This is a repository copy of *Inequality and welfare in quality of life among OECD countries: non-parametric treatment of ordinal data*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/132820/

Version: Accepted Version

Article:

Kobus, M, Półchłopek, O and Yalonetzky, G orcid.org/0000-0003-2438-0223 (2019) Inequality and welfare in quality of life among OECD countries: non-parametric treatment of ordinal data. Social Indicators Research, 143 (1). pp. 201-232. ISSN 0303-8300

https://doi.org/10.1007/s11205-018-1962-8

© Springer Nature B.V. 2018. This is a post-peer-review, pre-copyedit version of an article published in Social Indicators Research The final authenticated version is available online at: https://doi.org/10.1007/s11205-018-1962-8

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Inequality and welfare in quality of life among OECD countries: non-parametric treatment of ordinal data

June 26, 2018

Abstract

The last few years have witnessed an increasing emphasis on going beyond GDP per capita when measuring a nation's quality of life (QoL). Countries (e.g. UK, France, Canada, Germany, Italy, Japan, Korea, Spain) and international organizations (e.g. OECD) have been developing methods suitable for non-income indicators. However, this involves serious measurement challenges due to: (a) multidimensionality, and (b) ordinality (i.e. unlike income these indicators do not have a natural scale). This paper is the first summary of the methods developed in the last decade in the field of inequality and welfare measurement to address these challenges. Next, we utilize the presented methodology and provide evidence on the ranking of OECD countries in terms of welfare and inequality in education and happiness. We find that when dimensions are analysed separately, welfare dominance is frequent (42% of all comparisons in education and 31% in life satisfaction). The number drops to only $4,4\%$ for bivariate dominance, which highlights the empirical relevance of multidimensional analysis. Greece, Portugal and Hungary feature the lowest joint welfare. Northern European countries are most often dominating and Southern European countries are most often dominated in both inequality and welfare analyses.

Keywords: ordinal data, quality of life, inequality and welfare, partial order, majorization, education-happiness gradient
JEL codes: I31; D63

1 Introduction

In the last decade there has been an increasing interest among policymakers (e.g. in France, Germany, Italy, Japan, Korea, Spain, and the UK) in measuring quality of life in a more comprehensive way than via GDP per capita statistics. In 2010 British Prime Minister David Cameron announced that in evaluating people's quality of life the Government would rely not only on GDP growth but also on non-income indicators such as education, health and environment, and that this broad measure of well-being will steer government policy.[1] Cameron earlier described monitoring people's well-being as one of the central political issues of

[1]Source: The Guardian, http://www.theguardian.com/lifeandstyle/2010/nov/14/happiness-index-britain-national-mood

our time. Since then the Office for National Statistics has been carrying out the program "Measuring what matters" which develops new measures of national well-being. The Canadian government has been working with the UK since 2010 on the same problem. Mexico has been at the forefront of developing measurement methods for tracking social progress for several years now. In 2008 the French government created the renowned Stiglitz-Sen-Fitousi commission in order to better tackle the measurement of social and economic performance. The commission's final report (Stiglitz et al. 2009) underlines the multidimensionality of the concept of well-being, its relationship to non-income dimensions and the importance of including both objective and subjective measures. On page 14 the authors write: "To define what well-being means a multidimensional definition has to be used", whereas in Recommendation 7 they stress that "Quality-of-life indicators in all the dimensions covered should assess inequalities in a comprehensive way."

Not only governments but also international organisations and NGOs undertake similar initiatives. In May 2011, in its 50th Anniversary Week, the OECD launched the Better Life Index which allows citizens to compare lives across 34 countries, based on 11 dimensions such as housing, income, jobs, community, education, environment, governance, health, life satisfaction, safety, and work-life balance. The Social Progress Imperative that developed the Social Progress Index is an initiative that aims to "solve the world's most pressing challenges by redefining how the world measures success". Clearly, all these initiatives prove that policymakers are finally responding to the economists' calls for going beyond GDP in measuring a nation's quality of life (Sen 1973; Fleurbaey and Blanchet 2013).

Many non-income wellbeing dimensions are available in surveys in the form of ordinal data, i.e. variables not possessing a natural scale and whose only relevant information is the ordering of their categories. Thus any sequence of numbers consistent with a given ordering can be used as a scale.[2] This is the case of such well-known indicators as happiness, self-reported health, occupational status, and educational attainment. That these variables are purely ordinal poses serious measurement challenges since measures of welfare, inequality or poverty are typically based on the mean. Here the mean changes depending on the assigned scale and consequently the conclusions of distributional comparisons change too, which is undesirable (e.g. see examples in Allison and Foster 2004; Kobus 2015).

New measurement theories are being developed in various fields of economics to deal with the ordinality problem.[3] This article is to serve as a "map" of the state-of-the-art approaches to inequality and welfare measurement for ordinal data. We

[2]More precisely, we refer to data that are ordinal and discrete. By ordinal we mean total orders as opposed to cardinality, where numbers are meaningful. Discrete means that there is a fixed number of values, each carries a probability mass as opposed to continuous variables which accord a particular value with probability zero. In general, we can have variables that are: ordinal and discrete (such as health status); ordinal and continuous (such as the Body Mass Index (BMI) which, as the ratio of two continuous variables, is continuous but the differences between two BMI's are meaningful only in an ordinal sense); cardinal and discrete (e.g. the distribution of the number of cars in households, where there is a fixed number of values, and particular values are meaningful); cardinal and continuous (e.g. income).

[3]These theories often build upon statistics literature on ordinal data e.g. Blair and Lacy (2000), Berry and Mielke (1992).

summarise the available methods and present them to a larger social science audience. We review the measurement aspects of the quality of life when data are ordinal and/or multidimensional. This has been considered by various authors i.e. Allison and Foster 2004; Apouey 2007; Abul Naga and Yalcin 2008; Zheng 2011; Kobus and Miłoś 2012; Dutta and Foster 2013; Kobus 2015; Lv, Wang, and Xu 2015, Gravel, Magdalou, and Moyes 2015, Cowell and Flachaire 2017.[4] These authors deal mostly with the case when QoL is proxied by a single variable. The literature on multidimensional ordinal variables is very preliminary. Yalonetzky (2013) proposes conditions and a test for multidimensional welfare and we draw on his results. Other contributions in the area of multidimensionality have very limited applicability e.g. Sonne-Schmidt et al. (2016) and Makdisi and Yazbeck (2014) study only binary variables.

Next using the presented methods we provide new evidence on the ranking of OECD countries in terms of achievement and inequality in the quality-of-life (QoL) distribution. Thus we complete the analysis started by Balestra and Ruiz (2014). Based on the Gallup World Poll 2010 they study inequality and welfare comparisons also in OECD countries. Their results are largely consistent with ours but they apply only a subset of the tools we use. Furthermore, we address the two shortcomings acknowledged in their study, namely, the common median assumption in the AF procedure and the lack of multidimensional analysis. We offer many more insights and test for statistical significance where possible.

Firstly, starting with univariate distributions (education and life-satisfaction separately) we use first-order dominance for welfare comparisons and the ordinal approach developed by Allison and Foster (2004) (henceforth AF) for inequality comparisons. The latter states that a distribution which has more probability mass concentrated around the median than another distribution is more equal.[5] We use a generalisation of the AF inequality ordering which extends the scope of its applicability. More specifically, we require that the distribution does not have to be concentrated around the median necessarily, but it can be any other category; we call it the AF_α ordering. This generalisation comes in fact from Mendelson (1987), but was overlooked by the literature.[6] Abul Naga and Yalcin (2008) and Kobus and Miłoś (2012) construct family of indices consistent with the AF ordering. Some of the indices value more inequality below the median and some value more inequality above the median. We check how these different value judgements influence countries' ranking. We also use the family of indices constructed by Cowell and Flachaire (2017) for which the reference point is the highest category, not the median. Thanks to the recent results of Abul Naga and Stapenhurst (2015) and Cowell and Flachaire (2017), we can now test for statistical significance of the differences between inequality measures. Then, we analyse inequality and welfare *jointly* by using the H^+-dominance curves developed re-

[4] Recent theoretical advances to deal with ordinality are, however, broader and include also poverty measurement (Bennett and Hatzimasoura 2011; Yalonetzky 2012), dissimilarity (Andreoli and Zoli 2014), segregation (Cuhadaroglu 2013), inequality of opportunity (Silber and Yalonetzky, 2011), estimation of achievement gaps (Nielsen 2015), etc.

[5] In Section 2.2 we offer a brief discussion on whether bi-polarisation criterion such as AF is an appropriate inequality dominance condition. For more detailed discussion, we refer readers to Kobus (2015). Unarguably, AF approach has been so far the most widely used approach to measuring inequality in ordinal data.

[6] We thank Professor Brice Magdalou for referencing this work to us.

cently by Gravel et al. (2015) i.e. to compare, in the income setting, Lorenz curve takes into account only the spread, whereas Generalized Lorenz curves cares both about the spread and the mean income. Finally, we use the results of Yalonetzky (2013) to test for multidimensional welfare dominance. We thus shed light on how countries compare in terms of multivariate QoL when attributes' association is taken into account. Univariate dominance is a necessary but not a sufficient condition for a multivariate dominance.

The results we obtain offer a few observations. Firstly, the countries most often dominated are Southern European, whereas those most often dominating are Northern European countries. This holds for different forms of welfare and inequality dominance relationships tested in this paper, and for inequality measures. Bidimensional welfare dominance relationships are almost exclusively driven by a few countries being dominated by most countries. These are Greece, Portugal, Hungary and, to a lower extent, France and Spain. In France the outcome is likely driven by life satisfaction. Indeed, this is known as the "French unhappiness puzzle" (Senik 2014). On the other hand, Eastern European countries (Estonia, Hungary, Russia) appear to have one of the worst results in life satisfaction, for both welfare and inequality.

Secondly, the results highlight the importance of a multidimensional welfare analysis. There are many cases of first-order stochastic dominance (henceforth FSD) when dimensions are analysed separately. 42% of all comparisons in education exhibit univariate FSD and 31% in life satisfaction. A different pattern emerges from the multidimensional analysis: Only 4.4% of all comparisons show bivariate FSD. One should be therefore careful in drawing conclusions from single-indicator analyses. At least among OECD countries, unidimensional dominance relationships most often do not imply bidimensional dominance. Dimensions' dependence does matter substantially in welfare comparisons.

Thirdly, the extension of AF dominance to AF_α dominance increases the number of dominance tenfold, so the problem of incompleteness is significantly reduced this way.

Finally, value judgements about inequality do influence countries' rankings, but differently for education and life satisfaction. This can be studied via inequality indices, which are aggregating functions with different weights applied to inequality below or above the median. For life satisfaction, the countries' inequality ranking changes depending on whether inequality below or above the median gets higher weight in the index, whereas for education, the countries' inequality ranking changes when the reference point changes i.e. from the median (as in the AF approach) to the highest category (as in the Cowell-Flachaire family of indices).

The paper is organised as follows. In Section 2 we present the challenges of applying standard techniques to ordinal data. We then explain in a non-formal way the methodology developed to address these challenges. Then we introduce formal definitions. In Section 3 we apply this methodology to study education and happiness in OECD countries separately and jointly. Finally, we offer some discussion and conclusions (Section 4). The results are included in Tables and Figures in Appendix A. Appendix B includes brief presentation of the statistical inference for bivariate welfare dominance.

