-2014) which covers 37 OECD countries. The sample sizes range from 784 (New Zealand) to 2809 (South Africa) respondents. Educational attainment is an ordinal indicator with 9 categories which are answers to the question "What is the highest education level that you have attained?" - which we encoded as a 4-category variable. Life satisfaction is a 10-category variable answering the question "All things considered, how satisfied are you with your life as a whole these days?". We encoded two adjacent categories as one and got a 5-category indicator. Table 1 provides the detailed description. The number of possible pair-wise dominance relationships we consider is 666, that is, we consider all situations where there is a dominance of $p_1$ over $p_2$ but there is no dominance of $p_2$ over $p_1$.

As for unidimensional comparisons, the extension of $AF$ dominance to $AF_\tau$ dominance reduces the incompleteness of inequality rankings sixfold for each dimension (Table 2). Concerning education, Austria, New Zealand, Norway, Slovakia and USA are dominated by most countries, that is, they are the least educationally polarized. The most polarized countries are Slovenia, Turkey, Mexico and Luxembourg. For life satisfaction the best performing countries are generally Northern European countries and the worst are Eastern European countries and France (“the French unhappiness puzzle” - Senik (2014)). Except for the UK, as a region Northern Europe performs well in both dimensions.
Bidimensional polarization is rare. We find only 20 cases of $mAF_\tau$ dominance for $\tau = \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}$. We also find 14 cases when $mAF_\tau$ dominance practically comes down to two unidimensional $FSD$ dominances, because the crossing point is the last category. These 14 cases thus represent examples of welfare dominance, when welfare is understood as the joint occurrence of two univariate $FSD$ dominances. In those cases, the value of $P_{a,b,c} = \sum_{i<m} P(i) \left( \frac{m}{m-1} \right) \tau$ does not change with parameters because they cancel out, whereas the value of $P_{\alpha,\beta,\gamma} = \sum_{i<m} P(i) \left( \frac{\alpha}{m-1} \right)^{\gamma+1}$ does with $\alpha$ due to non-linearity. Furthermore, many $mAF_\tau$ dominances repeat, namely, we find dominance for both $\tau = \frac{1}{4}$ and for $\tau = \frac{1}{2}$. In those cases the same category is both the first quartile and the median, but neither of them is the first nor the last category. Index $P_{mAF_2}$ remains unchanged. Tables 3, 4 and 5 report the values of indices together with standard errors. Polarization scores differ significantly between countries - e.g. between 0.28 and 0.47 for the $P_{1,1,2}$ index (Table 3 second column).

Table 6 contains the values of a test statistic for differences in polarization measures. Generally Northern Europe performs best as a region. This is true for welfare dominance as well, so this regions enjoys both relatively high aggregate levels of achievement in both dimensions and a low level of polarization whatever scale considered. The results often resolve ambiguities in the BR study. For example, New Zealand is less bidimensionally polarized than Australia, but the comparison of dimensions separately gives ambiguous conclusions.

The results are typically highly significant. In the case of $P_{\alpha,\beta,\gamma}$ what makes the results less significant is putting less emphasis on polarization in the lower end of the distribution (fourth vs. second column of Table 6). For $P_{a,b,c}$ this happens too but to a lesser extent and when more emphasis is put on polarization at the higher end of the distribution. The results obtained due to $P_{mAF_2}$ are practically the same as the results from two $mAF_1$ indices. As already discussed, comparing to $P_{a,b,c}$ and $P_{a,b,c}$, $P_{mAF_2}$ takes into account the association between dimensions of well-being. Therefore this indicates that in the evaluation of OECD countries in terms of bidimensional polarization, there is no impact of association on whether a given country dominates the other given that it already dominates it in each dimension.

7 Concluding remarks

The Allison-Foster polarization ordering is the main approach in measuring inequality for ordinal data. In this paper, we studied its multidimensional counterparts and proved a set of standard results regarding the dominance relations and measures based on them. While Kobus (2015) relaxes the assumption of a unique median and even more conclusiveness can be reached with Mendelson’s generalization, still the reliance on a common crossing is a limiting feature of the AF approach, namely, distributions that do not cross cannot be compared. Abul Naga and Yalcin (2010) and Sarkar and Santra (2016) shed some light on median-independent orderings. This is definitely an important research direction for multidimensional polarization for ordinal data.

