On the Robustness of Multidimensional Poverty Orderings in the EU

Tomas Zelinsky (Faculty of Economics, Technical University of Kosice) and Francisco Azpitarte (University of Melbourne and Brotherhood of St Laurence)

Paper prepared for the 34th IARIW General Conference

Dresden, Germany, August 21-27, 2016

Session 4F: Meeting the Measurement Challenges of Official Statistics Offices II

Time: Tuesday, August 23, 2016 [Afternoon]

On the Robustness of Multidimensional Poverty Orderings in the EU

Tomáš Želinský
Faculty of Economics, Technical University of Košice

Francisco Azpitarte
University of Melbourne and Brotherhood of St Laurence

Paper prepared for the 34th General Conference of The International Association for Research in Income and Wealth [Dresden, Germany, August 21-27, 2016.]

Abstract

The goal of this paper is to assess the robustness of cross-country and cross-year comparisons in the EU using the official multidimensional poverty measurement framework. The previous research suggests that poverty comparisons are in general sensitive to weights. Different weighting vectors can hence yield different results, and one is not able to claim robustly whether poverty in country A is higher than in country B, or that poverty in a given country has unambiguously declined or increased. We find that approximately 50 per cent of all pair-wise country comparisons are not robust to changes in weights, i.e. it is always possible to find a set of weights for which ranks of countries reverse. Similar results are obtained, when assessing the robustness of poverty comparisons over time. The findings further indicate that rankings of countries are extremely sensitive to change in weights around the official definition of the composite indicator (i.e. weights of all dimensions and threshold equal one third). Our results hence suggest that evaluations of the progress made in alleviating multidimensional poverty in the EU are highly sensitive to the set of weights used to quantify poverty, and that more attention needs to be paid to the checks for sensitivity of poverty comparisons to changes in weights.

Keywords: Multidimensional poverty; counting measures; dominance conditions; EU-SILC
JEL Classification: I32.

This work was supported by the Slovak Scientific Grant Agency as part of the research project VEGA 2/0026/15 and 2015 Group of Eight (Go8) European Fellowship awarded to Tomáš Želinský for visiting the Melbourne Institute of Applied Economic and Social Research, Faculty of Business and Economics, University of Melbourne. This study is based on data from Eurostat, EU-SILC 2004-2013 (Eurostat, 2009a,b, 2010, 2011, 2015a,b,c,d,e,f). The responsibility for all conclusions drawn from the data lies entirely with the authors

1 Introduction

Among the researchers there is a consensus on the multidimensional nature of poverty. Poverty has been recognized as a multidimensional phenomenon by the transnational institutions since the late 1990's (UNDP, 1997; World Bank, 2001), and in a similar fashion Stiglitz et al. (2009) suggest that a multidimensional definition of well-being should be used. Multidimensionality of poverty was recognized even earlier as part of the basic (human) needs approach to development (UNCTAD, 1975) and further elaborated by Streeten et al. (1981). One of the first axiomatic conceptualizations of multidimensional poverty measurement was offered by Bourguignon and Chakravarty (2003), and despite the progress made in this field, there is no single generally accepted multidimensional measure of poverty. Moreover, there are opinions against constructing such measures (Ravallion, 2011).

In the sense of Sen's (1976) tradition conceptualization of poverty measurement is based on two main steps: identification of the poor, and aggregation of such information. In their recent paper Alkire and Foster (2011a) proposed an approach to multidimensional poverty measurement combining a 'counting' based method to identification of the poor (Atkinson, 2003) and FGT poverty measures family (Foster et al., 1984). This family of measures satisfies desirable axiomatic properties, and is easy to implement and interpret.

Identification of the poor within the methodology proposed by Alkire and Foster (2011a) (A-F hereafter) is based on a 'dual cutoff' which depends on the dimensions' weights. Multidimensional poverty measures are often criticized due to arbitrariness in setting dimensions' weights and poverty cutoffs. In this paper we focus on the role of weights in identification and aggregation of the poor, and thus the deprivation cutoffs are considered as exogenous – determined by an authority.

In general the weights should reflect the variables' importance in the index. Theoretical and empirical literature offers a number of approaches to setting weights of dimensions. A detailed classification of approaches to weights setting can be found in a study by Decancq and Lugo (2013) who distinguish between data-driven approaches (frequency-based, statistical and most favorable weights), normative approaches (equal or arbitrary, expert opinion and price-based weights), and hybrid approaches (stated preference and hedonic weights). As there is no universally accepted guideline for setting the weights, the selection should be open to criticism in order to gain reasonable public acceptance (Azevedo and Robles, 2013). The arbitrary equal weighting of all dimensions is the most commonly used approach in the empirical applications and is often defended by its simplicity (Alkire and Santos, 2014), but at the same time equal weighting is universally considered to be wrong (Chowdhury and Squire, 2006). It is thus obvious that ranking the alternatives based on multidimensional poverty indices depends on the weights of dimensions, and the chosen weights can have a significant effect on the ranking of various alternatives (Athanassoglou, 2015). Cox et al. (1992) defend the simple weighting schemes and methods of analysis, and they argue that many previously published weighting schemes were based on too complex multivariate methods concluding that sensitivity analyses were highly desirable.

The previous research suggests that poverty comparisons are in general sensitive to weights (Alkire et al., 2015). Different weighting vectors can hence yield different results, and one is not able to claim robustly whether e.g. poverty in country A is higher than in country B, or that poverty in a given country has unambiguously declined or increased. Thus, aggregation of the weighted dimensions should be followed by an examination of the given ranking robustness to changes in the weights. Methodological literature provides a number of approaches to robustness checks such as stochastic dominance concepts (Duclos et al., 2006; Mehmet Pinar, 2013), Monte Carlo simulations (Saisana et al., 2005), linear programming tools (Cherchye et al., 2008), models from decision theory suggested by Foster et al. (2013) and generalized by Permanyer (2011, 2012), approach based on Kemeny's (1959) rule from social choice theory (Athanassoglou, 2015).

The goal of this paper is to assess the extent to which the cross-country and cross-year comparisons in Europe are robust to changes in weights. Our empirical application is based on European Union Statistics on Income and Living Conditions (EU-SILC) microdata adopting the counting approach to multidimensional poverty measurement proposed by Alkire and Foster (2011a). Two main analyses are performed in the study. First, we check the robustness of cross-country and cross-year comparisons in the EU for all pair-wise comparisons (cross-country as well as cross-year) by testing the necessary and sufficient dominance conditions recently proposed in the literature by Gallegos et al. (2016). Secondly, we focus on the pair-wise comparisons for which dominance cannot be assumed. We analyse the maximum change in weights that preserves the initial ranks in pair-wise comparisons.

