INE PAN Working Paper Series

Paper number 48

Multidimensional inequality for ordinal data: further results and application to SES-health inequality

Martyna Kobus, Radosław Kurek

Warsaw, May 2019

wp@inepan.waw.pl

Multidimensional inequality for ordinal data: further results and application to SES-health inequality

Martyna Kobus [1*†], Radosław Kurek [2,†]

May 13, 2019

Abstract

Standard inequality measures are not well-suited for ordered response data e.g. self-reported health status. Literature addressing this problem has been growing, mostly around a bi-polarization relation (Allison and Foster 2004, AF). Recently (Kobus and Kurek 2018) have extended it to a multidimensional framework e.g. their results are applied to the case of measuring inequality in health and other variables (e.g. status). We further extend their results, in particular, we offer two multidimensional extensions of second-order AF (proposed originally by (Chakravarty and Maharaj 2015), propose classes of measures, study their properties and develop inference and estimation procedures. Second–order AF cares about spread of mass away from the median (like AF), but also about homogeneity of groups (so called increased bipolarity (Apouey 2007). We then compare joint distributions of two ordinal indicators, namely education and health, using Survey of Health, Ageing and Retirement in Europe. In 10% of all pair-wise comparisons we find unambiguous rankings; typically multivariate dominances are infrequent but robust. Israel emerges as the most unequal country. Dependence between education and health plays a role in overall inequality level, but not in the countries' rankings - these are the same when education and health are treated separately (i.e. as independent); similarly for increased bipolarity.

Keywords: multidimensional inequality; socio-economic inequalities in health; ordinal data

JEL classification: I1; I31; D63

[1]Institute of Economics, Polish Academy of Sciences; mkobus@inepan.waw.pl
[2]Institute of Economics, Polish Academy of Sciences; radek.kurek@inepan.waw.pl
[*]Current address: Nowy Swiat 72, 00-330 Warsaw, Poland
[†]These authors contributed equally to this work.

1 Introduction

Following the work of (Allison and Foster 2004, henceforth AF) a large body of theoretical literature has emerged to solve the problem of measuring inequality with ordered response data. Such data are very common in all social sciences. These are popular indicators such as self-reported health status, self-declared life satisfaction, educational attainment, and other. They have been more and more used in policy research analyses to evaluate prosperity e.g. in such high-profile initiatives as the report by the Stiglitz-Sen-Fitoussi Commission on the Measurement of Economic Performance and Social Progress, the OECD analyses based on the Better Life Index, or the very recent launch of the Global Multidimensional Poverty Index. The problem with such data is that standard way of cardinalizing an ordinal indicator, such as for example imposing a given scale and computing summary statistics such as mean, variance and inequality measures, suffers from a large degree of arbitrariness. In particular, when a different scale is used, conclusions reached with such procedures may change. There are many empirical examples of such reversals (Abul Naga and Yalcin 2008, Lazar and Silber 2013, Kobus 2015, Bond and Lang 2018).

A similar problem arises with the measurement of socio-economic inequalities in health. Such inequalities are extensively studied in economics, social epidemiology, medicine and health psychology, and in medical sociology (see (Smith 1999, O'Donnel et al. 2008, Cutler, Lleras-Muney and Vogl 2011, Evans, Wolfe and Adler 2012) for reviews). Generally, the lower socioeconomic status (SES) the lower the health status. This phenomenon is observed in many industrialised countries. When health is proxied by self-reported health status, standard measurement tools, such as concentration index (van Doorslaer et al. 1997, van Doorslaer and Koolman 2004) are ill–suited to such data (see (Makdisi and Yazbeck 2014)).

Therefore, for more than a decade a new methodology is being developed that is based on the distributions of ordinal variables and thus avoids scaling difficulties. Many authors have contributed to this line of research e.g. (Allison and Foster 2004, Apouey 2007, Abul Naga and Yalcin 2008, Kobus and Milos 2012, Apouey and Silber 2013, Lazar and Silber 2013, Abul Naga and Stapenhurst 2015, Gravel, Magdalou and Moyes 2017, Lv, Wang and Xu 2015, Kobus 2015, Cowell and Flachaire 2017), and others. The methodology is non-parametric and uses a minimal set of assumptions. So far the dominant approach in this field has been to measure inequality as bi-polarization as proposed by (Allison and Foster 2004) partial ordering. The most unequal distribution according to this relation is one for which half of the population occupies the lowest ordinal category and half the highest i.e. the most bi-polarized distribution.[1]

The results from this now fairly extensive literature, however, cannot be readily applied to SES – health distributions, because they are developed for a single ordinal indicator. They are suited for "pure" inequalities in health, and indeed (Madden 2010) compares the performance of ordinal and cardinal inequality indices to measure health inequality. SES-related inequalities in health require at least two indicators (a proxy for SES and a proxy for health) and thus a bidimensional framework. Very recently (Kobus and Kurek 2018) made a first attempt at developing a theory of multidimensional inequality for ordinal data. They extend the AF approach by offering two partial orderings, one which assumes independence of ordinal indicators and another one which takes dependence into account. They characterize these orderings by a set of axioms and develop classes of measures consistent with it. Here we further develop their results - please refer to the next paragraph. We then use these results to compare joint distributions of SES and health in terms of inequality. That is, if some criteria hold we judge one distribution as more bidimensionally equal than the other. Furthermore, by comparing results from methods that assume independence

[1] There has been significant controversies as to why inequality is measured as bi-polarization (see e.g. (Zheng 2008)) whereas in a cardinal framework the two are disjoint concepts (Esteban and Ray 1994), however, since many authors have already commented on this (see e.g. (Kobus 2015)), we leave this discussion behind and refer to original work by (Allison and Foster 2004) and use the word "inequality" instead of "polarization". It is important to note, however, that especially in the last years new approaches to inequality in ordinal data have emerged which deviate from AF setting (e.g. (Cowell and Flachaire 2017, Gravel, Magdalou and Moyes 2017)).

with the ones that do not, we can evaluate the role of dependence between SES and health in the overall inequality. Dependence may be viewed as capturing solely SES gradient in health, whereas joint comparisons as capturing both gradient and inequalities in marginal distributions ("pure" health and income inequalities). Here there are no assumptions as to the form of this dependence. Please note that standard approaches such as concentration index assumes a particular measure of dependence and such that are not distribution - free i.e. they take marginal distributions into account (mostly through cardinalizing an ordinal indicator (Makdisi and Yazbeck 2014).

While the AF inequality ordering is based on the concept of first order dominance, the ordering proposed by (Chakravarty and Maharaj 2015) (henceforth ChM or second – order AF (soAF)) is based on the concept of second order dominance.[2] While AF ordering is mostly characterized through how far the mass is away from the median (i.e. increased spread), second order AF additionally takes into account the distributions of these masses, namely, how much they are concentrated around particular points below and above the median (i.e. increased bipolarity). This is consistent with defining properties of polarization orderings, namely, that of high between–group heterogeneity and high within–group homogeneity (Esteban and Ray 1994). ChM implement these properties by a dominance relation which uses the sums of cumulative distribution functions (cdfs) to compare two distributions. In more detail, at the bottom end of the distribution they compare cumulative cdfs and at the top of the distribution (above the median) they compare cumulative survival functions. A distribution which is higher according to both is more polarized/more unequal. We extend their definition to a multidimensional setting, similarly to what (Kobus and Kurek 2018) have done recently with AF definition (their ordering is denoted by mAF from multidimensional AF). We propose two multidimensional orderings: the first one states that ChM relation holds for each dimension and the second uses ChM summation of cumulative proportions, but here these are proportions of the *joint* distribution. The first ordering, called $somAF1$ (second – order multidimensional AF) implicitly assumes dimensions' independence and is applicable to cases in which well-being attributes are only weakly dependent. As pointed out by (Fattore and Maggino 2014) this happens often with many quality of life indicators. The second ordering, called $somAF2$ acknowledges dimensions' dependence. We offer indices consistent with these two relations and study their properties (Lemma 1) and 2). By comparing the results from two families of measures, one can study the impact of dependence on joint inequality.

The approach presented in this papers gives very robust result *if* they hold; a typical limitation is inconclusiveness. Dominance curves may cross in which case nothing can be said about a given pair of distributions. In order to increase conclusiveness of our dominance orderings we do two things. From the beginning we remove the assumption of identical medians embedded in both AF and ChM definition, which means that we can compare distributions with different medians. The latter was the main limitation of the AF approach, recently removed by (Sarkar and Santra 2018) and we use their results. Allowing for differing medians causes problems with normalization of measures and we solve them.[3] We also allow for the concentration around different points than the median (this originally from criteria proposed by (Mendelson 1987)). This way we get what we call $somqAF$ relations, where q stands for any quantile.

Finally, we develop estimation and inference procedures for both $somAF$ – related measures and measures introduced in (Kobus and Kurek 2018) for mAF partial ordering (Theorem 1). We illustrate how these methods can be applied to distributions of health and socio–economic status. We use Wave 6 of the Survey of Health, Ageing and Retirement in Europe (SHARE), which contains rich information about health and socio–economic status of individuals aged 50 and above. We proxy SES by 7–category educational attainment, and

[2] Formally, these are not stochastic dominances, but partial sums. For differentiation please refer to the classic paper by (Fishburn and Lavalle 1995).

[3] (Sarkar and Santra 2018) are silent about these problems i.e. measures in Table 1, pp. 35 of their paper are not sufficiently normalized.

health by 5 – category self–reported health status. We find dominances in both health and education (i.e. one country dominates the other in both dimensions according to $somqAF1$) in 10% of the cases, which in comparison to e.g. (Kobus, Polchlopek and Yalonetzky 2018) who get 4% for OECD countries is significantly higher. For these 10% of the cases, the dominating countries are robustly more unequal when it comes to SES – health inequalities. These countries are mostly Israel followed by Greece. Here we find the highest education – health polarization. We observe two patterns. Firstly, the indices which take into account dependence between education and health show smaller values than what is obtained by just adding marginal inequalities in health and income. The direction of this change is to be expected (i.e. it would be the same only in the case of perfect dependence between health and education), however, the magnitude is significant; the decrease is from around 0.5 to around 0.3. Dependence does matter for the level of inequality, however, at least in this sample, not for the ranking of countries. These are the same according to both $somqAF1$ and $somqAF2$ measures. Secondly, in the comparison between $mqAF$ and $somqAF$ measures, rankings too remain stable. The added value of increased bipolarity embedded in $somAF1$ – type of orderings seems to matter for the magnitude of differences, but for the direction it is comparisons with respect increased spread property that are dominant.