2 Measurement framework

2.1 The non-robustness of location and dispersion statistics for ordinal data

The numerous measurement problems with ordinal data have been increasingly acknowledged in the economic literature, and the conclusions are far-reaching. In particular, Bond and Lang (2013) show that the estimates of the black-white test score gap in the US, a phenomenon widely studied in educational economics, are very fragile to the choice of scaling. Recently, the same authors (Bong and Lang 2014) highlight that binominal regression techniques are also a form of arbitrary cardinalisation of ordinal variables and are thus easily reversible in their results. In particular, they show that the Easterlin paradox (a key concept studied in happiness economics, whereby beyond a certain value more income does not appear to be correlated with more happiness) is very sensitive to the chosen distribution of happiness when the concept of happiness is assumed to be an underlying continuous variable with discrete representation.

The following example concerns standard inequality measures such as the Gini index and comes from Kobus and Miłoś (2012). Similar examples were given for the mean (Allison and Foster 2004), coefficient of variation (Lazar and Silber 2013), and measures of bi-polarisation (Kobus 2015). They all reflect the sensitivity of location and dispersion statistics to the scale.

Suppose the distributions of self-reported health status among men and women are, respectively: $h = (0.2, 0.2, 0.2, 0.2, 0.2)$ and $w = (0.3, 0.2, 0.1, 0.1, 0.3)$. That is, 20% of men are in each health category, 30% of women are in the first category, etc. By assumption, higher category number indicates better health status. We consider two scales: $c = (1, 2, 3, 4, 5)$ and $\tilde{c} = (1, 2, 3, 4, 100)$; note that both correspond to the same order of health categories. Then, under scale c the Gini index for the men's distribution is $GINI(h, c) = 0.26$ whereas for women's distribution we get $GINI(w, c) = 0.31$, hence health inequality is lower among men than women.[7] However, under scale $\tilde{c}$ the ranking is reversed: $GINI(h, \tilde{c}) = 0.72 > GINI(w, \tilde{c}) = 0.66$.

2.2 The solution: distribution-based approach

To deal with the presented problem, many authors in the last decade have started to develop the theory of inequality and welfare measurement that is scale free and that relies directly on the distribution of an ordinal indicator. We will now briefly describe main tools and results developed in this field.

Similarly to the cardinal case, welfare in the ordinal case can be compared using first-order stochastic dominance i.e. a distribution generates higher welfare

[7] We calculated the Gini index by assuming there are two men in each health category, three women in the first health category, two women in the second health category and so on. This is valid since the Gini index is replication invariant.

than another one if its cumulative distribution function (cdf) lies everywhere below the cdf of the other distribution. Then the percentage of the population in the i-th lowest category is not higher in dominating distribution than in the dominated one, for every i. Thus the dominating distribution gives more probability mass to higher categories than the dominated one, hence it has higher mean for any increasing scale.

For inequality comparisons, the most widely used criterion is the one offered by Allison and Foster (2004) (AF), according to which a more unequal distribution is the one with more probability mass further from the median. The most unequal distribution is the one that has 50% of the mass in the lowest category and 50% in the highest category i.e. the most bi-polarized distribution. One can transform the more equal distribution into more unequal through transfers that move probability mass away from the median. Thus Allison and Foster (2004) called this relation a median-preserving spread, however, the formal connection between such spreads and their definition is provided in Kobus (2015). AF criterion requires that only distributions with the same median can be compared, therefore it might be inconclusive. Then it is standard practice to resort to measures which give complete rankings of distributions. This comes at the expense of robustness. Indices take concrete functional forms expressing value judgements. For example, some indices may be more sensitive than others to inequality in the lower tail of the distribution. This translates into different weights that a given index attaches to inequality below and above the median. A few authors (Allison and Foster 2004, Apouey 2007, Abul Naga and Yalcin 2008, Kobus and Miłoś 2012) propose indices consistent with the AF relationship. Abul Naga and Stapenhurst (2015) develop inference procedures for these measures.

The AF approach has been criticised for being more about bi-polarisation than inequality (Zheng 2008). Polarisation refers to the "disappearing middle class" (Wolfson, 1994) and the emergence of a divided population. In a cardinal data context, polarisation and inequality are different concepts (Esteban and Ray 1994). In particular, polarisation may increase following the transfer of income from the richer to the poorer as long as it increases group homogeneity. This cannot happen with inequality measures. In a cardinal setting, inequality is measured as deviation from the perfectly equal distribution which is unique. In an ordinal context the most equal distribution is the one for which all mass is concentrated in one category, but there are as many such distributions as there are categories, so the question arises: "deviation from *which* distribution?" Therefore, inequality with ordinal data is measured as deviation from the perfectly unequal distribution and the most notable approach is the AF approach. To address the bi-polarization problem, a different approach, however, has been recently proposed by Gravel et al. (2015) and Cowell and Flachaire (2017).

Gravel et al. (2015) relate ordinal inequality reduction to the Hammond's principle of transfers (Hammond 1976). A Hammond transfer reduces the spread between two individuals being in different categories of an ordinal attribute. This is unrelated to whether the loss experienced by one individual equals the gain by another. The Pigou-Dalton Transfer in which the *same* amount of income is taken from the rich and given to the poor is thus a special case of Hammond transfer. In an ordinal data context, the magnitudes of gains and losses cannot be compared anyways, therefore Hammond transfer becomes the proper notion of inequality reduction. Gravel et al. (2015) propose inequality dominance criteria based on Hammond transfers. They take into account both the spread and the achievement i.e. both inequality and welfare.

The concept of status is central to the approach of Cowell and Flachaire (2017). It is quite common in the income distribution literature to use a person's location in the distribution as her status in society. In particular, they focus on *downward-looking status* i.e. the proportion of individuals in the classes below a given individual and *upward-looking status* i.e. the proportion of individuals in the classes above a given individual. Then they define the equality reference point e.g. mean or median status or the maximum value of status which is 1 for the two definitions of status mentioned. By way of examples, they show that changes in the mean or the median may have counterintuitive effects on inequality. Therefore, in applications, they focus on maximum status which does not have this undesirable property.

The so far presented methodology focuses on a single ordinal indicator. The literature on multidimensional ordinal data has been very scarce, but we use the results of Yalonetzky (2013) to compare dimensions of QoL *jointly*. Indeed, it has become common practice in the economics literature (Atkinson and Bourguignon 1982) to analyse dimensions of wellbeing jointly, namely, to acknowledge and account for their dependence. The latter may have different impact on the joint welfare and thus on countries' comparisons. If attributes are considered substitutes, then the positive impact of, say, better education on wellbeing decreases for higher levels of life satisfaction (and vice versa). Here goods substitute for each other, thus having more of one good makes an individual desire less of the other good. The reverse holds if attributes are treated as complementary goods. Then the impact of education on quality of life is strengthened for higher levels of declared life satisfaction. Having more of one good makes an individual desire even more of the other good if the two complement each other. In the case of substitutes, the higher the association between attributes the more inequality and less welfare, because there is higher likelihood that a given individual is deprived in both dimensions. Distributions with lower association will thus be considered better. Univariate dominance is a necessary but not a sufficient condition for a multivariate dominance. That is, it may happen that although country A dominates country B in both dimensions, there is no joint dominance. Univariate comparisons may thus not be very robust when a broader view of welfare and inequality is taken into account. This is indeed the case in our study of OECD countries.

Yalonetzky (2013) shows that in the case of two variables we can identify up to four first-order stochastic dominance conditions, resembling those derived by Atkinson and Bourguignon (1982) for continuous variables. These criteria differ in how they evaluate relationship between two dimensions of QoL. In one case two dimensions are treated as substitutes and in another as complementary goods. In the case of neutrality, having more of one good is unrelated to desiring the other good. Yalonetzky (2013) also develops tests for multivariate stochastic dominance conditions for ordinal variables, so we can test for joint welfare dominance between each pair of countries.

To sum up, below we present the formal definitions and conduct empirical analyses for OECD countries in terms of:
(i) welfare via $\leqslant_{FSD}$
(ii) inequality via $\leqslant_{AF_\alpha}$ and $I_{\alpha,\beta}, I_{a,b}$ families of measures
(iii) joint welfare and inequality via H^+-dominance curves
(iv) welfare in joint distribution of education and life satisfaction via bivariate dominance criteria.

Dominance comparisons may turn out inconclusive. Then we cannot obtain a complete ranking of countries; however, we can rank countries based on their tendency to outperform other countries. Copeland's (1951) score measures this tendency. Each country is assigned a normalised difference between the number of countries it dominates and the number of countries it is dominated by. If x is the number of countries considered, then the Copeland score ranges between $(-(x-1), x-1)$ which is normalised to $(-1, 1)$. Copeland's score is one of many multi-criteria decision analysis tools. It has been increasingly applied to evaluate population well-being (Arndt et al. 2016, Siersbaeck et al. 2017) and we use it here too.

2.3 Formal definitions and notation

For presentation purposes, we define our framework and conditions in terms of bivariate distributions; however, this can be extended to an arbitrary number of dimensions. We define a numerical representation of categories of ordinal variables $\mathbb{I} := \mathbb{I}_1 \times \mathbb{I}_2 = \{1, \ldots, n_1\} \times \{1, \ldots, n_2\}$ which are ordered (i.e. $1 < \ldots n_s$, $s = 1, 2$), but particular values are not important; they are only labels/names of the categories. Let f be a probability distribution on the set $\mathbb{I}$. Obviously we require $\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} f_{ij} = 1$ and for all $(i, j) \in \mathbb{I}$, $f_{ij} \geq 0$. We define marginal distributions by:

$$f_i^1 := \sum_{j=1}^{n_2} f_{ij} \quad f_j^2 := \sum_{i=1}^{n_1} f_{ij}, \tag{1}$$

cumulative distributions by:

$$F_i^1 := \sum_{k=1}^{i} f_k^1 \quad F_j^2 := \sum_{l=1}^{j} f_l^2, \tag{2}$$

and a multidimensional cumulative distribution function F at (i, j) by:

$$F_{ij} := \sum_{k=1}^{i} \sum_{l=1}^{j} f_{kl} \tag{3}$$

For each dimension s we define a median m_s which is the (numbered) category for which $F_{m_s-1}^s < 1/2$ and $F_{m_s}^s \geq 1/2$.[8] Let Λ denote a set of probability distributions. Finally, let an inequality index be denoted by $I : \Lambda \mapsto \mathbb{R}_+$.

Univariate first-order stochastic dominance is defined in the following way.

[8] We define the median category slightly differently than Allison and Foster (2004) to avoid dealing with multiple medians due to empty categories. Kobus (2015) extends the AF relation to a setting with multiple medians, so in fact AF_α can be extended too in order to allow for multiple quantiles, however this is mostly technical.

Definition 1. ***First order stochastic dominance (FSD)*** *Fixing $n_s \geq 1$ for $s = 1, 2$ and allowing f^s, g^s to be two probability distributions on $\mathbb{I}_s$; we say that f^s first-order dominates g^s i.e.*

$$f^s \leqslant_{FSD} g^s \text{ if and only if } F_i^s \leq G_i^s \quad \text{for all } i \in (1, \ldots, n_s),$$

with at least one strict inequality.

In Section 3 we use the $\leqslant_{FSD}$ ordering to compare countries in terms of welfare. FSD is closely related to the following type of transfers that move probability mass upward:

Definition 2. ***Upward shift*** *Let $\epsilon > 0$. Fixing $n_s \geq 1$ for $s = 1, 2$ and allowing f^s, g^s to be two probability distributions on $\mathbb{I}_s$; we say that f^s is obtained from g^s via an upward shift if there exist categories $1 \leq m < l < n_s$ such that*

$$f_i^s = g_i^s \quad \text{for all } i \text{ different than } m, l$$

$$f_m^s = g_m^s - \epsilon, \quad f_l^s = g_l^s + \epsilon$$

For inequality comparisons, as already mentioned the AF relation (Allison and Foster 2004) is typically used. Allison and Foster (2004) introduce the following partial ordering $\leqslant_{AF}$.