Still, other generalizations of the AF approach are worth considering. For example, one could take different definitions of the multidimensional median - e.g. in the case of the geometric median one considers transformations of the sample points that minimize the sum of Euclidean distances without changing the argument minimum. Another extension is to consider the convex set spanned by the points $(n_1, \ldots, m_j, \ldots, n_k)$, $j = 1, \ldots, k$ and
Then, multidimensional $AF$ can be defined as first-order dominance below this set and either survival or reverse first-order dominance above this set.

Furthermore, Theorem 1 shows explicitly the types of transfers that $mAF1$ is sensitive to and is salient about. Given this theorem it seems natural to combine this ordering with the relation that treats association separately. Such separation is consistent with the view of a multidimensional distribution as a collection of the marginals and dependence structure (e.g. Nelsen (1999)). This is an interesting direction because it almost automatically gives an attribute decomposability of polarization measures into univariate polarizations and a measure of dependence. One should bear in mind, however, technical problems related to dependence for discrete distributions (Carley 2002). This is because the concept of dependence is not clear for ordinal variables, and more general for discrete variables (Genest and Neslehova 2007). Correlation is certainly not a good measure as it is not scale-free and has many other drawbacks (e.g. it only picks up linear dependence). It is therefore through such comparisons as in Table 6, among others, i.e. the comparisons of scale-free measures that treat association differently, that the impact of association can be detected for ordinal indicators.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Abul Naga, R.H., Stapenhurst, C.H.: Estimation of inequality indices of the cumulative distribution function. Econ. Lett. 130, 109–112 (2015)
Abul Naga, R.H., Yalcin, T.: Inequality measurement for ordered response health data. J. Health Econ. 27(6), 1614–1625 (2008)
Abul Naga, R.H., Yalcin, T.: Median independent inequality orderings, University of Aberdeen Business School Working Paper Series No. 3–125 (2010)
Allison, R.A., Foster, J.E.: Measuring health inequality using qualitative data. J. Health Econ. 23(3), 505–524 (2004)
Apouey, B.: Measuring health polarization with self-assessed health data. Health Econ. 16, 875–894 (2007)
Apouey, B., Silber, J.: Inequality and bi-polarization in socioeconomic status and health: ordinal approaches, PSE Working Papers No. 2013–30 (2013)
Arrighi, Y., Abu-Zaineh, M., Ventelou, B.: To count or not to count deaths: reranking effects in health distribution evaluation. Health Econ. 24(2), 193–205 (2015)
Atkinson, A.B.: On the measurement of inequality. J. Econ. Theory 2(3), 244–263 (1970)
Atkinson, A.B., Bourguignon, F.: The comparison of multidimensioned distributions of economic status. Rev. Econ. Stud. 49(2), 183–201 (1982)
Ballestra, C., Ruiz, N.: Scale-invariant measurement of inequality and welfare in ordinal achievements: An application to subjective well-being and education in oecd countries. Soc. Indic. Res. 123, 479–500 (2015)
Carley, H.: Maximum and minimum extensions of finite subcopulas. Communications in Statistics - Theory and Methods 31(12), 2151–2166 (2002)
Cowell, F., Flachaire, E.: Inequality with ordinal data. Economica 84(334), 290–321 (2017)
Cowell, F., Kobus, M., Kurek, R.: Welfare and inequality comparisons for Uni- and multidimensional distributions of ordinal data, LSE PEP Discussion Paper No. 31 (2017)
Crosci Angelini, E., Michelangeli, A.: Axiomatic measurement of multidimensional well-being inequality: Some distributional questions. J. Socio-Econ. 41(5), 548–557 (2012)
Dutta, I., Foster, J.: Inequality of happiness in US: 1972-2008. Rev. Income Wealth 59(3), 393–415 (2013)
Epstein, L.G., Tanny, S.M.: Increasing generalized correlation: a definition and some economic consequences. Can. J. Econ. 13(1), 16–34 (1980)
Esteban, J.M., Ray, D.: On the measurement of polarization. Econometrica 62(4), 819–851 (1994)
Fattore, M., Maggino, F.: Partial orders in socio-economics: A practical challenge for poset theorists or a cultural challenge for social scientists?. In: Brüggemann, R., et al. (eds.) Springer Science + Business Media New York (2014)