The paper is organized as follows. Section 2 contains basic definitions and notation. Section 3 offers multidimensional polarization orderings and Section 4 develops measures consistent with these orderings. Section 5 develops inference and estimation for these measures and measured introduced in (Kobus and Kurek 2018). Section 6 offers a generalization of all these measures to any quantile. Then, in Section 7 we study distributions of education and health, and finally, Section 8 concludes.

2 Basic definitions and notation

We define $\mathbb{I} := \{1, \ldots, n_1\} \times \{1, \ldots, n_2\} \times \ldots \times \{1, \ldots, n_k\}$ 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 $j \in \{1, \ldots, k\}$. $\mathbb{I}$ gives labeling of ordinal categories; the results are the same if such labeling is transformed monotonically. Let $\mathbb{i} = (i_1, \ldots, i_k)$ denote the element of $\mathbb{I}$. Throughout the article $\mathbb{I}, k, n_i$ are fixed unless we explicitly state otherwise. Now let $\mathbb{p}$ be a probability distribution on the set $\mathbb{I}$.[4]

Obviously we require $\sum_{\mathbb{i} \in \mathbb{I}} \mathbb{p}(\mathbb{i}) = 1 \quad \text{and} \quad \mathbb{p}(\mathbb{i}) \geq 0 \quad \forall_{\mathbb{i} \in \mathbb{I}}$. Let $\mathbb{p}$ be a probability distribution on $\mathbb{I}$ as above. For $j \in \{1, 2, \ldots, k\}$ we define

$$p^j(i) := \sum_{\mathbb{i} \in \mathbb{I} \text{ such that } i_j = i} \mathbb{p}(\mathbb{i}), \quad l \in \{1, 2, \ldots, n_j\}. \tag{1}$$

We notice that p^j is a unidimensional distribution for which we define the cumulative distribution function

$$P^j(i) = \sum_{h \leq i} p^j(h), \quad j \in \{1, 2, \ldots, k\}. \tag{2}$$

Let

$$\bar{P^j}(i) = \sum_{h > i} p^j(h), \quad j \in \{1, 2, \ldots, k\}. \tag{3}$$

denote the survival function for dimension j-th.

In a similar manner we define a multidimensional cumulative distribution function by

$$\mathbb{P}(\mathbb{i}) = \sum_{\mathbb{h} \preceq \mathbb{i}} \mathbb{p}(\mathbb{h}). \tag{4}$$

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

and a multidimensional survival function by

$$\bar{\mathbb{P}}(\mathbb{i}) = \sum_{\mathbb{h} \succ \mathbb{i}} \mathbb{p}(\mathbb{h}). \tag{5}$$

Let λ, Λ 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 $\mathbb{m} = (m_1, \ldots, m_k)$ denote the vector of unidimensional medians. We often call such defined multidimensional median simply the median. It is unique. This assumption can be relaxed, but it is mostly technical.[5]

Finally, let a multidimensional polarization index be denoted by $\mathcal{P} : \Lambda \to \mathbb{R}$.

3 Polarization partial orderings

Here we propose two polarization partial orderings which are multidimensional extensions of ChM ordering. Before we do this, however, we need to introduce a definition of ChM ordering and a definition of AF ordering which is a foundation of ChM. We start with AF.

(Allison and Foster 2004) proposed the following relation to measure inequality for ordinal data.

Definition 1. ***First-order unidimensional AF (AF, Allison and Foster 2004)***
Let p_1, p_2 be two distributions and let m denote the median. We write $p_1 \precsim_{AF} p_2$ if and only if the following conditions hold

(AF1) *p_1, p_2 have a unique and common median m,*

(AF2) *$P_1(i) \leq P_2(i)$ for any $i < m$,*

(AF3) *$P_1(i) \geq P_2(i)$ for any $i \geq m$,*

We keep the notation of (Allison and Foster 2004), so the dominating distribution is worse in the sense of the polarization relation. The interpretation of the AF ordering is intuitive. In particular, we have that $p_1 \precsim_{AF} p_2$ 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 distribution according to this relation. The most equal distribution, on the other hand, has all probability mass in one category.

The unidimensional ChM ordering we define below relates with AF in that it distinguishes between below and above median part of the distribution. It double sums the probability distribution, starting from the median. It compares partial sums of cumulative distribution functions below the median and partial sums of survival functions above the median. It is consistent with both transfers that move mass away from the median (i.e. increased spread (Apouey 2007)) or, equivalently median-preserving spread (Kobus 2015)) and transfers that increase within – group homogeneity (i.e. increased bipolarity (Apouey 2007, Chakravarty and Maharaj 2015)).

Definition 2. ***Second-order unidimensional AF (ChM, Chakravarty and Maharaj 2015)***
Let $CS(P, i, m)$ denote cumulative sum of P up to i-th category, starting from the median m, i.e. $CS(P, i, m) = \Sigma_{i \leq \mathbb{h} \prec m} P(h)$ for $i < m$ and $CS(\bar{P}, i, m) = \Sigma_{m \leq h \leq i} \bar{P}(h)$ for $i \geq m$. Let us note that CS includes value in median category for $i \geq m$ and does not for $i < m$.

We write $p_1 \precsim_{soAF} p_2$ if and only if the following conditions hold

(AF1) *p_1, p_2 have a unique and common median m,*

(AF2) *$CS(P_1, i) \leq CS(P_2, i)$ for any $i < m$,*

[5] (Kobus 2015) shows how the definition of AF can be extended to cover the case of several medians and only one common.

(AF3) $CS(\bar{P}_1, i) \leq CS(\bar{P}_2, i)$ *for any* $i \geq m$.

Here $p_1 \precsim_{soAF} p_2$ means that distribution p_2 has more mass away from the median and/or this mass is concentrated around a given category below and/or above the median. That is, p_2 forms fewer groups further from the median than p_1. We now introduce two multidimensional extensions of ChM (or unidimensional AF), we call them *somAF1* and *somAF2*, where *som* denotes "second-order multidimensional". *somAF1* is a straightoforward extension in which we postulate that Definition 2 holds for each dimension. In this case, dimensions are treated as independent. *somAF2* accounts for dependence between dimensions in that it double sums multidimensional cumulative distribution functions and multidimensional survival functions in a manner similar to Definition 2.

Definition 3. ***Second Order Multidimensional AF (somAF1)***
Let $\mathbb{p}_1, \mathbb{p}_2$ *be two probability distributions with a unique and common median* $\mathbb{m}$. *We say that* $\mathbb{p}_1 \precsim_{somAF1} \mathbb{p}_2$ *if and only if the following two conditions hold*

(1) $CS(P_1^j, i) \leq CS(P_2^j, i)$ *for* $i < m_j$ *and for all* j,

(2) $CS(\bar{P}_1^j, i) \leq CS(\bar{P}_2^j, i)$ *for* $i \geq m_j$ *and for all* j, *where* $\bar{P}_1^j, \bar{P}_2^j$ *denote survival functions of* j*-th marginals of* $\mathbb{p}_1$ *and* $\mathbb{p}_2$, *respectively.*

In order to introduce the type of summation embedded in Definition 2 for a multidimensional distribution we define the following. Let $CS(\mathbb{P}, \mathbb{i}, \mathbb{m})$ denote cumulative sum of $\mathbb{P}$ up to $\mathbb{i}$-th category vector, starting from the multidimensional median $\mathbb{m}$, i.e. $CS(\mathbb{P}, \mathbb{i}, \mathbb{m}) = \Sigma_{\mathbb{i} \preceq \mathbb{h} \prec \mathbb{m}} \mathbb{P}(\mathbb{h})$ for $\mathbb{i} \prec \mathbb{m}$ and $CS(\bar{\mathbb{P}}, \mathbb{i}, \mathbb{m}) = \Sigma_{\mathbb{m} \preceq \mathbb{h} \preceq \mathbb{i}} \bar{\mathbb{P}}(\mathbb{h})$ for $\mathbb{i} \succ \mathbb{m}$.

Definition 4. ***Second Order Multidimensional AF (somAF2)***
Let $\mathbb{p}_1, \mathbb{p}_2$ *be two probability distributions with a unique and common median* $\mathbb{m}$. *We say that* $\mathbb{p}_1 \precsim_{somAF2} \mathbb{p}_2$ *if and only if the following two conditions hold*

(1) $CS(\mathbb{P}_1, \mathbb{i}) \leq CS(\mathbb{P}_2, \mathbb{i})$ *for* $\mathbb{i} \prec \mathbb{m}$

(2) $CS(\bar{\mathbb{P}}_1, \mathbb{i}) \leq CS(\bar{\mathbb{P}}_2, \mathbb{i})$ *for* $\mathbb{i} \succeq \mathbb{m}$ *where* $\bar{\mathbb{P}}_1, \bar{\mathbb{P}}_2$ *denote survival functions of* $\mathbb{p}_1$ *and* $\mathbb{p}_2$, *respectively.*

Both *somAF1* and *somAF2* are extensions to second-order unidimensional AF relation (Definition 2. The way we extend a unidimensional relation is directly linked to multidimensional extension to first - order unidimensional AF (Definition 1) proposed by (Kobus and Kurek 2018). These are the following.

Definition 5. ***Multidimensional AF (mAF1)***
Let $\mathbb{p}_1, \mathbb{p}_2$ *be two probability distributions with a unique and common median* $\mathbb{m}$. *We say that* $\mathbb{p}_1 \precsim_{mAF1} \mathbb{p}_2$ *if and only if* $p_1^j \precsim_{AF} p_2^j$ *for all* $j \in \{1, 2, \ldots, k\}$.

According to *mAF1*, AF holds on each dimension. 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 $\mathbb{m}$), then so is the probability mass on each marginal. This is the least polarized distribution according to $\precsim_{mAF1}$.

Definition 6. ***Multidimensional AF (mAF2)***
Let $\mathbb{p}_1, \mathbb{p}_2$ *be two probability distributions with a unique and common median* $\mathbb{m}$. *We say that* $\mathbb{p}_1 \precsim_{mAF2} \mathbb{p}_2$ *if and only if the following two conditions hold*

(1) $\mathbb{P}_1(\mathbb{i}) \leq \mathbb{P}_2(\mathbb{i})$ *for* $\mathbb{i} \prec \mathbb{m}$

(2) $\bar{\mathbb{P}}_1(\mathbb{i}) \leq \bar{\mathbb{P}}_2(\mathbb{i})$ *for* $\mathbb{i} \succeq \mathbb{m}$ *where* $\bar{\mathbb{P}}_1, \bar{\mathbb{P}}_2$ *denote survival functions of* $\mathbb{p}_1$ *and* $\mathbb{p}_2$, *respectively.*

In other words, for $\mathbb{i} \prec \mathbb{m}$ relation $mAF2$ increases according to first - order stochastic dominance and for $\mathbb{i} \succeq \mathbb{m}$ relation $mAF2$ increases according to survival dominance. Thus the more dependence between dimensions, the more multidimensional polarization as measured by $mAF2$. Please note that nothing is imposed for the case when when $i_1 \leq m_1$

and $i_2 \geq m_2$ or when $i_1 \geq m_1$ and $i_2 \leq m_2$. In such a case transfers of probability mass change polarization on respective dimensions in different directions. Such transfers are not characterized by any known relation on distributions. For $\mathbb{i} \prec \mathbb{m}$ and $\mathbb{i} \succeq \mathbb{m}$, however, the probability mass is moved towards, respectively, lower and higher categories jointly on both dimensions. Scuh transfers are consistent with either first-order stochastic dominance or with survival dominance, which are both well-known partial orderings on distributions.