Definition 3. ***AF relation (Allison and Foster (2004))***
Fixing $n_s \geq 1$ for $s = 1, 2$ and allowing f^s, g^s to be two probability distributions on $\mathbb{I}_s$:

$$f^s \leqslant_{AF} g^s \text{ if and only if}$$

(AF1) *f^s, g^s have a common median m^s,*

(AF2) *$F_i^s \leq G_i^s$ for any $i < m^s$,*

(AF3) *$F_i^s \geq G_i^s$ for any $i \geq m^s$,*

with at least one strict inequality.

Interpreting this ordering is intuitive; in particular, $f^s \leqslant_{AF} g^s$ when f^s is more concentrated around the median than g^s, and is therefore less unequal. The drawback of the AF relationship is that it hinges on the common median, but Mendelson (1987) extends it to any common quantile.

Definition 4. *AF_α* ***relation (Mendelson (1987))*** *Fixing $n_s \geq 1$ for $s = 1, 2$ and allowing f^s, g^s to be two probability distributions on $\mathbb{I}_s$, let $\alpha \in [0, 1]$ denote percentiles and $m(\alpha)$ denote the respective quantile (such that, for instance, $m = m(0.5)$):*

$$f^s \leqslant_{AF_\alpha} g^s \text{ if and only if}$$

(AF1) *f^s, g^s have a common $m(\alpha)$,*

(AF2) *$F_i^s \leq G_i^s$ for any $i < m(\alpha)$,*

(AF3) *$F_i^s \geq G_i^s$ for any $i \geq m(\alpha)$,*

with at least one strict inequality.

$f^s \leqslant_{AF_\alpha} g^s$ when f^s is more concentrated around the $m(\alpha)$ quantile than g^s and is therefore considered more equal. In Section 3 we check whether allowing for more crossings significantly reduces inconclusiveness of AF relation. Note that the AF relationship is a special case of the AF_α relationship for $\alpha = 0.5$. For instance, if we used quartiles, we would consider $m(0.25), m, m(0.75)$.

We use two families of measures consistent with the AF. Kobus and Miłoś (2012) propose the following family.

$$I_{a,b} := \frac{a\sum_{i<m^s} F_i^s - b\sum_{i\geq m^s} F_i^s + b(n_j + 1 - m^s)}{\left(a(m^s-1) + b(n_j - m^s)\right)/2}; \;\; a, b \geq 0. \tag{4}$$

When $a > b$ the index is more sensitive to inequality below the median, whereas the opposite is true if $a < b$ and more weight is attached to inequality above the median. If $a = 1$ and $b = 1$, both types of inequality are treated in the same manner. In Section 3 we use this family of indices with $a = 3, b = 1; a = 1, b = 1; a = 1, b = 3$.

Abul Naga and Yalcin (2008) propose the following family of measures.

$$I_{\alpha,\beta} := \frac{a\sum_{i<m^s} (f^s)_i^\alpha - b\sum_{i\geq m^s} (f^s)_i^\beta + n_j + 1 - m^s}{\left((m^s-1)\left(\frac{1}{2}\right)^\alpha + (n_j - m^s)\left(\frac{1}{2}\right)^\beta\right) + n_j - m^s}; \;\; \alpha, \beta \geq 1. \tag{5}$$

Similarly to the $I_{a,b}$ family, α and β measure sensitivity towards inequality below and above the median, e.g. for a given value of β, the index is more sensitive to the inequality below the median as $\alpha \to 1$. Conversely, when $\alpha \to \infty$ the index ignores dispersion in the bottom of the distribution. Recently, Abul Naga and Stapenhurst (2015) provide estimation procedures and obtain explicit standard errors formulas for both the $I_{\alpha,\beta}$ and $I_{a,b}$ families.

We now turn to approaches that are different than AF bi-polarization criterion. Gravel et al. (2015) propose a novel inequality dominance criterion called H^+-dominance which is closely related to the notion of a Hammond transfer.

Definition 5. ***Hammond's principle of transfers (Hammond 1976)*** *Let $\epsilon > 0$. Fixing $n_s \geq 1$ for $s = 1, 2$ and allowing f^s, g^s to be two probability distributions on $\mathbb{I}_s$; we say that distribution f^s is obtained from distribution g^s via a Hammond's transfer if there exist categories $1 \leq t < h \leq j < l < n_s$ such that*

$$f_i^s = g_i^s \quad \textit{for all } i \textit{ different than } t, h, j, l$$

$$f_m^s = g_t^s - \epsilon, \quad f_h^s = g_h^s + \epsilon$$
$$f_j^s = g_j^s + \epsilon, \quad f_l^s = g_l^s - \epsilon.$$

A Hammond transfer brings two individuals closer in the distribution of an ordinal attribute. H^+-dominance combines both Hammond transfers and upward shifts of probability mass. Thus it is a criterion for joint inequality and welfare comparisons, akin to Generalised Lorenz dominance in the cardinal setting.

Definition 6. ***H^+-dominance (Gravel et al. 2015)*** *Fixing $n_s \geq 1$ for $s = 1, 2$ and allowing f^s, g^s to be two probability distributions on $\mathbb{I}_s$; we define:*

$$H^+_{f^s}(i) := \sum_{k=1}^{i} (2^{i-k}) f^s_k \tag{6}$$

We say that distribution f^s H^+-dominates distribution g^s if and only if

$$H^+_{f^s}(i) \leq H^+_{g^s}(i) \quad \text{for all } i = 1, \ldots, n_s,$$

with at least one strict inequality.

The H^+-dominance curve has a very simple recursive structure:

Remark 1.

$$H^+_{f^s}(i) = 2H^+_{f^s}(i-1) + f^s_i \tag{7}$$

Another approach which is not bi-polarisation has been proposed by Cowell and Flachaire (2017). They measure inequality as a deviation from maximum status. For downward-looking status (i.e. the proportion of individuals in the classes below a given individual) this gives rise to the following measures of inequality for $\alpha \neq 0, 1$. Please recall that n_s is the number of categories for dimension s (either $s = 1$ or $s = 2$ in our framework).

$$I_\alpha = \frac{1}{\alpha(\alpha-1)} \left(\sum_{i=1}^{n_s} f_i \left(\sum_{j=1}^{i} f_j \right)^{\alpha} - 1 \right) \tag{8}$$

and for $\alpha = 0$ they get

$$I_\alpha = -\sum_{i=1}^{n_s} f_i \log \left(\sum_{j=1}^{i} f_j \right). \tag{9}$$

Yalonetzky (2013) develops conditions for welfare comparisons of bivariate ordinal data. Let $U(\mathbb{I}_1, \mathbb{I}_2) \mapsto \mathbb{R}_+$ be an individual utility function. Let W be a social utility function mapping into the real line from all the individual utility functions in a particular society. W is symmetric with respect to individuals, additively decomposable, and satisfies the population principle i.e. it attains the same value when all individual utility functions are replicated. Consider also the partial-difference $U_1(i, j) \equiv U(i, j) - U(i-1, j)$ (with similar definition for $U_2(i, j)$) and the cross-partial difference $U_{12}(i, j) \equiv U(i, j) - U(i-1, j) - U(i, j-1) + U(i-1, j-1,)$. Let $S^A_{12}(i, j) \equiv \Pr[\mathbb{I}_1 \geq i \wedge \mathbb{I}_2 \geq j]$ be the joint survival function. Finally, in any comparison of statistics between society A and B consider the following definition of differences, e.g. for the case of W: $\Delta W \equiv W^A - W^B$. We have the following first-order dominance criteria.

Definition 7. ***FSD BIVARIATE (FSDB)*** *$\Delta W > 0$ for all U characterised by $U_1(i, j) > 0$, $U_2(i, j) > 0$ and $U_{12}(i, j) \leq 0$ or $U_{12}(i, j) \geq 0$ for all (i, j) if and only if $\Delta S_{12}(i, j) \geq 0$ for all (i, j) (with at least one strict inequality) and $\Delta F_{12}(i, j) \leq 0$ for all (i, j) (with at least one strict inequality).*

Definition 8. ***FSD ALEP S (FSDAS)*** *$\Delta W > 0$ for all U characterised by $U_1(i, j) > 0$, $U_2(i, j) > 0$, $U_{12}(i, j) \leq 0$ for all (i, j) if and only if $\Delta F_{12}(i, j) \leq 0$ for all (i, j) (with at least one strict inequality).*

We call Definition 8 "FSD ALEP S" because it applies only to individual utility functions characterised by the so-called Auspitz-Lieben-Edgeworth-Pareto (ALEP) substitutability (i.e. $U_{12}(i,j) < 0$). Kannai (1980) gives a description of ALEP concepts of substitution and complementarity for discrete variates and we use them here. In this case two goods are treated as substitutes.

Definition 9. ***FSD ALEP C (FSDAC)*** $\Delta W > 0$ *for all* U *characterised by* $U_1(i,j) > 0$, $U_2(i,j) > 0$, $U_{12}(i,j) \geq 0$ *for all* (i,j) *if and only if* $\Delta S_{12}(i,j) \geq 0$ *for all* (i,j) *(with at least one strict inequality).*

We call Definition 9 "FSD ALEP C" because it applies only to individual utility functions characterised by ALEP complementarity (i.e. $U_{12}(i,j) > 0$). In this case two goods are treated as complementary goods.

Definition 10. ***FSD ALEP N (FSDAN)*** $\Delta W > 0$ *for all* U *characterised by* $U_1(i,j) > 0$, $U_2(i,j) > 0$, $U_{12}(i,j) = 0$ *for all* (i,j) *if and only if* $\Delta F_1(i) \leq 0$ *for all* (i,j) *(with at least one strict inequality) and* $\Delta F_2(j) \leq 0$ *for all* (i,j) *(with at least one strict inequality).*

We call Definition 10 "FSD ALEP N" because it applies only to individual utility functions characterised by ALEP neutrality (i.e. $U_{12}(i,j) = 0$). In this case two goods are considered separately i.e. their association does not matter.

There is an interesting logical connection between these four first-order bivariate dominance conditions (Remark 2).

Remark 2. *(i) ($A \geq_{FSDB} B$) if and only if ($A \geq_{FSDAS} B$ and $A \geq_{FSDAC} B$)*
(ii) ($A \geq_{FSDAS} B$ or $A \geq_{FSDAC} B$) implies ($A \geq_{FSDAN} B$)
(iii) ($A \geq_{FSDB} B$) implies ($A \geq_{FSDAN} B$).

The tests of these conditions come from Yalonetzky (2013) (please see Appendix).

3 Results

The non-income dimensions chosen in this study are education and life satisfaction. Education is included in many measures of multidimensional welfare e.g. Human Development Index (UNDP 2014) and Multidimensional Poverty Index (Alkire et al. 2015). Several governments are now measuring happiness. Moreover, some countries like Bhutan even try to maximise it. The high-profile report by Stiglitz et al. (2009) underlies the need to promote survey measures of wellbeing to inform policy. The United Nations passed a resolution in 2011 encouraging countries to evaluate nations' happiness, and in 2012 it published the World Happiness Report for the first time. As Richard Layard writes: "happiness should become the goal of policy, and the progress of national happiness should be measured and analysed as closely as the growth of GDP" (Layard 2005, pp. 147).