A number of empirical studies analysing multidimensional poverty/well-being indicators include robustness assessments, however they are usually based on a very limited number of weighting vectors. Alkire and Santos (2014) present a multidimensional poverty index with equally weighted dimensions and perform robustness checks comparing three alternative weighting structures, applying a 25%-25%-50% weight on each dimension obtaining the rank correlation across all three alternative weighting systems no lower than 0.83. As further pointed by Alkire and Santos (2013), the global MPI has been released annually by UNDP since 2010 and has been subjected to numerous robustness tests suggesting that 85 % of pair-wise comparisons were the same if the weights on each dimension varied between 25 % and 50 % of the total. The similar types of analyses were done also by Alkire and Foster (2011a,b); Alkire et al. (2015), Weziak-Bialowolska (2016). Foster et al. (2012) assess robustness of 2004 HDI to weights using the robustness method proposed by Foster et al. (2013) and Permanyer (2011, 2012) and they conclude that for the entire dataset 70% of the pair-wise comparisons are fully robust while about 92% have robustness levels of 25% or higher. A similar analysis was performed by Foster et al. (2013) who examined the empirical prevalence of robust comparisons for three widely used indices: the Human Development Index (HDI), the Index of Economic Freedom (IEF), and the Environmental Performance Index (EPI). The authors concluded that the rank robustness of the HDI was found to be the highest (73% of pair-wise 1998 country rankings were fully robust), while the EPI was the least robust (no more than 6.5% of its pairwise rankings being fully robust).

Robustness assessments based on a limited number of weighting structures is defensible if the considered weighting vectors are generally agreed and/or accepted. If that is not the case, and there is no such agreement, the limited robustness assessment analysis may be insufficient. We contribute to the robustness to weights empirical literature by assessing the robustness of the official multidimensional poverty index used in the European Union. The arbitrarily determined equal dimensional weights were not justified, and thus the question regarding the relative importance of the dimensions remains open. If in such a situation one's aim is to claim unambiguously that poverty in a given country has declined or increased over time (or is higher/lower than in other country), the results have to be fully robust to weights. In our empirical study we find that approximately 50 % of all pair-wise country comparisons are not robust to changes in weights, i.e. it is always possible to find a set of weights for which ranks of countries reverse. Similar results are obtained, when assessing the robustness of poverty comparisons over time. Analysis of the relationship between maximum change in the weights (preserving the initial ranks) and threshold indicates that the highest probability or preserving ranks is attained in the case of the union and intersection approaches. The findings further indicate that rankings of countries are extremely sensitive to the change in weights around the official definition of the composite indicator (i.e. weights of all dimensions and threshold equal to one third). Our results hence suggest that evaluations of the progress made in alleviating multidimensional poverty in the EU are highly sensitive to the set of weights used to quantify poverty, and that more attention needs to be paid to the checks for sensitivity of poverty comparisons to changes in weights.

The rest of our paper is organized as follows: the second section presents the measurement framework summarizing the counting approach to poverty measurement followed by the description of necessary and sufficient condition of comparisons robust to the weights. In the third section the official EU at-risk-of-poverty or social exclusion indicator, as a special case of A-F multidimensional poverty measure, EU-SILC microdata (as the source of data for our analyses) and methods used are described. Finally, in the last section the main results are presented.

2 The Measurement Framework

We consider a population with N individuals and $D > 1$ indicators of wellbeing. Let x_{nd} denote the level of attainment by individual n on dimension d. If $x_{nd} < z_d$, where z_d is a deprivation line for dimension d from a D-dimensional vector of deprivation lines, Z, then we say that individual n is deprived in indicator d.

In order to account for the breadth of deprivations, counting measures rely on individual deprivation scores defined as a weighted count of deprivations. Let $W := (w_1, w_2, ..., w_D)$ denote the vector of dimensional weights such that $w_d \geq 0 \wedge \sum_{d=1}^{D} w_d =$

1. The deprivation score for individual n is given by

$$c_n \equiv \sum_{d=1}^{D} w_d \mathbb{I}(x_{nd} < z_d),$$

where $\mathbb{I}$ is the indicator function that takes value 1 if the argument in parenthesis is true, and 0 otherwise. There is only one vector of possible values of c_n for each particular choice of deprivation lines and weights. Moreover it is easy to show that the maximum number of possible values is given by: $\sum_{i=0}^{D} \binom{D}{i} = 2^D$. The vector of possible values is defined as: $V := (v_1, v_2, ..., v_l)$, where $\max l = 2^D$, $v_i < v_{i+j}$, $v_1 = 0$ and $v_l = 1$.

Following Alkire and Foster (2011a) we characterise the set of multidimensionally poor with an identification rule $\rho_k(c_n)$ that equals 1 when the individual is poor and 0 otherwise. The indicator function ρ_k compares individuals' c_n with a multidimensional cut-off $k \in [0, 1] \subset \mathbb{R}_+$ so that any person n is deemed to be poor if and only if: $c_n \geq k$. As shown in Lasso de la Vega (2010), the function ρ_k is the only identification rule that satisfies the property of poverty consistency which requires $\rho_k(c_{n'}) = 1$ whenever $\rho_k(c_n) = 1$ and $c_n \leq c_{n'}$.

Let $P(C)$ denote a social poverty counting measure depending on the vector of poverty scores, C. We consider a broad family of social poverty measures satisfying standard axioms in the literature on poverty measurement including:

Axiom 1 *Focus (FOC): P should not be affected by changes in the deprivation score of a non-poor person as long as for this person it is always the case that: $c_n < k$.*

Axiom 2 *Monotonicity (MON): P should increase whenever c_n increases and n is poor.*

Axiom 3 *Symmetry (SYM): P should not be affected by permutations in the vector of poverty scores C, i.e., $P(C, \rho_k) = P(C', \rho_k)$ where C' is any permutation of C.*

Axiom 4 *Population-replication invariance (PRI): $P(C, \rho_k) = P(C_R, \rho_k)$ where $C_R = (C, C, ..., C)$ is any replication of the vector of scores C.*

Axiom 5 *Progressive deprivation transfer (PROG): A rank-preserving transfer of a deprivation from a poorer individual to a less poor individual, such that both are deemed poor, should decrease P.*

2.1 Necessary and sufficient conditions with variable weights

Analyses performed in this study are based on a new dominance conditions proposed by Gallegos et al. (2016) to examine the robustness of poverty orderings to the choice of weighting schemes.