4 Multidimensional polarization measures

We recall three families of multidimensional polarization measures that are consistent with $mAF1$ and $mAF2$ proposed by (Kobus and Kurek 2018). The following two indices are consistent with $mAF1$ relation.

$$\mathcal{P}_{\alpha,\,\beta,\gamma}(\mathbb{p}) = \left[\frac{(\mathcal{P}_{\alpha_1,\beta_1}(\mathbb{p}^1))^\gamma + (\mathcal{P}_{\alpha_2,\beta_2}(\mathbb{p}^2))^\gamma + \cdots + (\mathcal{P}_{\alpha_k,\beta_k}(\mathbb{p}^k))^\gamma}{k}\right]^{\frac{1}{\gamma}}, \alpha, \beta \geq 1, \quad (6)$$

where $\mathcal{P}(p)$ is the (Abul Naga and Yalcin 2008) α, β index, namely

$$\mathcal{P}_{\alpha,\beta}(p) = \frac{\Sigma_{i<m} P(i)^\alpha - \Sigma_{i\geq m} P(i)^\beta + n + 1 - m}{(m-1)(\frac{1}{2})^\alpha - \left[1 + (n-m)(\frac{1}{2})^\beta\right] + (n+1-m)}, \alpha, \beta \geq 1$$

Here α, β 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 β. $\mathcal{P}$ increases in γ. 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.

$$\mathcal{P}_{a,\,b,\,c}(\mathbb{p}) = \frac{c_1 \mathcal{P}_{a_1,b_1}(\mathbb{p}^1) + c_2 \mathcal{P}_{a_2,b_2}(\mathbb{p}^2) + \cdots + c_k \mathcal{P}_{a_k,b_k}(\mathbb{p}^k)}{\Sigma_{i=1}^k c_i}, \quad (7)$$

where $\mathcal{P}(p)$ is the (Kobus and Milos 2012) a, b index, namely

$$\mathcal{P}_{a,b}(p) = \frac{a\Sigma_{i<m} P(i) - b\Sigma_{i\geq m} P(i) + b(n+1-m)}{\frac{a(m-1)}{2} + \frac{b(n-m)}{2}}; \; a, b \geq .$$

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.

The following index is consistent with $mAF2$ relation.

$$\mathcal{P}_{mAF2}(\mathbb{p}) = \frac{\Sigma_{\mathbb{i}\prec\mathbb{m}} \mathbb{P}(\mathbb{i}) + \Sigma_{\mathbb{i}\succeq\mathbb{m}} \bar{\mathbb{P}}(\mathbb{i})}{\frac{\#\{\mathbb{i}:\mathbb{i}\prec\mathbb{m}\} + \#\{\mathbb{i}:\mathbb{i}\succ\mathbb{m}\}}{2}} \quad (8)$$

The index sums cumulative distribution function below the median and the survival function above the median and is normalized to ensure that the index is between 0 and 1.

Before we introduce measures consistent with *somAF1* and *somAF2* we propose a set of axioms with which these measures should be consistent. Then we prove their consistency.

CON $\mathcal{P} : \lambda \to \mathbb{R}$ is a continuous function.

NORM The range of $\mathcal{P}$ ($Ran(\mathcal{P})$) is the closed interval $[0, 1]$.

DECOMP There exist $f : Ran(\mathcal{P}) \times Ran(\mathcal{P}) \times (0, 1) \to \mathbb{R}$ continuous and strictly increasing with respect to the first two coordinates such that for any $\mathbb{p}_1, \mathbb{p}_2 \in \lambda$, $\alpha \in (0, 1)$

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

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 $\mathbb{p}_1(\mathbb{i})$ to category $\mathbb{i}$ and $\mathbb{p}_2$ assigns mass $\mathbb{p}_2(\mathbb{i})$, then the probability mass attributed to $\mathbb{i}$ in $\alpha\mathbb{p}_1 + (1-\alpha)\mathbb{p}_2$ is $\alpha\mathbb{p}_1(\mathbb{i}) + (1-\alpha)\mathbb{p}_2(\mathbb{i})$.

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

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

where

$$\mathcal{P}(p^j) = \mathcal{P}(\underbrace{p^j \otimes p^j \otimes \cdots \otimes p^j}_{k \text{ times}})$$

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

ADDSEP Let $\mathcal{P}$ fulfill ATTRDECOMP. Thus $\mathcal{P}(\mathbb{p}) = f(\mathcal{P}(p^1), \mathcal{P}(p^2), \dots, \mathcal{P}(p^k))$. ADDSEP further requires that function f has the form $f(x) = \sum_{j=1}^{k} f_j(x_j)$.

EQUALsomAF1 Let $\mathcal{P}$ be consistent with $somAF1$ relation i.e. $\mathbb{p}_1 \prec_{somAF1} \mathbb{p}_2 \implies \mathcal{P}(\mathbb{p}_1) < \mathcal{P}(\mathbb{p}_2)$.

EQUALsomAF2 Let $\mathcal{P}$ be consistent with $somAF2$ relation i.e. $\mathbb{p}_1 \prec_{somAF2} \mathbb{p}_2 \implies \mathcal{P}(\mathbb{p}_1) < \mathcal{P}(\mathbb{p}_2)$.

CATADD Let $\mathbb{p}_2, \mathbb{p}_3$ be such distributions that $p_2^j(1) = 0$, $\mathbb{p}_2((i_1, \dots, i_j + 1, \dots, i_k)) = \mathbb{p}_1((i_1, \dots, i_j, \dots, i_k))$ for $1 \le i_j \le n_j$, $p_3^j(n_j + 1) = 0$, $\mathbb{p}_3((i_1, \dots, i_j, \dots, i_k)) = \mathbb{p}_1((i_1, \dots, i_j, \dots, i_k))$ for $1 \le i_j \le n_j$ and let $\mathbb{q}_1, \mathbb{q}_2, \mathbb{q}_3$ be obtained in the same way then $\mathcal{P}(\mathbb{p}_1) \le \mathcal{P}(\mathbb{q}_1) \iff \mathcal{P}(\mathbb{p}_2) \le \mathcal{P}(\mathbb{q}_2) \iff \mathcal{P}(\mathbb{p}_3) \le \mathcal{P}(\mathbb{q}_3)$.

SLIDE Let $\mathbb{p}_1$ be such distribution that $p_1^j(1) = 0$, let $\mathbb{p}_2((i_1, \dots, i_j, \dots, i_k)) = \mathbb{p}_1((i_1, \dots, i_j + 1, \dots, i_k))$ for $i_j < n_j$ and $p_2^j(n_j) = 0$, then $\mathcal{P}(\mathbb{p}_1) \le (\ge)\mathcal{P}(\mathbb{q}) \iff \mathcal{P}(\mathbb{p}_2) \le (\ge)\mathcal{P}(\mathbb{q})$.

CON is a natural technical assumption. NORM means that the index achieves lowest value (zero) for the most equal distribution (i.e. all mass in one category) and highest value (one) for the most unequal distribution (i.e. most bi-polarized distribution). DECOMP means that the index is decomposable by population subgroups, namely, that it is a function of the weighted mean of the value of indices in subgroups, with weights corresponding to subgroups' population size (Shorrocks 1984, Kobus and Milos 2012). ATTRDECOMP, on the other hand, means that the index is decomposable into unidimensional indices on each dimension. The notion originated from the contribution by (Abul Naga and Geoffard 2006) and allows for the evaluation of each dimension's contribution to overall inequality. (Zhong 2009) applies this decomposition in the health – income context. In addition, ADDSEP states that the index is an additive function of unidimensional indices. Two EQUAL axioms ensure consistency with, respectively, Definitions 3 and 4.

The next two axioms involve operations that apply to a single dimension and such that they preserve the original dominance relations on (multidimensional) distributions. CATADD involves an operation which adds empty category to one dimension either below the lowest category or above the highest category, whereas SLIDE moves probability mass to empty categories so the medians of two distributions agree (provided there are enough empty categories). CATADD and SLIDE have been recently introduced by (Sarkar and Santra 2018) who show that these axioms allow to compare distributions with different medians through AF criterion. Please note that a significant restriction in Definition 1 is that there has to be a common median. $mAF1$ and $mAF2$ can be extended by incorporating SLIDE, namely, by moving a whole distribution along a coordinate if $\mathbb{m}_2 \succ \mathbb{m}_1$, $\mathbb{h}^1 + \mathbb{h}^2 \succeq \mathbb{m}_2 - \mathbb{m}_1$, where $\mathbb{h}_j^1 = \sup\{i \le n_j | p_1^j(i) \ne 0\}$ and $\mathbb{h}_j^2 = \inf\{i \ge 1 | p^j(i)_2 \ne 0\}$ i.e. the number of empty categories on respective sides of distributions must be greater than the difference between medians. Here empty categories must exist, one cannot add them like in CATADD because it potentially removes mAF-type of dominance. It is a technical issue, but in case of $somAF$, where the summation of survival function starts from the median adding zero category at the end does not matter. Therefore, for $somAF1$ and $somAF2$, both CATADD and SLIDE can be used, and any two distributions with differing medians can be compared.

We now introduce two indices consistent with, respectively, $somAF1$ and $somAF2$ and show that they have the desired properties.

Lemma 1. *The following index*

$$\mathcal{P}^{somAF1}_{a,b,c}(\mathbb{p}) = \frac{c_1\mathcal{P}^{so}_{a_1,b_1}(\mathbb{p}^1) + c_2\mathcal{P}^{so}_{a_2,b_2}(\mathbb{p}^2) + \cdots + c_k\mathcal{P}^{so}_{a_k,b_k}(\mathbb{p}^k)}{\Sigma_{i=1}^k c_i}, \tag{9}$$

where $\mathcal{P}^{so}_{a_i,b_i}(p)$ *is*

$$\mathcal{P}^{so}_{a,b}(p) = \frac{a(n-m)\Sigma_{i<m}P(i) + a\Sigma_{i<m}CS(P,i,m) + b(n+m)\Sigma_{i\geq m}\bar{P}(i) - b\Sigma_{i\geq m}CS(\bar{P},i,m)}{\sup_m\{\frac{a(m-1)(2n-m)}{4} + \frac{b(n-m)(n+m+1)}{4}\}}. \tag{10}$$

fulfills CON, NORM, DECOMP, ATTRDECOMP, EQUALsomAF1, ADDSEP and SLIDE. It fulfills CATADD if $n = \inf\{i|P(i) = 1\} - \sup\{i \geq 0|P(i) = 0\}$.