The data we use come from merging the World Values Survey and European Values Survey databases. We evaluate the years 2006-2012 (Waves 5 and 6) for which the data covered 37 OECD countries. Education and life satisfaction. in these surveys are multi-level indicators described in detail in Table 1. The sample sizes range from 784 (New Zealand) to 2809 (South Africa) respondents (a complete overview is in Table 2). The number of possible pair-wise dominance relationships we consider is 666 i.e. we consider all situations where there is a dominance of A over B but no dominance of B over A.

We start with welfare analysis according to Definition 1. We find substantial number of dominance relationships: 42% pair-wise comparisons exhibit dominance in education and 31% in life satisfaction (Table 3). The leading countries in both dimensions are similar and these are mostly Northern European countries, Australia, South Korea and New Zealand. Yet the correlation between countries' ranks in both dimensions is very low i.e. 0.146. This is because there are many countries which are high up in one dimension but significantly lower in the other e.g. Brazil (21st in education and 5th in life satisfaction) or Germany (5th in education and 16th in life satisfaction). The most dominated countries in education are Southern European countries (Italy, Greece, Portugal) and South American countries (Chile, Brazil). This tendency for Northern Europe to outperform and for Southern Europe to underperform exhibits stability throughout further comparison. The FSD results for life satisfaction are consistent with Balestra and Ruiz (2014), but less so for education i.e. their dominating countries occupy middle positions here (United States, United Kingdom and the Netherlands).

Comparing countries in terms of inequality according to Definition 4 (i.e. the dominated countries are more equal) we note that the AF_α is a significant improvement over AF relation (Table 4 and 5). The number of dominances increases tenfold when other points than the median are considered i.e. values of α different than $\frac{1}{2}$. Here too Northern European countries occupy relatively high ranks. Interestingly, for inequality, the rankings of countries in education and life satisfaction are negatively correlated. The value of the correlation coefficient equals -0.107. That is, countries that are more equal in one dimension are slightly more likely to be less equal in the other dimension. On the other hand, the rankings of countries in welfare and inequality dominance are significantly positively correlated, both in education (0.412) and in life satisfaction (0.556). That is, for both dimensions there is significant likelihood that the country which is dominant/dominated in terms of welfare is also dominant/dominated in terms of inequality.

Differences between countries in inequality values as measured by both $I_{a,b}$ and $I_{\alpha,\beta}$ are almost always statistically significant at 5% level of significance.[9] Generally, for most countries educational inequality is higher than life satisfaction inequality as measured by both $I_{a,b}$ and $I_{\alpha,\beta}$ (Figure 1 and 2). Life satisfaction inequality changes more than educational inequality with differing weights attached to inequality below and above the median. In particular, for education the average inequality ranges from $\tilde{I}_{3,1} = 0.46$ to $\tilde{I}_{1,3} = 0.52$ and for life satisfaction from $\tilde{I}_{3,1} = 0.29$ to $\tilde{I}_{1,3} = 0.41$. Another way to see this is to study the stability of countries' rankings. For education, the ranking of countries does not change much with different weights; the correlation between $I_{1,1}$ and $I_{3,1}$ is 0.82, whereas for life satisfaction, the rankings are much less stable when different value judgements are employed (i.e. the correlation between $I_{1,1}$ and $I_{1,3}$ is significantly lower at 0.54). Northern European countries have lower average values of inequality in education for $I_{3,1}$ ($\tilde{I}_{3,1} = 0.38$) and higher for $I_{1,3}$, ($\tilde{I}_{1,3} = 0.55$), whereas the reverse is true

[9] These are large tables of comparisons of each country against each other. They are available on demand.

for Southern European countries ($\tilde{I}_{3,1} = 0.61 > \tilde{I}_{1,3} = 0.52$). This suggests that in Northern Europe the key driver of educational inequality is the higher end of the distribution, whereas in Southern Europe it is inequality in the lower end of the distribution that is the most important. Consistent with Balestra and Ruiz (2014), Austria is the most equal country in education and the Netherlands and Belgium in life satisfaction. However, for education, we obtain very different results than theirs when the reference point is not the median, but maximum status i.e. the Cowell-Flachaire family of indices (Figure 3)[10] For life satisfaction, the results remain the same as with AF approach. Thus life satisfaction inequality shows more stability than educational inequality with respect to changes in the reference point.

OECD countries appear very heterogenous when it comes to joint inequality and welfare in education (Figure 4 (a)). New Zealand and the USA dominate the vast majority of OECD countries, 33 and 32 of them respectively, and the two are not comparable with each other i.e. their H^+ curves cross. For life satisfaction inequality and welfare, the differences between countries' curves appear smaller (Figure 4 (b)). Some countries (Portugal, Greece, Mexico) that have the highest educational inequality and lowest welfare combined, fare very well in life satisfaction dimension. Eastern Europe has the worst outcomes for joint welfare and inequality in life satisfaction. It is important to note that we should naturally except more dominance relationships based on H^+ dominance than FSD dominance. This is because the former criterion requires robustness across welfare measures that react with higher values both to upward shifts and Hammond transfers, whereas FSD requires agreement over *all* welfare function increasing on upward shifts (regardless of how they react to Hammond transfers). Indeed, in Table 3 we see that the USA does not dominate any country in either education or life satisfaction, whereas according to Definition 6 it dominates most other OECD countries in education. Secondly, if country A first-order dominates B then A will also dominate B in terms of Definition 6, yet the reverse is not true. Therefore all FSD relationships that we found in Table 3 get "pasted" onto Figures 4 (a) and (b).[11]

In $4,4\%$ of comparisons we find bivariate dominance (B, S, C or N together) (Table 6). As can be seen the bivariate dominance is much rarer than univariate dominance relationships. Bivariate dominance is mainly driven by three countries which are dominated by most other countries (Greece, Hungary and Portugal). The importance of bivariate analysis is best seen in 7 cases in which despite univariate dominance in each dimension (N in Table 6), no bidimensional dominance is found. Unidimensional dominance relationships are only a necessary but insufficient condition for bidimensional dominance, therefore analysing dimensions separately one may wrongly conclude joint welfare being higher or lower. Southern Europe appears to have the lowest multidimensional welfare among OECD countries.

Since we analyse the joint distribution of education and life satisfaction in

[10] An individual enjoys a maximum status if everyone else is either in categories below hers or in her same category.

[11] We can also visualise the unidirectional relationship between FSD and Definition 6 by noting that: $H^+_{f^s}(i) = H^+_{F^s}(i) = \sum_{k=1}^{i-1} 2^{i-k-1} F^s_k + F^s_i$. Hence FSD implies Definition 6, but the reverse is not true.

OECD countries, our paper also relates to the literature on the education-happiness relationship. The relationship between education and happiness is admittedly important, because if education and happiness are positively associated then a seemingly natural hypothesis is that it might be possible to influence a nation's happiness through the design of educational policies.[12] Therefore, this relationship has been broadly analysed in various domains of social science (e.g. Michalos 1985; Di Tella et al. 2001; Becchetti et al. 2006; Cunado and de Garcia 2012 amongst others) and medicine (e.g. Stewart-Brown et al. 2015 and papers cited therein). Already thirty years ago, in a first meta-analysis on the topic, Witter et al. (1984) found that the so-called education-happiness gradient is positive, i.e. education is significantly positively related to subjective well-being. However, the literature that followed is inconclusive. Becchetti et al. (2006) find a positive effect of education on happiness, whereas other studies find insignificant (e.g. Inglehart and Klingemann 2000) or negative results (Clark and Oswald 1996). More recent studies are inconclusive too. Using a broader spectrum of education levels for Spain, Cunado and de Garcia (2012) find that controlling for labour and income situation, as well as other socio-economic variables, education has a positive relationship to happiness. On the other hand, utilising a British Cohort Study, Layard et al. (2014) find that, among socio-economic controls, education is the least important predictor of life-satisfaction. Such inconclusiveness may be partially due to the arbitrary cardinalisation of education and happiness indicators imposed by probit or logit regressions. This is essentially the problem highlighted by Bond and Lang (2014).

Here we analyse the implications of the education-happiness relationship on cross-country welfare comparisons independent of cardinalisation. As pointed out by Duclos and Echevin (2011) in the health-income context, dominance comparisons are more general than the analysis of education-happiness gradients as they take into account changes in the whole distribution. Findings in Table 6 are consistent with the inconclusive relationship between education and happiness found in the literature. If any form of association between education and happiness, positive or negative, is not robust to alternative scales in most countries then it should not be a surprise to find lack of agreement in ranking them among two sets of bivariate welfare indices: those that increase when the two variables are more positively associated (i.e. characterised by ALEP complementarity), and those that decrease in the same situation (characterised by ALEP substitutability). This is exactly reflected in Table 6 where B, S, and C dominance relationships are very scarce, in fact less than $4,4\%$ as the latter number includes 7 cases of multiple univariate dominance unmatched by bivariate joint dominance.

4 Conclusions and discussion

We collect in one place and review the theoretical results concerning welfare and inequality measurement for ordinal data. Then, using these results we compare OECD countries in terms of univariate welfare and inequality separately, inequality and welfare together, and joint bivariate welfare. This has been recently studied by Balestra and Ruiz (2014), but our analysis uses a much more extensive toolkit and thus offers a more in-depth view. We use several methods and measures, but

[12] Obviously, causality may run in both directions, therefore the use of the word "hypothesis".

some patterns emerge. Northern European countries occupy most often better positions in the rankings. This is true for FSD, AF_α, H^+ dominance relationships and for inequality measures. On the contrary, Southern European countries appear most often in the least favourable parts of the rankings. These groups of countries have also different patterns of educational inequality; namely, in Northern Europe above-median inequality is higher than its average value for all OECD countries, whereas in Southern Europe below-median inequality is higher than its average value for all OECD countries. Greece and Portugal have also the lowest bidimensional welfare, followed by Spain. Eastern Europe (Estonia, Hungary, Russia) appears to have one of the worst results in life satisfaction, for both welfare and inequality. Varying the reference point in the inequality dominance approach (AF) substantially improves the conclusiveness of the method. Life satisfaction inequality takes place mostly in the higher end of the distribution. We stress the importance of multidimensional analysis: in most cases unidimensional dominance relationships do not imply bidimensional dominance. Our welfare dominance results enable us to unambiguously declare that some countries have lower welfare than other countries, and this conclusion is very robust.

Most of the authors in social choice theory have turned to the distribution approach presented in this paper to solve the problem of scaling when measuring welfare and inequality with ordinal data. This approach has certainly its limitations. Firstly, FSD, AF_α, and H^+ dominance relationships may yield inconclusive results when no dominance is found between two countries. For example, 42% of comparisons in education are found to be FSD dominance. This provides limited information about the welfare of all countries. That is, we are unable to provide a complete ranking of countries in terms of educational welfare. This can be partially remedied by using Copeland scores which rank countries according to their tendency to outperform other countries in dominance relationships. However, some countries have the same Copeland score so they still remain incomparable.

Secondly, in order to obtain full conclusiveness and complete rankings we resort to indices. This comes at the cost of additional assumptions underlying these measures. The rankings obtained in this way rely on a weighting scheme. Thirdly, as expected, inconclusiveness becomes much more pronounced with increasing number of dimensions. For two dimensions we find welfare dominance only in $4,4\%$ of the cases. Although this is inconvenient, it also reveals that one should be careful drawing conclusions about welfare based on a single dimension or separate treatment of dimensions. Such conclusions usually do not hold when outcomes are treated in a truly joint manner. Dependence between dimensions significantly changes welfare comparisons.

Finally, the dominance approach does not provide information on whether a given country is much or only slightly better than the dominated country. Therefore, country A may dominate country B for education and the reverse might be true for life satisfaction, but we will not know how large is the difference in each case unless we resort to indices (with their aforementioned concomitant costs).