The authors derive the new dominance conditions (necessary, sufficient and necessary and sufficient) to assert the robustness of counting poverty orderings within the classes of poverty measures $\mathbb{P}_1$ and $\mathbb{P}_2$. These conditions build on the dominance results derived in Lasso de la Vega (2010). Let P^A and P^B refer to the social poverty indices of populations

A and B, respectively, and let $H^A(k)$ and $H^B(k)$ refer to the multidimensional headcount of each of those populations. The following result sets out the conditions for unambiguous poverty orderings within the class $\mathbb{P}_1$:

Condition 1 *$P^A < P^B$ for all P in $\mathbb{P}_1$ and any identification cut-off, k, if and only if $H^A(k) \leq H^B(k) \quad \forall k \in [0, v_2, ..., 1] \quad \wedge \exists k | H^A(k) < H^B(k)$.*

For the proof see Lasso de la Vega (2010).

Condition (1) states that poverty comparisons of A and B are robust to the choice of the poverty function satisfying FOC, MON, SYM, and PRI only when the ordering of headcount measures is the same *for every relevant value of k*.

The following condition is both necessary and sufficient to guarantee unambiguous poverty orderings within the class of measures $\mathbb{P}_1$ for any possible combination of identification cut-off and dimensional weights.

Condition 2 *Consider the class of poverty measures $\mathbb{P}_1$. The following three statements are equivalent:*

1. *$P^A < P^B$ for all $P \in \mathbb{P}_1$ for any weighting vector, W, and poverty threshold, k.*
2. *For any vector of weights, W, $H^A(k) \leq H^B(k) \quad \forall k \in [0, v_2, ..., 1] \wedge \exists k | H^A(k) < H^B(k)$.*
3. *For all $\gamma_{W,k} \in \Gamma$, $\Pi(W, k)$ in A is no greater than in B, and at least once strictly lower.*

The proof and the further details can be found in Gallegos et al. (2016).

3 Description of the Data and Methods

3.1 Definition of the Aggregated Indicator

The 'people at risk of poverty or social exclusion' multidimensional poverty measure is one of the Europe 2020 headline indicators created by Eurostat in order to monitor progress towards the Europe 2020 strategy targets adopted by the European Council in 2010 (Eurostat, 2015h). The indicator is defined as the sum of persons who are: at-risk-of-poverty and/or severely materially deprived and/or living in households with very low work intensity as a share of the total population, expressed in numbers or shares of the population (European Commission, 2013). The composite indicator has thus three dimensions:

Monetary poverty (at-risk-of-poverty) defined as an equivalised disposable income below 60 % of the national equivalised median income (after social transfers)[1]. The

[1] Prior to 2000, poverty estimates in the EU were based on 50 % of mean equivalised disposable income (Besharov and Couch, 2012), in 1998 the Statistical Programme Committee (CPS, 1998) recommended to use 60 % of median income cut-off as the main reference point.

equivalised income is defined as the household's total disposable income divided by its equivalent size, while the equivalisation is made on the basis of the OECD modified scale (giving a weight of 1.0 to the first person aged 14 or more, a weight of 0.5 to other persons aged 14 or more and a weight of 0.3 to persons aged 0-13). The modified OECD scale adopted by the European Union was proposed in a study by Hagenaars et al. (1994), later published by de Vos and Zaidi (1997), and it is questionable whether a single equivalence scale should be applied to all countries[2].

A person is considered to be *severely materially deprived* if she lives in a household that cannot afford to pay for at least four out of nine items (i. to face unexpected expenses; ii. one week annual holiday away from home; iii. to pay for arrears (mortgage or rent, utility bills or hire purchase installments); iv. a meal with meat, chicken or fish every second day; v. to keep home adequately warm, or could not afford (even if wanted to): vi. a washing machine; vii. a colour TV; viii. a telephone; ix.) a personal car.)[3]

From the perspective of the last dimension, *living in very low work intensity (quasi-jobless) households*, a person is considered to be deprived if she lives in a household, where workinig-age adults (18-59) worked less than 20 % of their total work potential during the past year. When adopting the indicator, in the Social Protection Committee discussion (SPC, 2010) several national delegations expressed a preference for composite indicator based only on two dimensions, excluding the quasi-jobless household indicator, most of them nevertheless accepted it. SPC (2010) admitted that defining jobless households on the basis of the EU-SILC may require further methodological refinements. This can thus be perceived as a political indication, that weight assigned to this sub-indicator should be lower than of the previous two.

From the definitions of the indicators it is obvious that the thresholds have been chosen arbitrarily, and, moreover, application of the same equivalence scale to all countries is questionable. However, both these aspects are beyond the scope of this paper. For a more detailed discussion and description of the evaluation of the set of EU poverty and social exclusion indicators see e.g. Daly (2010).

[2] The authors warned that '...more research efforts should be devoted to the choice of equivalence scales which can be used for cross-country comparisons. One principal issue to be resolved is whether in the cross-country comparisons we should use a single equivalence scale for all the Member States, or whether a single methodology should be applied to estimate equivalence scales which can be different across different countries.' (de Vos and Zaidi, 1997, p. 194)

[3] Before adopting the Europe 2020 strategy headline indicators, the previously agreed EU material deprivation indicator was defined as the share of people who were concerned with at least three out of the nine items. The change in the indicator definition was initiated by the Social Protection Committee (SPC, 2010) without any further justification.

3.2 Data Sources

The calculations in this paper are based on European Union Statistics on Income and Living Conditions (EU-SILC)[4] 2004–2013 microdata (Eurostat, 2009a,b, 2010, 2011, 2015a,b,c,d,e,f). In accordance with the EU methodology (European Commission, 2009), calculations are performed on individual (personal) level using personal cross-sectional weights and the determination of poverty status in each dimension is based on the variable *RX070*[5] (a three-digit value, where 1st digit depicts monetary poverty, 2nd digit captures severe material deprivation and 3rd digit low work intensity, i.e. three binary variables).

The complete sample sizes range from 307,577 observations for 2004 microdata to 588,608 observations for 2013 microdata (and from 8,545 observations for Iceland in 2009 to 61,542 observations for Italy in 2009 at country level). German microdata were not included in the dataset, as the German National Statistical Institute has not given permission to use the German microdata in this research project, and thus German data are excluded from the analyses.

3.3 Statistical Inference

Our testing strategy is based on comparing subdimensional ratios. We set the following null hypothesis: $Ho : z(r) = 0 \ \forall r = 1, 2, ..., R$, against the alternative: $Ha : z(r) < 0 \ \forall r = 1, 2, ..., R$. We reject the null hypothesis in favour of this particular alternative if $\max\{z(1), z(2), ..., z(R)\} < z_\alpha < 0$, where z_α is a left-tail critical value, and α is both the size of a single-comparison test as well as the overall level of significance of the multiple-comparison test. It is not difficult to show that, generally, the overall size of the test will be lower than α. Given the nature of the conditions, if we reject the null in favour of the alternative hypothesis then A dominates B in the sense of being deemed less poor for a broad class of poverty measurement choices (which depends on the condition in question).