Proof. The fact that $\mathcal{P}^{somAF2}_{a,b,c}$ fulfills CON and DECOMP is obvious since index is a linear function. It is also straightforward to check that it fulfills ATTRDECOMP. From the fact that $\mathcal{P}^{so}_{a,b}$ is increasing with $CS(P)$ we can conclude that it fulfills EQUALsomAF1. Let us check NORM axiom. Let p and q be, respectively, the best and the worst distributions according to AF, that is q has all probability mass in median category, while p has half minus infinitesimal of the mass in first category, infinitesimal mass in median and half in the last. We have $\mathcal{P}^{so}_{a,b}(q) = \frac{0}{\sup_m\{\frac{a(m-1)(2n-m)}{4}+\frac{b(n-m)(n+m+1)}{4}\}} = 0$ and $\sup_m \mathcal{P}^{so}_{a,b}(p) = \sup_m \frac{a\Sigma_{i<m}(n-m+i)\frac{1}{2}+b\Sigma_{i\geq m}(n+m-i)\frac{1}{2}-\frac{bm}{2}}{\sup_m\{\frac{a(m-1)(2n-m)}{4}+\frac{b(n-m)(n+m+1)}{4}\}} = \frac{\sup_m\{\frac{a(m-1)(2n-m)}{4}+\frac{b(n-m)(n+m+1)}{4}\}}{\sup_m\{\frac{a(m-1)(2n-m)}{4}+\frac{b(n-m)(n+m+1)}{4}\}} = 1$. Since $\mathcal{P}^{somAF1}_{a,b,c}$ is weighted mean and it is also normalized. Let us now take distributions p and q such that $p(1) = q(n) = 0$ and $p(i+1) = q(i)$. We have that $P(i+1) = Q(i)$ and $\bar{P}(i+1) = \bar{Q}(i)$ and we obtain

$$\begin{aligned}
\mathcal{P}^{so}_{a,b}(p) &= \frac{a\Sigma_{i<m}(n-m+i)P(i)+b\Sigma_{i\geq m}(n+m-i)\bar{P}(i)}{\sup_m\{\frac{a(m-1)(2n-m)}{4}+\frac{b(n-m)(n+m+1)}{4}\}} = \\
&= \frac{a\Sigma_{i+1<m}(n-m+i+1)P(i+1)+b\Sigma_{i+1\geq m}(n+m-i-1)\bar{P}(i+1)}{\sup_m\{\frac{a(m-1)(2n-m)}{4}+\frac{b(n-m)(n+m+1)}{4}\}} = \\
&= \frac{a\Sigma_{i<m-1}(n-(m-1)+i)Q(i)+b\Sigma_{i\geq m-1}(n+(m-1)-i)\bar{Q}(i)}{\sup_m\{\frac{a(m-1)(2n-m)}{4}+\frac{b(n-m)(n+m+1)}{4}\}} = \mathcal{P}^{so}_{a,b}(q).
\end{aligned}$$

□

$\mathcal{P}^{somAF1}_{a,b,c}$ is a weighted mean of unidimensional indices $\mathcal{P}^{so}_{a,b}(p)$ that agree with second-order AF (Definition 2). It is easier to interpret this index re-writing it in the following form.

Corollary 1. *We can also write* $\mathcal{P}^{so}_{a,b}(p)$ *as*

$$\mathcal{P}^{so}_{a,b}(p) = \frac{a\Sigma_{i<m}(n-m+i)P(i) + b\Sigma_{i\geq m}(n+m-i)\bar{P}(i)}{\sup_m\{\frac{a(m-1)(2n-m)}{4} + \frac{b(n-m)(n+m+1)}{4}\}}. \tag{11}$$

$\mathcal{P}^{so}_{a,b}(p)$ combines cdfs and survival functions with appropriate weights. Combining SLIDE with NORM posed the biggest challenge in deriving these weights. The most common procedure in fulfilling NORM is subtracting the lowest value and dividing it by the range of an index. The denominator then depends on the weights. SLIDE requires weights at $P(i)$ to be of the form $f(i-m)$. Then the denominator also depends on m and SLIDE changes m, which may potentially reverse the direction of the ranking of distributions implied by indices. To solve this problem we use a global norm of the index, that is, we take a supremum of the values of indices for all medians and get rid of dependence on m in the denominator. CATADD requires that weights do not depend on the total number of categories but they depend on the number of categories between the first non-empty category and the last non-empty category (which is defined formally in Lemma 1. Finally, to fulfill EQUALsomAF1 we need weights to be increasing/decreasing and nonnegative.

Lemma 2. *The following index*

$$\mathcal{P}^{somAF2}_{a,b}(\mathbb{p}) = \frac{\Sigma_{\mathbb{i}\prec\mathbb{m}} a_{\mathbb{i}}\mathbb{P}(\mathbb{i}) + \Sigma_{\mathbb{i}\succeq\mathbb{m}} b_{\mathbb{i}}\bar{\mathbb{P}}(\mathbb{i})}{C} \tag{12}$$

where $a_{\mathbb{i}} = a\Pi^k_{j=1}(n_j - m_j + i_j)$*,* $b_{\mathbb{i}} = b\Pi^k_{j=1}(n_j + m_j - i_j)$*,* $C = \sup_{\mathbb{m}}\{\frac{1}{2}(a\Sigma_{\mathbb{i}\prec\mathbb{m}}\Pi^k_{j=1}(n_j - m_j + i_j) + b\Sigma_{\mathbb{m}\preceq\mathbb{i}\prec\mathbb{n}}\Pi^k_{j=1}(n_j + m_j - i_j))\}$ *fulfills CON, NORM, DECOMP, SLIDE, and EQUALsomAF2. It also fulfills CATADD (with respect to attribute j) if* $n_j = \inf\{i|\bar{P}^j(i) = 0\} - \sup\{i \geq 0|P^j(i) = 0\}$*.*

Proof. The fact that $\mathcal{P}^{somAF2}_{a,b}$ fulfills CON and DECOMP is obvious since it's linear function. From the fact that $\mathcal{P}^{somAF2}_{a,b}$ is increasing with $CS(\mathbb{P})$ (it follows from the fact that weights at $\mathbb{P}(\mathbb{i})$ are increasing and at $\bar{\mathbb{P}}(\mathbb{i})$ they are decreasing) we can conclude that it fulfills EQUALsomAF2. Let us check NORM axiom. Let $\mathbb{p}$ and $\mathbb{q}$ be the best and the worst distributions according to AF, that is $\mathbb{q}$ has all probability mass in median category, while $\mathbb{p}$ has half minus infinitesimal of the mass in first category, infinitesimal mass in median and half of the mass in the last (according to $\preceq$). We have $\mathcal{P}^{somAF2}_{a,b}(\mathbb{q}) = \frac{0}{C} = 0$ and $\sup_{\mathbb{m}} \mathcal{P}^{somAF2}_{a,b}(\mathbb{p}) = \frac{\sup_{\mathbb{m}}\{\frac{1}{2}(a\Sigma_{\mathbb{i}\prec\mathbb{m}}\Pi^k_{j=1}(n_j - m_j + i_j) + b\Sigma_{\mathbb{m}\preceq\mathbb{i}\prec\mathbb{n}}\Pi^k_{j=1}(n_j + m_j - i_j))\}}{C} = \frac{C}{C} = 1$. We note that by $\mathbb{i}+1 = (i_1, \ldots, i_j + 1, \ldots, i_k)$. Let us now take distributions $\mathbb{p}, \mathbb{q}$ such that $p^j(1) = 0$, $\mathbb{q}_{\mathbb{i}} = \mathbb{p}_{\mathbb{i}+1}$ and $q^j(n_j) = 0$. We have that $a^{\mathbb{p}}_{\mathbb{i}+1} = a(n_j - m_j + i_j + 1)\Pi^k_{l\neq j}(n_l - m_l + i_l) = a(n_j - (m_j - 1) + i_j)\Pi^k_{l\neq j}(n_l - m_l + i_l) = a^{\mathbb{q}}_{\mathbb{i}}$, $b^{\mathbb{p}}_{\mathbb{i}+1} = b(n_j + m_j - i_j - 1)\Pi^k_{l\neq j}(n_l + m_l - i_l) = b(n_j + (m_j - 1) - i_j)\Pi^k_{l\neq j}(n_l + m_l - i_l) = b^{\mathbb{q}}_{\mathbb{i}}$ and finally

$$\begin{aligned}
\mathcal{P}^{somAF2}_{a,b}(\mathbb{p}) &= \frac{\Sigma_{\mathbb{i}\prec\mathbb{m}} a^{\mathbb{p}}_{\mathbb{i}}\mathbb{P}(\mathbb{i}) + \Sigma_{\mathbb{i}\succeq\mathbb{m}} b^{\mathbb{p}}_{\mathbb{i}}\bar{\mathbb{P}}(\mathbb{i})}{C} = \\
&= \frac{\Sigma_{\mathbb{i}+1\prec\mathbb{m}} a^{\mathbb{p}}_{\mathbb{i}+1}\mathbb{P}(\mathbb{i}+1) + \Sigma_{\mathbb{i}+1\succeq\mathbb{m}} b^{\mathbb{p}}_{\mathbb{i}+1}\bar{\mathbb{P}}(\mathbb{i}+1)}{C} = \\
&= \frac{\Sigma_{\mathbb{i}\prec\mathbb{m}-1} a^{\mathbb{q}}_{\mathbb{i}}\mathbb{Q}(\mathbb{i}) + \Sigma_{\mathbb{i}\succeq\mathbb{m}-1} b^{\mathbb{q}}_{\mathbb{i}}\bar{\mathbb{Q}}(\mathbb{i})}{C} = \mathcal{P}^{somAF2}_{a,b}(\mathbb{q})
\end{aligned}$$

□

$\mathcal{P}^{somAF2}_{a,b}$ sums multidimensional cdfs and survival functions with weights being a multiplication of weights that would hold on each dimension if it were a unidimensional index consistent with second-order unidimensional AF or, equivalently ChM.