Acknowledging the limitations of the current approach, it is worth mentioning that a different approach to handling ordinal data has been proposed in a series of papers by M. Fattore and F. Maggino and other col- leagues (Fattore et al. 2011a, 2011b, 2012, Fattore and Maggino 2014, Fattore 2017).13 It exploits the fact that the multidimensional dataset has a partial order structure. The authors use the results from partial-order theory to build syn- thetic indicators that achieve full comparability without scaling ordinal variables and aggregating them. If a given profile (e.g. educational attainment takes the value “high school” and life satisfaction equals “bad”) is considered a benchmark/a reference point, then one counts over linear extensions (i.e. complete ranking of alternatives, preserving the orders of the original partial order) how frequently a profile is classified below or above a reference point. For each linear extension one can use unidimensional inequality indices. Some regularity needs to be preserved (i.e. in line with inequality measurement axioms) when either the structure of the partial order or the probability distribution over it changes. To offer brief comparison with the approach presented here, as already mentioned, most of the results presented in this paper are for the unidimensional case with Yalonetzky (2013) as the only exception so far. The standard approach in the social choice theory (and the one followed by Yalonetzky 2013) is to take the partial-order structure as given and to exploit various partial-order relations defined on distributions (or variables in a cardinal setting) that preserve inequal- ity/welfare axioms. Complete conclusiveness is achieved via constructing measures of inequality that are consistent with a partial order. Robustness is provided by the characterization of measures. Different measures will produce different results; however, characterization ensures that one understands the properties and value judgements behind these measures. That is, normative valuation is inherently embedded in the measurement exercise.

Multidimensionality remains the most important open problem in the field presented in this article. This concerns not only ordinal variables but also the combination of ordinal and cardinal variables. Another problem is to formally differentiate between inequality and polarization in an ordinal setting. Given the recent contributions of Gravel et al. (2015), Cowell and Flachaire (2017), and Cowell, Kobus and Kurek (2017) who develop approaches to ordinal inequality different than Allison and Foster (2004) bi-polarization ordering, it seems likely that this problem will be dealt with soon. A different problem within the AF approach is to offer definitions of AF criterion that are independent of the median. AF_α is one such definition, but others have been very recently offered too (Sarkar and Satra 2017).

That said, the results developed in the last decade in the field of welfare and inequality measurement for ordinal data allow us to analyse inequality and welfare without the assumptions of a parametric form for the distributions underlying the observed data. As pointed out by Bond and Lang (2014) these assumptions are essentially another form of arbitrary scaling of data that are scale free. Thus the conclusions we obtain using dominance concepts and measures for ordinal data are more robust than imposing some form of scaling.

References

Abul Naga R H, Yalcin T. 2008. Inequality measurement for ordered response health data, Journal of Health Economics; 27(6):1614-1625.
Abul Naga R H, Stapenhurst Ch. 2015. Estimation of inequality indices of the cumulative distribution function, Economics Letters 130:109-112.

Alkire S, Foster J, Seth S, Santos M E, Roche J M, Ballon P. 2015. Multidimensional Poverty Measurement and Analysis, Oxford, Oxford University Press.
Allison R A, Foster J E. 2004. Measuring health inequality using qualitative data, Journal of Health Economics; 23(3):505-524.
Andreoli F, Zoli C. 2014. Measuring dissimilarity, WP Series Department of Economics University of Verona WP 23.
Apouey B. 2007. Measuring health polarisation with self-assessed health data, Health Economics 16:875-894.
Atkinson A B, Bourguignon F. 1982. The Comparison of MultiDimensioned Distributions of Economic Status, Review of Economic Studies, 49(2):183-201.
Balestra C, Ruiz N. 2015. Scale-Invariant Measurement of Inequality and Welfare in Ordinal Achievements: An Application to Subjective Well-Being and Education in OECD Countries, Social Indicators Research 123: 479-500.
Becchetti L, Castriota S, Londono D. 2006. Income, relational goods and happiness, CEIS WP No. 227.
Bennett C, Hatzimasoura C. 2011. Poverty measurement with ordinal data. Institute for International Economic Policy, IIEP-WP-2011-14.
Berry K J, Mielke P W. 1992. Indices of ordinal variation, Perceptual and Motor Skills 74:576-578.
Blair J, Lacy M G. 2000. Statistics of ordinal variation, Sociological methods and research, 28: 251-279.
Bond T, Lang K. 2013. The Evolution of the Black-White Test Score Gap in Grades K–3: The Fragility of Results, The Review of Economics and Statistics 95(5):1468-1479.
Bond T, Lang K. 2014. The Sad Truth About Happiness Scales, NBER Working Papers 19950, National Bureau of Economic Research, Inc.
Carlsen, L. 2018. Happiness as a sustainability factor. The world happiness index: a posetic - based data analysis. Sustainability Science 13:549-571.
Clark A E, Oswald A J. 1996. Satisfaction and comparison income, Journal of Public Economics 61.
Copeland A. H. 1951. A "reasonable" social welfare function, University of Michigan Seminar on Applications of Mathematics to the Social Sciences.
Cowell F, Flachaire E. 2017. Inequality with ordinal data, Economica 84:290-321.
Cuhadaroglu T. 2013. My group beats your group: evaluating non-income inequalities, University of St Andrews Discussion Paper np. 1308.
Cunado J, Gracia F. 2012. Does Education Affect Happiness? Evidence for Spain, Social Indicators Research 108(1): 185-196.
Di Tella R, MacCulloch R, Oswald A. 2001. Preferences over inflation and unemployment: evidence form surveys of happiness. The American Economic Review 91 (1), 335 – 341.
Duclos, J.Y. and D. Echevin. 2011. Health and Income: a Robust Comparison of Canada and the US, Journal of Health Economics 30 (2): 293-302.
Dutta I, Foster J. 2013. Inequality of happiness in the U.S.: 1972-2010, Review of Income and Wealth 59(3): 393-415.
Esteban J, Ray D. 1994. On the measurement of polarization, Econometrica 62:819-852.
Fattore M, Maggino F. 2014. Partial Orders in Socio-economics: A practical challenge for poset theorists or a cultural challenge for social scientists?, In: Brüggemann R et al. (eds.), Multi-indicator Systems and Modelling in Partial Order, Chapter 9, Springer Science + Business Media New York.

Fattore M, Brüggemann R, Owsiński J. 2011. Using poset theory to compare fuzzy multidimensional material deprivation across regions. In: Ingrassia S et al. (eds.), New perspectives in statistical modelling and data analysis, Springer: Berlin, Heidelberg.
Fattore M, Maggino F, Greselin F. 2011. Socio-economic evaluation with ordinal variables: integrating counting and poset approaches, Statistica & Applicazioni, Special Issue 31-42.
Fattore M, Maggino F, Colombo E. 2012. 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.
Fattore M. 2017. Functionals and synthetic indicators over finite posets", In M. Fattore, R. Bruggemann (eds.), Partial Order Concepts in Applied Sciences, Springer.
Fleurbaey M, Blanchet D. 2013. Beyond GDP. Measuring welfare and assessing sustainability, Oxford University Press.
Gravel N, Magdalou B, Moyes P. 2015. Ranking distributions of an ordinal attribute, https://hal.archives-ouvertes.fr/halshs-01082996/document
Hammond P J. 1976. Equity in two person situations: Some consequences. Econometrica 47: 1127–1135.
Inglehart R, Klingemann H D. 2000. Genes, culture, Democracy and Happiness. In E. Diener & E. Suh, (Eds.) Subjective well-being across cultures (pp. 165-183). Cambridge: MIT Press.
Kannai Y. 1980. The ALEP definition of complementarity and least concave utility functions, Journal of Economic Theory 22(1):115-117.
Kobus M, Miłoś P. 2012. Inequality decomposition by population subgroups for ordinal data, Journal of Health Economics 31:15-21.
Kobus, M. 2015. PolariSation measurement for ordinal data, Journal of Economic Inequality 13(2): 275-297.
Layard R. 2005. Happiness: lessons from a new science, New York, Penguin.
Layard, R, Clark A E, Cornaglia F, Powdthavee N, Vernoit J. 2014. What Predicts a Successful Life? A Life-course Model of Well-being, Economic Journal 124: F720-738.
Lazar A. Silber J. 2013. On the cardinal measurement of health inequality when only ordinal information is available on individual health status, Health Economics 22(1):106-113.
Lv G, Wang Y, Xu Y. 2015. On a new class of measures for health inequality based on ordinal data, Journal of Economic Inequality 13:465-477.
Makdisi P, Yazbeck M. 2014. Measuring socioeconomic health inequalities in the presence of multiple categorical information, Journal of Health Economics 34: 84-95.
Mendelson H. 1987. Quantile-preserving spread, Journal of Economic Theory 42: 334-351.
Michalos A C. 1985. Multiple Discrepancies Theory, Social Indicators Research 16: 347-413.
Nielsen E.R. 2015. Achievement Gap Estimates and Deviations from Cardinal Comparability, Finance and Economics Discussion Series 2015-040. Washington: Board of Governors of the Federal Reserve System,
http://dx.doi.org/10.17016/FEDS.2015.040.
Sarker S, Santra S. 2017. Extending the approaches to inequality ordering of ordi-

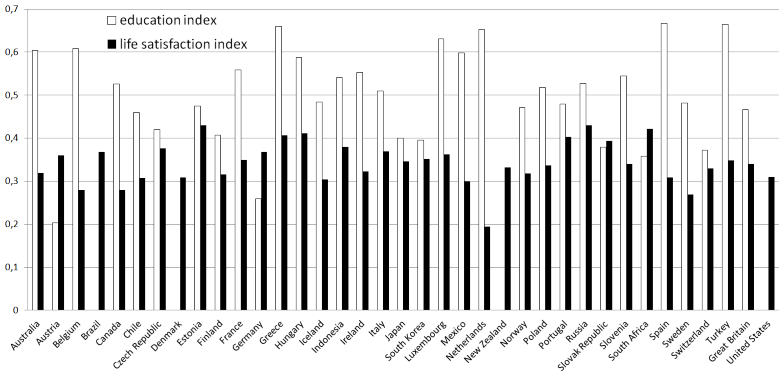
nal variables, http://www.ecineq.org/ecineq_nyc17/FILESx2017/CR2/p105.pdf
Sen A K. (1973). On economic inequality. Oxford, UK: Clarendon Press.
Shorrocks A F. 1984. Inequality decomposition by population subgroups, Econometrica 52(6):1369-85.
Sonne Schmidt C, Osterdal L P, Tarp F. 2016. Ordinal bivariate inequality: concepts and application to child deprivation in Mozambique, Review of Income and Wealth 62(3):559-573.
Silber J, Yalonetzky G. 2011. Measuring Inequality in Life Chances with Ordinal Variables, in Juan Gabriel Rodríguez (ed.) Inequality of Opportunity: Theory and Measurement (Research on Economic Inequality, Volume 19) Emerald Group Publishing Limited, 77 - 98.
Stewart-Brown S, Samaraweera P C, Taggart F,, Kandala N B, Stranges S. 2015. Socioeconomic gradients and mental health: implications for public health, Br J Psychiatry 206(6): 461-465.
Stiglitz J E, Sen A K, Fitoussi J P. 2009. Report by the Commission on the Measurement of Economic Performance and Social Progress, www.stiglitz-sen-fitoussi.fr
UNDP. 2014. Human development report 2014 - sustaining human progress: reducing vulnerabilities and building resilience, United Nations Development Programme, http://hdr.undp.org/en/global-reports
Yalonetzky G. 2012. Poverty measurement with ordinal variables: a generalisation of a recent contribution, ECINEQ WP 2012-246.
Yalonetzky G. 2013. Stochastic Dominance with Ordinal Variables: Conditions and a Test, Econometric Reviews, 32:1, 126-163
Wagstaff, A., P. Paci, and E. van Doorslaer. 1991. On the measurement of inequalities in health. Social Science and Medicine 33:545–557.
Wagstaff, A. and E. van Doorslaer. 1994. Measuring inequalities in health in the presence of multiple-category morbidity indicators. Health Economics 3(4):281–291.
Witter R, Okun M A, Stock W, Haring M. 1984. Education and Subjective Well-Being: A Meta-Analysis, Educational Evaluation and Policy Analysis 6(2): 165-173.
Wolfson M C. 1994. When inequalities diverge, American Economic Review P&P 94:353-358.
Zheng B. 2008. Measuring inequality with ordinal data: a note, Research on Economic Inequality 16:177-188.
Zheng B. 2011. A new approach to measure socioeconomic inequality in health, Journal of Economic Inequality 9:555-577.