Fattore, M., Brüggemann, R., Owsinski, J., et al.: Using poset theory to compare fuzzy multidimensional material deprivation across regions. In: Inggrassia, S. (ed.) New perspectives in statistical modelling and data analysis. Springer, Berlin (2011)

Fattore, M., Maggino, F., Greselin, F.: Socio-economic evaluation with ordinal variables: integrating counting and poset approaches, Statistica & Applicazioni, Special Issue 31–42 (2011)

Fattore, M., Maggino, F., Colombo, E.: From composite indicators to partial order: evaluating socio-economic phenomena through ordinal data. In: Maggino, F., Nuvolati, G. (eds.) Quality of life in Italy: research and reflections, Social indicators research series, vol. 48. Springer, the Netherlands (2012)

Gravel, N., Magdalou, B., Moyes, P.: Ranking distributions of an ordinal attribute. HALSH Discussion Paper, 01082996 (2015)

Hammond, P.J.: Equity in two person situations: some consequences. Econometrica 47(5), 1127–1135 (1976)

Hardy, G.H., Littlewood, J.E., Pólya, P.: Inequalities. Cambridge University Press, Cambridge (1934)

Jones, A., Rice, N., Robone, S., Rosa Dias, P.: Inequality and polarization in health systems responsiveness: a cross-country analysis. J. Health Econ. 30(4), 616–625 (2011)

Kannai, Y.: The ALEP definition of complementarity and least concave utility functions. J. Econ. Theory 22(1), 115–117 (1980)

Kobus, M., Mikoł, P.: Inequality decomposition by population subgroups for ordinal data. J. Health Econ. 31(1), 15–21 (2012)

Kobus, M.: Polarization measurement with ordinal data. J. Econ. Inequa. 13(2), 275–297 (2015)

Lazar, A., Silber, J.: On the cardinal measurement of health inequality when only ordinal information is available on individual health status. Health Econ. 22(1), 106–113 (2013)

Lv, G., Wang, Y., Xu, Y.: On a new class of measures for health inequality based on ordinal data. J. Econ. Inequa. 13(3), 465–477 (2015)

Madden, D.: Ordinal and cardinal measures of health inequality: an empirical comparison. Health Econ. 19(2), 243–250 (2010)

Makdisi, P., Yazbeck, M.: Measuring socioeconomic health inequalities in the presence of multiple categorical information. J. Health Econ. 34, 84–95 (2014)

Mendelson, H.: Quintile-preserving spread. J. Econ. Theory 42(2), 334–351 (1987)

Nelsen, R.B.: An introduction to copulas. Springer-Verlag, New York (1999)

Genest, C., Neslehova, J.: A primer on copulas for count data. ASTIN Bull. 37, 475–515 (2007)

Tchen, A.H.: Inequalities for distributions with given the marginals. Ann. Probab. 8(4), 814–827 (1980)

Tsui, K.Y.: Multidimensional inequality and multidimensional generalized entropy measures: an axiomatic derivation. Soc. Choice Welf. 16(1), 145–157 (1999)

Sarkar, S., Santra, S.: On quasi approaches to inequality ordering of ordinal variables, http://www.isid.ac.in/epu/acegd2016/papers/SandipSarkar.pdf (2016)

Senik, C.: The French unhappiness puzzle: The cultural dimension of happiness. J. Econ. Behav. Organ. 106, 379–401 (2014)

Shorrocks, A.F.: Inequality decomposition by population subgroups. Econometrica 52(6), 1369–85 (1984)

Sonne Schmidt, C., Osterdal, L.P., Tarp, F.: Ordinal bivariate inequality: concepts and application to child deprivation in Mozambique. Rev. Income Wealth 62(3), 559–573 (2016)

United Nations General Assembly Resolution 65/309: Happiness: towards a holistic approach to development (2011)

Yalonetzky, G.: Stochastic dominance with ordinal variables: Conditions and a test. Econ. Rev. 32(1), 126–163 (2013)

Zhong, H.: A multivariate analysis of the distribution of individual’s welfare in China: What is the role of health? J. Health Econ. 28(6), 1062–1070 (2009)