5 Statistical inference and Estimation

In this section we present how to estimate three multidimensional polarization indices introduced by (Kobus and Kurek 2018) for consistency with $mAF1$ and $mAF2$ i.e. $\mathcal{P}_{\alpha,\,\beta,\gamma}$(6), $\mathcal{P}_{a,\,b,\,c}$ (7), $\mathcal{P}_{mAF2}$ (8), and two indices introduced here for consistency with $somAF1$ and $somAF2$, namely $\mathcal{P}^{somAF1}_{a,b,c}$ (9) and $\mathcal{P}^{somAF2}_{a,b}$ (12). We establish their large sample distribution. We consider the following setting.

Each individual chooses one and only one answer in every category, which gives a unique response in the set of all states $\mathbb{I}$. We will assume that there are N independent responses $(N_{1\ldots11}, N_{1\ldots12}, \ldots, N_{n_1\ldots n_{k-1}n_k})$ defining the vector of frequencies $\hat{\mathbb{p}}$. In such a model, $\mathbb{I}$ has joint multinomial distribution with parameters $p = (p_{1\ldots11}, p_{1\ldots12}, \ldots, p_{n_1\ldots n_{k-1}n_k})$ and N. Empirical probability distribution is the following

$$\hat{p}_{i_1i_2\ldots i_k} = \frac{1}{N}\Sigma^N_{t=1}\mathbf{1}(I^t_1 = i_1, I^t_2 = i_2, \ldots, I^t_k = i_k) = \frac{N_{i_1i_2\ldots i_k}}{N},$$

In (Formby et al. 2004) it is shown that multinomial distributions are asymptotically distributed as normal with mean p and covariance matrix Σ, where (for $\mathbb{I} = \{1, 2, \ldots, n_1\} \times$

$\{1,2,\ldots,n_2\} \times \cdots \times \{1,2,\ldots,n_k\}$) we assume the following notation

$$p = \begin{pmatrix} p_{1\ldots1} \\ p_{1\ldots12} \\ \vdots \\ p_{n_1\ldots n_{k-1}n_k} \end{pmatrix},$$

$$\Sigma = \begin{pmatrix} p_{1\ldots11}(1-p_{1\ldots11}) & -p_{1\ldots11}p_{1\ldots12} & \cdots & -p_{1\ldots11}p_{n_1\ldots n_{k-1}n_k} \\ -p_{1\ldots12}p_{1\ldots11} & p_{1\ldots12}(1-p_{1\ldots12}) & \cdots & -p_{1\ldots12}p_{n_1\ldots n_{k-1}n_k} \\ \vdots & \vdots & \ddots & \vdots \\ -p_{n_1\ldots n_{k-1}n_k}p_{1\ldots11} & -p_{n_1\ldots n_{k-1}n_k}p_{1\ldots12} & \cdots & p_{n_1\ldots n_{k-1}n_k}(1-p_{n_1\ldots n_{k-1}n_k}) \end{pmatrix}.$$

That is, lth row/column corresponds to $(i_1, i_2, \ldots i_k)$ category, where $l = i_k + \Sigma_{j=1}^{k-1}(i_j - 1)n_{j+1}n_{j+2}\ldots n_k$. We have $i_k = l \bmod n_k + n_k \mathbf{1}_{n_k|l}$, $l_k = \frac{l-i_k}{n_k}$ and $i_j - 1 = l_{j+1} \bmod n_j$, $l_j = \frac{l_{j+1}-(i_j-1)}{n_j}$ for $1 \le j < k$. For example, for $k = 3$ and $(n_1, n_2, n_3) = (6, 5, 8)$, the 28th category is $28 = 4 + 3 \times 8$ i.e. $(0+1, 3+1, 4) = (1, 4, 4)$, and the 134th category is $134 = 6 + 16 \times 8 = 6 + 1 \times 8 + 3 \times (5 \times 8)$ i.e. $(3+1, 1+1, 6) = (4, 2, 6)$.

The multivariate, multinomial distributions of two samples can be compared by stacking the probabilities in respective probability vectors

$$\hat{\mathcal{E}} = \hat{\mathbb{p}} - \hat{\mathbb{q}} \tag{13}$$

Under the null hypothesis of homogenous distributions,

$$\hat{\mathcal{E}} \xrightarrow{d} \mathcal{N}\left(0, \frac{N_{\mathbb{p}} + N_{\mathbb{q}}}{N_{\mathbb{p}}N_{\mathbb{q}}}\Sigma\right). \tag{14}$$

Using Continuous Mapping Theorem we get that under appropriate assumptions $F(\hat{\mathbb{P}})$ will be a consistent estimator of $F(\mathbb{P})$. Let us define $n_1n_2\ldots n_k$-dimensional Jacobian (horizontal) vector of the transformation F,

$$J = \left(\frac{\partial F}{\partial \mathbb{P}_{1\ldots11}}, \frac{\partial F}{\partial \mathbb{P}_{1\ldots12}}, \ldots \frac{\partial F}{\partial \mathbb{P}_{n_1\ldots n_{k-1}n_k}}\right).$$

N independent responses $(N_{1\ldots11}, N_{1\ldots12}, \ldots, N_{n_1\ldots n_{k-1}n_k})$ are jointly distributed from a multinomial distribution with parameters N and p and such that $\mathrm{cov}(\sqrt{n}\hat{\mathbb{p}}) = \Sigma(\mathbb{p})$, where $\Sigma(\mathbb{p})$ is a finite positive semi-definite matrix.

Theorem 1. *For multinomial distribution $\mathbb{P}$ and function F which is continuously differentiable at $\mathbb{P}$ and does not involve total number of observations N, we have that*

$$\sqrt{n}(F(\hat{\mathbb{P}}) - F(\mathbb{P})) \xrightarrow{d} \mathcal{N}(0, JL\Sigma L^T J^T)$$

where $J = J(\mathbb{P})$ is Jacobian vector evaluated at $\mathbb{P}$, and $L = (l_{ij})_{1 \le i,j \le n_1n_2\ldots n_k}$ is such that $l_{ij} = 1$ if $j \equiv (j_1, j_2 \ldots j_k) \preceq (i_1, i_2, \ldots i_k) \equiv i$ and 0 otherwise where $\preceq$ is lexicographical order.[6]

[6] The result is unchanged when $\mathbb{m}$ is replaced by τ.

Proof. For $k = 2$ and $(n_1, n_2) = (3, 3)$, L is the following

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

In general case let L^t be square matrix of dimension t satisfying $l^t_{ij} = 1$ if $i \geq j$ and $l^t_{ij} = 0$ otherwise. We define L recursively by setting $L^{n_{k-t}}_{t+1} = L^{n_{k-t}} \otimes L^{n_{k-t+1}}_t$, where $\otimes$ denotes Kronecker product (for any $a \times b$ matrix A and $c \times d$ matrix B, $A \otimes B$ is $ac \times bd$ matrix C such that $c_{(b(i-1)+k)(d(j-1)+l)} = a_{ij} b_{kl}$). Next we let $L^{n_k}_1 = L^{n_k}$ and finally $L = L^{n_1}_k$.

From Central Limit Theorem we have that $\sqrt{n}(\hat{\mathbb{p}} - \mathbb{p}) \xrightarrow{d} \mathcal{N}(0, \Sigma)$. It is straightforward to check that for

$$\hat{\mathbb{P}} = L\hat{\mathbb{p}}, \mathbb{P} = L\mathbb{p}$$

we have $\sqrt{n}(\hat{\mathbb{P}} - \mathbb{P}) \xrightarrow{d} \mathcal{N}(0, L\Sigma L^T)$. From Law of Large Numbers $\hat{\mathbb{P}}$ converges in probability to $\mathbb{P}$ and from Continuous Mapping Theorem we see that $J(\hat{\mathbb{P}})$ converges in probability to $J(\mathbb{P})$ (due to continuous differentiability). Finally, from the delta method we obtain

$$\sqrt{n}(F(\hat{\mathbb{P}}) - F(\mathbb{P})) \xrightarrow{d} \mathcal{N}(0, JL\Sigma L^T J^T).$$

□

A consistent estimator for Σ is

$$\hat{\sigma}_{(i_1 i_2 \ldots i_k)(i_1 i_2 \ldots i_k)} = \frac{N_{i_1 i_2 \ldots i_k}}{N}(1 - \frac{N_{i_1 i_2 \ldots i_k}}{N})$$

and

$$\hat{\sigma}_{(i_1 i_2 \ldots i_k)(j_1 j_2 \ldots j_k)} = -\frac{N_{i_1 i_2 \ldots i_k} N_{j_1 j_2 \ldots j_k}}{N^2} \text{ if } i_t \neq j_t \text{ for some } 1 \leq t \leq k.$$

It follows that standard error σ is equal to

$$\sigma = \sqrt{\frac{J(\hat{\mathbb{P}}) L \hat{\Sigma} L^T J(\hat{\mathbb{P}})^T}{n}}.$$

We would get similar result if we substituted matrix L by any matrix M with constant coefficients. For our indices using $\bar{\mathbb{P}}$ it will be useful to present following remark.

Remark 1. *For multinomial distribution $\mathbb{P}$, with survival function $\bar{\mathbb{P}}$ and function F which is continuously differentiable at $\mathbb{P}$ (which implies that it is also differentiable at $\bar{\mathbb{P}}$) and does not involve total number of observations N, we have that*

$$\sqrt{n}(F(\hat{\mathbb{P}}, \hat{\bar{\mathbb{P}}}) - F(\mathbb{P}, \bar{\mathbb{P}})) \xrightarrow{d} \mathcal{N}(0, JM\Sigma M^T J^T)$$

where $J = J(\mathbb{P}, \bar{\mathbb{P}})$ is Jacobian vector evaluated at $\mathbb{P}, \bar{\mathbb{P}}$, $L = (l_{ij})_{1 \leq i,j \leq n_1 n_2 \ldots n_k}$ is such that $l_{ij} = 1$ if $j \equiv (j_1, j_2 \ldots j_k) \preceq (i_1, i_2, \ldots i_k) \equiv i$ and 0 otherwise, $U = (u_{ij})_{1 \leq i,j \leq n_1 n_2 \ldots n_k}$ is such that $l_{ij} = 1$ if $j \equiv (j_1, j_2 \ldots j_k) \succeq (i_1, i_2, \ldots i_k) \equiv i$and 0 otherwise where $\preceq$ is lexicographical order and $M = \begin{pmatrix} L \\ U - I \end{pmatrix}$. [7]

The Jacobian vector for $F = \mathcal{P}_{a,b}$ is

$$J = \frac{2}{a(m-1) + b(n-m)}(a_{(1)}, a_{(2)}, \ldots, a_{(m-1)}, -b_{(m)}, -b_{(m+1)}, \ldots, -b_{(n)})$$