For each poverty set we need to perform comparisons with respective z-statistics. In the case of poverty set $\Pi(W, k)$ we use z-statistics of the form:

$$T_{w,k} = \frac{\Pi^A(W,k) - \Pi^B(W,k)}{\sqrt{\frac{\sigma^2_{\Pi^A(W,k)}}{N^A} + \frac{\sigma^2_{\Pi^B(W,k)}}{N^B}}}, \tag{1}$$

where:

$$\sigma^2_{\Pi^A(W,k)} = \Pi^A(W,k)[1 - \Pi^A(W,k)] \tag{2}$$

[4]EU-SILC is an instrument for the systematic production of EU statistics on income and living conditions, encompassing comparable and timely cross-sectional and longitudinal data on income and on the level and composition of poverty and social exclusion at national and European levels (European Union, 2003).

[5]Poverty status for data prior to 2008 had to be estimated in accordance with Eurostat methodology and algorithms (Eurostat, 2015g).

3.4 Maximum deviation from equal weights

For pairs of countries for which dominance cannot be assumed (i.e. countries A and B are identified as a pair of countries for which it is possible to find a set of weights for which country A dominates country B and another set of weights for which country B dominates country A), we perform further analysis. Following the approach proposed by Permanyer (2012) the analysis is aimed at answering the question how far we can deviate from equal weights (i.e. by what value of $\delta \in \left(0, \frac{2}{3}\right]$ weight of each dimension can be increased) while preserving the initial ranks of a pair of countries.

The following algorithm is used:

1. For each pair an initial ranking is detected for equal weights (i.e. $w_1 = w_2 = w_3 = \frac{1}{3}$) at each value of $k, k \in \{\frac{1}{240}, \frac{2}{240}, \ldots, 1\}$.

2. For each value of k the weights are redistributed increasing weight of one of the dimensions by δ and decreasing weights of the two remaining dimensions by $\frac{\delta}{2}, \delta \in \{\frac{1}{120}, \frac{2}{120}, \ldots, \frac{2}{3}\}$, i.e.:

   2.1. $w_{1_i} = \frac{1}{3} + \delta_i, w_{2_i} = w_{3_i} = \frac{1}{3} - \frac{\delta_i}{2}$ for $i = 1, 2, \ldots, 80$,

   2.2. $w_{2_i} = \frac{1}{3} + \delta_i, w_{1_i} = w_{3_i} = \frac{1}{3} - \frac{\delta_i}{2}$ for $i = 1, 2, \ldots, 80$,

   2.3. $w_{3_i} = \frac{1}{3} + \delta_i, w_{1_i} = w_{2_i} = \frac{1}{3} - \frac{\delta_i}{2}$ for $i = 1, 2, \ldots, 80$.

3. For each of the dimensions a value of $\delta_j \in [\frac{1}{120}, \frac{2}{3}], (j = 1, 2, 3)$ is identified as $\min\{\delta_i\}, (i = 1, 2, \ldots, 80)$ for which the ranks differ from the initial ranks for the particular value of k.

4. Finally, the maximum value of δ which preserves the initial rankings for the particular k is identified as $\delta_k = \min\{\delta_j\} - \frac{1}{120}$. If $\forall \delta_i$ for all dimensions the change in rankings is not observed, then $\delta_k = \frac{2}{3}$.

As a result of this analysis, for each pair of country we have the vector of maximum δ's preserving the initial rankings for each value of k. If $\delta = 0 \ \forall k$, even a small change in equal weights results in change in rankings, if $\delta = \frac{2}{3} \ \forall k$, we can assume that one of the countries dominates the other regardless of weights and threshold.

4 Results

Before proceeding to the main results we will replicate the commonly used approach to sensitivity to weights analysis using a very limited number of weighting vectors. Following that approach, i.e. comparing the original rankings (based on the official approach assigning equal weights to each dimension) to alternative rankings (assigning a weight of 0.5 to one dimension and 0.25 to the remaining two) we get changes in ranks as depicted in Fig 1. The red points in Fig. 1 indicate the initial rankings of the countries and the arrows indicate minimum and maximum ranks based on the three weighting schemes described above. In case of around two thirds of observations the maximum differences between the initial rank and the new ranks are between 2 and 4. But e.g. in case of Hungary, different weighting schemes can change the country's position from rank 26 to as low as 14, which is more than 10 ranks difference. Although the variability of changes in alternative ranks is not negligible, the correlation between the original and alternative ranks is high (values of Spearman's ρ between 0.91 and 0.96), which would suggest that changes in weights can influence rank only to a very small extent.