Appendix A

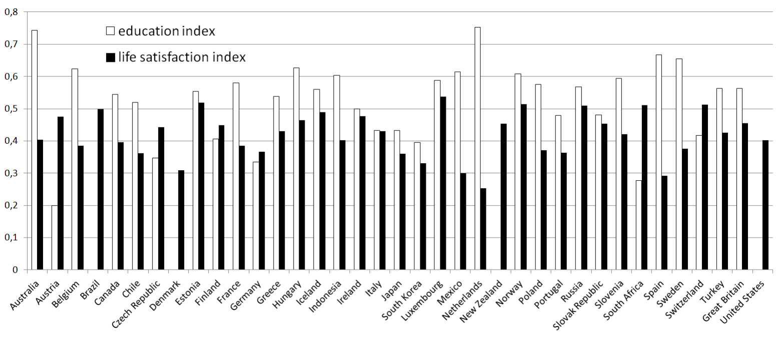
Table 1: Description of dimensions

Dimension	Indicator	Level	Construction
Education	Highest educational level attained	1	Primary school or no education
		2	Technical school
		3	Secondary school
		4	University
Life satisfaction	Self-reported satisfaction with life on a scale of 1 to 5	1	Completely dissatisfied
		2	
		3	
		4	
		5	Completely satisfied

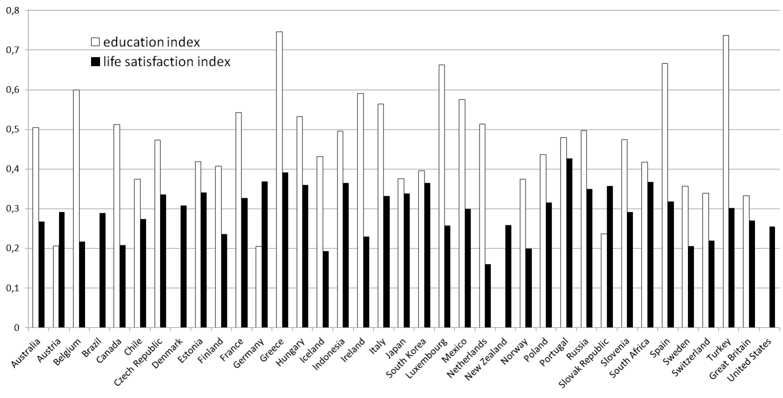
Figure 1: The values of $I_{a,b}$ indices



(a) for a=1, b=1



(b) for a=1, b=3



(c) for a=3, b=1

Notes: Figure gives the values of $I_{a,b}$ family of indices for the cases when inequality below and above the median is treated equally $I_{1,1}$, when inequality below the median matters more $I_{3,1}$ and when inequality above the median matters more $I_{1,3}$.

Table 2: List of abbreviations

Abbreviation	Country	sample size
AUS	Australia	1034
AUT	Austria	1508
BEL	Belgium	1507
BRA	Brazil	1468
CAN	Canada	2132
CHL	Chile	969
CZE	Czech Republic	1800
DNK	Denmark	1485
EST	Estonia	1527
FIN	Finland	1116
FRA	France	1495
DEU	Germany	2066
GRC	Greece	1496
HUN	Hungary	1002
ISL	Iceland	786
IDN	Indonesia	1884
IRE	Ireland	992
ITA	Italy	1470
JPN	Japan	2331
KOR	South Korea	1180
LUX	Luxembourg	1584
MEX	Mexico	1926
NLD	Netherlands	1858
NZL	New Zealand	784
NOR	Norway	1087
POL	Poland	961
PRT	Portugal	1541
RUS	Russia	2452
SVK	Slovakia	1480
SVN	Slovenia	1049
ZAF	South Africa	2809
ESP	Spain	1151
SWE	Sweden	1176
CHE	Switzerland	1259
TUR	Turkey	1510
GBR	United Kingdom	1477
USA	United States	2208

Table 3: Cumulative outcomes of FSD comparisons and Copeland score

	EDUCATION				LIFE SATISFACTION			
	dominates	dominated	score	rank	dominates	dominated	score	rank
AUS	17	2	0.405	2	5	0	0.135	5
AUT	1	0	0.027	10	5	4	0.027	8
BEL	6	5	0.027	10	9	0	0.243	2
BRA	0	25	-0.676	21	7	2	0.135	5
CAN	7	5	0.054	9	11	0	0.297	1
CHL	1	10	-0.243	19	5	0	0.135	5
CZE	1	0	0.027	10	5	5	0.000	9
DNK	5	0	0.135	7	9	0	0.243	2
EST	12	0	0.324	3	0	17	-0.459	18
FIN	17	0	0.459	1	8	0	0.216	3
FRA	6	5	0.027	10	2	7	-0.135	13
DEU	7	0	0.189	5	1	11	-0.270	16
GRC	1	9	-0.216	18	0	13	-0.351	17
HUN	1	11	-0.270	20	0	19	-0.514	19
ISL	9	3	0.162	6	8	0	0.216	3
IDN	5	4	0.027	10	1	7	-0.162	14
IRE	2	6	-0.108	15	8	0	0.216	3
ITA	1	7	-0.162	17	1	3	-0.054	11
JPN	5	3	0.054	9	1	2	-0.027	10
KOR	12	0	0.324	3	0	2	-0.054	11
LUX	3	5	-0.054	13	2	0	0.054	7
MEX	2	7	-0.135	16	2	0	0.054	7
NLD	3	4	-0.027	12	3	0	0.081	6
NZL	9	0	0.243	4	1	0	0.027	8
NOR	6	0	0.162	6	7	0	0.189	4
POL	1	3	-0.054	13	1	1	0.000	9
PRT	0	8	-0.216	18	0	4	-0.108	12
RUS	3	0	0.081	8	0	8	-0.216	15
SVK	0	0	0.000	11	0	1	-0.027	10
SVN	0	2	-0.054	13	0	0	0.000	9
ZAF	0	3	-0.081	14	0	1	-0.027	10
ESP	0	4	-0.108	15	0	1	-0.027	10
SWE	1	1	0.000	11	0	0	0.000	9
CHE	1	1	0.000	11	1	0	0.027	8
TUR	0	1	-0.027	12	0	0	0.000	9
GBR	0	1	-0.027	12	0	0	0.000	9
USA	0	0	0.000	11	0	0	0.000	9

Notes: For a list of abbreviations refer to Table 2. Table 3 gives the outcomes of FSD comparisons and Copeland scores for each country (a lower rank means greater tendency to outperform other countries). The Copeland score is the number between $(-(x-1), x-1)$, where x is the number of countries a given country might dominate. Here the score is normalised to $(-1, 1)$. Countries that have the same value of the score are assigned the same rank.

Table 4: Outcomes of α-quantile dominance

	EDUCATION									LIFE SATISFACTION								
Percentiles	1/5	1/4	1/3	2/5	1/2	3/5	2/3	3/4	4/5	1/5	1/4	1/3	2/5	1/2	3/5	2/3	3/4	4/5
AUS	1	1	1	1	1	1	2	7	11	1	1	4	4	4	4	3	3	6
AUT	1	3	4	7	14	16	18	12	6	1	1	1	1	1	1	2	6	9
BEL	2	2	1	1	1	2	1	3	5	1	3	6	6	6	6	1	7	11
BRA	5	4	2	1	1	2	1	1	1	1	1	1	1	1	1	2	5	9
CAN	3	3	2	1	2	3	2	3	7	2	4	7	7	7	6	1	9	14
CHL	7	7	6	6	3	2	2	1	1	7	3	4	6	6	6	5	2	4
CZE	11	9	9	7	2	5	9	6	4	1	1	2	2	2	2	1	2	3
DNK	3	3	2	1	1	1	1	3	5	2	2	2	2	1	2	6	10	14
EST	3	3	1	2	2	2	1	6	10	1	1	1	1	1	2	2	2	2
FIN	1	1	1	1	1	2	4	10	16	1	1	1	1	1	1	4	10	14
FRA	1	1	1	1	1	1	1	3	4	2	2	1	2	2	2	2	2	1
DEU	2	4	5	8	11	10	10	4	4	1	1	1	1	1	1	1	1	1
GRC	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
HUN	1	1	2	3	1	2	2	2	2	2	1	1	1	1	1	1	1	1
ISL	2	2	1	1	1	2	1	7	11	3	3	3	3	3	1	7	14	18
IDN	1	1	1	1	1	1	1	4	7	3	3	2	1	1	1	1	1	1
IRE	2	2	2	1	2	2	3	1	2	1	1	1	1	1	1	6	13	18
ITA	3	3	2	1	5	7	10	8	5	1	1	2	2	2	2	1	1	3
JPN	1	1	2	3	4	5	5	2	4	4	4	2	2	2	2	2	2	1
KOR	1	1	1	1	1	2	5	11	17	6	6	4	1	1	1	1	1	1
LUX	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3	7	11
MEX	1	1	1	2	1	1	1	1	1	1	1	1	1	1	2	5	9	13
NLD	1	1	1	1	1	1	1	2	3	7	11	20	25	25	24	17	7	2
NZL	4	4	4	4	1	4	8	15	22	1	1	2	2	2	2	3	8	12
NOR	1	1	1	2	2	1	4	10	14	1	1	1	1	1	3	10	19	24
POL	3	3	3	2	1	1	2	1	1	6	6	2	3	3	3	3	3	1
PRT	6	5	3	2	2	1	1	1	1	2	2	2	1	1	1	1	1	1
RUS	6	6	4	1	2	3	2	4	6	1	1	1	1	1	1	1	1	1
SVK	17	16	15	12	3	2	7	5	3	1	1	1	1	1	1	1	1	2
SVN	1	1	1	1	1	1	1	1	1	5	2	2	2	2	2	1	4	8
ZAF	1	1	1	2	5	7	9	6	6	1	1	1	1	1	1	1	1	2
ESP	5	4	2	1	1	1	1	1	1	12	10	5	2	3	3	3	3	2
SWE	1	1	1	1	1	1	3	7	12	2	4	7	7	7	7	2	8	13
CHE	1	1	2	3	6	7	7	2	4	2	2	2	2	2	1	6	12	16
TUR	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	1	2	5
GBR	7	7	6	5	2	1	2	1	2	1	1	2	2	2	2	1	5	8
USA	1	1	1	1	3	6	10	16	23	2	1	3	3	3	3	1	6	9
SUM	110	108	95	92	90	109	141	170	225	89	88	102	103	103	103	110	190	262
PERCENT	0,17	0,16	0,14	0,14	0,14	0,16	0,21	0,26	0,34	0,13	0,13	0,15	0,15	0,15	0,15	0,17	0,29	0,39

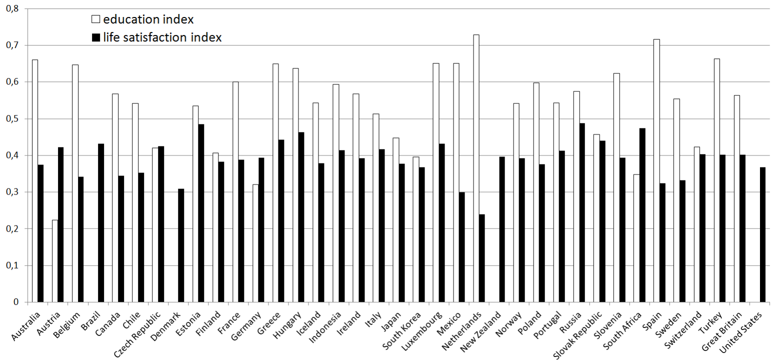
Notes: For a list of abbreviations refer to Table 2. The table gives for each value of α the number of countries a given country dominates in terms of AF_α dominance (Definition 4). PERCENT gives the number of dominances for a given α divided by the number of all potential dominances i.e. 666. If country 1 dominates country 2 for at least one value of α, then country 2 does not dominate country 1 for any α.