[7] The result is unchanged when $\mathbb{m}$ is replaced by τ.

and for $F=\mathcal{P}_{a,b,c}$ it is

$$J=(j_l)_{1\le l\le n_1n_2\dots n_k},$$

where

$$j_{n_1n_2\dots n_{t-1}i_tn_{t+1}\dots n_k}=\frac{c_t}{\Sigma_{i=1}^k c_i}\frac{2}{a_t(m_t-1)+b_t(n_t-m_t)}\left(a_t\mathbf{1}_{i_t<m_t}-b_t\mathbf{1}_{i_t\ge m_t}\right)$$

and is equal to 0 for all other values of j_l. Similarly, the Jacobian vector for $F=\mathcal{P}_{\alpha,\beta}$ is

$$J=C(\alpha\mathbb{P}_1^{\alpha-1},\alpha\mathbb{P}_2^{\alpha-1},\dots,\alpha\mathbb{P}_{m-1}^{\alpha-1},-\beta\mathbb{P}_m^{\beta-1},-\beta\mathbb{P}_{m+1}^{\beta-1}\dots,-\beta\mathbb{P}_n^{\beta-1}),$$

where

$$C=\frac{1}{(m-1)(\frac{1}{2})^\alpha-(n-m)(\frac{1}{2})^\beta+n-m},$$

and for $F=\mathcal{P}_{\alpha,\beta,\gamma}$ it is

$$J=(j_l)_{1\le l\le n_1n_2\dots n_k},$$

where

$$j_{n_1n_2\dots n_{t-1}i_tn_{t+1}\dots n_k}=D_t\left(\alpha_t\mathbb{P}^{\alpha_t-1}_{n_1n_2\dots n_{t-1}i_tn_{t+1}\dots n_k}\mathbf{1}_{i_t<m_t}-\beta_t\mathbb{P}^{\beta_t-1}_{n_1n_2\dots n_{t-1}i_tn_{t+1}\dots n_k}\mathbf{1}_{i_t\ge m_t}\right)$$

and $D_t=\frac{1}{k}C_t(\mathcal{P}_{\alpha,\beta,\gamma}(\mathbb{P}))^{1-\gamma}(\mathcal{P}_{\alpha_t,\beta_t}(\mathbb{P}^t))^{\gamma-1}$ and is equal to 0 for all other values of j_l.

Jacobian vector for $F=\mathcal{P}_{mAF2}$ is vector J such that $J_i=\frac{2}{\#\{\mathbb{i}:\mathbb{i}\prec\mathbb{m}\}+\#\{\mathbb{i}:\mathbb{i}\succ\mathbb{m}\}}$ for $i\equiv\mathbb{i}\prec\mathbb{m}$ or $i-n_1n_2\dots n_k\equiv\mathbb{h}\succeq\mathbb{m}$ and $J_i=0$ otherwise.

Similarly for $F=\mathcal{P}^{somAF2}_{a,b}$ Jacobian vector is given by

$$J_i=\frac{a_{\mathbb{i}}}{C}$$

for $i\equiv\mathbb{i}\prec\mathbb{m}$,

$$J_i=\frac{b_{\mathbb{i}}}{C}$$

for $i-n_1n_2\dots n_k\equiv\mathbb{h}\succeq\mathbb{m}$ and $J_i=0$ otherwise, where $a_{\mathbb{i}},b_{\mathbb{i}},C$ are defined as in Lemma 2.

Finally for $F=\mathcal{P}^{somAF1}_{a,b,c}$ Jacobian vector is given by

$$J=(j_l)_{1\le l\le n_1n_2\dots n_k},$$

where

$$j_{n_1n_2\dots n_{t-1}i_tn_{t+1}\dots n_k}=\frac{c_t}{\Sigma_{i=1}^k c_i}\frac{4}{b_t(n+2)(n-1)}\left(a_t(n_t-m_t+i_t)\mathbf{1}_{i_t<m_t}-b_t(n_t+m_t-i_t)\mathbf{1}_{i_t\ge m_t}\right)$$

and is equal to 0 for all other values of j_l

6 Generalization to any quantile

Median, as a half-point of distribution is definitely most interesting of the quantiles, but there may be other points of interest, for example quartiles. mAF and $somAF$ relations can be generalised to account for this. First such generalisation of unidimensional AF was in fact proposed by Mendelson in the 1980s (Mendelson 1987) but was largely overlooked by the literature. (Kobus and Kurek 2018) apply his extension to $mAF1$ and $mAF2$ (Section 6 in their paper). In fact, such generalisation makes even more sense for second-order relations because there we measure not only concentration around the median, but also around some other categories below and above the median. To this end, we note that in the definition of

$CS(\mathbb{P}, \mathbb{i}, \mathbb{m})$ a vector of medians can be replaced by a vector of arbitrary quantiles and the definition remains valid.

Definition 7. ***Quantile somAF1 (somqAF1)***
Let $\mathbb{p}_1, \mathbb{p}_2$ be two probability distributions with a unique and common τ-quantile $\mathbb{q}$ for $\tau = (\tau_1, \tau_2, \dots, \tau_k)$, that is $\mathbb{q}_1 = \mathbb{q}_2$ where $(\mathbb{q}_1)_i = \inf\{j|\mathbb{P}_1(j) \geq \tau_i\}$ and $(\mathbb{q}_2)_i = \inf\{j|\mathbb{P}_2(j) \geq \tau_i\}$. We say that $\mathbb{p}_1 \precsim_{somqAF1} \mathbb{p}_2$ if and only if the following two conditions hold

(1) $CS(P_1^j, i, q_j) \leq CS(P_2^j, i, q_j)$ for $i < q_j$ and for all j,

(2) $CS(\bar{P}_1^j, i, q_j) \leq CS(\bar{P}_2^j, i, q_j)$ for $i \geq q_j$ and for all j, where $\bar{P}_1^j, \bar{P}_2^j$ denote survival functions of jth marginals of $\mathbb{p}_1$ and $\mathbb{p}_2$, respectively.

Definition 8. ***Quantile somAF2 (somqAF2)***
Let $\mathbb{p}_1, \mathbb{p}_2$ be two probability distributions with a unique and common τ-quantile $\mathbb{q}$. We say that $\mathbb{p}_1 \precsim_{somqAF2} \mathbb{p}_2$ if and only if the following two conditions hold

(1) $CS(\mathbb{P}_1, \mathbb{i}, \mathbb{q}) \leq CS(\mathbb{P}_2, \mathbb{i}, \mathbb{q})$ for $\mathbb{i} \prec \mathbb{q}$

(2) $CS(\bar{\mathbb{P}}_1, \mathbb{i}, \mathbb{q}) \leq CS(\bar{\mathbb{P}}_2, \mathbb{i}, \mathbb{q})$ for $\mathbb{i} \succeq \mathbb{q}$ where $\bar{\mathbb{P}}_1, \bar{\mathbb{P}}_2$ denote survival functions of $\mathbb{p}_1$ and $\mathbb{p}_2$, respectively.

Please note that this is different from extension brought about by CATADD and SLIDE. In order for two distributions to be compared according to *somqAF*1 or *somqAF*2 they have to have the same quantiles in $CS(\mathbb{P}, \mathbb{i}, \mathbb{q})$ and CATADD and SLIDE ensure that we can do this. In this section what we do is to say that the choice of $\mathbb{q}$ can be any, not necessarily the median.

The following indices are consistent with Definitions 7 and 8.

Remark 2. *The following index*

$$\mathcal{P}_{a,b,c}^{somqAF1}(\mathbb{p}, \mathbb{q}) = \frac{c_1 \mathcal{P}_{a_1,b_1}^{so}(p^1, q_1) + c_2 \mathcal{P}_{a_2,b_2}^{so}(p^2, q_2) + \cdots + c_k \mathcal{P}_{a_k,b_k}^{so}(p^k, q_k)}{\Sigma_{i=1}^k c_i}, \tag{15}$$

where $\mathcal{P}_{a_i,b_i}^{so}(p, q)$ is

$$\mathcal{P}_{a,b}^{soq}(p, q) = \frac{a(n-q)\Sigma_{i<q} P(i) + a\Sigma_{i<q} CS(P, i, q) + b(n+q)\Sigma_{i\geq q}\bar{P}(i) - b\Sigma_{i\geq q} CS(\bar{P}, i, q)}{\sup_q\{\tau \frac{a(q-1)(2n-q)}{2} + (1-\tau)\frac{b(n-q)(n+q+1)}{2}\}}.$$

We can also write

$$\mathcal{P}_{a,b}^{soq}(p, q) = \frac{a\Sigma_{i<q}(n-q+i)P(i) + b\Sigma_{i\geq q}(n+q-i)\bar{P}(i)}{\sup_q\{\tau \frac{a(q-1)(2n-q)}{2} + (1-\tau)\frac{b(n-q)(n+q+1)}{2}\}}.$$

fulfills CON, NORM, DECOMP, ATTRDECOMP, EQUALsomqAF1 (i.e its quantile version), ADDSEP and SLIDE. It fulfills CATADD if $n = \inf\{i|P(i) = 1\} - \sup\{i \geq 0|P(i) = 0\}$.

Remark 3. *The following index*

$$\mathcal{P}_{a,b}^{somqAF2}(\mathbb{p}, \mathbb{q}) = \frac{\Sigma_{\mathbb{i}\prec\mathbb{q}} a_{\mathbb{i}} \mathbb{P}(\mathbb{i}) + \Sigma_{\mathbb{i}\succeq\mathbb{q}} b_{\mathbb{i}} \bar{\mathbb{P}}(\mathbb{i})}{C} \tag{16}$$

where $a_{\mathbb{i}} = \Pi_{j=1}^k(n_j - q_j + i_j)$, $b_{\mathbb{i}} = \Pi_{j=1}^k(n_j + q_j - i_j)$, $C = \sup_{\mathbb{q}}\{\min \tau_j \Sigma_{\mathbb{i}\prec\mathbb{q}} \Pi_{j=1}^k(n_j - q_j + i_j) + (1 - \max \tau_j)(\Sigma_{\mathbb{q}\preceq\mathbb{i}\prec\mathbb{n}} \Pi_{j=1}^k(n_j + q_j - i_j) - 1)\}$ fulfills CON, NORM, DECOMP, SLIDE, and EQUALsomqAF2 (i.e. its quantile version). It also fulfills CATADD (with respect to attribute j) if $n_j = \inf\{i|\bar{P}^j(i) = 0\} - \sup\{i \geq 0|P^j(i) = 0\}$.

7 Empirical application

In this section we illustrate the use of multidimensional polarization relations and measures with an empirical application. The application is the measurement of socio – economic inequalities using data from Survey of Health, Ageing and Retirement in Europe (SHARE).