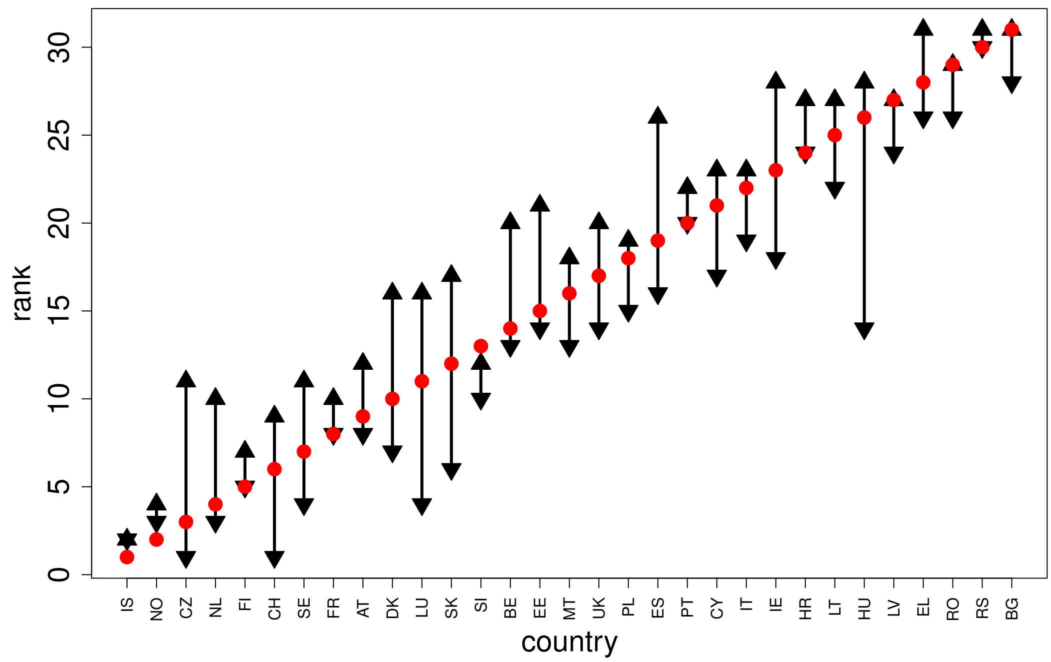

Figure 1: Changes in ranks as a result of alternative weighting scheme

4.1 Dominance in pair-wise comparisons

Applying the approach described in the previous sections we perform all the possible pair-wise comparisons. We are thus able to claim unambiguously which countries dominate other countries and which are dominated by other countries. Cross-country pair-wise comparisons for a particular year can be simply presented by a dominance table (see the example in Tab. 1). This easy-readable table reports all dominance-relationships. The countries in the highlighted row are dominated by the countries above, and at the same time, the countries in the highlighted row dominate the countries below. One can thus unambiguously claim that for instance in 2004, Austria was dominated by Iceland and Luxembourg (i.e. multidimensional poverty was higher in Austria than in Iceland and Luxembourg), and that Austria dominated countries such as Belgium, Estonia, Greece, Spain etc. At the same time the results suggest that one cannot unambiguously claim whether Iceland dominates Luxembourg or vice-versa, and hence for some sets of weights one could conclude that Iceland dominates Luxembourg and for other sets of weights that Luxembourg dominates Iceland.

For each year included in the analysis a table similar to Tab. 1 can be produced, and thus instead of describing 10 such tables (for the period of 2004-2013) the aggregated proportions of dominant pair-wise comparisons are summarized in Tab 2.

According to the results 42 – 58 % of the pair-wise cross-country comparisons can be considered as unambiguous, and thus the dominance conditions are fulfilled. Similar results are obtained if only complete data sets are taken into consideration (i.e. country data that are complete and available for the period of 2008-2013.

Summarizing the partial tables resulting from the complete data sets we are able to identify countries which are mainly 'dominatees' or 'dominators' over the period of 2008 – 2013 (Fig. 2). Taking into account the full data sets during the given period Bulgaria and Ireland were the only two countries that never dominated any other country, while Iceland was the only country which was never dominated by any other country. Tab. 2 thus reports relative positions of countries in terms of dominance – countries in the upper part of the table are more likely to be dominated by other country and less likely to dominate any other country, while the opposite is true for countries in the lower part of the table.

Results in Fig. 2 also provide information on the changes in the relative position of countries in terms of dominance, taking into account both directions (i.e. 'dominator' and 'dominatee' positions). It is e.g. obvious how the relative position of Greece worsened between 2010-2011 in both directions (i.e. number of countries by which Greece was dominated increased significantly, and at the same time the number of countries dominated by Greece decreased to zero after 2010). On the other hand number of countries by which Romania was dominated decreased considerably between 2009-2010, but the number of countries dominated by Romania did not change.

However, the ranking of countries is only relative and one has to be very careful when claiming that poverty level in country A is higher/lower than in country B. Poverty in the EU is measured in relative terms, and thus the poverty levels are not necessarily comparable across the countries.

The proposed approach allows us to assess the poverty patterns across years, enabling us to answer the question whether one can unambiguously claim that poverty decreased or increased over time (regardless of dimensional weights and cutoffs for the given set of poverty measures, and the fixed poverty thresholds for the given dimensions). Tab. 3 gives an example of a country (Poland) dominance table, and it suggests that poverty in the given country decreased over time (e.g. years 2005, 2006 and 2007 were dominated by all the consecutive years). Such an evolution is rather monotonous with a clear trend, while 80.6% of pair-wise year comparisons are unambiguously ranked. However, a high percentage of robust comparisons across years within a country does not necessarily indicate a monotonous trend in poverty. The question of monotonous trend in poverty (in terms of robust cross-year comparisons) can be analysed using directed graphs. We get the paths of poverty trends, and the examples of patterns for selected countries (the longest robust trends available) are reported in Tab. 4. Those results e.g. allow us to claim robustly that multidimensional poverty level in Poland declined between 2005 and 2013, however the comparisons for 2010 and 2011 were unclear. Similarly, the results suggest that multidimensional poverty in Malta and Spain increased over time.

Such comparisons provide valuable information to the policy-makers, as they enable them to claim robustly whether e.g. anti-poverty policies were effective and whether poverty levels (in multidiomensional terms) decreased (regardless of weights and multidiomensional cutoffs).

4.2 Maximum Deviation from Equal Weights

For the group of countries for which the dominance conditions do not hold and re-ranking can appear we search for the maximum deviation from equal weights, denoted as δ. As the δ's are calculated for each value of threshold, the results can be presented as a relationship between the δ and the threshold. Several patterns of relationship between δ and k have been identified, of which four (see Fig. 3) are the most frequent and account for around 94 % of all the identified patterns for cross-country comparisons and around 83 % for cross-year comparisons.

In our interpretations we focus on the patterns of the δ behaviour at cutoff set at 1/3, which corresponds to the official definition of the multidimensional cutoff adopted in the EU.[6] The results suggest that if the multidimensional cutoff is set as 1/3, approximately 21 % of all non-robust pair-wise comparisons are extremely sensitive to a change in the weights, suggesting that even a small deviation from the equal weights results in re-ranking. However, for approximately 20 % of all non-robust pair-wise comparisons the initial rankings were robust to the changes in weights not deviating by more than $\delta = \frac{1}{3}$ from the equal weights.

A similar assessment (i.e. multidimensional cutoff set at 1/3 and using equal weights) was performed for cross-year comparisons at country level and according to the results 12 % of pair-wise comparisons are sensitive to a change in the weights, while, for approximately 14 % of pair-wise comparisons the initial rankings were robust to the changes in weights not deviating by more than $\delta = \frac{1}{3}$ from the equal weights.

5 Concluding Remarks

The paper presents the first preliminary results of assessment of robustness of multidimensional poverty orderings in the European Union. Our empirical application is based on European Union Statistics on Income and Living Conditions (EU-SILC) microdata adopting the counting approach to multidimensional poverty measurement proposed by Alkire and Foster (2011a). Two main analyses are performed in the study. First, we check the robustness of cross-country and cross-year comparisons in the EU for all

[6] More precisely, the multidimensional poverty in the EU is defined as union approach, i.e. a person is deemed poor if she is deprived on at least one of the dimensions. In terms of our measurement framework - the multidimensional cutoff set as 1/3 and using equal weights.

pair-wise comparisons (cross-country as well as cross-year) by testing the necessary and sufficient dominance condition and secondly, we focus on the pair-wise comparisons for which dominance cannot be assumed. We further investigate how far we can deviate from equal weights while preserving the initial ranks in pair-wise comparisons, i.e. we search for the maximum change in weights that preserves the initial ranks.