Table 5: Outcomes of AF_α comparisons (for all α) and Copeland score

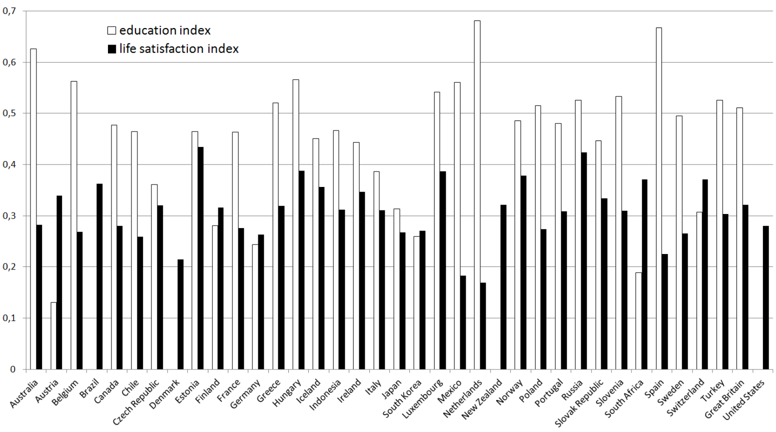
	EDUCATION				LIFE SATISFACTION			
	dominates	dominated	score	rank	dominates	dominated	score	rank
AUS	10	10	0.000	16	8	13	-0.139	19
AUT	21	0	0.583	3	8	16	-0.222	22
BEL	6	17	-0.306	22	15	5	0.278	9
BRA	5	0	0.139	13	8	10	-0.056	17
CAN	10	14	-0.111	18	19	2	0.472	4
CHL	8	20	-0.333	23	14	5	0.250	10
CZE	21	1	0.556	4	3	22	-0.528	25
DNK	6	0	0.167	12	14	0	0.389	6
EST	13	4	0.250	10	1	0	0.028	15
FIN	15	2	0.361	7	13	3	0.278	9
FRA	3	21	-0.500	26	2	24	-0.611	28
DEU	17	2	0.417	5	0	3	-0.083	18
GRC	1	9	-0.222	21	0	31	-0.861	31
HUN	3	8	-0.139	19	1	0	0.028	15
ISL	12	9	0.083	15	19	0	0.528	3
IDN	6	20	-0.389	24	2	23	-0.583	27
IRE	4	19	-0.417	25	17	2	0.417	5
ITA	11	3	0.222	11	3	23	-0.556	26
JPN	9	9	0.000	16	4	2	0.056	14
KOR	16	2	0.389	6	5	0	0.139	12
LUX	0	26	-0.722	28	10	7	0.083	13
MEX	1	9	-0.222	21	12	1	0.306	8
NLD	2	20	-0.500	26	24	0	0.667	1
NZL	24	0	0.667	2	12	10	0.056	14
NOR	14	6	0.222	11	23	0	0.639	2
POL	3	23	-0.556	27	7	16	-0.250	23
PRT	5	1	0.111	14	1	1	0.000	16
RUS	12	2	0.278	8	0	2	-0.056	17
SVK	25	0	0.694	1	1	25	-0.667	29
SVN	0	28	-0.778	29	12	5	0.194	11
ZAF	9	0	0.250	10	1	29	-0.778	30
ESP	4	7	-0.083	17	13	0	0.361	7
SWE	11	5	-0.167	12	18	1	0.472	4
CHE	11	8	0.083	15	16	2	0.389	6
TUR	0	28	-0.778	29	5	21	-0.444	24
GBR	8	14	-0.167	20	8	15	-0.194	20
USA	22	1	0.583	3	11	11	0.000	16

Notes: For a list of abbreviations refer to Table 2. Table gives Copeland scores for AF_α comparisons. The Copeland score is the number between $(-(x-1), x-1)$, where x is the number of countries a given country might dominate. Here the score is normalised to $(-1, 1)$. Countries that have the same value of the score are assigned the same rank.

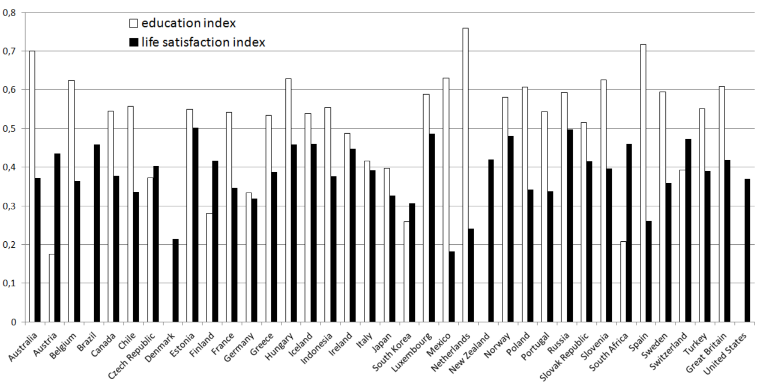
Figure 2: The values of $I_{\alpha,\beta}$ indices



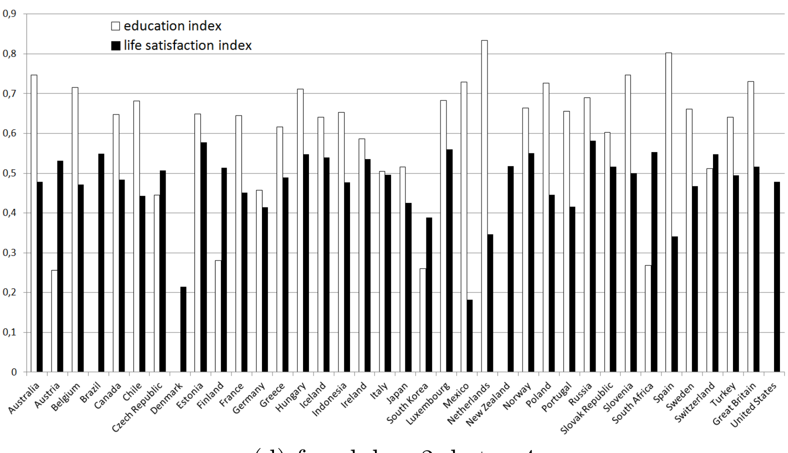
(a) for alpha=1, beta=2



(b) for alpha=2, beta=1



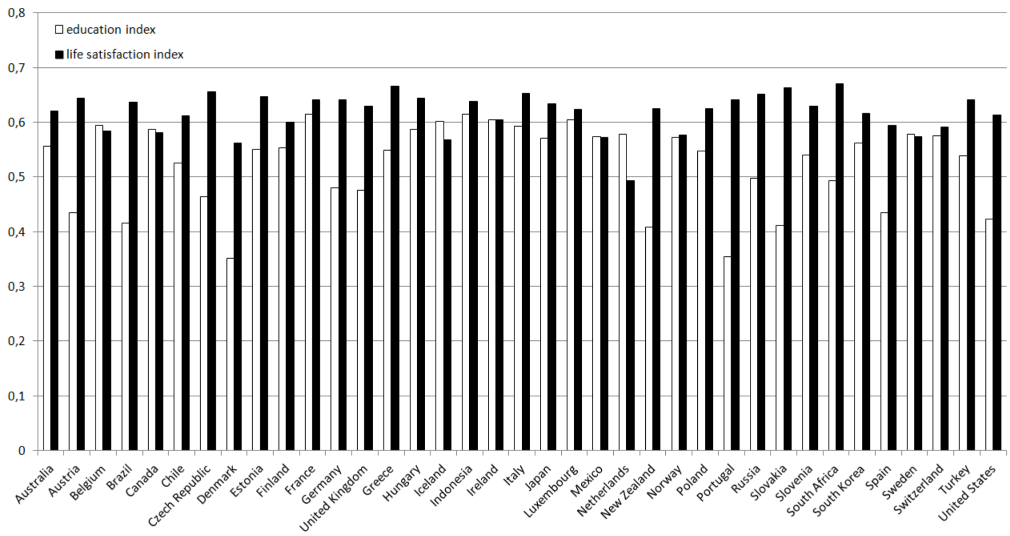
(c) for alpha=2, beta=2



(d) for alpha=2, beta=4

Notes: Figure gives the values of $I_{\alpha,\beta}$ family of indices for the cases when either inequality below or above the median is given higher weight.

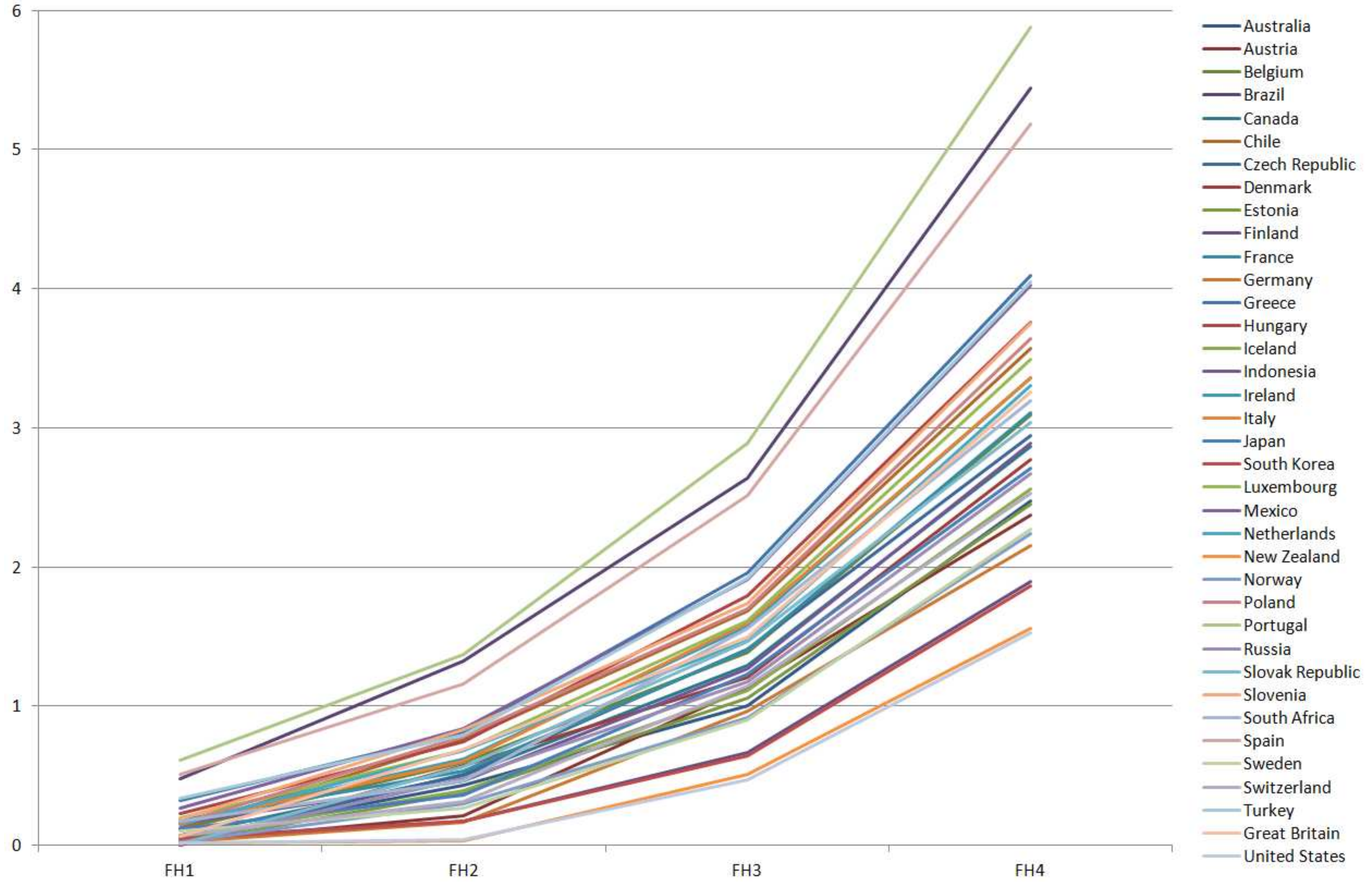
Figure 3: The values of I_α indices



Notes: Figure gives the values of I_α index for the case when $\alpha = 0$.