SHARE is a multi-country representative survey of non-institutionalized individuals aged 50 and and spouses regardless of age. It is a cross-national panel database of micro data on health, socio-economic status and social and family networks. The first wave of SHARE was collected in 2004/5. It is a longitudinal survey with new cohorts of participants being added over time. Individual interviews are conducted approximately every two years. The analysis is based on date from Wave 6 that was collected in year 2015. We measure SES-health inequalities, when SES is proxied by educational attainment and health by self-reported health status. The specific indicators are presented in Table 1.

Table 1: Ordinal indicators of health and education

Dimension	Indicator	Level	Construction
Education	Highest educational level attained	1	Pre-primary education
		2	Primary education
		3	Lower secondary education
		4	(Upper) secondary education
		5	Post-secondary non-tertiary education
		6	First stage of tertiary education
		7	Second stage of tertiary education (research)
Health	Self-reported health status	1	Excellent
		2	Very good
		3	Good
		4	Fair
		5	Poor

Notes: Data come from Wave 6 of SHARE.

18 countries took part in Wave 6. For each pair of countries we look for a bidimensional dominance $somqAF1$ at either $q = .25$, $q = .5$ or $q = .75$. Please note that due to CATADD and SLIDE we can always compare two countries in terms of $somqAF1$. We find 15 cases (out of 171 possible) in which $somqAF1$ dominance holds. This gives around 10% of cases; in fact, there is a bit more because in case there was dominance for, say, $q = .25$ and $q = .5$ we choose a lower quantile. The results are summarized in Table 2. Isreal emerges as the country which is dominant in most cases, that is, it is the most bidimensionally unequal.

Table 2: Results of $B \precsim_{somqAF1} A$

Dominance A-B	q Education	q Health
GRC - AUT	.5	.75
ISR - AUT	.5	.5
ISR - GER	.25	.5
SWE - SWZL	.5	.5
SWE - POR	.25	.5
ISR - SPN	.25	.75
SPN - CZR	.75	.75
FRA - IT	.25	.75
ISR - IT	.25	.75
GRC - LUX	.75	.75
GRC - EST	.5	.5
ISR - BEL	.25	.5
ISR - CZR	.25	.75
ISR - SVN	.5	.75
ISR - EST	.5	.5

Notes: Own calculations based on SHARE data (Wave 6).

Table 3 presents the values of indices consistent with $somqAF1$. The subscripts inform about different weights (i.e. 1, 10) attached to categories below and above a given quantile (Table 2. The country A dominates country B, therefore its value is higher in each case. In case of comparisons such as Sweden – Czech Republic (SWE – CZE), Greece – Luxembourg (GRC – LUX), or Israel – Belgium (ISR – BEL) sensitivity to bottom and top inequality plays no role; the results are more or less the same no matter the weights. For another group of comparisons such as Greece – Estonia (GRC – EST) or Spain – Czech Republic (ESP – CZE) the differences are the smallest when mostly top of the distribution is taken into account and the highest when the bottom is. This suggests that it is mostly in lower categories of health and education where these countries differ in terms of inequality. On the other hand, for Greece – Austria (GRC – AUT), Sweden – Portugal (SWE – PRT) or Israel – Italy (ISR – ITA) a different pattern emerges: the differences are the smallest when the whole distribution is considered with equal weight ($\mathcal{P}^{somqAF1}_{1,1,1}$) and they increase both at the bottom ($\mathcal{P}^{somqAF1}_{10,1,1}$) and at the top ($\mathcal{P}^{somqAF1}_{1,10,1}$). This might suggest that not only the amount of mass away from a given quantile is important, but also its concentration. Overall, we observe highest differences in the last column, which indicates that it is most often inequality in the bottom of the distribution where countries differ.

Table 3: Indices consistent with $\precsim_{somqAF1}$ and standard errors

Dominance	$\mathcal{P}^{somqAF1}_{1,1,1}$		$\mathcal{P}^{somqAF1}_{10,1,1}$		$\mathcal{P}^{somqAF1}_{1,10,1}$	
A-B	$\mathcal{P}_A$ (SE)	$\mathcal{P}_B$ (SE)	$\mathcal{P}_A$ (SE)	$\mathcal{P}_B$ (SE)	$\mathcal{P}_A$ (SE)	$\mathcal{P}_B$ (SE)
GRC-AUT	.56 (.0037)	.44 (.0044)	.62 (.0039)	.42 (.0039)	.47 (.0061)	.40 (.0076)
ISR-AUT	.52 (.0052)	.39 (.0043)	.47 (.0066)	.36 (.0049)	.56 (.0094)	.42 (.0069)
ISR-DEU	.54 (.0052)	.35 (.0037)	.53 (.0073)	.27 (.0040)	.56 (.0082)	.41 (.0053)
SWE-CHE	.47 (.0039)	.34 (.0044)	.50 (.0053)	.37 (.0059)	.44 (.0061)	.31 (.0063)
SWE-PRT	.48 (.0041)	.39 (.0077)	.52 (.0056)	.35 (.0083)	.44 (.0052)	.41 (.0098)
ISR-ESP	.54 (.0057)	.40 (.0033)	.56 (.0065)	.51 (.0044)	.55 (.0101)	.41 (.0060)
ESP-CZE	.50 (.0034)	.37 (.0032)	.51 (.0031)	.34 (.0030)	.52 (.0079)	.45 (.0083)
FRA-ITA	.51 (.0045)	.40 (.0036)	.48 (.0047)	.40 (.0042)	.51 (.0067)	.39 (.0054)
ISR-ITA	.54 (.0057)	.40 (.0036)	.56 (.0065)	.40 (.0042)	.55 (.0101)	.39 (.0054)
GRC-LUX	.56 (.0033)	.48 (.0060)	.57 (.0033)	.48 (.0054)	.51 (.0078)	.45 (.0135)
GRC-EST	.50 (.0035)	.36 (.0034)	.55 (.0047)	.28 (.0032)	.47 (.0053)	.43 (.0055)
ISR-BEL	.54 (.0052)	.42 (.0034)	.53 (.0073)	.40 (.0042)	.56 (.0082)	.43 (.0043)
ISR-CZE	.54 (.0057)	.39 (.0034)	.56 (.0065)	.36 (.0035)	.55 (.0101)	.43 (.0065)
ISR-SVN	.53 (.0053)	.38 (.0035)	.49 (.0057)	.39 (.0034)	.56 (.0110)	.40 (.0073)
ISR-EST	.52 (.0052)	.36 (.0034)	.47 (.0066)	.28 (.0032)	.56 (.0094)	.43 (.0055)

Notes: Own calculations based on SHARE data (Wave 6).

Table 4 shows the results for indices consistent with $somqAF2$ relation (i.e. which takes into account dependence between dimensions). Interestingly, the rankings obtained by $\mathcal{P}^{somqAF2}$ indices coincide with the ones obtained by $\mathcal{P}^{somqAF1}$ indices (Table 3), suggesting that the dimensions may be weakly dependent among these countries. We observe more reversals of this trend with higher weight given to top of the distribution. These are GRC – AUT, SWE – PRT, ESP – CZE, and GRC – EST. Important to note is that the values of $somqAF2$ indices are all significantly lower than the values of $somqAF1$ indices. This indicates that probability mass is distributed in parts of the joint distribution in which unidimensional dominances give conflicting results (please recall our discussion below Definition 6). Univariate indices (and a combination of those such as $\mathcal{P}^{somqAF1}$) attributes all this mass to marginal distributions, whereas from the point of view of *joint* inequality, the presence of mass in these parts lowers inequality. To sum up, dependence between health and education seems to matter for the level of inequality, but not for the ranking of distributions.

Table 4: Indices consistent with $\precsim_{somqAF2}$ and standard errors

Dominance	$\mathcal{P}^{somqAF2}_{1,1}$		$\mathcal{P}^{somqAF2}_{10,1}$		$\mathcal{P}^{somqAF1}_{1,10}$	
A-B	$\mathcal{P}_A$ (SE)	$\mathcal{P}_B$ (SE)	$\mathcal{P}_A$ (SE)	$\mathcal{P}_B$ (SE)	$\mathcal{P}_A$ (SE)	$\mathcal{P}_B$ (SE)
GRC-AUT	.29 (.0051)	.21 (.0040)	.35 (.0070)	.23 (.0048)	.18 (.0058)	.19 (.0073)
ISR-AUT	.21 (.0055)	.16 (.0033)	.16 (.0066)	.12 (.0034)	.26 (.0092)	.19 (.0059)
ISR-DEU	.33 (.0086)	.23 (.0040)	.23 (.0094)	.13 (.0032)	.38 (.0113)	.29 (.0058)
SWE-CHE	.18 (.0035)	.13 (.0030)	.19 (.0050)	.13 (.0041)	.18 (.0052)	.13 (.0047)
SWE-PRT	.25 (.0050)	.27 (.0092)	.22 (.0064)	.20 (.0082)	.27 (.0066)	.30 (.0122)
ISR-ESP	.38 (.0111)	.33 (.0055)	.33 (.0143)	.46 (.0093)	.43 (.0178)	.22 (.0069)
ESP-CZE	.23 (.0031)	.16 (.0024)	.27 (.0039)	.16 (.0027)	.14 (.0042)	.16 (.0049)
FRA-ITA	.32 (.0080)	.28 (.0056)	.25 (.0082)	.28 (.0077)	.37 (.0121)	.28 (.0077)
ISR-ITA	.38 (.0111)	.28 (.0056)	.33 (.0143)	.28 (.0077)	.43 (.0178)	.28 (.0077)
GRC-LUX	.28 (.0040)	.22 (.0056)	.32 (.0048)	.25 (.0067)	.16 (.0038)	.12 (.0051)
GRC-EST	.19 (.0033)	.14 (.0021)	.20 (.0049)	.09 (.0019)	.18 (.0046)	.21 (.0046)
ISR-BEL	.33 (.0086)	.26 (.0040)	.23 (.0094)	.17 (.0039)	.38 (.0113)	.30 (.0053)
ISR-CZE	.38 (.0111)	.27 (.0051)	.33 (.0143)	.20 (.0054)	.43 (.0178)	.35 (.0091)
ISR-SVN	.27 (.0071)	.20 (.0034)	.26 (.0087)	.21 (.0042)	.29 (.0121)	.15 (.0052)
ISR-EST	.21 (.0055)	.14 (.0021)	.16 (.0066)	.09 (.0019)	.26 (.0092)	.21 (.0046)

Notes: Own calculations based on SHARE data (Wave 6).