We find that approximately 50 % of all pair-wise country comparisons are not robust to changes in weights, i.e. it is always possible to find a set of weights for which ranks of countries reverse. Similar results are obtained, when assessing the robustness of poverty comparisons over time. Analysis of the relationship between maximum change in weights (preserving initial ranks) and threshold indicates that the highest probability or preserving ranks is in case of union and intersection approaches. The findings further indicate that rankings of countries are extremely sensitive to change in weights around the official definition of the composite indicator (i.e. weights of all dimensions and threshold equal one third).

Our results hence suggest that evaluations of the progress made in alleviating multi-dimensional poverty in the EU are highly sensitive to the set of weights used to quantify poverty, and that more attention needs to be paid to the checks for sensitivity of poverty comparisons to changes in weights.

References

Alkire, S. and Foster, J. (2011a). Counting and Multidimensional Poverty Measurement. *Journal of Public Economics*, 95(7-8):476–487.

Alkire, S. and Foster, J. (2011b). Understandings and Misunderstandings of Multidimensional Poverty Measurement. *Journal of Economic Inequality*, 9(2):289–314.

Alkire, S., Foster, J. E., Seth, S., Santos, M. E., Roche, J. M., and Ballon, P. (2015). Robustness analysis and statistical inference. In *Multidimensional Poverty Measurement and Analysis*, pages 233–255. Oxford University Press, Oxford.

Alkire, S. and Santos, M. E. (2013). A Multidimensional Approach: Poverty Measurement and Beyond. *Social Indicators Research*, 112(2):239–257.

Alkire, S. and Santos, M. E. (2014). Measuring Acute Poverty in the Developing World: Robustness and Scope of the Multidimensional Poverty Index. *World Development*, 59:251–274.

Athanassoglou, S. (2015). Multidimensional Welfare Rankings under Weight Imprecision: A Social Choice Perspective. *Social Choice and Welfare*, 44(4):719–744.

Atkinson, A. B. (2003). Multidimensional deprivation: Contrasting social welfare and counting approaches. *Journal of Economic Inequality*, 1(1):51–65.

Azevedo, V. and Robles, M. (2013). Multidimensional Targeting: Identifying Beneficiaries of Conditional Cash Transfer Programs. *Social Indicators Research*, 112(2):447–475.

Besharov, D. J. and Couch, K. A. (2012). Introduction. In Besharov, D. J. and Couch, K. A., editors, *Counting the Poor New Thinking About European Poverty Measures and Lessons for the United States*, pages 3–23. Oxford University Press, Oxford, UK.

Bourguignon, F. and Chakravarty, S. R. (2003). The measurement of multidimensional poverty. *Journal of Economic Inequality*, 1(1):25–49.

Cherchye, L., Moesen, W., Rogge, N., Puyenbroeck, T. V., Saisana, M., Liska, A. S. R., and Tarantola, S. (2008). Creating Composite Indicators with DEA and Robustness Analysis: The Case of the Technology Achievement Index. *Journal of the Operational Research Society*, 59(2):239–251.

Chowdhury, S. and Squire, L. (2006). Setting Weights for Aggregate Indices: An Application to the Commitment to Development Index and Human Development Index. *The Journal of Development Studies*, 42(5):761–771.

Cox, D. R., Fitzpatrick, R., Fletcher, A. E., Gore, S. M., Spiegelhalter, D. J., and Jones, D. R. (1992). Quality-of-Life Assessment: Can We Keep It Simple? *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 155(3):353–393.

CPS (1998). Recommendations on social exclusion and poverty statistics cps 98/31/2. 31st Meeting of the Statistical Programme Committee, Luxembourg, 26-27 November 1998.

Daly, M. (2010). Assessing the EU Approach to Combating Poverty and Social Exclusion in the Last Decade. In Marlier, E. and Natali, D., editors, *Europe 2020: Towards a More Social EU?* , pages 143–161. Peter Lang International Academic Publishers, Brussels.

de Vos, K. and Zaidi, M. A. (1997). Equivalence Scale Sensitivity of Poverty Statistics for the Member States of the European Community. *Review of Income and Wealth*, 43(3):319–333.

Decancq, K. and Lugo, M. A. (2013). Weights in Multidimensional Indices of Wellbeing: An Overview. *Econometric Reviews*, 32(1):7–34.

Duclos, J.-Y., Sahn, D. E., and Younger, S. D. (2006). Robust Multidimensional Poverty Comparisons. *The Economic Journal*, 116(514):943–968.

European Commission (2009). Portfolio of Indicators for the Monitoring of the European Strategy for Social Protection and Social Inclusion - 2009 update. European Commission, Employment, Social Affairs and Equal Opportunities DG, Social protection and social integration, Social and demography analysis, Brussels.

European Commission (2013). EU social indicators - Europe 2020 poverty and social exclusion target. European Commission, Social Protection Committee, Indicators Sub-group, Brussels.

European Union (2003). Regulation (EC) No 1177/2003 of the European Parliament and of the Council of 16 June 2003 concerning Community statistics on income and living conditions (EU-SILC) . Official Journal of the European Union, L 165, Vol. 46, 3 July 2003, P. 1-9.

Eurostat (2009a). EU-SILC 2004 UDB, cross-sectional data ver. 2004-4 from 01-08-09. European Commission, Eurostat, Luxembourg.

Eurostat (2009b). EU-SILC 2005 UDB, cross-sectional data ver. 2005-5 from 01-08-09. European Commission, Eurostat, Luxembourg.

Eurostat (2010). EU-SILC 2006 UDB, cross-sectional data ver. 2006-4 from 01-03-10. European Commission, Eurostat, Luxembourg.

Eurostat (2011). EU-SILC 2007 UDB, cross-sectional data ver. 2007-6 from 01-08-11. European Commission, Eurostat, Luxembourg.

Eurostat (2015a). EU-SILC 2008 UDB, cross-sectional data ver. 2008-7 from 01-03-15. European Commission, Eurostat, Luxembourg.

Eurostat (2015b). EU-SILC 2009 UDB, cross-sectional data ver. 2009-7 from 01-03-15. European Commission, Eurostat, Luxembourg.

Eurostat (2015c). EU-SILC 2010 UDB, cross-sectional data ver. 2010-6 from 01-03-15. European Commission, Eurostat, Luxembourg.

Eurostat (2015d). EU-SILC 2011 UDB, cross-sectional data ver. 2011-5 from 01-03-15. European Commission, Eurostat, Luxembourg.