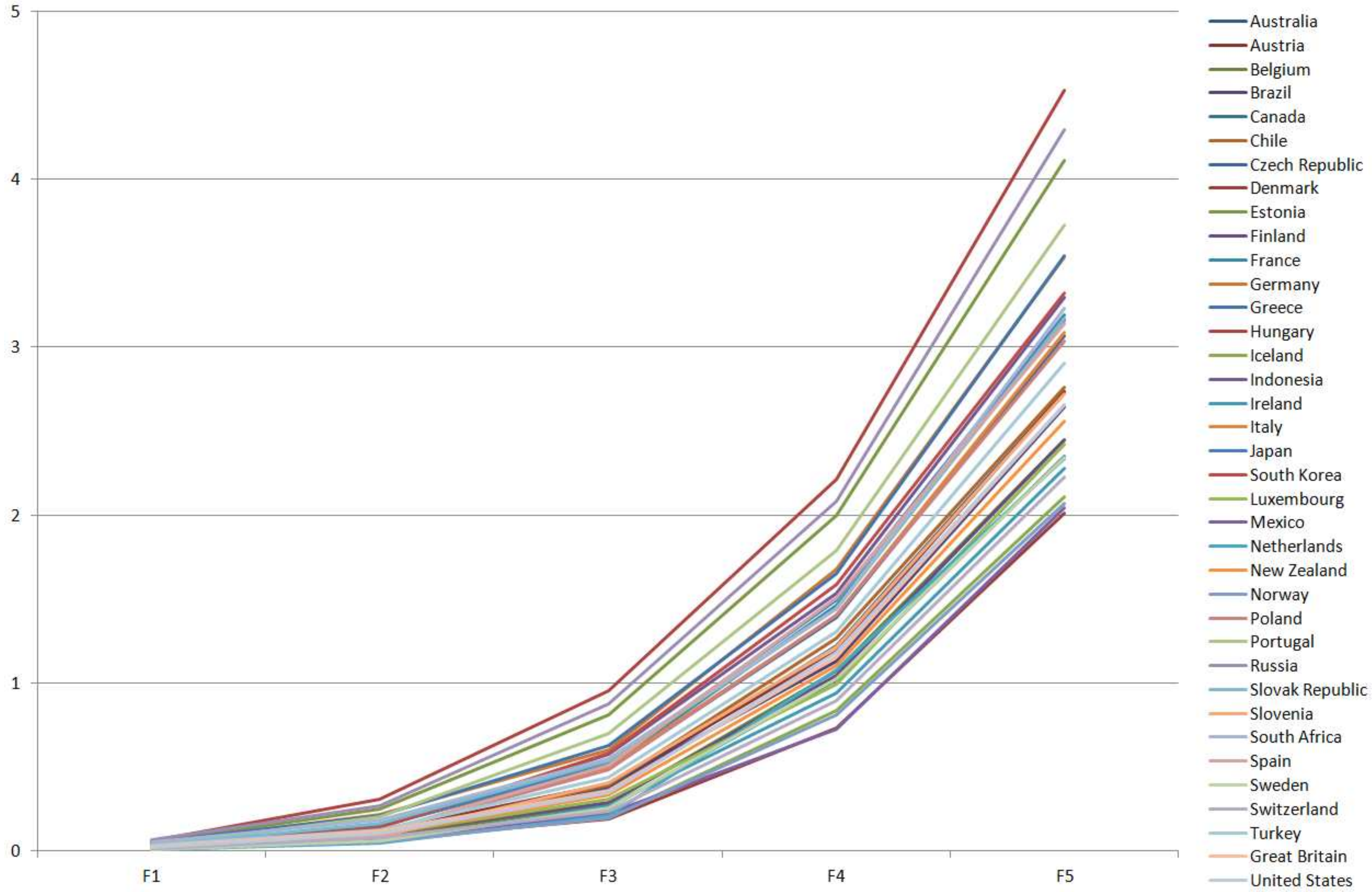
Notes: Figure gives the results of H^+ dominance (Definition 6) which measures both inequality and welfare (Gravel et al. 2015).

Figure 4: H^+-dominance curves for education (a) and life satisfaction (b) for OECD countries



(a) Education

Notes: Figure gives the results of H^+ dominance (Definition 6) which measures both inequality and welfare (Gravel et al. 2015).



(b) Life satisfaction

Table 6: Bivariate dominance

	AUS	AUT	BEL	BRA	CAN	CHL	CZE	DNK	EST	FIN	FRA	DEU	GRC	HUN	ISL	IDN	IRE	ITA	JPN
AUS													NS	NS					
AUT																			
BEL														NC					
BRA																			
CAN													NB	NB					
CHL																			
CZE																			
DNK																			
EST																			
FIN											NS		NB	NS				NS	
FRA														NS					
DEU																			
GRC																			
HUN																			
ISL				NS									NS	NS				NS	
IDN																			
IRE														N					
ITA																			
JPN														NS					
KOR														NB					
LUX																			
MEX																			
NLD																			
NZL											NS		NS	NS		NS		NS	
NOR				NS							NS		NB	NB				NS	
POL																			
PRT																			
RUS																			
SVK																			
SVN																			
ZAF																			
ESP																			
SWE											NS		NB	NB		NB			
CHE													NB	NB				N	
TUR																			
GBR																			
USA											NS		NB	NB					

	KOR	LUX	MEX	NLD	NZL	NOR	POL	PRT	RUS	SVK	SVN	ZAF	ESP	SWE	CHE	TUR	GBR	USA
AUS																		
AUT																		
BEL							N	NS										
BRA																		
CAN							N	NS										
CHL								NS										
CZE																		
DNK																		
EST																		
FIN								NS										
FRA																		
DEU																		
GRC																		
HUN																		
ISL							N	NS								N		
IDN																		
IRE								NS										
ITA																		
JPN																		
KOR																		
LUX																		
MEX																		
NLD								NB					NS					
NZL																		
NOR							NS	NS								NS	N	
POL																		
PRT																		
RUS																		
SVK																		
SVN								NS										
ZAF																		
ESP																		
SWE							NS	NS					NS					
CHE								NS				NB				NS		
TUR																		
GBR																		
USA								NS										

Notes: For a list of abbreviations refer to Table 2. Table gives the results of bivariate dominance comparisons (Section ??). A letter in a cell indicates that a row country dominates the column country according to a given type of dominance. B refers to bivariate dominance in terms of Definition 7 of all forms of ALEP interactions; S refers to Definition 8 involving ALEP substitutability, C refers to Definition 9 involving ALEP complementarity, and N refers to Definition 10 involving ALEP neutrality.

Appendix B

The tests of bivariate dominance conditions developed in Section 2.3 rely on the following statistics: $\Delta S_{12}(ij)$, $\Delta F_{12}(ij)$, $\Delta F_1(i)$ and $\Delta F_2(j)$. One straightforward method is an intersection-union test of the form proposed by Yalonetzky (2013). In the case of the dominance relationship in $FSDAS$, the test's null hypothesis is that $\Delta F_{12}(ij) = 0 \; \forall (i,j) \neq (n_1, n_2)$. The alternative hypothesis is that $\Delta F_{12}(ij) < 0 \; \forall (i,j) \neq (n_1, n_2)$. Now, let $SE(\Delta F_{12}(ij))$ be the standard error of $\Delta F_{12}(ij)$ (i.e. we are comparing samples).[13] Then we know that when the central limit theorem applies, i.e. when the sample size is large enough, $z^{F12}(ij) \equiv \frac{\Delta F_{12}(ij)}{SE(\Delta F_{12}(ij))}$ converges in distribution to a standard normal distribution. We reject the null hypothesis in favour of the alternative if $\max\{z^{F12}(ij)\} \leq -z_\alpha$ where $-z_\alpha$ is the left-tail critical value in a one-tailed test. The overall level of significance of this test is also bound to be α although its size (i.e. the error type I) is very likely to be lower.

In the case of the relationship in $FSDAC$ we use the statistics $z^{S12}(ij) \equiv \frac{\Delta S_{12}(ij)}{SE(\Delta S_{12}(ij))}$, and the alternative hypothesis is of the form: $\Delta S_{12}(ij) > 0 \; \forall (i,j) \neq (0,0)$. We reject the null in favour of the alternative if $\min\{z^{S12}(ij)\} \geq z_\alpha$ where z_α is the right-tail critical value in a one-tailed test. Again, the overall level of significance is also α even though the actual size is lower.

For the relationship in $FSDB$ we test a null hypothesis of $\Delta F_{12}(ij) = 0 \; \forall (i,j) \neq (n_1, n_2)$ and $\Delta S_{12}(ij) = 0 \; \forall (i,j) \neq (0,0)$ against an alternative of $\Delta F_{12}(ij) < 0 \; \forall (i,j) \neq (n_1, n_2)$ and $\Delta S_{12}(ij) > 0 \; \forall (i,j) \neq (0,0)$. We reject the null hypothesis if $\max\{z^{F12}(ij)\} \leq -z_\alpha$ and $\min\{z^{S12}(ij)\} \geq z_\alpha$, again, with an overall level of significance of α.

Finally, in the case of the relationship in $FSDAN$ we use statistics of the form $z^{F1}(i) \equiv \frac{\Delta F_1(i)}{SE(\Delta F_1(i))}$. We test a null hypothesis of $\Delta F_1(i) = 0 \; \forall (i) \neq (n_1)$ and $\Delta F_2(j) = 0 \; \forall (j) \neq (n_2)$ against an alternative of $\Delta F_1(i) < 0 \; \forall (i) \neq (n_1)$ and $\Delta F_2(j) < 0 \; \forall (j) \neq (n_2)$. We reject the null hypothesis if $\max\{z^{F1}(i)\} \leq -z_\alpha$ and $\max\{z^{F2}(j)\} \leq -z_\alpha$, again, with an overall level of significance of α.

[13] For the actual formula of the standard errors of this section's statistics see Yalonetzky (2013).