In Table 5 we report the values of test statistics for the comparisons presented in Tables 3 and 4. Most of the comparisons is highly significant. This is also true for $mqAF1$ and $mqAF2$, for which we report the values of test statistic in Table 6. We note that the rankings imposed by $somqAF1$ and $mqAF1$ agree and the same holds for $somqAF2$ and $mqAF2$. Sometimes the values of statistics changes e.g. it is small for $mqAF$ for weights (10, 1, 1) and high in $somqAF$. This shows the added value of differences in mass concentration between distribution (i.e. increased bipolarity), but in general for the ranking itself it is the first order comparison (i.e. increased spread) which is dominant.

Table 5: The values of test statistic for countries comparisons ($somqAF1$ and $somqAF2$)

Dominance	$\mathcal{P}_{a,b,c}^{somqAF1}$			$\mathcal{P}_{a,b}^{somqAF2}$		
A-B	1,1,1	10,1,1	1,10,1	1,1	10,1	1,10
GRC-AUT	2.93	35.25	6.99	11.93	14.47	[-0.15]
ISR-AUT	18.95	13.54	12.00	9.24	6.30	6.60
ISR-DEU	28.96	31.62	14.94	11.06	[1.52]	7.09
SWE-CHE	22.18	16.58	14.60	11.27	8.71	6.79
SWE-PRT	[1.72]	17.90	2.55	[-1.18]	2.36	-2.09
ISR-ESP	22.33	6.65	12.46	4.53	-7.30	1.99
ESP-CZE	28.46	40.01	5.53	19.24	23.35	-2.61
FRA-ITA	18.40	13.62	13.47	4.41	-2.09	6.45
ISR-ITA	2.43	21.31	14.17	8.17	3.50	7.86
GRC-LUX	11.14	13.75	3.58	9.14	8.55	6.02
GRC-EST	29.61	46.79	4.95	13.23	2.70	-4.39
ISR-BEL	18.52	15.41	13.91	7.58	5.96	6.14
ISR-CZE	23.36	27.48	[1.02]	9.32	9.16	4.31
ISR-SVN	23.07	17.06	12.36	9.72	5.29	[1.70]
ISR-EST	26.49	25.42	11.93	13.44	11.34	5.31

Notes: Own calculations based on SHARE data (Wave 6). [..] denotes insignificant values of test statistic.

Table 6: The values of test statistic for countries comparisons ($mqAF1$ and $mqAF2$)

Dominance	$\mathcal{P}_{a,b,c}^{mqAF1-KM}$			$\mathcal{P}_{a,b}^{mqAF2}$		
A-B	1,1,1	10,1,1	1,10,1	1,1	10,1	1,10
GRC-AUT	21.18	4.49	6.16	13.53	16.23	[-0.38]
ISR-AUT	18.20	29.86	12.17	9.49	6.55	6.75
ISR-DEU	28.94	35.38	14.30	11.18	[1.92]	6.91
SWE-CHE	21.46	36.53	15.16	11.71	8.81	7.43
SWE-PRT	12.94	35.41	2.18	[-1.41]	2.60	-2.46
ISR-ESP	24.13	67.78	15.05	5.32	-6.77	11.91
ESP-CZE	3.61	44.85	5.85	21.54	25.17	[-1.75]
FRA-ITA	17.56	37.79	14.56	4.44	-2.56	7.12
ISR-ITA	22.06	62.42	16.90	9.15	2.92	9.56
GRC-LUX	11.84	15.12	3.99	9.91	9.43	7.70
GRC-EST	34.91	56.97	3.89	15.71	23.07	-6.06
ISR-BEL	17.92	27.26	14.03	8.33	6.22	6.93
ISR-CZE	24.68	52.80	12.40	[1.24]	9.26	5.19
ISR-SVN	23.19	38.72	14.21	9.84	6.43	11.15
ISR-EST	29.31	41.50	12.43	15.36	12.64	5.33

Notes: Own calculations based on SHARE data (Wave 6). [..] denotes insignificant values of test statistic.

8 Conclusions

In this paper we develop the set of tools available to researchers who want to measure inequality in several ordinal indicators. Given numerous results concerning inequality in single ordinal indicators that have appeared mostly in the last decade, it is natural to develop inequality measurement theory for ordinal data into a multidimensional setting. This paper together with (Kobus and Kurek 2018) constitute an important step in this direction.

From the perspective of the particular application such as SES – health what seems as a very desirable property and an open question for future research is the decomposition of multidimensional inequality measures. Since the dependence between health and income (or education) receives a lot of attention in the literature on socio – economic inequality in health, it would be very useful to be able to distinguish *in the numerical sense* (here we do this in a more qualitative sense) the impact of dependence on joint inequality from the impact of health and income inequalities. The added value of the framework presented here and developed in the last decade is zero assumptions on the form of dependence, which is not true for more parametric approaches. This of course always comes at the expense of conclusiveness. This line of research, but in a cardinal framework, has been started by (Abul Naga and Geoffard 2006) who introduce a notion of attribute decomposability, namely, decomposability of inequality measures into inequalities in given attributes and a measure of dependence between attributes.

Funding: This research was funded by National Science Centre, Poland, grant number 2017/26/M/HS4/00905 awarded to Dr. Kobus. R. Kurek acknowledges the receipt of MNiD grant from Polish Ministry of Science and Higher Education awarded to INE PAN.

References

Abul Naga R H, Geoffard P Y. Decomposition of bivariate inequality indices by attributes. *Econ Lett* **2006**, *90(3)*, 362-367.

Abul Naga R H, Stapenhurst Ch. Estimation of inequality indices of the cumulative distribution function. *Econ Lett* **2008**, *130*, 109-112.

Abul Naga R H, Yalcin T. Inequality measurement for ordered response health data. *JHE* **2008**, *27(6)*, 1614-1625.

Allison R A, Foster J E. Measuring health inequality using qualitative data. *JHE* **2004**, *23(3)*, 505-524.

Apouey B. Measuring health polarization with self-assessed health data. *HE* **2007**, *16*, 875-894.

Apouey B, Silber J. Inequality and bi-polarization in socioeconomic status and health: ordinal approaches. In *Health and Inequality*; P.R.Dias, O. O'Donnell, Eds.; Emerald Group Publishing Limited: Bingley, UK, 2013; Vol. 21, Research on Economic Inequality.

Bond T, Lang K. The Evolution of the Black-White Test Score Gap in Grades K–3: The Fragility of Results. *REStat* **2013**, *95(5)*, :1468-1479.

Bond T, Lang K. 2018. The Sad Truth About Happiness Scales. (Previously) *WP* **2018**, *NBER No. 1995.*.

Chakravarty S R, Maharaj B. Generalized Gini polarization indices for an ordinal dimension of human well-being. *IJET* **2015**, *11(1)*, 231-246.

Cowell F, Flachaire E. Inequality with ordinal data. *Economica* **2017**, *84(334)*, 290-321.

Cutler D, Lleras-Muney A, Vogl T. Socioeconomic status and health: dimensions and mechanisms. In *The Oxford Handbook of HE*; Glied S and Smith P C, Eds.; Oxford University Press, UK, 2011; ISBN.

Esteban J M, Ray D. On the measurement of polarization. *Econometrica* **1994**, *62(4)*, 819-851.

Evans W, Wolfe B, Adler N. The SES-health gradient: a brief review of the literature. In *The Biological Consequences of Socioeconomic Inequalities*; Wolfe B, Evans W, Seeman T, Eds.; Russell Sage Foundation, UK, 2012; Chapter 1, ISBN.

Fattore M, Maggino F. 2014. Partial Orders in Socio-economics: A practical challenge for poset theorists or a cultural challenge for social scientists?. In *Multi-indicator Systems and Modelling in Partial Order*; Brüggemann R et al., Eds.; Springer Science + Business Media New York, US, 2014; Chapter 9, ISBN.

Fishburn P C, Lavalle I H. Stochastic dominance on unidimensional grids. *Math. Oper.*

Res. **1995**, *20(3)*, 513-525.
Formby J, Smith J, Zheng B. Mobility measurement, transition matrices and statistical inference. *J. Econom.* **2004**, *120*, 181–205.
Gravel N, Magdalou B, Moyes P. Ranking distributions of an ordinal attribute. (Previously) *WP* **2017**, *HALSH No. 01082996.*.
Kobus M, Miłoś P. Inequality decomposition by population subgroups for ordinal data. *JHE* **2012**, *31(1)*, 15-21.
Kobus M. Polarization measurement with ordinal data. *JoEI* **2015**, *13(2)*, 275-297.
Kobus M, Kurek R. Multidimensional polarization for ordinal data. *JoEI* **2018**, doi: 10.1007/s10888-018-9402-1.
Kobus M, Polchlopek O, Yalonetzky G. Inequality and welfare in quality of life among OECD countries: non-parametric treatment of ordinal data. *Soc. Indic. Res.* **2018**, doi:10.1007/s11205-018-1962-8.
Lazar A, Silber J. On the cardinal measurement of health inequality when only ordinal information is available on individual health status. *HE* **2013**, *22(1)*, 106-113.
Lv G, Wang Y, Xu Y. On a new class of measures for health inequality based on ordinal data. *JoEI* **2015**, *13(3)*, 465-477.
Madden D. Ordinal and cardinal measures of health inequality: an empirical comparison. *HE* **2010**, *19(2)*, 243–25.
Makdisi P, Yazbeck M. Measuring socioeconomic health inequalities in the presence of multiple categorical information. *JHE* **2014**, *34*, 84-95.
Mendelson H. Quintile-preserving spread. *JET* **1987**, *42(2)*, 334-351.
O'Donnell O, van Doorslaer E, Wagstaff A, Lindelow M. Analyzing health equity using household survey data: a guide to techniques and their implementation, Washington DC, World Bank, **2008**.
Sarkar S, Santra S. Extending the approaches to polarization ordering of ordinal variables. Working Paper presented at ECINEQ 2017/CR2, **2018**.
Shorrocks A F. Inequality decomposition by population subgroups. *Econometrica* **1984**, *52(6)*,1369-85.
Smith J P. Healthy bodies and thick wallets: The dual relation between health and economic status. *JEP* **1999**, *13(2)*, 145-166.
van Doorslaer E, Koolman X. Explaining the differences in income-related health inequalities across European countries. *HE* **2004**, *13(7)*, 609-628.
van Doorslaer E, Wagstaff A, Bleichrodt H, Calonge S, Gerdtham U, Gerfin M, Geurts J, Gross L, Hakkinen U, Leu R. Income-related inequalities in health: some international comparisons. *JHE* **1997**, *16(1)*, 93–112.
Zheng B. Measuring inequality with ordinal data: a note. *Research on Economic Inequality* **2008**, *16*, 177–188.
Zhong H. A multivariate analysis of the distribution of individual's welfare in China: What is the role of health?. *JHE* **2009**, *28(6)*, 1062-107.