Eurostat (2015e). EU-SILC 2012 UDB, cross-sectional data ver. 2012-3 from 01-03-15. European Commission, Eurostat, Luxembourg.

Eurostat (2015f). EU-SILC 2013 UDB, cross-sectional data ver. 2013-2 from 01-08-15. European Commission, Eurostat, Luxembourg.

Eurostat (2015g). EU statistics on income and living conditions (EU-SILC) methodology. `http://ec.europa.eu/eurostat/statistics-explained/index.php/EU_statistics_on_income_and_living_conditions_%28EU-SILC%29_methodology`. (Accessed: 2015-12-29). European Commission, Eurostat, Luxembourg.

Eurostat (2015h). *Smartner, greener, more inclusive? Indicators to support the Europe 2020 strategy.* Publications Office of the European Union, Luxembourg.

Foster, J., Greer, J., and Thorbecke, E. (1984). A Class of Decomposable Poverty Measures. *Econometrica*, 52(3):761–766.

Foster, J. E., McGillivray, M., and Seth, S. (2012). Rank Robustness of Composite Indices: Dominance and Ambiguity. OPHI Working Paper No. 26b. Oxford Poverty and Human Development Initiative (OPHI).

Foster, J. E., McGillivray, M., and Seth, S. (2013). Composite Indices: Rank Robustness, Statistical Association, and Redundancy. *Econometric Reviews*, 32(1):35–56.

Gallegos, J. V., Yalonetzky, G., and Azpitarte, F. (2016). On the Robustness of Multidimensional Counting Poverty Orderings. Melbourne Institute Working Paper Series Working Paper No. 22/15.

Hagenaars, A. J. M., de Vos, K., and Zaidi, M. A. (1994). *Poverty Statistics in the Late 1980s: Research Based on Micro-data.* Office for Official Publications of the European Communities, Luxembourg.

Kemeny, J. (1959). Mathematics without Numbers. *Daedalus*, 88(4):577–591.

Lasso de la Vega, C. (2010). Counting poverty orderings and deprivation curves. In Bishop, J., editor, *Research on Economic Inequality*, pages 153–172. Emerald.

Mehmet Pinar, Thanasis Stengos, N. T. (2013). Measuring Human Development: A Stochastic Dominance Approach. *Journal of Economic Growth*, 18(1):69–108.

Permanyer, I. (2011). Assessing the Robustness of Composite Indices Rankings. *The Review of Income and Wealth*, 57(2):306–326.

Permanyer, I. (2012). Uncertainty and Robustness in Composite Indices Rankings. *Oxford Economic Papers*, 64(1):57–79.

Ravallion, M. (2011). On multidimensional indices of poverty. *Journal of Economic Inequality*, 9(2):235–248.

Saisana, M., Saltelli, A., and Tarantola, S. (2005). Uncertainty and Sensitivity Analysis Techniques as Tools for the Quality Assessment of Composite Indicators. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 168(2):307–323.

Sen, A. (1976). Poverty: An Ordinal Approach to Measurement. *Econometrica*, 44(2):219–231.

SPC (2010). Europe 2020 Strategy - SPC Contribution. Council of the European Union, The Social Protection Committee, Brussels.

Stiglitz, J. E., Sen, A., and Fitoussi, J.-P. (2009). *Report by the Commission on the Measurement of Economic Performance and Social Progress.* Commission on the Measurement of Economic Performance and Social Progress, Paris.

Streeten, P., Burki, S. J., ul Haq, M., Hicks, N., and Stewart, F. (1981). *First Things First: Meeting Basic Human Needs in the Developing Countries.* Oxford University Press.

UNCTAD (1975). The Cocoyoc Declaration. *International Organization*, 29(3):893–901.

UNDP (1997). *Human Development Report 1997.* Oxford University Press, New York and Oxford.

Weziak-Bialowolska, D. (2016). Spatial Variation in EU Poverty with Respect to Health, Education and Living Standards. *Social Indicators Research*, 125(2):451–479.

World Bank (2001). *World Development Report 2000/2001: Attacking Poverty.* Oxford University Press, Oxford.

6 Appendix

Table 1: Dominance table (based on 2004 EU-SILC data)

AT		AT	AT	AT		AT	AT		AT			AT		
	DK		DK		DK	DK	DK	DK		DK				
	FI		FI		FI		FI	FI		FI				
			FR											
IS	IS	IS	IS	IS	IS	IS	IS	IS		IS		IS	IS	IS
LU	LU		LU	LU	LU		LU	LU		LU			LU	
	NO		NO		NO		NO	NO		NO				
	SE		SE				SE	SE		SE				
Countries above dominate the highlighted countries below														
AT	BE	DK	EE	EL	ES	FI	FR	IE	IS	IT	LU	NO	PT	SE
Countries below are dominated by the highlighted countres above														
									AT		AT			
BE		BE				BE			BE		BE	BE		BE
									DK					
EE		EE				EE	EE		EE		EE	EE		EE
EL									EL		EL			
ES		ES				ES			ES		ES	ES		
		FI							FI					
FR		FR				FR			FR		FR	FR		FR
IE		IE				IE			IE		IE	IE		IE
IT		IT				IT			IT		IT	IT		IT
									NO					
PT									PT		PT			
									SE					

Table 2: Proportions of dominant pair-wise comparisons [%]

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
All data sets	52	51	44	47	42	42	49	56	56	59
Complete sets	x	x	x	x	42	42	47	54	55	55

The following countries are included in the complete data sets: Austria (AT), Belgium (BE), Bulgaria (BG), Cyprus (CY), Czech Republic (CZ), Denmark (DK), Estonia (EE), Greece (EL), Spain (ES), Finland (FI), France (FR), Hungary (HU), Switzerland (CH), Ireland (IE), Iceland (IS), Italy (IT), Lithuania (LT), Luxembourg (LU), Latvia (LV), Malta (MT), Netherlands (NL), Norway (NO), Poland (PL), Portugal (PT), Romania (RO), Sweden (SE), Slovenia (SI), Slovakia (SK), United Kingdom (UK).

dominated_by							dominated					
2008	2009	2010	2011	2012	2013		2008	2009	2010	2011	2012	2013
19	14	10	20	21	23	BG	0	0	0	0	0	0
6	3	4	19	20	23	EL	2	3	3	0	0	0
5	13	19	20	18	17	LT	3	2	1	1	2	2
8	16	25	21	19	16	LV	0	0	0	0	1	1
7	9	10	8	11	14	PT	2	4	2	5	5	2
12	10	12	13	15	13	IT	0	0	2	3	2	3
9	7	6	11	11	12	HU	0	0	1	1	0	1
9	8	8	9	12	12	UK	0	0	1	3	2	2
8	9	11	11	10	11	MT	3	1	3	7	7	6
4	5	11	12	12	10	EE	3	4	2	3	4	5
4	5	7	9	10	10	ES	2	1	1	0	0	1
10	13	10	12	14	10	IE	0	0	0	0	0	0
9	9	12	10	12	9	BE	0	0	1	0	0	1
3	2	1	1	3	9	CY	7	7	6	6	6	4
3	1	3	5	2	7	SI	5	10	12	10	12	11
13	10	9	7	6	6	PL	2	3	3	4	5	5
2	3	4	2	3	4	DK	10	5	5	9	9	8
19	15	6	6	7	4	RO	0	0	0	1	0	0
5	4	5	6	3	3	AT	4	10	10	13	14	16
1	2	4	3	3	3	SK	5	6	3	5	5	5
4	4	4	5	3	2	FI	8	4	11	12	13	15
5	4	6	6	4	2	FR	6	4	5	9	13	15
1	1	0	0	0	1	LU	17	13	15	21	15	13
1	1	1	2	2	1	NL	10	9	15	15	15	13
1	1	1	1	0	1	NO	15	16	20	21	23	22
2	1	0	1	0	1	SE	20	17	17	18	17	14
0	0	0	0	1	0	CZ	5	11	11	14	15	15
1	1	0	0	0	0	CH	15	14	16	16	13	20
0	0	0	0	0	0	IS	27	27	23	23	24	24

Figure 2: Frequencies of 'dominated' vs. 'dominated by' relations

Table 3: Example of Dominance table (Poland)

	2006								
	2007	2007							
	2008	2008	2008						
	2009	2009	2009	2009					
	2010	2010	2010						
	2011	2011	2011			2011			
	2012	2012	2012	2012	2012	2012			
	2013	2013	2013	2013		2013		2013	
Years above dominate the highlighted years below									
PL	2005	2006	2007	2008	2009	2010	2011	2012	2013
Years below are dominated by the highlighted years above									
		2005	2005	2005	2005	2005	2005	2005	2005
			2006	2006	2006	2006	2006	2006	2006
				2007	2007	2007	2007	2007	2007
					2008			2008	2008
								2009	
							2010	2010	2010
									2012

Table 4: Example of poverty patterns over time in a sample of countries

PL : 2013 < 2012 < 2009 < 2008 < 2007 < 2006 < 2005
AT : (2004, 2005) < 2006 < 2012 < 2013 < 2010 < 2011
CZ : 2009 < 2010 < 2013 < (2008, 2012) < (2006, 2007) < 2005
MT : 2008 < 2011 < 2010 < 2012 < 2013
ES : 2007 < 2008 < 2009 < 2010 < 2012 < 2013
SK : (2008, 2009) < 2007 < 2006 < 2005

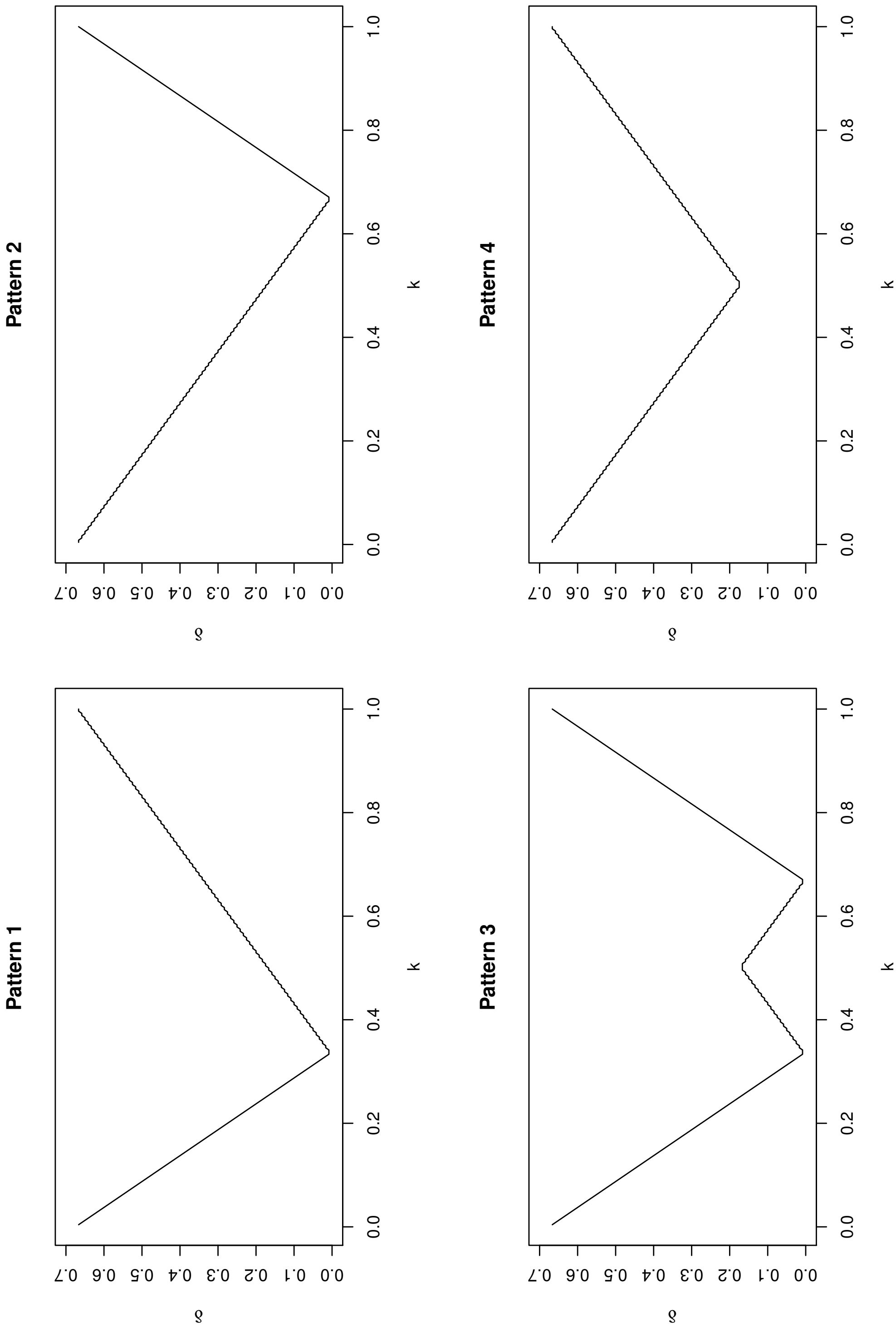

Figure 3: Patterns of δ vs k relationship